# Financing Constraints and Firm Dynamics

by

Andrea Caggese

Submitted to the University of London
in partial fulfillment of the requirements for the degree of

Ph.D. Economics

at the

LONDON SCHOOL OF ECONOMICS

August 2001



Signature of Author ..........................................................
University of London
December 21, 2001

Accepted by ..........................................................

**Financing Constraints and Firm Dynamics**

by

Andrea Caggese

Submitted to the University of London
on December 21, 2001, in partial fulfillment of the
requirements for the degree of
Ph.D. Economics


## Abstract

How important are financing constraints in explaining the cyclical behaviour of investment? How do they affect the investment responses to macroeconomic shocks? This thesis answers these questions by developing a structural model of investment with financing and irreversibility constraints and by analysing its implications both theoretically and empirically.

After briefly reviewing the recent advancements of the investment literature in chapter 1, in chapter 2 we present a preliminary empirical analysis of the links between financial structure and firm dynamics. In chapter 3 we develop a basic structural model which analyses optimal investment and saving choices of entrepreneurs in the presence of uncertainty as well as of financing constraints. We show that future expected financing constraints generate a precautionary saving behaviour which affects the optimal allocation between risky investment and saving.

In chapter 4 we extend the basic model to include both fixed and variable capital as well as financing constraints and irreversibility of fixed capital. We show that the interactions between financing and irreversibility constraints amplify the effects of financing constraints on the cyclical fluctuations of investment and production. This interaction together with the precautionary saving behaviour is essential in explaining a number of stylised facts about investment dynamics: i) aggregate inventory investment is very volatile and procyclical, especially in recessions; ii) it leads the business cycle, while fixed capital investment lags it; iii) fixed and especially inventory investment are sensitive to net worth; iv) output and inventories are more volatile and procyclical for small firms than for large ones.

In chapter 5 we verify empirically the theoretical results derived in chapter 4. We use our panel of balance sheet data on Italian manufacturing firms to test and not reject the financing constraints hypothesis. This hypothesis is also strongly supported by the direct qualitative information about the problems faced by the entrepreneurs in financing new investment projects.

# Contents

# List of Figures

# Introduction

In order to explain the aggregate behaviour of investment and production, it is important to understand the factors that determine the investment decisions at firms level. Most of the neoclassical investment theory studies the determinants of the investment in fixed capital, like plants and equipment. A common feature of these studies is the difficulty in explaining the cyclical movements of aggregate investment: they predict that fixed investment should be driven by the cost of capital and by the expected marginal profitability of capital. However in practice both factors explain only a small fraction of investment, while measures of liquidity and cash flow have a much higher explanatory power.

Some authors argue that financial factors may be important in explaining these phenomena. Financial conditions may affect the ability of firms to invest when financiers are unwilling to fund their profitable investment opportunities. This could happen because once the funds have been handed to the firms, contractual and/or informational problems may prevent the financiers appropriating their share of the revenues from the investment's output. But if firms are unable to raise external financing, they only invest when internally generated funds become available. A literature started by Fazzari, Hubbard and Petersen (1988) shows 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 first motivation of this thesis is the consideration that the majority of these studies focus on the estimation of the reduced form investment equation, without explicitly solving the structural model. This reduced form approach is subject to the Kaplan and Zingales (1997) critique: there is no theoretical support for the claim that the cash flow-investment sensitivity is monotonously increasing in the intensity of financing constraints. This implies that such empirical evidence is not conclusive, because the criteria used in these studies to select firms "more likely to face capital-markets imperfections" are themselves arbitrary. More importantly, the reduced form approach cannot answer the main questions: how important are the financing constraints in explaining the cyclical behaviour of investment? How do they affect the investment responses to macroeconomic shocks?

The second motivation of this thesis is the consideration that the investment literature analyses separately fixed capital investment and inventory investment. This approach makes it difficult to explain a series of stylised facts about investment: i) aggregate inventory investment is very volatile and procyclical; ii) its decline accounts for a large part of the GDP decline in recessions iii) it is contemporarily correlated with sales; iv) it leads the business cycle, while fixed capital investment lags it; v) output and inventories are more volatile and procyclical for small firms than for large ones.

Our thesis answers these questions by developing a structural theory of investment in both fixed capital and variable capital in the presence of financing constraints and irreversibility of fixed capital, and by analysing its implications both theoretically and empirically.

In chapter 3 we develop a simple structural model of investment with financing constraints[1] where the entrepreneur has the opportunity to invest either in a risky technology that requires physical capital or in safe financial assets. Because of an enforceability problem, the entrepreneur[2] can obtain external financing only if she secures it with collateral. The only collateral accepted by the lenders is the physical capital used in the production. Because capital depreciates and only its residual value after production is valuable as collateral, the entrepreneur needs some downpayment to finance her investment. Therefore her borrowing capacity depends on her financial wealth, which is endogenous being a function of past investment, saving and consumption decisions. We determine the conditions under which, because of the uncertainty about productivity, the entrepreneur has a positive probability of facing future financing constraints. When this happens, the entrepreneur keeps some financial assets, or spare borrowing capacity, as a precautionary saving motive. The amount of this precautionary saving is proportional to the intensity of future expected financing constraints.

This precautionary saving is important because it affects the allocation of wealth between investment and saving: the more the entrepreneur is likely to face future financing constraints, the more she reduces the investment in the risky technology to increase her

[1] Other recent works in this direction by Gross (1994) and Cooley, Marimon and Quadrini (2000) are briefly revewed in chapter 1.

[2] We use in this thesis the term "entrepreneur" rather than "firm". However the maximising agent in our models can be interpreted either as an entrepreneural household or as the management of a firm.

holding of safe assets. This implies that financing constraints may be quantitatively important for aggregate investment dynamics even if only a small share of entrepreneurs face binding financing constraints at any point in time. For example a macroeconomic shock which reduces the net worth of all entrepreneurs would also on average increase their precautionary saving and hence reduce their risky investment.

Moreover by engaging in precautionary saving entrepreneurs prevent the financing constraint from binding in most states of nature. We show that this implies that investment can be more sensitive to internal finance in response to future expected financing constraints rather than to current ones. In this sense our theoretical analysis supports the Kaplan and Zingales (1997) critique that the correlation between investment and cash flow is not monotonously increasing in the intensity of financing constraints. However it also provides a new interpretation of their empirical findings. Kaplan and Zingales (1997) find that the investment-cash flow correlation is stronger for firms which are profitable and financially very wealthy than for firms that are less profitable and more levered[3]. Our analysis in chapter 3 shows that this is consistent with the view that cash flow-investment sensitivities are related to future expected financing constraints. We compare two identical entrepreneurs subject to different realisations of their productivity shocks: a productive and profitable entrepreneur who increases her financial wealth and reduces her borrowing and an unproductive loss making one who decreases her financial wealth and increases her borrowing. In both cases investment is sensitive to cash flow and in both cases the sensitivity is determined by the change in future expected financing constraints.

We also simulate a simple aggregate partial equilibrium economy to show that the precautionary saving effect implies large fluctuations of capital in response to macroeconomic shocks. This is especially true for entrepreneurs with smaller businesses, because the heterogeneity of entrepreneurs with respect to their net worth implies that smaller businesses are on average more financially constrained, even though all are ex ante identical regarding their ability to access external finance. This is because small entrepreneurs tend to have smaller net worth relative to output, either because they have younger growing businesses, or because they became small after experiencing recent low productive periods.

[3]Similar evidence is produced by Cleary (1999), who studies a larger sample of 1317 US firms.

The formalisation of the precautionary saving effect and of its consequences for aggregate investment is an original contribution of this thesis[4]. In fact precautionary saving has been extensively studied in consumption literature as an important determinant of the intertemporal consumption allocation, but no study analyses its effects on investment decisions. The importance of this analysis is confirmed by a recent empirical work by Hubbard and Gentry (2000), which finds evidence of a higher amount of precautionary saving of entrepreneurial households with respect to non entrepreneurial ones.

The model developed in chapter 3 is the basis for the analysis in chapter 4, which is the main theoretical contribution of this thesis. In chapter 4 we extend the basic model by considering an environment in which entrepreneurs have access to a multifactor technology with both fixed and variable capital and cannot sell fixed capital without liquidating their whole business. This extended model allows us to examine how the interaction between fixed capital investment, variable investment and saving of financial assets leads to rich dynamics when fixed capital investment is irreversible and the firm is subject to financing constraints. In particular, we show that the effects of financing imperfections on investment are amplified by the presence of irreversibility of fixed capital.

The intuition is that, when fixed capital is irreversible, the entrepreneur knows that in the case of an economic downturn it will take some time to reduce the stock of fixed capital to the new optimal level. During the adjustment period expected profits drop and the expected rate of wealth accumulation drops as well relatively to a situation where fixed capital is reversible. As a result the entrepreneur has a higher chance of facing future financing constraints: i) if the contraction period is long enough, then the drop in wealth can be so severe that she does not have enough funds to invest in variable capital; ii) if the contraction period ends, she will have not enough resources to invest in both fixed and variable capital and will be financially constrained for some time. In order to compensate these higher costs of future expected financing constraints, the entrepreneur engages in precautionary saving ex ante, more than she would have done with reversible fixed capital.

We explicitly solve and obtain the policy functions of the dynamic investment prob-

[4] A similar effect is studied by Gross (1994), even though in a different theoretical setting. However he mainly focuses on the consequences for firms financial policy rather than for aggregate investment.

lem and derive the consequences of financing an irreversibility constraints for aggregate investment dynamics. We simulate an artificial economy with heterogeneous entrepreneurs and with both idiosyncratic and aggregate uncertainty[5]. The amplification effect described before implies not only that the precautionary saving effect is stronger, but also that during downturns it mainly affects variable capital, fixed capital being irreversible. Because we assume a time lag for the investment to produce output, changes in the level of variable capital in the model can be interpreted as investment in input inventories, such as raw materials and work in progress. Thus our model explains the high volatility and procyclicality of inventories: "*Changes in business inventories, which constitute but a small fraction of total GDP, account for one-fourth of the cyclical movements in GDP*" in the US. (Stock and Watson (1998)). Ramey (1989), Blinder and Maccini (1991) and Ramey and West (1999) show that this is especially true during recessions, when the drop in inventory investment accounts for a large part of the GDP decline. They also provide evidence in support of our approach to model inventories as a production factor, because they emphasise that input inventories[6] are quantitatively more important and more volatile than finished goods inventories.

Our model can also explain why inventories are contemporaneously correlated with sales (Ramey and West (1999)) and very sensitive to financial conditions: US data show that "*inventory investment for small firms absorb from 15% to 40% of cash flow fluctuations*" (Carpenter, Fazzari and Petersen, 1998).

In our artificial economy aggregate uncertainty is given by a combination of recurrent transitory and persistent aggregate shocks that generate an exogenous stochastic business cycle. This means that the cross-sectional distribution of net worth and fixed capital among entrepreneurs is determined by both idiosyncratic and aggregate uncertainty and it affects the way aggregate output and investment react to aggregate shocks.

In particular the heterogeneity of entrepreneurs with respect to fixed capital together

[5] This methodology is analogous to the one adopted by some authors in the nonconvex adjustment cost literature (see for example Bertola and Caballero (1994) and Caballero, Hengel and Haltiwanger (1995), reviewed in chapter 1), which in the past decade has been very successful in providing a microfoundation of the observed nonlinearities in fixed investment dynamics. In this respect our model can be considered an extension to this literature, since we show that the interactions between the two problems are important in explaining the stylised facts about not only fixed capital but also variable capital investment dynamics.

[6] Raw materials and work in progress.

with the fact that precautionary saving effect affects variable capital implies that aggregate fixed investment has a lagged reaction to persistent shocks, which affect variable capital first. This helps to explain why inventory investment leads the business cycle, while fixed investment lags it (Stock and Watson, 1998). Moreover the heterogeneity of entrepreneurs with respect to their net worth, which implies that smaller businesses are on average more financially constrained, explains why small firms are more procyclical than large ones in inventories, output, and short term debt (Bernanke, Gertler and Gilchrist, 1996).

Chapter 5 examines empirically the implications of the model analysed in chapter 4, by developing a new procedure to detect the presence of financing constraints at firm level. The motivation for this empirical analysis is twofold: first, we verify that the data do not reject the financing constraints hypothesis, which is essential for the aggregate results derived in chapter 4. Second, we show that our method is more efficient than the cash flow-investment correlation in detecting financing constraints at firm level.

The model developed in chapter 4 implies that, because variable capital investment is reversible, the "premium" of expected marginal productivity over user cost of variable capital reflects the tightness of current and future expected financing constraints. We call this premium the "excess" expected marginal productivity of variable capital. More specifically the financing constraints hypothesis implies that this premium is monotonously decreasing in the financial wealth of the entrepreneurs, conditional on fixed capital stock and productivity shock. We test this hypothesis by estimating empirical measures of the productivity shocks and of the expected marginal productivity of variable capital.

This test has two important properties: i) it is based on an indicator which, according to our structural theory, is monotonously increasing in the intensity financing constraints. Therefore it is robust to the Kaplan and Zingales critique mentioned before; ii) it 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 convex adjustment costs and the misspecification of the stochastic process for the productivity shock. This is because we argue that 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".

These properties make our test more efficient than the tests based on the cash flow-investment correlation. In fact, as we argued before, both the Kaplan and Zingales critique and our analysis in chapter 3 argue that there is no theoretical foundation for the monotonous relationship between the cash flow-investment correlation and the intensity of financing constraints. Moreover the cash flow-investment correlation is likely to be biased upwards when future expected profitability is not properly estimated. This is because cash flow is very sensitive to current profitability, which in turn is correlated to future expected profitability. Therefore the positive cash flow-investment correlation could simply be caused by the fact that the former absorbs the positive effect of the unobserved future expected profitability on the latter. This empirical problem can explain why many authors (including Fazzari, Hubbard and Petersen (1988) and Kaplan and Zingales (1997)) estimate positive cash flow coefficients for very large firms. Such firms have direct access to equity and bond markets, and it is very difficult to argue that they can be financially constrained, in the sense of being unable to obtain external financing for profitable investment opportunities.

We use the balance sheet information on a panel data of 561 Italian manufacturing firms for 11 years to estimate the excess expected marginal productivity of variable capital for each firm-year observation. A unique feature of our dataset 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. This information is a direct proxy for financing constraints, unique in the empirical investment literature. We use it to show that the value of our indicator of financing constraints, the excess expected marginal productivity of variable capital, is strongly positively correlated with the likelihood that the entrepreneurs state financing problems. We then use the estimated indicator to verify the prediction of the model, and we show that the test does not reject the financing constraints hypothesis for all the firms in the sample but the larger ones.

The thesis is organised as follows: chapter 1 briefly reviews the recent advancements in the investment literature. Chapter 2 proposes a preliminary empirical analysis of our dataset of Italian firms, using both qualitative and quantitative information. Chapter 3 illustrates and solves the basic theoretical framework with financing imperfections and discusses the implications of current and future expected financing constraints for firm level and aggregate investment. Chapter 4 illustrates and solves the extended model with both fixed and variable capital and both financing and irreversibility constraints. We discuss the qualitative features of the solution and calibrate an artificial economy with many entrepreneurs to study the effects of financing and irreversibility constraints on the cyclical fluctuations of output and investment. Chapter 5 illustrates the empirical test of our theory. We combine qualitative and quantitative information to test and not reject the financing constraints hypothesis.

# Acknowledgements

I have been working on the ideas developed in this thesis since 1998. However my previous work at the University of Rome "Tor Vergata" with Michele Bagella and Leonardo Becchetti between 1994 and 1996 has been on related themes and it has been useful in helping me to focus on the relevant empirical issues analysed in this thesis.

Since October 1997 I worked on my thesis at the Financial Markets Group at LSE and I greatly benefited of the "material" support from the Group as well as of the academic interaction with the other members. Among the faculty in the FMG I especially wish to thank David Webb, Bob Nobay and Margareth Bray and all my colleagues, both in the FMG and in the LSE economics Department, that helped to create a nice and relaxed working atmosphere. In particular Alex Muermann, Vicente Cunat, Andrew Ellul, Hesky Bar Isaac, Matteo Iacoviello, Michela Cella, Michele Arslan, Paolo Ramezzana, Cecilia Testa and Valentino Larcinese. I wish also to thank my former flatmate and colleague Lorenzo Coppi, for all the discussions about dynamic stochastic models that animated many London winter nights.

For the empirical work developed in this thesis I used the dataset produced by Mediocredito Centrale. In particular Chapter 2 of the thesis is based on a research project for Mediocredito Centrale, whose financial support is gratefully acknowledged. I also wish to thank the other two members of the project, Sandra Bulli and Marcello Messori. I am also grateful to the faculty members in the LSE economics department and the University of London for the useful comments I received during the completion of my work, in particular from Luca Deidda, Francois Ortalo Magne' and Steve Pischke. I am especially grateful to my supervisor Nobu Kiyotaki. His comments and discussions have been very useful in all the stages of this project, and his supervision very valuable in guiding me in developing my ideas in the theoretical section of the thesis. The usual disclaimer applies.

A special thanks if for my mother Gabriella and my father Vittorio, who have been helping me first in my decision to pursue a research degree in economics and then in all the difficult phases in the last four years.

This thesis is dedicated to my wife Eva, who has always been supporting me during this journey.

# Chapter 1

# A selected review of investment theory

## 1.1 Introduction

Comprehensive reviews of the recent advancements in investment theory are already present in the literature. In particular, Caballero (1997) focuses mainly on the recent contributions concerning the investment theory with nonconvex adjustment costs, and Hubbard (1998) considers the literature about capital market imperfections and investment. Therefore this chapter simply provides a brief overview of these two literatures, emphasising their achievements as well as their limits, in order to motivate the empirical and theoretical analysis proposed in this thesis.

As Caballero (1997) points out, both literatures are based on the consideration that "*the quintessential problem of investment is that it is almost always sunk, possibly along many dimensions. That is, the number of possible uses of resources is reduced dramatically once they have been committed or tailored to a specific project or use [...]. To invest often means opening a vulnerable flank. Funds which were ex-ante protected against certain realisation of firm or industry specific shocks, for example, are no longer so*". The literature about non convex adjustment costs considers physical barriers to investment, while the literature about capital markets imperfections considers problems which limit the willingness of financiers to fund profitable investment opportunities: once the funds have been handed to the entrepreneurs, contractual and/or informational problems may prevent the financiers to appropriate their share of the revenues from the investment's output.

Since both approaches focus mainly on fixed investment, in the last section of this chapter we briefly review the empirical literature about inventory investment, especially focusing on those stylised facts on inventories dynamics which are important in understanding the business cycle.

The outline of this chapter is as follows: in section 1.2 we introduce the neoclassical investment theory. In section 1.3 we briefly revise the main results of the recent literature about investment with nonconvex adjustment costs. In section 1.4 we critically revise the literature about capital markets imperfections. In section 1.5 we briefly review the literature about inventory investment. Section 1.6 presents some conclusions.

## 1.2 Neoclassical investment theory

Jorgenson (1963) is the first to derive the investment function as the first order condition of the optimisation problem of a firm. Assuming perfect competition, no adjustment costs, and a constant return to scale Cobb-Douglas technology, he shows that the optimal capital is equal to the ratio between the capital share of output and the cost of capital.

The idea that not only current but also future expected profitability should affect current investment is put forward by Tobin (1969), who argues that investment in physical capital $K$ should be a function of, among other things, its available rate of return $r_k$. Such a rate of return is equal to:

$$r_k = \frac{R}{p_k} \tag{1.1}$$

where $R$ is the "*marginal efficiency of capital relative to reproduction cost*" and $p_k$ is the current market price of capital goods. By definition $R$ includes the net present value of all future profits generated by a unit of capital installed today. Such relation implies the so called $Q - theory$ : investment is a function of average $Q$, the ratio between the market value and the replacement cost of capital. Abel (1979) and Hayashi (1982) show that the Jorgenson model with convex adjustment costs is equal to a marginal $q - model$. The problem is that marginal $q$, the ratio between marginal productivity and replacement cost of capital, is not directly observable, and the conditions under which the marginal $q$

and average $Q$ are equal (Hayashi, 1982) are very restrictive. In any case the empirical success of the estimated $Q-$ based models is modest (Von Furstenberg, 1977; Summers, 1981; Blanchard and Wyplosz, 1981). Abel and Blanchard (1986) take the alternative approach to directly estimate marginal $q$ and show that, as is the case for average $Q$, also the estimated marginal $q$ explains only a small fraction of the fluctuations in investment[1].

## 1.3 Investment with non convex adjustment costs

The quadratic adjustment cost model implies decreasing returns in the adjustment technology. This implies that it is optimal, at firm level, to increase the stock of capital smoothly and gradually towards the desired level. Such smooth investment at firm level implies smooth aggregate investment. In reality adjustment costs are more complicated than the quadratic cost assumption. In fact it is reasonable to assume that new projects that involve investment in plant and equipment also include relevant fixed costs and a certain degree of irreversibility. The presence of fixed costs implies increasing returns in the adjustment technology, and this is sufficient to generate "lumpiness" in investment, which implies an advantage in bunching rather than in smoothing new investment over time. Caballero (1997) provides a simple version of the investment model with fixed adjustment costs: when the only uncertainty in the investment decision is a productivity shock that follows a continuous stochastic process, the desired stock of fixed capital $k_t^*$ follows a stochastic process with similar characteristics. The actual level of capital $k_t$ instead remains constant and changes only when the absolute difference $|k_t^* - k_t|$ is large enough. That is when the fixed cost in adjusting capital from $k_t$ to $k_t^*$ is smaller than the loss of value in keeping $k_t$ constant. When such adjustment takes place, it is optimal to invest up to the optimal level $k_t^*$. Intuitively the bigger the fixed cost, the larger the difference $|k_t^* - k_t|$ before the entrepreneur finds optimal to invest or disinvest. In this case we should observe more lumpy investment at firm level.

Models with fixed costs have been successful in explaining why we observe lumpy investment at plant level. Among the empirical evidence, Doms and Dunne (1998) analyse

[1] Caballero (1997) provides a review of the recent literature that shows how cost of capital and $q$ are more relevant in determining investment in two cases: i) long-run relationship; ii) in presence of tax adjustments that determine large swings in the cost of capital.

the investment patterns of over 13,700 plants in the U.S. manufacturing sector, drawn from over 300 four digit industries, for the period 1972-1988. They find that over half the plants in their sample experience a 1 year capital adjustment of at least 37%, and that smaller plants and plants that changed ownership have lumpier investment patterns. They also show that the simulated investment models that best fit with the observed capital adjustment patterns are those in which plants mainly invest only when the difference between the desired and the actual capital stocks is large. More importantly they show that such lumpiness at plant level does not wash out at aggregate level: 25% of expenditures on new equipment and structures goes into plants that are increasing their real capital stock by more than 30%, and that makes up only 8% of the sample.

This evidence suggests that microeconomic lumpiness has important consequences for aggregate investment dynamics. Caballero, Engel and Haltiwanger (1995) use information on a similar dataset of approximately 7000 U.S. manufacturing plants for the same period and estimate the hazard function, that is the probability to adjust the capital level conditional on the imbalance between desired and actual capital. They find that the hazard function has a value of approximately zero in presence of zero or positive imbalance (actual capital greater than desired capital), while it is increasing for negative imbalances. This means that plants do not disinvest when actual capital is too high, suggesting the presence of irreversibility, while they are more likely to adjust when actual capital is too low.

On the theoretical side Caballero and Engel (1999) derive the aggregate investment function in presence of non convex adjustment costs, to show that in general aggregate investment depends on the cross sectional density of plants capital imbalances. The dynamic implications of this result are particularly important: in fact a sequence of positive aggregate shocks not only causes some plants to invest, but also changes the cross sectional distribution of imbalances, meaning that more and more plants are likely to adjust at each subsequent aggregate shock. Caballero, Hengel and Haltiwanger (1995) show that such mechanism implies much higher investment volatility in presence of large macroeconomic shocks.

Another source of nonconvexity in adjustment costs is the irreversibility of fixed capital, which can be seen as an infinite fixed cost of disinvestment. Why is fixed investment irreversible? Dixit and Pindyck (1994) argue that investment expenditures are sunk costs when they are firm specific. Also industry specific investment can be irreversible if the industry is sufficiently competitive. In this case, even though in principle a firm could sell its plants to another firm in the industry, this would not happen because all firms in the industry would want to disinvest at the same time as a consequence of a negative aggregate shock. Other sources of irreversibility are informational problems: potential buyers may be unwilling to buy from a firm that wants to sell its own fixed capital because they cannot observe its quality.

Irreversibility implies that investment is "cautious": the rate of return required to convince firms to invest is higher than when fixed investment is reversible. The difference is a premium that compensates the expected cost of being unable to reduce the capital in the future conditional on a negative productivity shock[2]. Bertola and Caballero (1994) build and calibrate a model of firm fixed investment with irreversibility and uncertainty, and show that irreversibility is useful in explaining why fixed investment is much smoother at aggregate level than at firm level.

## 1.4 Investment with capital markets imperfections

All the papers reviewed in the previous section assume the absence of any capital market imperfection. In this case the Modigliani and Miller (1958) theorem states that, in absence of tax distortions, financial structure and financial policy are irrelevant for real investment decisions. The choice of financing is irrelevant because internal finance, debt and equity are equivalent sources of funds with identical opportunity costs. Many authors have disagreed with the view that financial structure is irrelevant, and have argued that financial factors may have an impact on real business cycles (Fisher, 1933; Gurley and Shaw, 1955 and 1960). Yet models of optimal firm investment have been maintaining the perfect market hypothesis for some time. Since the 70s many papers have shown that under asymmetric information or contract incompleteness (imperfect enforceability) adverse selection and moral hazard problems limit the availability of debt (Stiglitz and Weiss, 1981; Besanko and

[2] A formal illustation of this effect is provided in chapter 4.

Thakor, 1986; Milde and Riley, 1988; Hart and Moore, 1998; Albuquerque and Hopenhayn, 2000). Adverse selection can also increase equity financing costs (Myers and Majluf, 1984).

Such models imply either that external finance is not available, or that there is a cost differential between internal and external financing sources, which may itself depend on the financial wealth of the borrower. In both cases the implication is that investment is inefficiently low when internal finance is not available. An avenue of empirical research, henceforth referred to as the "*investment-cash flow literature*", focuses on the sensitivity of firm investment to internal finance availability as a way of testing for the effect of financing imperfections[3]. The idea is that, if firms face high cost, or rationing, of external funds due to capital markets imperfections, then investment should be more sensitive to internal funds than neoclassical models predict.

As Hubbard (1998) points out in his review: "*The principal findings of these studies are that: (1) all else being equal, investment is significantly correlated with proxies for changes in net worth or internal funds; and (2) that correlation is most important for firms likely to face information-related capital-market imperfections*".

In order to derive these results, most of this literature is based on a joint test of two distinct hypotheses (henceforth H1 and H2): H1) Some observable characteristics of firms (size, age, affiliation to group, etc.) are related to how likely they are to be financially constrained. H2) Financially constrained firms, selected according to these characteristics, can invest optimally only when internal finance is available. As a consequence they reject the neoclassical model because of excess sensitivity of investment to cash flow.

All these papers adopt a quadratic adjustment cost assumption in order to derive a testable investment equation. As we mentioned in the previous section, this assumption implies that future expected productivity affects current investment choices. Therefore it becomes of crucial importance to find a robust method of estimating future investment opportunities. Otherwise, since current productivity is correlated to future productivity, and since cash flow is correlated to current productivity, we would be unable to distinguish

[3]Fazzari, Hubbard and Petersen (1988), Devereux and Schiantarelli (1989), Hoshi, Kashyap and Sharfstein (1992), Withed (1992), Bond and Meghir (1994), Hubbard, Kashyap and Withed (1995), Gilchrist and Himmelberg (1995) and (1998), Hu and Schiantarelli (1998), Bagella, Becchetti and Caggese, (2001), and many others.

H1 and H2 from the following alternative joint hypotheses: H1a) *a priori* information is not related to financing constraints; H2a) cash flow just captures the significance of an omitted variable, the unobservable future investment opportunities.

Since the quadratic adjustment cost model implies that $marginal - q$ is a sufficient statistic for the effect of future expected productivity on current investment, many authors assume that the Hayashi (1982) conditions are satisfied and estimate an investment demand function where $marginal - q$ is proxied by $average - Q$ (Fazzari, Hubbard and Petersen, 1988; Devereaux and Schiantarelli, 1989; Hoshi, Kashyap and Sharfstein, 1992). The problem with this approach is that the Hayashi (1982) conditions are very restrictive[4]. Moreover, there are also substantial problems in measuring $average - Q$, both regarding the numerator (market value of firms) and the denominator (replacement value of capital). An alternative approach is to follow Abel and Blanchard (1986) and use a vector autoregression (VAR) forecasting framework that directly estimates future profitability (Gilchrist and Himmelberg, 1995 and 1998).

Another strategy is to solve the first order condition of the quadratic adjustment cost model backwards instead of forward, and to estimate an Euler equation that does not include $marginal - q$ among the regressors, because expectations are valued at realised values (Withed, 1992; Bond and Meghir, 1994; Hubbard, Kashyap and Withed, 1995; Hu and Schiantarelli (1998); Bagella, Becchetti and Caggese, 2001). Among the results provided by this approach, Hubbard Kashyap and Withed (1995) show that the Euler equation based on an alternative model with exogenously imposed financing imperfections[5] is not rejected for firms *a priori* selected as financially constrained. Bagella, Becchetti and Caggese (2001) use direct qualitative information regarding the financing problems faced by the entrepreneurs to identify financially constrained firms, and show that the neoclassical model is not rejected for non financially constrained firms, while it is rejected for financially constrained ones because of an excess investment sensitivity to

[4]They are: perfect competitions in the factor and product market, homogeneity of fixed capital, linear homogeneity of technologies for production and adjustment costs.

[5]"Exogenously imposed" means that the authors assume a certain functional form that links the cost of external finance to observables, rather than providing the explicit microfoundation of the financing problems, as it is done in Albuquerque and Hopenhayn (2000) or in the models illustrated in chapters 3 and 4 of this thesis.

cash flow. These encouraging results are however limited because of the low reliability of the Euler equation tests. As Caballero (1997) notes: *"...unlike the case in basic finance and consumption applications, these* [Euler equation] *procedures are a form of morphine rather than a remedy: their lack of statistical power allows us to sometimes not see the problem".*

In addition to these methodological problems, two additional criticisms more seriously question the results of the cash flow-investment literature: i) the empirical success of the investment models with non convex adjustment costs implies that the quadratic adjustment cost assumption is most likely not correct. Therefore all the cash flow-investment models, which adopt this assumption to derive a testable investment equation, are misspecified regarding the way $marginal - q$ enters in the investment equation. This misspecification is likely to be important. In fact non convex adjustment costs imply that the monotonous relationship between investment and $marginal - q$ no longer holds (Caballero and Leahy, 1996) and even for the subclass of model for which it holds the relation becomes nonlinear (Abel and Eberly, 1994).

ii) Kaplan and Zingales (1997 and 2000) argue that there is no theoretical foundation for the claim that the correlation between investment and cash flow is monotonously increasing in the intensity of financing constraints, and that there is no empirical evidence that such correlation is more intense for financially constrained firms. The latter statement is based on the in-depth analysis of the 49 firms classified as financially constrained by Fazzari, Hubbard and Petersen (1988). Kaplan and Zingales find that the investment-cash flow correlation is stronger for firms which are financially very wealthy and surely not financially constrained[6]. Similar evidence is produced by Cleary (1999), who studies a larger sample of 1317 US firms with complete financial information available for the

[6] Among the points raised by Fazzari, Hubbard and Petersen (2000) in their reply there is the objection that the criteria chosen by Kaplan and Zingales (1997) are such that financially constrained firms are distressed firms, which may be *"restricted by creditors from using internal funds for investment"*. This reflects what we think is a weakness of the Kaplan and Zingales analysis: in their classification scheme they implicitly assume a monotonous relation between the financial wealth of firms and the intensity of their financing constraints. We instead think that the intensity of financing constraints is rather monotonic in the shadow value of money for the firm. The higher such value, the higher the loss for not being able to invest in profitable projects, the higher the intensity of the constraints on external financing. Therefore such intensity depends on the available wealth but also on the investment opportunities, while Kaplan and Zingales consider only the former.

1987-94 period.

Even rejecting all these criticisms and accepting that investment-cash flow sensitivities are useful in detecting financing constraints, it is finally important to note that the cash flow-investment literature does not answer the most important question, which instead receives maximum attention in the non-convex adjustment cost literature: what are the implications of these constraints for aggregate investment dynamics? Following the tradition of Fisher (1933) and Gurley and Shaw, (1955) and (1960), a recent theoretical literature has emphasised three channels which can, in the presence of financing imperfections, amplify and propagate the effects of initial real and monetary shocks (Bernanke and Gertler, 1989 and 1990; Greenwald and Stiglitz, 1993; Kiyotaki and Moore, 1997; Bernanke, Gertler and Gilchrist, 1998): i) *the financial accelerator effect*: constrained firms can only invest if internal finance is available. Hence at the beginning of a downturn the reduction in profits depresses investment; ii) *the asset price effect*: when the borrowing capacity of a firm depends on the collateral value of its assets, at the beginning of a downturn the drop in asset prices reduces borrowing and investment ; iii) *the flight to quality effect*: during a downturn banks increase collateral requirements, thereby reducing loans to borrowers facing financing constraints. All three effects have opposite direction during an upturn.

Recent empirical work provides some evidence supporting this view. As in the cash flow-investment literature, the first step is the identification of firms that are more likely to be constrained in the access of external finance. The second step is the inspection of the behaviour of constrained versus unconstrained firms across different phases of the business cycle regarding sales, debt and inventories. Kashap, Stein and Wilcox (1993), Gertler and Gilchrist (1994) and Oliner and Rudebusch (1996) compare the behaviour of small versus large manufacturing firms after Romer dates, that represent episodes of tight monetary policy that led to a recession. Size is used as a proxy of financing problems. Bernanke Gertler and Gilchrist (1996) conduct a similar study. They inspect quarterly data disaggregated at firm level, and are able to control for industry effects and to use bank dependence as an alternative criterion to identify financially constrained firms. They observe that after a monetary policy tightening short term debt increases for large firms, which increase the supply of commercial papers, while it decreases for small firms. Moreover during the downturn that follows such monetary action sales drop earlier for small firms, which also substantially decrease inventories, while large firms maintain inventories at a higher level. As a result sales, short term debt, inventories and the inventory/sales ratio are more procyclical for small than for large firms. This behaviour is observed also during business cycle fluctuations not directly related to monetary actions. Bernanke Gertler and Gilchrist's (1996) computations show that one third of aggregate fluctuations can be accounted for by the difference between small and large firms.

The limit of these empirical studies on aggregate data is that they lack of a proper theoretical microfoundation. In fact the key assumptions that certain categories of firms (smaller firms and bank dependent firms) are more financially constrained, and that as a consequence they should be more procyclical, are imposed rather than derived from a structural theory of individual firm behaviour. As we argue in the introduction of the thesis the investment-cash flow literature, because of its reduced form approach, is not able to provide such microfoundations. The basic model in chapter 3 of this thesis is a first step in this direction, as are some recent works by Gross (1994), Albuquerque and Hopenhayn (2000) and Cooley, Marimon and Quadrini (2000). Gross (1994) considers a dynamic investment problem similar[7] to our basic problem in chapter 3, with one reversible factor of production and only debt financing available. He obtains the explicit solution of the problem using a numerical method and discusses the implications of financing problems for the investment and financing decisions of the firm. Albuquerque and Hopenhayn (2000) consider an investment model with imperfect enforceability where the bank can precommit to a long term lending contract with the entrepreneur. This allows for long term entrepreneur - bank relationships where the former is financially constrained in the initial phases of her business. Later she accumulates financial wealth, reduces the borrowing from the bank and becomes unconstrained. The Albuquerque-Hopenhayn framework is used by Cooley, Marimon and Quadrini (2000) to derive the aggregate consequences of these financing constraints. This framework is similar to our basic framework with financing constraints only. Indeed our financing constraints assumption is a special case of the

[7] See footnote n.4 in the introduction.

Albuquerque-Hopenhayn contract when the bank cannot precommit. Their approach has the advantage of endogenising the differences in the financial contracts between small and large firms, and this has interesting implications for the growth dynamic of firms. Our approach focuses more on business cycle dynamics rather than on growth dynamics, and in our opinion is better suited to explain the effects of financing constraints on the cyclical behaviour of firms, especially thanks to the precautionary saving effect derived in our model.

## 1.5 Inventory investment

All the literature cited so far, with only a few exceptions, focuses exclusively on fixed investment. This is quantitatively much more important than inventory investment, that in developed countries averages for about only 0.5% of the GDP (Ramey and West (1997)). Nonetheless inventory investment theory has received a growing interest in recent years. The main reason is that inventory investment is very volatile and procyclical, and is a very important component of the business cycle fluctuations during recessions. What follows is a summary of the main empirical evidence about inventories.

1. "*Changes in business inventories, which constitute but a small fraction of total GDP, account for one-fourth of the cyclical movements in GDP*" in the US (Stock and Watson (1998)).

2. Quarterly US data from 1948 to 1991 show that inventory divestment accounts for more than 100% of the fall in GDP in three recessions (from 48:4 to 49:2; from 60:1 to 60:4; from 69:3 to 70:1) and between 21% to 59% in the others. In the early 90s recession inventory divestment accounted for 12 to 71% of the fall in GDP in the G7 countries[8] (Ramey and West (1997)). Hence inventories are procyclical. The correlation between inventory investment and sales is around +0.1/+0.2 in G7 countries (Ramey and West (1997)). In the US it is also positive both at the aggregate level (around +0.4) and in single industries (Blinder (1986)).

[8]This is based on annual data, that at least in the US case underestimated the more accurate quarterly data figure.

3. Inventories are much more procyclical for small than for large firms. The difference between small and large firms cumulative growth rates of inventories and of the inventory/sales ratio is highly procyclical (Bernanke, Gertler and Gilchrist (1996)).

4. Raw materials make up approximately 40% of total inventories and are much more volatile than work in progress and finished goods inventories (Blinder and Maccini (1991)). Ramey (1989) shows that raw materials and input inventories are more volatile than finished goods inventories during recent US recessions. Ramey and West (1997) show that, for G7 countries, finished goods inventories only account for 13% of total inventories, and are not particularly volatile.

5. Production is more volatile than sales at aggregate level (Blinder (1986), Ramey and West (1997)). This is confirmed also by firm level studies (Blanchard (1983), Kashap and Wilcox (1993), Schuh (1996); opposite evidence by Krane and Braun (1991)).

6. Inventory movements are persistent, even conditional on sales[9]. Ramey and West (1997) find the autoregression coefficient of a stationary linear combination of inventories and sales to be around 0.8. Schuh (1996) also finds an autoregression coefficient around 0.6 using disaggregated firm level data.

7. At the aggregate level, the ratio of sales to stocks (production + beginning of the period finished goods inventories) is highly persistent and procyclical. The sales-stocks ratio decreases in each recession, typically by 5% or 10%. This means that although inventory changes are correlated to sales, aggregate inventories are less procyclical than aggregate sales (Bils and Kahn, 2000).

While most of the literature focuses on finished good inventories and on explaining facts 5-7, facts 1-4 emphasise that inventories dynamics are important to understand business cycle fluctuations, and that input inventories are especially important in this respect. Ramey (1989) is, to our knowledge, the only paper to focus exclusively on input

[9]This is confirmed by a very slow estimated adjustment speed in models of optimal inventory management.

inventories. She notes that inventories can have several roles in the production process: i) reduce transaction and fixed costs; ii) reduce uncertainty about materials supply lines; iii) reflect the time element of production. These roles motivate the modelling of inventories as factors of productions. Starting from a restricted form cost function she derives and estimates factor demands, with the following results: inventories are very elastic with respect to output. 40% of the inventory investment is explained by a change in output (accelerator effect). This is stronger for work in progress than for raw materials or finished goods inventories. Ramey also finds that 60% of the decline in inventory investment during a recession is due to shifts in the demand for inventories. Hence it remains an open question whether this large shift is due to technological factors or it has financial explanations.

The idea that financial imperfections can help explaining inventories behaviour is present in the work by Kashap, Stein and Wilcox (1993). They note that after a tight monetary policy there is a change in firms financing regarding the mix between bank lending and commercial papers, and they argue that this confirms the presence of a credit channel of monetary policy. Their interpretation is that at the beginning of a recession banks reduce lending and firms are forced to substitute it with commercial papers (larger firms) or internal finance (smaller firms). In fact the financing mix (bank credit/commercial papers) is at the same time a very effective business cycle predictor and a powerful explanatory variable of inventory and partly also of fixed investment. As we mentioned earlier, also Bernanke Gertler and Gilchrist (1996) argue that the credit channel, and especially the inability of smaller firms to substitute bank lending with other forms of external financing at the beginning of a recession, is responsible for the heterogeneous behaviour of small versus large firms across the business cycle, in terms of sales and inventory investment.

The link between financing constraints and inventory investment is explored at the microeconomic level by Carpenter, Fazzari and Petersen (1994 and 1998). They note that corporate profits (internal finance flows) are extremely procyclical, and tend to lead the business cycle. Sales and revenues fall, often very sharply, just before and during a recession; hence for a financially constrained firm a contraction in investment is required if there is a negative shock in internal finance. Which assets would such firms liquidate? This depends on relative liquidation and adjustment costs. Hence the larger effect should be on inventories, whose liquidation costs are considerably lower. The authors employ a standard model of inventories, augmented with measures of internal finance estimated using quarterly US manufacturing data, from 1981 to 1992, available at firm level. Their findings strongly support the view that financial conditions affect inventory investment, especially for smaller firms.

## 1.6 Conclusions

In the last 15 years two independent strands of the investment literature have been analysing investment dynamics, one focusing on financing imperfections, the other on adjustment costs. Despite some similarities, they have been developing along different levels of analysis: the non convex adjustment cost literature aims at providing a microfoundation of aggregate investment behaviour. It analyses structural models with heterogeneous agents and studies their aggregate implications both theoretically and empirically. The literature about capital markets imperfections and investment instead provides evidence on the influence of financial factors on firm level investment, but does not quantify these effects, nor explain their aggregate implications. The few empirical papers that try to address the question of the aggregate consequences of financial imperfections lack a proper microfoundation. Chapters 3 and 4 aim at filling this gap in investment literature, developing a structural model of investment with financing constraints and deriving its consequences for firm level and aggregate investment dynamics. Chapter 5 shows that it is possible to use such a structural model to derive a new test for the presence of financing constraints on firm investment which is robust to the criticisms to the cash flow-investment type of tests.

Another limit of these literatures is that the effects of financing and real constraints are analysed separately: all the models with non convex adjustment costs assume perfect markets, and all the models with financing imperfections assume convex adjustment costs. One of the main theoretical contributions of this thesis is to show that by studying the interactions between real and financial constraints it is possible to improve our understanding of investment dynamics. Indeed Hubbard (1998) concludes his review article mentioning, among the direction for future research, the challenge of "*incorporating* *financing constraints in models of irreversible investment*", which is the aim of chapter 4.

Also regarding inventories, economic literature mainly analyses them separately from the fixed investment decisions of the firms. However input inventories and fixed capital are used together in production activities, and therefore non convex adjustment costs on fixed capital are likely to have an impact on the dynamics of variable capital as well. In chapter 4 we show that such interactions are particularly important in the presence of both non convexities and financing constraints and are useful in explaining a number of stylised facts regarding the cyclical fluctuations of investment.

# Chapter 2

# Financial structure and firm dynamics: a preliminary empirical analysis

## 2.1 Introduction

In chapter 1 we mention that a theoretical literature emphasises that capital markets imperfections could amplify the business cycle fluctuations of investment and production. However, the empirical evidence on the links between financial structure, financing constraints and firm dynamics is relatively scarce (see section 1.4). Therefore this chapter aims at finding further evidence of these links, conducting a preliminary empirical analysis on our sample of Italian manufacturing firms. We investigate on two aspects of firm dynamics: i) the dynamics of capital structure and output in the business cycle; ii) the importance of financial factors for the growth and the volatility of growth of small firms.

We think that this empirical analysis is important because of the unique feature of our dataset, which includes both balance sheet data as well as a rich dataset of qualitative financial information directly provided by the entrepreneurs (see appendix 2 for details). Such dataset represents a unique opportunity to explore the relevance of financial factors for the real activity of firms. Similar studies conducted on US manufacturing firms, like the one by Bernanke, Gertler and Gilchrist (1996), use indirect and arbitrary criteria, like firm's size, to identify firms with imperfect access to financial markets, while we have the advantage of using the direct information from the entrepreneurs about their financing problems and about what sources of funds they mostly used to finance their growth[1].

The main findings of this chapter are as follows: i) financing constraints affect cyclical fluctuations of firms. Those entrepreneurs who state financing problems are more procyclical in terms of both output and short term debt. They are able to increase their borrowing from banks only in expansion periods, while they are unable to maintain stable leverage levels during contraction periods. Moreover their growth is very volatile: among the smaller firms (turnover between 3 and 15 millions US$), the growth of the financially constrained ones is very sensitive to expansion and contraction phases, while the growth of the unconstrained ones is not affected by them. This result confirms and reinforces the evidence produced by Bernanke, Gertler and Gilchrist (1996) about US manufacturing firms. We also show that this result is robust to the criticisms that financially constrained firms could simply be inefficient firms, or firms who belong to more procyclical industrial sectors. ii) Small Italian manufacturing firms finance their growth mostly with short term banking debt and retained earnings. However these two financing sources show a certain degree of substitution: entrepreneurs who mainly use retained earnings are less able to maintain their borrowing levels during economic contractions. iii) The access to long term banking debt positively affects firms growth and profitability.

This chapter is important in motivating the theoretical and empirical analysis of the remaining of this thesis, for at least two reasons: i) it investigates on the links between financial structure and real activity of firms in the business cycle, in order to identify the relevant stylised facts. The essential motivation for the structural models developed in chapters 3 and 4 is that financial factors are important in order to explain the cyclical behaviour of investment. This idea is supported by the empirical evidence cited in chapter 1 regarding the asymmetric behaviour of small as opposed to large firms in the business cycle fluctuations, especially after changes in monetary policy (Bernanke, Gertler and Gilchrist, 1996). We believe that our sample of Italian manufacturing firm, with the unique

[1] In appendix 2 we also describe our sample of Italian manufacturing firms and we show that its characteristics are comparable with the samples of US and UK firms most frequently used in the empirical investment literature. This supports the generality of the empirical results obtained from the structural model's estimation in chapter 5.

combination of quantitative and qualitative data, provides additional empirical evidence on the importance of financial factors on firm dynamics. ii) This chapter also investigates on how the heterogeneity of firms in terms of their financial structure contributes to the empirical evidence cited before. Because we estimate the dynamics of the whole distribution of firms, and not simply of its first moment, our analysis leads to interesting insights especially regarding the dynamics of leverage, which are considerably different in recession and expansion phases especially for the group of firms with financing problems.

Therefore we are not only able to conclude that financing imperfections and financial structure matter, but also that there is considerable heterogeneity of firms both with respect to the level and the volatility of leverage, and that the latter is linked with the heterogeneity in the volatility of growth. These findings motivate the theoretical analysis in the remaining chapters of this thesis, where we investigate on how the heterogeneity of firms with respect to current and future expected financing constraints affect aggregate output and investment dynamics.

The outline of this chapter is as follows: in section 2.3 we illustrate aggregate dynamics of capital structure and production of firms. We compare aggregate statistics for small and large firms and for constrained and unconstrained ones, selected according to the direct information provided by the entrepreneurs. In section 2.4 we illustrate the distribution dynamics of firms. We estimate the transition dynamics of the distributions of firms' size, leverage and profitability. We select firms in subgroups according to stated financing problems and to the sources of funds mostly used by them to finance their growth.

## 2.2 Aggregate dynamics of financial structure and output

Basic descriptive statistics of our dataset are reported in appendix 2. For the empirical analysis of this section we use all the 5289 firms with complete balance sheet data from 1982 to 1992, as well as the subsample with 897 firms which also includes cross sectional qualitative information from the First Mediocredito Centrale survey on Small and Medium Manufacturing Firms (see appendix 2). We use two criteria to separate firms in subgroups: i) size. We select firms in three groups according to the average size of total assets in the period considered. Figure 2-1 shows that the group of 772 small firms is almost 20 times smaller than the group of 1351 large firms.

Figure 2-1: Number of firms and dimensional classes

**Dimensional classes**

| | N. Of firms | Total Assets |
|---|---|---|
| Small Firms | 772 (14,59%) | 2.880 |
| Medium Firms | 3166 (59,86%) | 8.577 |
| Large Firms | 1351 (25,54%) | 48.271 |
| Total | 5289 | |

**Values in US Dollars, constant 1997 prices**

ii) Financing problems. We consider the answers to the questions about the financing problems faced by entrepreneurs from the first Mediocredito Centrale Survey (see appendix 2). This information is cross-sectional, because it refers to the 1989-1991 period as a whole. Figure 2-2 shows that 23% of the entrepreneurs stated some kind of financing problem, and their average size is half of the size of the complementary subgroup of firms without financing problems: this is reflected in the fact that the share of firms with financing problems is almost double for small firms with respect to large ones (37.5% versus 19.9%). The most important problems are the cost of debt and the lack of financing, while lack of collateral is considered a problem only by 2% of the firms[2].

### 2.2.1 Dynamics of financial structure

The sources of funds, as a percentage of total assets, are reported in figures 2-3 and 2-4. The three main sources of funds, net wealth, bank loans and trade debt, ensure the financing of 75%-80% of total assets. Other liabilities have a marginal importance. The main differences between small and large firms are that small firms use more trade debt than large firms, which depend more on bank loans[3]. Another difference in the behaviour

[2]This may be caused by the fact that this problem is implied by the other two questions. In fact a firm with an adequate level of collateral could, in principle, always borrow secured long term debt at low interest rates.

[3]The fact that small firms finance less with short term bank loans does not contradict the typical finding of other studies, that larger firms borrow less from banks (see, among others, Gertler and Gilchrist, 1994). This is true also for Italian firms, but it is not found here because the "Centrale dei Bilanci" sample does not include very large firms. The more heterogeneous sample from Mediocredito Centrale shows that the relation between dimension and short term banking debt is concave[4]038<p type="texpara" tag="Body Text" >These additional statistics are available upon request., increasing from small to medium firms and decreasing from medium-large to very large firms, with direct access to capital markets.

Figure 2-2: Firms distribution according to financing problems

**Dimensional classes and financing problems**

| | N. of firms | Avg. Total Assets |
|---|---|---|
| Firms without financing problems | 688 (77%) | 37.77 |
| Firms with financing problems | 209 (23%) | 20.29 |
| Firms with high cost of debt | 136 (15%) | 15.70 |
| Firms with lack of medium-long term financing | 127 (14%) | 23.44 |
| Firms with lack of collateral | 21 (2%) | 17.09 |
| Firms with high cost of debt and lack of m./l. t. debt | 56 (6%) | 20.25 |
| Firms with high cost of debt and lack of collateral | 14 (1.5%) | 14.56 |
| Firms with lack of collateral and of m./l. term debt | 13 (1.5%) | 16.38 |
| Firms with all three problems | 10 (1%) | 14.95 |
| Total | 897 | 33.71 |

| | Small | Medium | Large |
|---|---|---|---|
| Without financing problems | 62.5% | 74.8% | 80.1% |
| With financing problems | 37.5% | 25.2% | 19.9% |

Values in US Dollars, constant 1997 prices

Figure 2-3: Capital structure - small firms

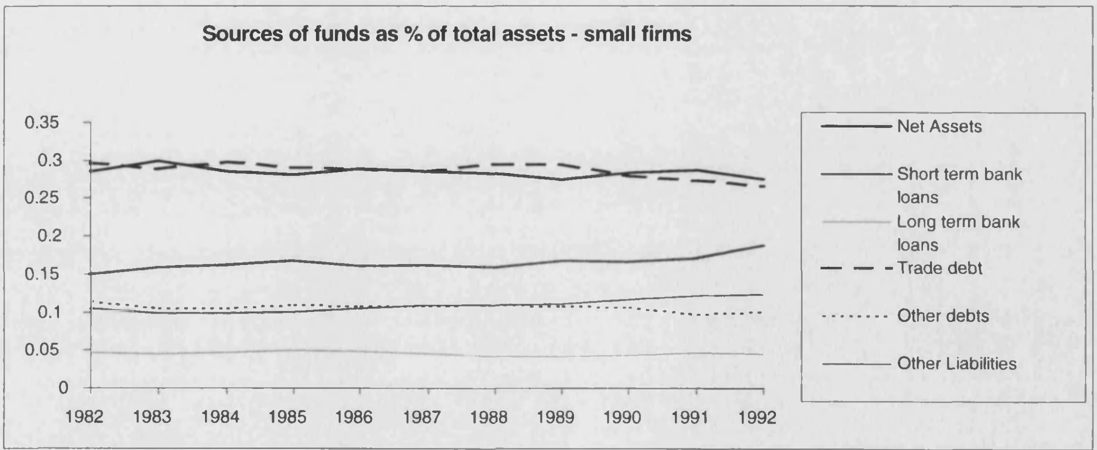
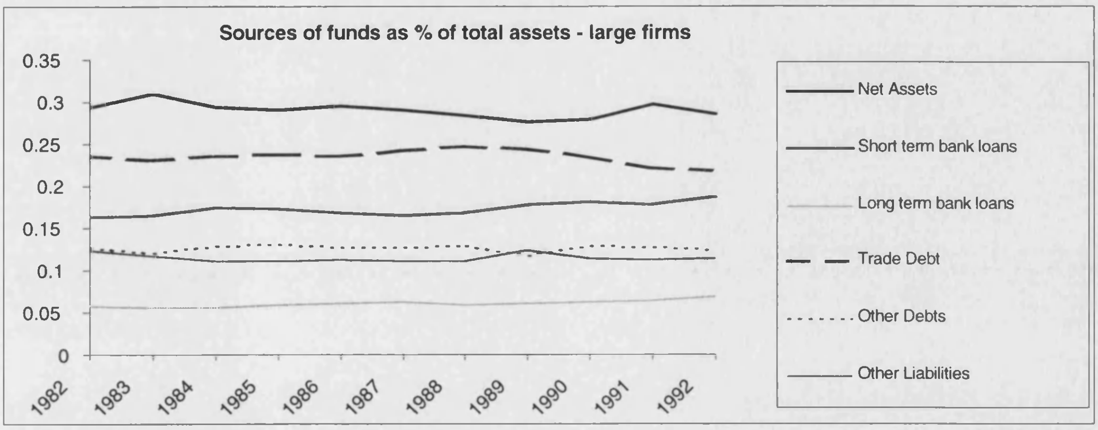


Figure 2-4: Capital structure - large firms




of small and large firms regards internal finance. Figure 2-5 shows that small firms retain more income in all the sample years.

Figure 2-5: Earnings retention - small and large firms


Figures 2-6, 2-7 and 2-8 analyse financial structure for firms selected using direct information about financing problems. Firms with financing problems borrow more from banks in all the sample years and have a lower wealth, measured as net assets over total assets. The figures also include the values of the test statistic for the equality of the subgroups means, computed for each year. They show that the differences in average year values are statistically significant for most of the sample years. In particular the difference in net wealth levels are strongly significant in all the years from 1984 on.

This result suggests that entrepreneurs with low collateral are the ones who stated

Figure 2-6: Short term banking debt - firms with and without financing problems

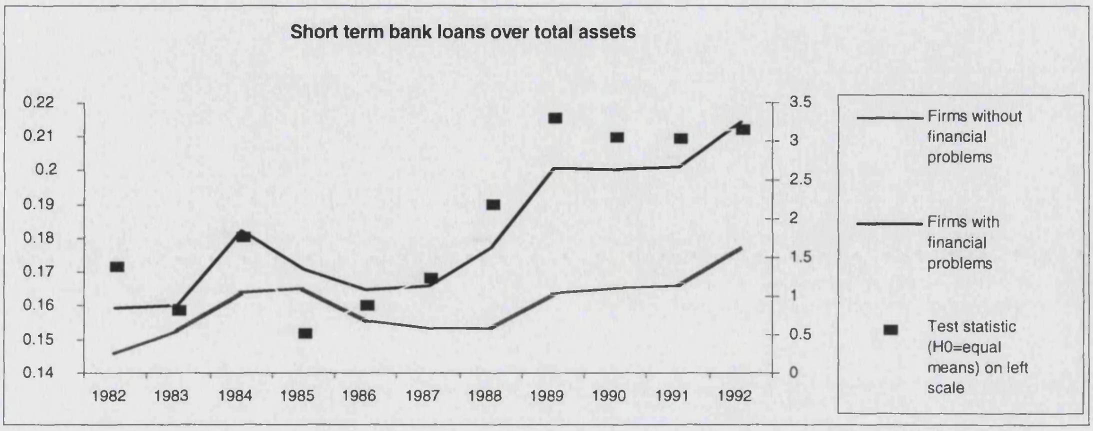



Figure 2-7: Long term banking debt - firms with and without financing problems

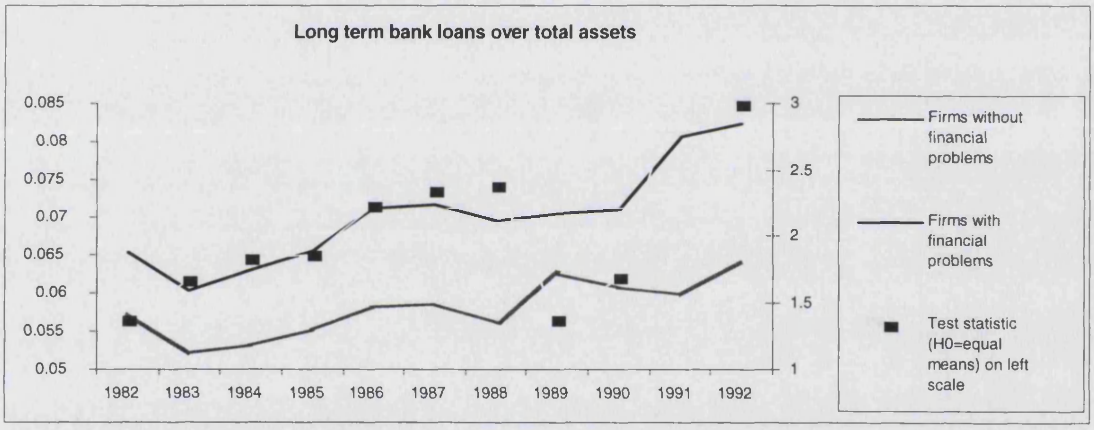
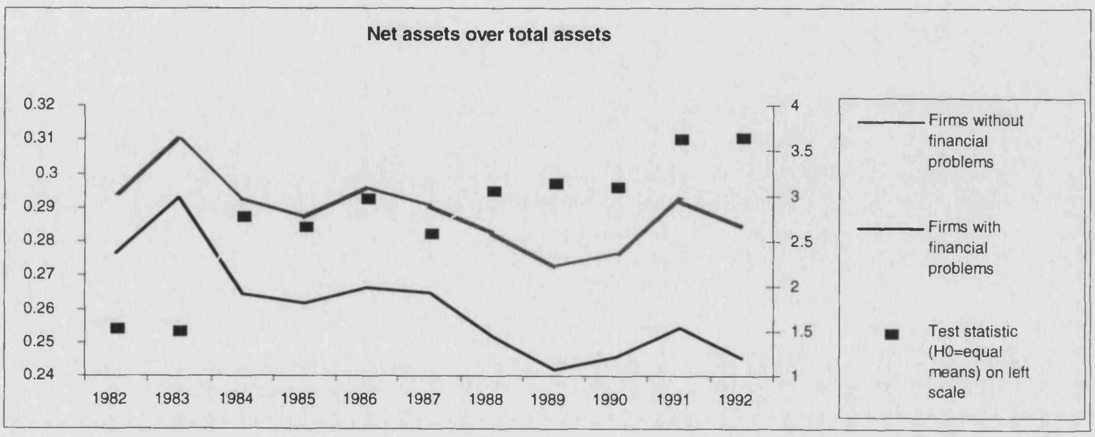



Figure 2-8: Net worth - firms with and without financing problems



financing problems. However this does not necessarily imply the presence of collateral constraints on borrowing. In fact the direct information used to identify firms with financing problems could have selected firms with poor performance, that needed additional finance to cover current losses rather than to finance new investments. This issue is considered in the next subsection.

### 2.2.2 Performance and volatility over the business cycle

Figure 2-9 shows the real rate of growth of total sales for the full sample of 5289 firms. It emphasises the fact that the sample period covers both expansion and contraction phases. A boom starts in 1984, after the deep depression of the beginning of the 80s, and lasts until 1989, with a relative decline in 1986. The 1990-1992 years experienced a contraction with a real growth rate of sales close to 0 in 1990 and negative in 1991 and 1992. The differences in the growth rates of output and short term debt between firms with and without financing problems (figures 2-10 and 2-11) are consistent with the "credit cycles" view, and in particular with the empirical findings of Bernanke, Gertler and Gilchrist (1996). Firms with financing problems exhibit higher procyclicality of both short term debt and output. Figure 2-10 shows that they have higher growth rates of output in expansion periods (1984, 1988 and 1992), and lower rates in contraction periods (1985, 1986, 1990 and 1991). The variance around these data is however quite high, and the mean values of subgroups are not significantly different, with the exception of 1985, 1990 and 1991. A similar pattern, but with greater differences between firms groups, is observed for the growth rate of short term debt (figure 2-11).

A possible critique to this result is that financially constrained firms are more procyclical simply because they are on average smaller, and smaller firms have procyclical technologies, either because they are more flexible than large firms, or because they belong to more procyclical sectors. Our data reject both objections. Figures 2-12 and 2-13 show that smaller firms are not more procyclical than larger ones, both in terms of output and short term debt[5]. Moreover figure 2-14 shows that financially constrained firms are

[5] This is in contrast with Bernanke, Gertler and Gilchrist (1996), who use dimension as a proxy for financing constraints. The difference is probably that we do not have enough size heterogeneity among the firms in our sample, which are almost exclusively small and medium firms below 500 employees and

Figure 2-9: Sales growth - all firms

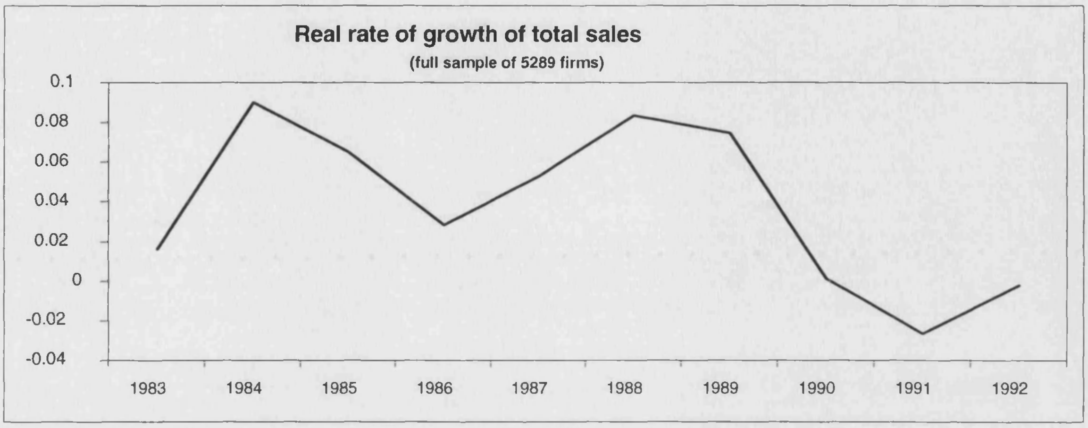
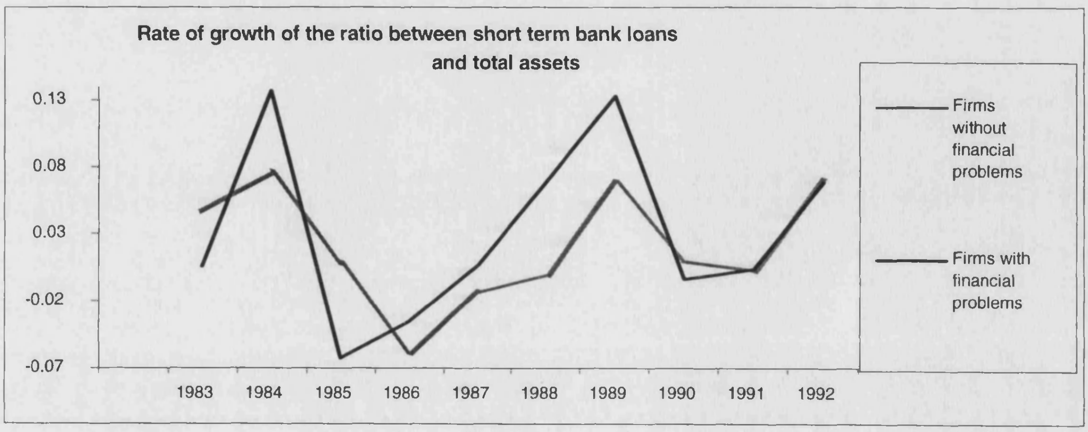



Figure 2-10: Sales growth - firms with and without financing problems


Figure 2-11: Leverage growth - firms with and without financing problems



present in similar shares in the different industrial sectors (the figure reports all sectors with more than 50 firms in the sample). The only exceptions are the "Shoes and clothes" and "Paper, printing and publishing" sectors.

Figure 2-12: Leverage growth - small and large firms

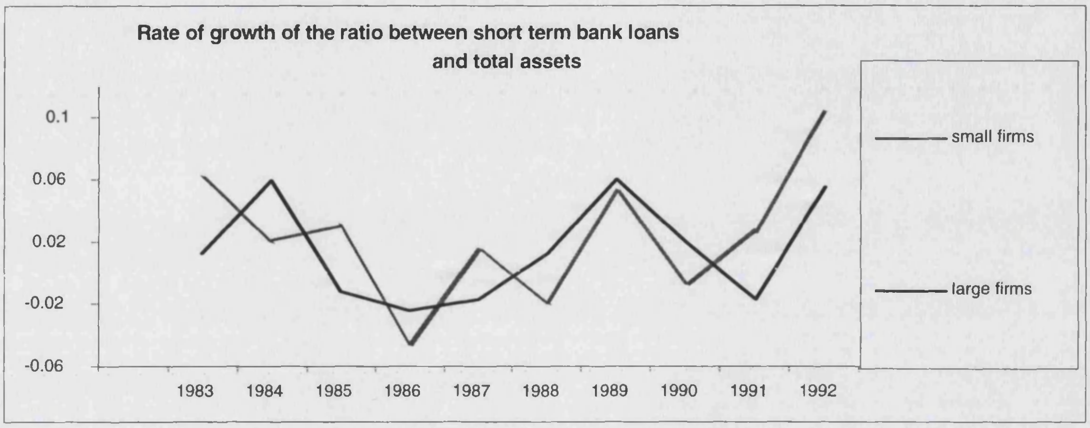



Figure 2-13: Sales growth - small and large firms


The last potential critique to this result is that we are analysing firms with poor performance that complain about the denied renewal of loans rather than profitable firms that face financing constraints. In fact, as figure 2-15 shows, financially constrained firms have lower performance, measured as the ratio of gross income over total sales[6]. We

without stock market quotation.

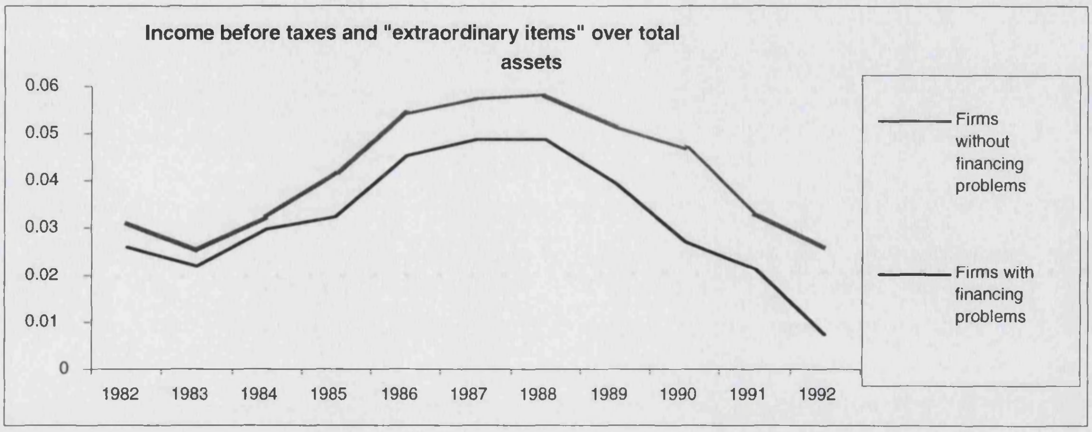
[6]I considered income before taxes and "extraordinary items". Such items are gains or losses caused by transactions not linked with the main activity of the firm. For example the revenue that comes from selling a building, for a firm that produces shoes, is considered extraordinary revenue. Hence such items are not related to firm's performance, even because are frequently used to cover losses in negative periods.

Figure 2-14: Distribution of financing problems across industrial sectors

**Industrial sectors and financing problems**

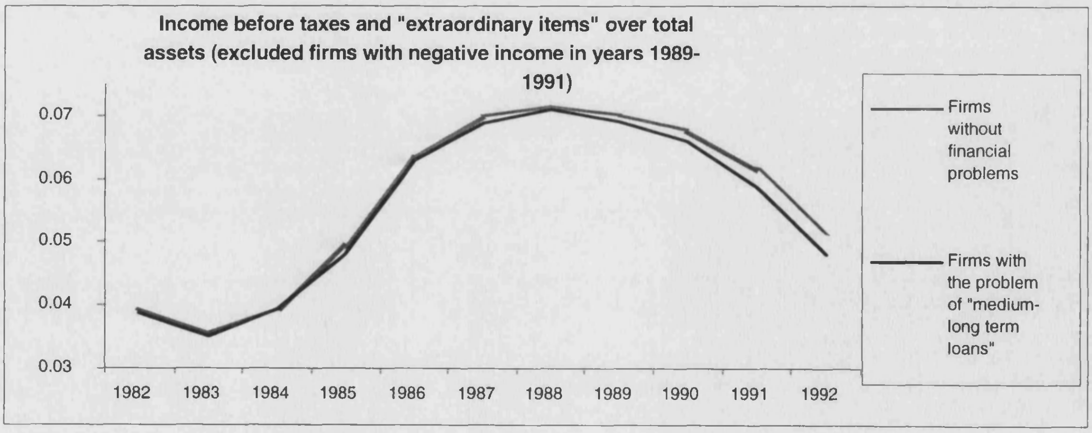
| Sector | N firms | % with financing problems | % without financing problems |
|---|---|---|---|
| Whole sample | 897 | 23 | 77 |
| Mechanic materials and machineries | 139 | 21.6 | 78.4 |
| Metallic products | 130 | 26.9 | 73.1 |
| Textiles | 87 | 21.8 | 78.2 |
| Shoes and clothes | 65 | 10.8 | 89.2 |
| 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 |

Figure 2-15: Net income - firms with and without financing problems



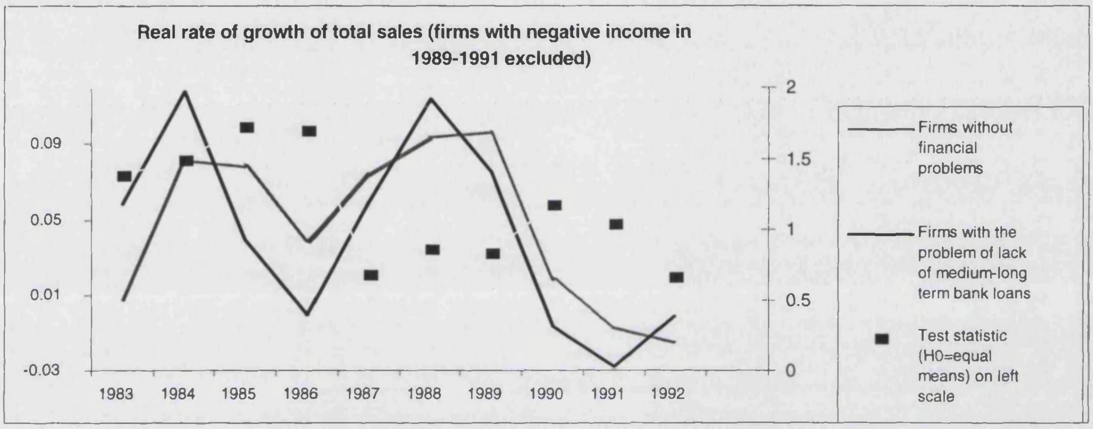

separate good firms with a binding liquidity constraint from bad firms in financial distress by selecting only firms with positive gross income (income before taxes and extraordinary items) in each of the three years considered by the survey, 1989, 1990 and 1991. The filter selects out 39.7% of the firms with financing problems and 29% of the firms without financing problems. Figure 2-16 shows that once all the companies with negative gross income in 1989-1991 are selected out, firms with the "lack of financing" problem have almost the same performance as firms without financing problems. Nevertheless financially constrained firms still exhibit higher variance in the rate of growth of total sales.

Figure 2-16: Net income - firms with and without the problem of lack of medium-long term bank loans




This is confirmed by figure 2-17. Firms with financial problems exhibit greater variations in sales over the cycle. Not only are their growth rates higher in booms and lower in recessions, but also they seem to lead the decline of sales in the recessions. In fact sales growth rates of these firms declined sharply in 1985 and 1989, while the decline for firms without financial problems begun one year later, in 1986 and 1990. Mean differences are very high, but not very significant. This is because also the standard error of the difference is high, as the group with the problem of lack of financing is made of only 48 firms, once the positive-performance filter is applied.

Figure 2-17: Sales growth - firms with and without the problem of lack of medium-long term bank loans




## 2.3 Financial structure and growth: distributions dynamics

This section[7] investigates the distribution dynamics of Italian manufacturing firms, with particular reference to the interaction between real and financial variables in the growth process. The analysis is conducted using a dataset which includes all the three Mediocredito Centrale surveys on small and medium Italian manufacturing firms (see appendix 2). They provide information on the financing problems faced by the firms in funding investment and also about the composition of financing sources used to fund new investment projects. This information is available for each firm for three distinct sub-periods (1989-91, 1992-94 and 1995-97). The sample is made of 372 firms for which also eight years of balance sheet data, from 1990 to 1997, are available (see appendix 2 for details). The

[7]This section is based on a joint work with Sandra Bulli.

econometric technique employed is the estimation of the distribution dynamics of firms' size, leverage and profitability[8].

Using the qualitative information described above we estimate such distribution dynamics for subgroups of firms selected according to the following criteria: a) the degree with which firms finance investment using internal funds; b) the degree with which firms finance investment using long term debt; c) the financing problems stated by the firms. The variables considered in the analysis are the following: *(i)* size, measured as sales at constant 1990 prices[9]; *(ii)* leverage, measured as the ratio of bank loans over total assets[10]; *(iii)* productivity, measured as the ratio of operative profits (income before taxes and financial expenses) over total assets. The analysis also distinguishes between periods of economic expansion and contraction. Using aggregate data on Italian industrial production in the 90s (figure 2-18, source: Datastream), we define "contraction" periods the years 1990-1993 and 1996, and "expansion" periods the years 1994-1995 and 1997.

### 2.3.1 Firms selected on the basis of the degree of self-financing

We select as high "self-financing" those firms that used retained earnings to finance at least 66% of fixed investments in each of the three surveys[11], for a total of 81 firms (21.8% of the sample considered). Figure 2-19 shows the distributions of size, measured in logarithms of total sales at constant 1990 prices, of these two groups of firms in 1990 and 1997. Firms with high self-financing are on average bigger in 1990 compared with low self-financing firms. This difference increases from 1990 to 1997. Information about growth expectations conditional on the initial size are reported in figure 2-20. This figure corresponds to a nonparametric estimation of the relation between size and growth. The

[8]The distribution dynamic method, explained in appendix 4, has been firstly used to analyse the dynamics of per capita income across countries (Quah, 1993 and 1997) and has subsequently been applied to different areas of economic research, like industrial organisation, finance and international trade (Koopmans and Lamo, 1995; Stanca, 1999; Redding, 1999). All estimations in this section are performed using D. Quah's econometric package tsrf.

[9]Similar results were obtained using the variable "total assets".

[10]The analysis of alternative measures of leverage that include commercial loans and other short term non banking debt has produced qualitatively similar but quantitatively less important results.

[11]The other firms are only relatively less self financed, as retained earnings is the main source of funding for almost all firms in the sample. For example among the sample of 897 Italian manufacturing firms (see appendix 2) very few of them distribute dividends. We observe nonzero dividends in only 28.6% of the total number of firm-year observations, and in only 2.7% of the observations the ratio between dividends and revenues is greater than 5%. The percentages fall to 27.1% and to 2.2% for "family firms", which report one entrepreneur plus one or more members of the family working full time in the firm.

Figure 2-18: Total output - Italian manufacturing sector

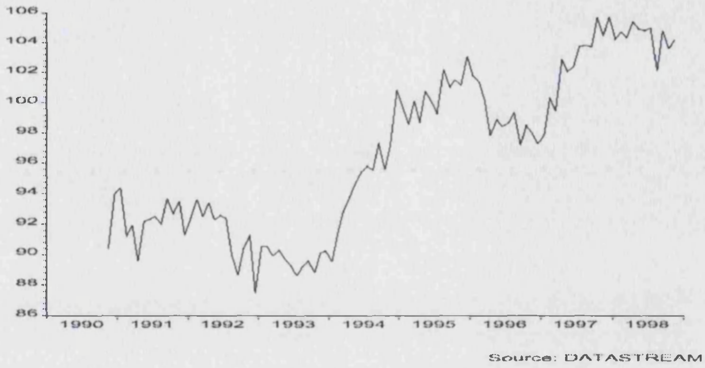



Figure 2-19: Size distributions - high and low self-financing firms

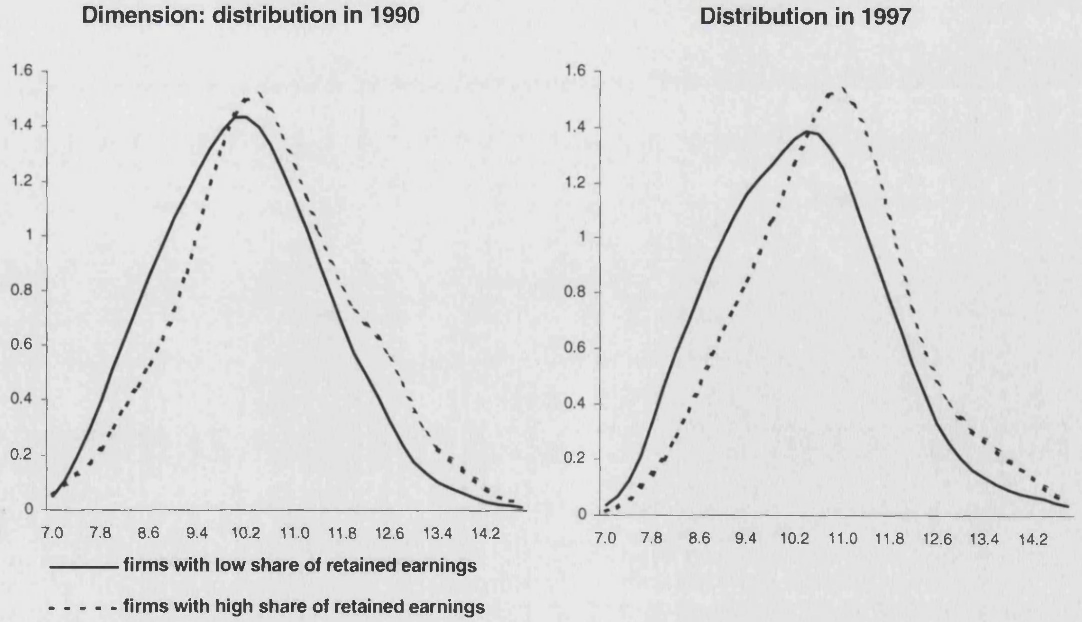


figure shows that, among smaller firms, high self-financing ones have a relatively higher expected growth than low self-financing ones. Therefore self-financing seems the preferred source of funds for smaller high growing firms.

Figure 2-20: Conditional growth expectations - high and low self-financing firms

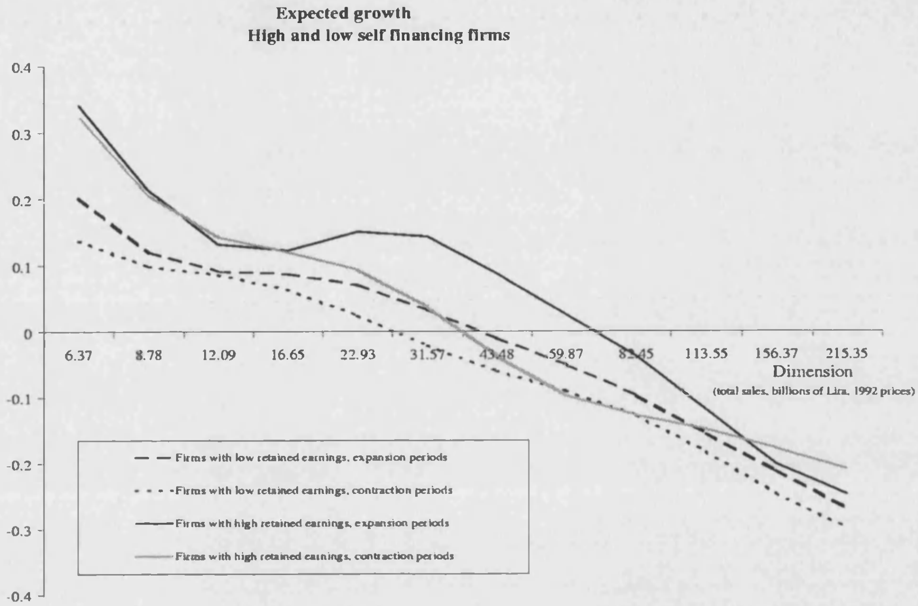
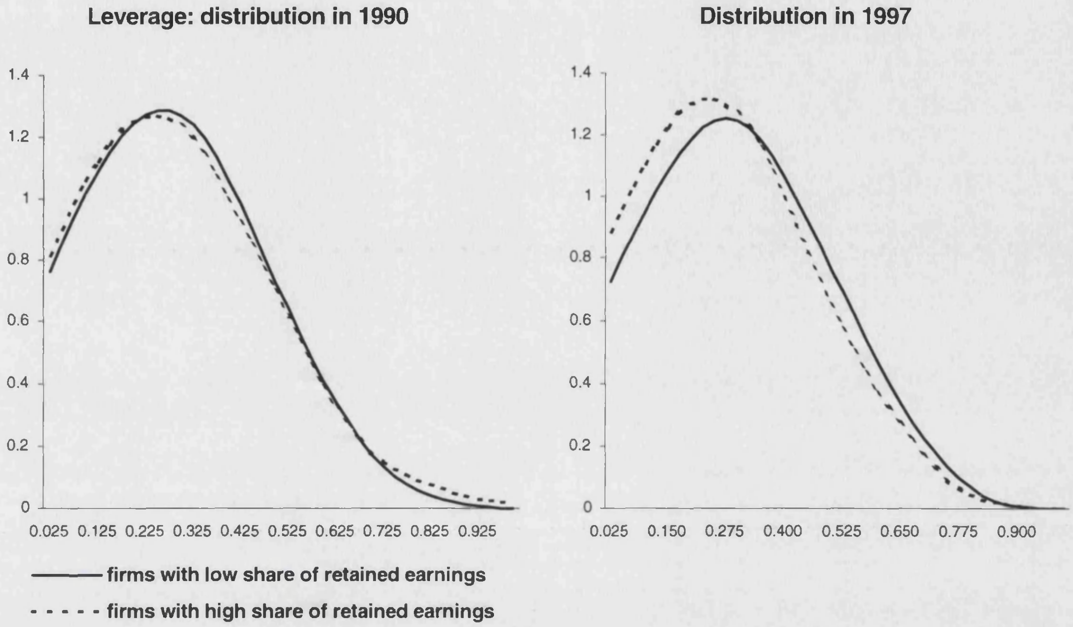



Figure 2-21 shows the distributions of leverage in 1990 and 1997. In 1990 the two distributions are similar, with a peak of density relative to a leverage of approximately 25%, but in 1997 they differ: the density of firms with low leverage (0%-20%) increases for high self-financing firms. This might suggest that for these firms internal funds substituted bank loans in the sample period.

Figures 2-22 and 2-23 show the distribution dynamics of leverage[12]. Data are in natural logarithms, and hence absolute variations represent percentage variations of the original variables. The higher the "mountain range" along the diagonal, the more persistent over

[12]The axes report the value of the variable at the generic years $t$ and $t+1$. Therefore a section of the graph from a point $x$ on the $t$ axis conducted parallel to the $t+1$ axis represents the probability density function of the value at time $t+1$, conditional on the value at time $t$ being $x$. The transition function, therefore, maps each portion of the distribution from one period to the other and thus describes the law of motion of the distribution. A density concentrated along the main diagonal would indicate high persistence of the current values. Conversely, if the density is more equally spread over the base this means more intra-distribution mobility.

Figure 2-21: Leverage distributions - high and low self-financing firms




time is the value of the variable analysed. It can be seen, especially in figure 2-23, that we have "two peaks" corresponding respectively to low levels (0-10%) and high levels (40%-50%) of leverage. This means that the probability of maintaining a stable leverage level is higher for very low or very high values of the ratio. This finding can be interpreted as the existence of two alternative optimal structures of capital: one with low leverage and occasional bank loans which are repaid in the short run, another with high leverage, where bank loans represent the main component of the firm's funding.

Figure 2-22 shows a higher leverage volatility of high self-financing firms. They present a lower persistence of the two peaks, especially in contraction periods. This means that high self-financing firms are more likely to vary the levels of leverage over time. Low self-financing firms (figure 2-23) present a higher persistence of leverage, both during expansion and contraction periods. The analysis of profitability[13] reveals that self-financing firms are slightly more profitable, but this does not seem sufficient to explain the major differences in the financial structure between high and low self financing firms. It seems therefore that the choice of high self-financing is motivated by higher earnings but also by the lower

[13] Not reported here, but available upon request.

Figure 2-22: Leverage distributions dynamics - high self-financing firms

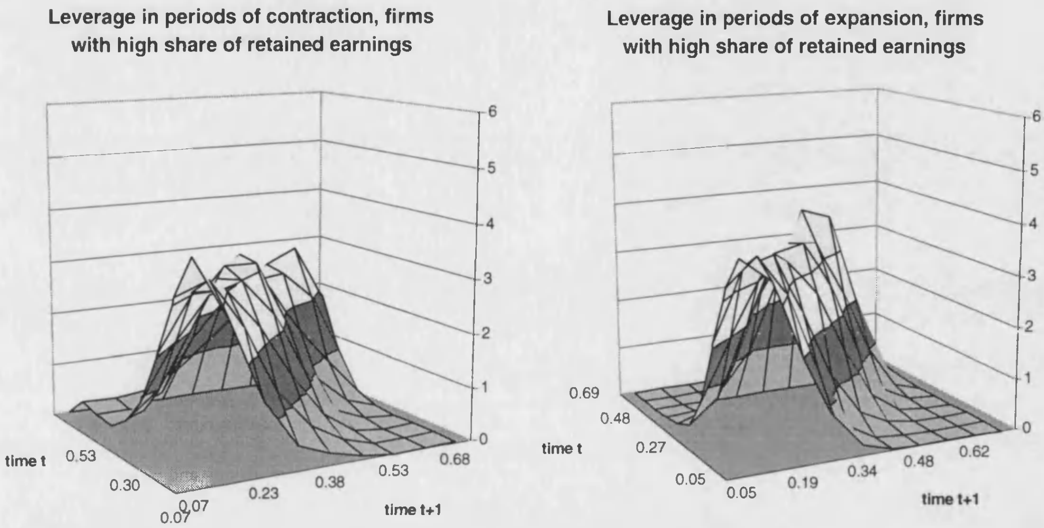
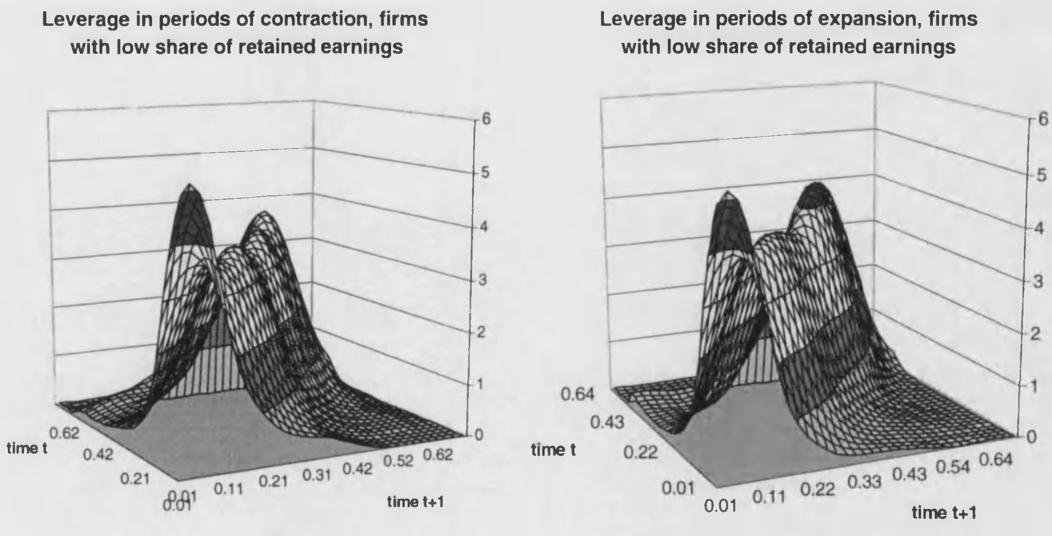



Figure 2-23: Leverage distributions dynamics - low self-financing firms

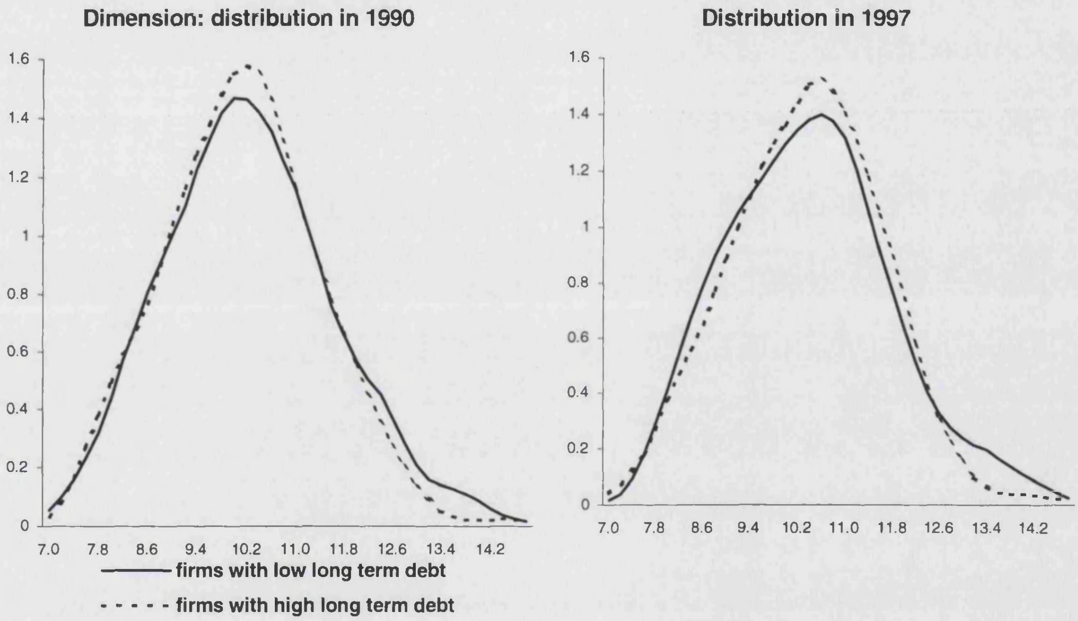


ability to keep stable leverage levels, especially during contraction periods.

### 2.3.2 Firms selected on the basis of their access to long term debt.

We select as "high debt" those firms who have financed at least 30% of their investment with medium or long term debt[14] in at least two of the three survey periods, for a total of 140 firms (37.6% of the sample). Figure 2-24 shows the size distributions in 1990 and 1997. It shows that access to long term debt is positively related to firms' growth. This

Figure 2-24: Size distributions - firms with high and low debt




is confirmed by figure 2-25, which shows the growth expectations of firms. Among smaller firms, "high debt" firms have a higher growth expectation than "low debt" firms, both in expansion and recession periods. This result could be caused by the fact that small very profitable firms at the same time grow faster and get more long term bank financing, especially in expansion periods. Figure 2-26, regarding the distribution of firms according to their profitability (net income over total assets), rejects this view. In fact "high debt" firms are slightly less profitable than "low debt" firms in 1990, while they are slightly more profitable in 1997. Therefore it seems that the access to long term debt has a positive

[14] Defined as debt repayable after more than 1 year.

Figure 2-25: Conditional growth expectations - high and low debt firms

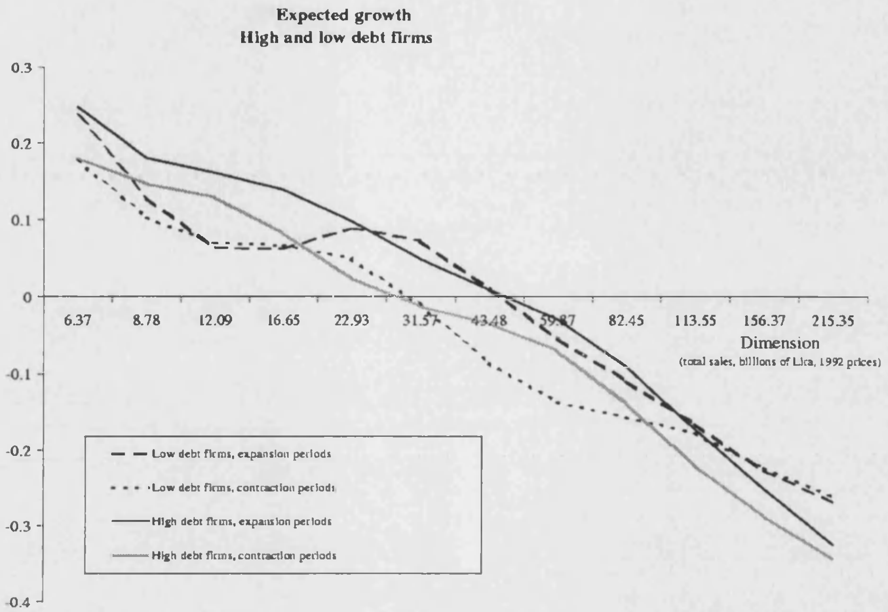
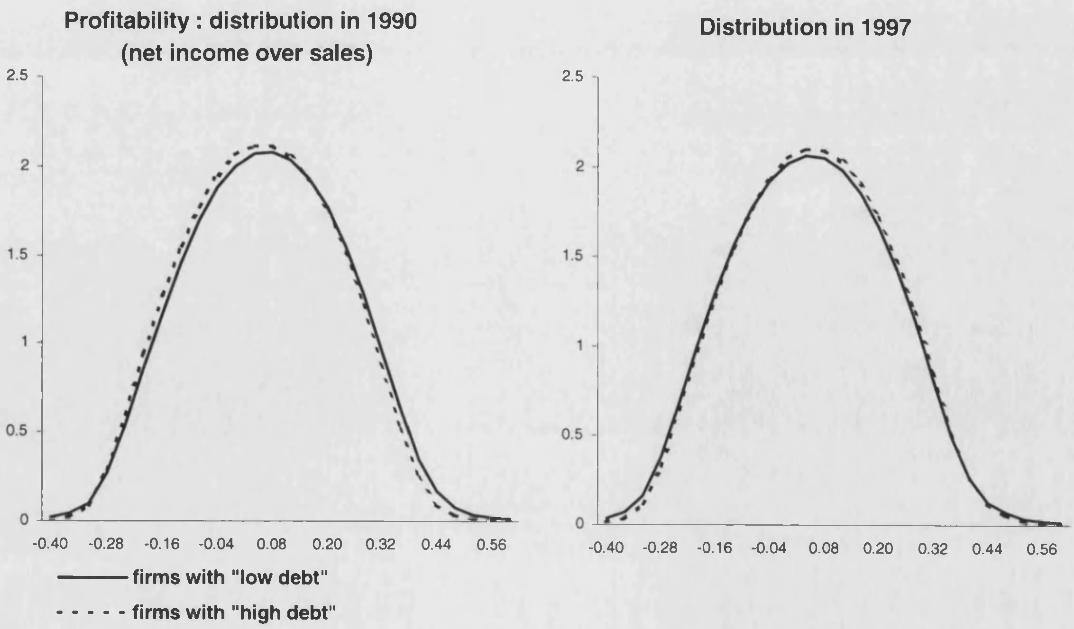
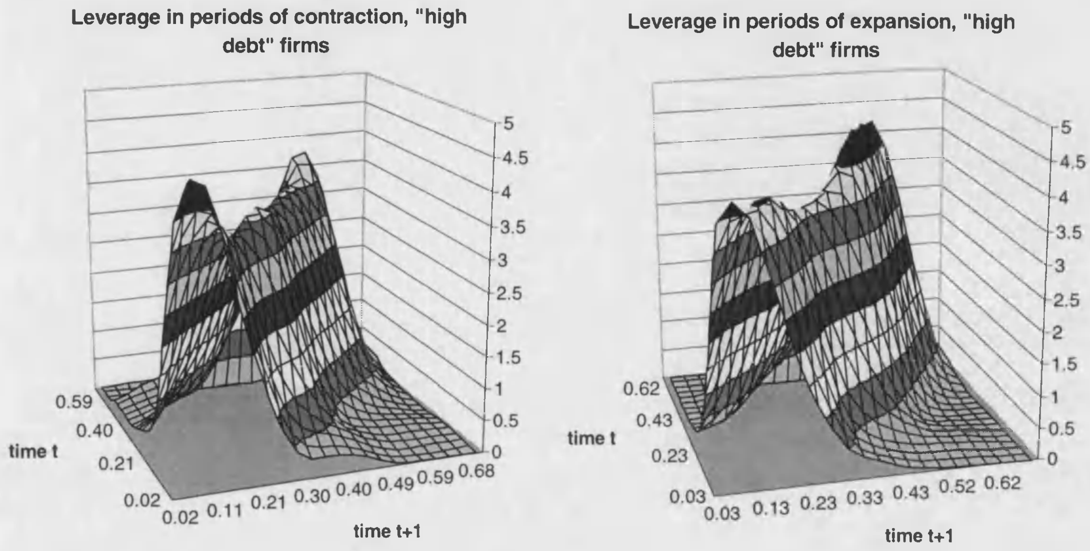



Figure 2-26: Net income distributions - high and low debt firms



effect not only on firms' growth but also on their profitability.

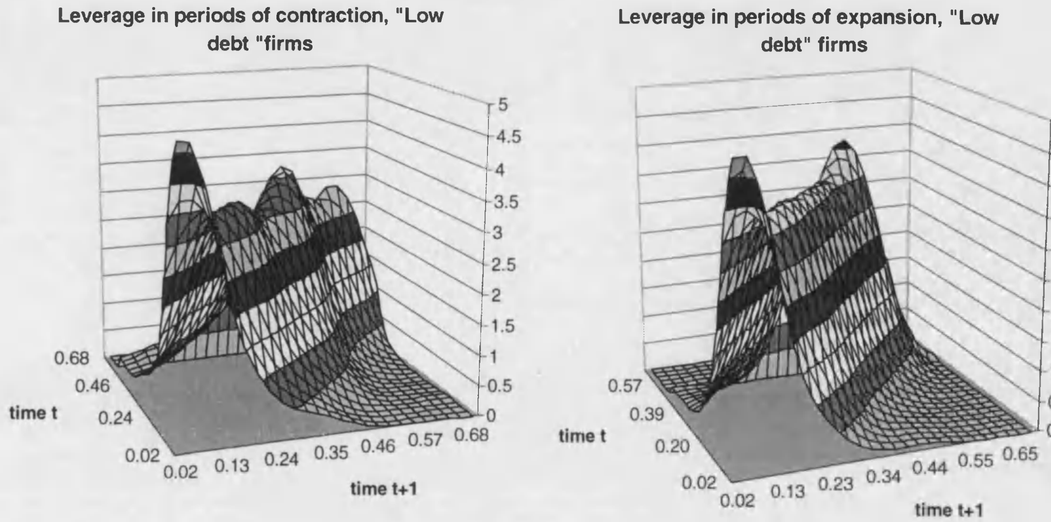
Figure 2-27: Leverage distributions dynamics - high debt firms




Figures 2-27 and 2-28 regard the dynamics of the distribution of leverage. The persistence of high levels of leverage is particularly strong for "high debt" firms (figure 2-27). This happens both in expansion and recession periods. In particular, the probability of reducing the leverage at time $t+1$, conditional on an high level at time $t$, is very small. Conversely "low debt" firms (figure 2-28) are very sensitive to the economic cycle. Their probability of maintaining high leverage levels drops dramatically during contraction periods (the "mountain range" on the left graphic is lower). This evidence seems to point out to a more stable relationship between high-debt firms and financial intermediaries, which allows such firms to maintain high levels of short-term debt financing even in recession periods[15]. For smaller high-debt firms this implies higher expected growth and also an increase in profitability.

[15]It is important to note that while leverage includes all banking debt, firms are selected according to long-term banking debt only, which is a very small fraction of the former (see figures 2-3 and 2-4), but it is an important signal of long-term bank-firm relationships.

Figure 2-28: Leverage distributions dynamics - low debt firms




### 2.3.3 Firms selected on the basis of directly revealed financing constraints

We select in the financially constrained groups all firms whose entrepreneurs stated some financing problem in any of the three Mediocredito Centrale surveys, for a total of 123 firms (33.1% of total firms). The distributions of size is shown in figure 2-29. Financially constrained firms exhibit a distribution skewed towards small firms. This skeweness diminishes from 1990 to 1997, indicating that the presence of financing constraints does not prevent small firms to grow in the sample period. Figure 2-30 shows the expectations of growth, conditional on initial size: it shows that smaller (turnover less than 20 billion lira) financially constrained firms have very different growth expectations across expansion and contraction periods. Conversely small unconstrained firms are almost not affected by the business cycles. The leverage distributions (not reported here) confirm the higher leverage of financially constrained firms shown in figures 2-6 and 2-7.

The figures relative to the distribution dynamics show that in expansion periods the peak at high levels of leverage is more pronounced for firms with financing constraints (figure 2-31, right graphic), which have a low probability of reducing those levels. Therefore firms with financing constraints have, in expansion periods, a dynamic of leverage very

Figure 2-29: Size distributions - firms with and without financing problems

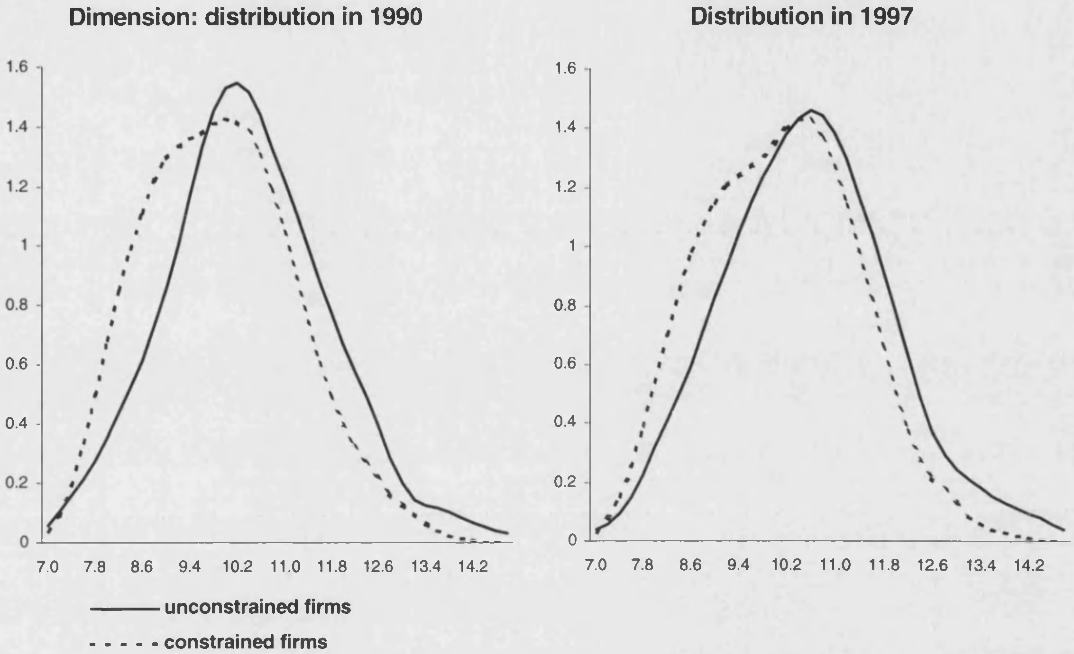



Figure 2-30: Conditional growth expectations - firms with and without financing problems

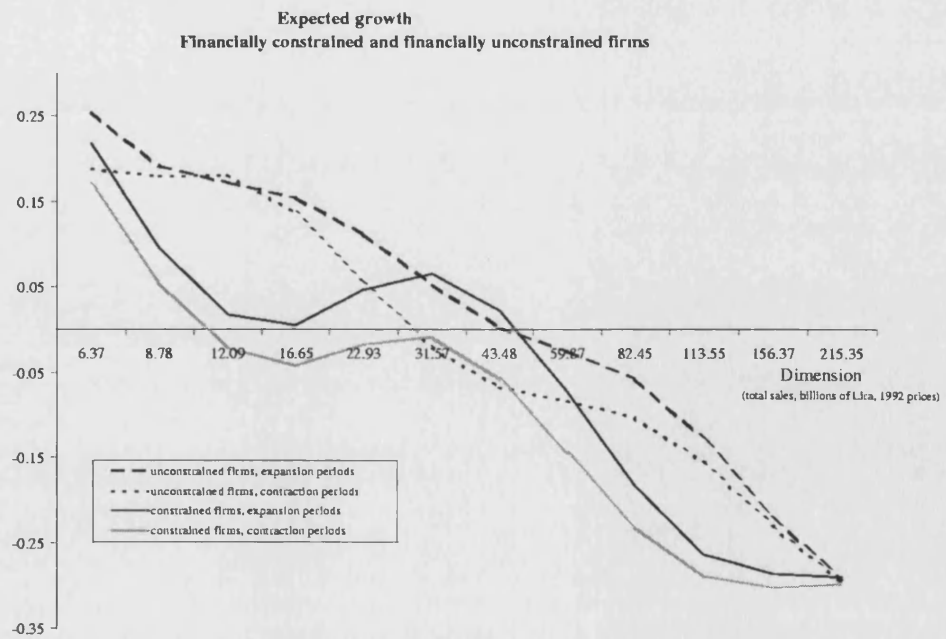
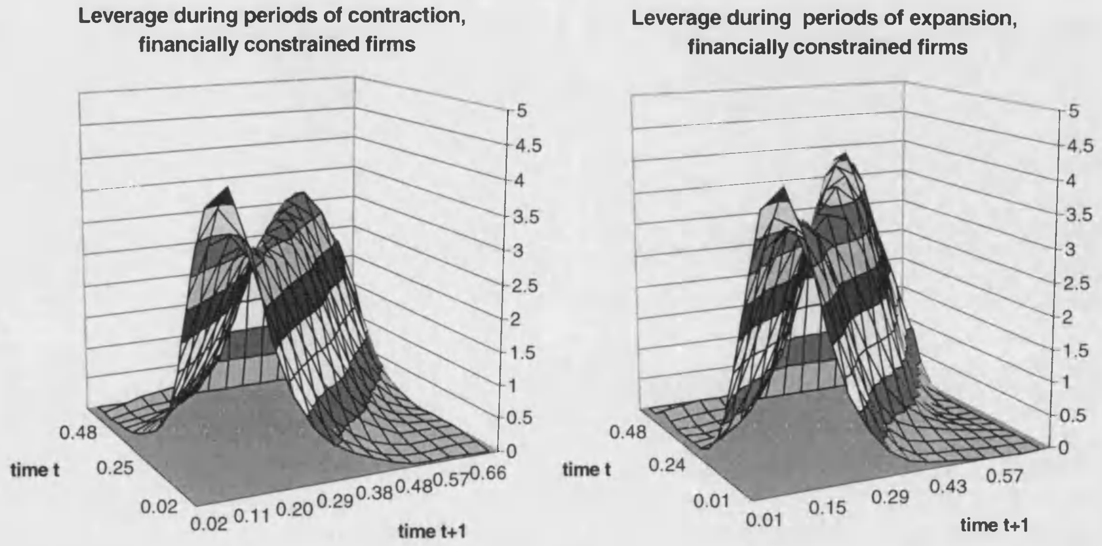


Figure 2-31: Leverage distributions dynamics - firms with financing problems



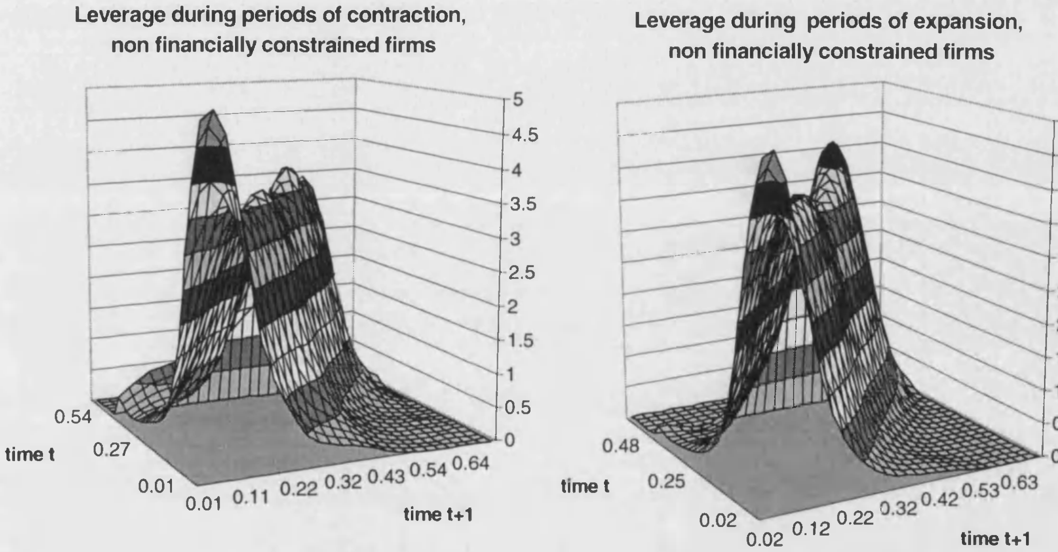

similar to "high debt" firms (figure 2-27, right graphic). In contraction periods, however, they behave differently: "high debt" firms maintain stable leverage levels (figure 2-27, left graphic), while the opposite happens to firms with financing problems (figure 2-31, left graphic), which have a higher probability of reducing them. This result confirms the findings of section 2.2: firms with financing problems seem unable to maintain stable leverage levels in periods of contraction and at the same time their growth is more sensitive to the business cycle fluctuations.

## 2.4 Conclusions & related literature

In this chapter we analyse our dataset on small and medium Italian manufacturing firms to find evidence of the links between financial structure, financing constraints and firm dynamics. The main results of this chapter reinforce the impression that financial structure matters: i) firms that have stronger links with financing institutions (measured as the amount of long term banking debt used to finance new investment projects) are able to maintain more stable borrowing levels during contraction phases and at the same time they grow more and increase their profitability more than the other firms. ii) Financing constraints affect cyclical fluctuations of firms. Those entrepreneurs who state financing problems are more procyclical in terms of both output and short term debt. Among the smaller firms (turnover between 3 and 15 millions US$), the growth of the financially constrained ones is very sensitive to the business cycle fluctuations, while the growth of the unconstrained ones is not affected by them. These results reinforce the evidence produced by Bernanke, Gertler and Gilchrist (1996) about US manufacturing firms. Our analysis has the advantage of using direct qualitative information provided by the entrepreneurs about their investment financing policy and their financing problems. Our results are confirmed by both the simple analysis of aggregate time series and by the distribution dynamics estimated for groups of firms selected according to the direct qualitative information.

In the empirical literature about financing constraints and investment of Italian firms two papers are related to the analysis of this chapter. Rondi et al. (1998) use aggregate annual balance sheet data for two subsamples of large and small private italian companies. They show that after episodes of monetary tightening smaller firms report a sharper decrease in short term debt and a steeper fall in both sales and inventories. Bagella, Becchetti and Caggese (2001), using the same dataset employed in chapter 5 (see appendix 2) investigate if the cash flow investment sensitivity is related to financing constraints. The authors estimate a reduced form investment equation following the Euler Equation approach (Bond and Meghir, 1994), and combine a priori information and direct qualitative information (see appendix 2) to consistently estimate for each firm the probability of being financially constrained. Their main finding is that, when financially constrained firms are properly identified, the neoclassical model is rejected only for unconstrained firm. Moreover financially constrained firms show positive correlation between investment and lagged cash flow.

Figure 2-32: Leverage distributions dynamics - firms without financing problems



# Chapter 3

# A dynamic model of investment with financing constraints: the precautionary saving effect

## 3.1 Introduction

This chapter develops a structural model of investment with financing constraints. The motivation for this analysis is that, as mentioned in chapter 1, existing studies about financing constraints and firm investment focus mainly on the estimation of reduced form investment equations, without explicitly solving the dynamic investment problems. We believe that the structural approach is instead important because it allows to study the effects of current and future expected financing constraints on the investment and saving choices of the entrepreneurs and to derive their consequences for aggregate investment dynamics.

The model proposed in this chapter has two distinctive features. First, output is produced by an entrepreneur who operates a concave risky technology using capital as a factor of production. Capital takes one period to become productive. Second, the entrepreneur's only source of external finance is debt secured by collateral. In this environment, anticipating a risk of a binding financing constraint in the future, the entrepreneur reduces the investment spending in the risky technology, and keeps some financial assets (or spare borrowing capacity) as a precautionary saving motive. We solve the dynamic investment problem and simulate the investment path of an entrepreneur to show that with financing constraints investment is more volatile and more sensitive to cash flow than with perfect markets. However this sensitivity is mainly driven by future expected financing constraints rather than by currently binding ones. This is because the entrepreneur anticipates them by reducing investment and increasing her holding of safe financial assets. This implies that it is not necessarily true that the investment sensitivity to internal finance is stronger for firms facing current financing constraints. Therefore our analysis, even though it confirms that the cash flow-investment sensitivity can be caused by financing constraints, also supports the Kaplan and Zingales (1997) critique that there is no theoretical foundation for the claim that such sensitivity is monotonous in the intensity of financing constraints.

Our theory also implies that, despite only a small fraction of entrepreneurs having a binding financing constraint at any point in time, financing constraints can have a big impact on aggregate investment dynamics, because all firms engage in precautionary saving. We simulate an aggregate economy and we show that the precautionary saving effect implies that aggregate investment can have large and persistent responses to temporary macroeconomic shocks. Moreover the heterogeneity of entrepreneurs with respect to the net worth implies that smaller businesses are on average more financially constrained, even though all are ex ante identical regarding their ability to access external finance. This is because small entrepreneurs tend to have smaller net worth relative to output, either because they have younger growing businesses, or because they became small after experiencing low productivity periods recently. This implies that investment of smaller businesses is much more volatile in response to temporary and permanent macroeconomic shocks.

This chapter is structured as follows: in section 3.2 we discuss the importance of precautionary saving in the context of the literature on investment with financing imperfections. In section 3.3 we illustrate the model. In section 3.4 we obtain a numerical solution of the problem and illustrate the properties of the policy function. In section 3.5 we provide some simple aggregate simulations. In section 3.6 we summarise the conclusions.

## 3.2 Financing constraints, precautionary saving and investment.

Before illustrating the formal model, let us explain the motivation behind it: in the context of dynamic investment models a binding financing constraint implies a corner solution to the optimal investment problem. The entrepreneur (henceforth $E$) would like to invest more in her business, but cannot find the necessary financing. When such corner solution holds $E$ reinvests in her business all internally generated finance, and the neoclassical investment model is rejected because of excess investment sensitivity to cash flow. The cash flow-investment literature tests for the positive correlation between investment and changes in internal finance for firms who *a priori* are more likely to have a binding financing constraint. This implicitly assumes that investment is not (or less) sensitive to cash flow when the constraint is not binding. But how plausible is this assumption? Intuitively an optimising agent would try to anticipate future financing problems by saving more when the economic conditions are favourable. Such "precautionary saving" could prevent financing constraints to be binding in most of the states of nature.

Yet this does not necessarily imply that financing constraints have a negligible impact on firm level and aggregate investment. Let's consider an entrepreneur who allocates wealth between consumption, risky investment in her business and safe investment in financial assets (precautionary saving). If $E$ wants to increase the precautionary saving to counter future expected financing problems, she must compensate by reducing either investment or consumption. Hence if the increase in saving is counterbalanced by a reduction in investment in the business rather than by a reduction in consumption, then financing constraints may have a large impact on investment dynamics, even if they are seldom binding at the individual level. This may happen in reality as consumption may be difficult to reduce, either because $E$'s household has a minimum consumption requirement, or because of a preference for consumption smoothing.

This discussion clarifies that, in order to study the importance of financing constraints for investment dynamics, it is important to focus on the link between financing problems and the allocation of wealth between consumption, saving and investment. Yet so far precautionary saving has been studied exclusively[1] in the consumption literature. Leland (1968) defines precautionary saving as the difference between consumption when income is certain and when it is uncertain but with the same mean. The literature that followed has extended Leland's results in the multiperiod case and for different sources of uncertainty. Caballero (1991) notes that "*The chief conclusion of this research agenda is that whenever preferences can be characterized by a separable utility function with convex marginal utility, the slope of the consumption path rises as the level of income uncertainty increases*". The steeper the consumption path, the more individuals engage in precautionary saving and postpone consumption to the future. Zeldes (1989) notes that precautionary saving can be also the result of liquidity constraints: *"However, even if the current constraint is not binding, so that the Euler equation between time t and time t+1 is satisfied, the presence of constraints that will bind in the future with some positive probability will lower the current consumption of any risk averse individual".*

Recent empirical studies have found evidence of precautionary saving on households behaviour (see, among others, Guiso, Jappelli and Terilzzese, 1992; Carrol, 1994; Hubbard, Skinner and Zeldes, 1994; Carrol and Samwick, 1998). Even assuming that liquidity constraints do not matter and that such precautionary saving is only generated by income uncertainty, one would expect that precautionary saving should be larger for entrepreneurial households, who bear a lot of the income risk. Yet no systematic study has been done, until recently, regarding the relationship between entrepreneurship and savings: Hubbard and Gentry (2000) analyse data from the 1983 and 1989 Federal Reserve Board Surveys of Consumer Finances, and find the following: i) entrepreneurial households own a substantial share of households wealth and income, and this share increases throughout the distributions of wealth and income; ii) wealth-income ratios are higher for entrepreneurial households and saving-income ratios are higher for entrants and continuing entrepreneurs, even after controlling for age and demographic variables; iii) the portfolios of entrepreneurial households, even wealthy ones, are undiversified, with the bulk of assets held within active businesses.

These findings suggest that entrepreneurship has a positive influence on precautionary

[1] An exception is Gross (1994). See footnote n.4 in the introduction to the thesis.

saving and wealth accumulation, and that it is important to study the effects of precautionary saving on investment. This is the aim of the basic model illustrated in the remaining sections of this chapter.

## 3.3 A model of investment with financing constraints

We consider an economy composed of many entrepreneurs, some active and some retired, and many competitive banks. We assume that each active $E$ chooses consumption and investment in order to maximise the expected value of her lifetime utility function. All entrepreneurs have the same preferences and have access to the same risky technology.

- **Preferences**

*p1) E has a utility function linear in consumption; p2) her subjective discount rate is equal to* $1/R$, *where* $R = 1 + r$, *and* $r$ *is the lending/borrowing risk free interest rate.*

Assumptions *p1* and *p2* are those implicitly or explicitly adopted in the cash-flow investment literature, which usually assumes that the firm maximises the discounted sum of future expected profits. If markets are perfect and the Modigliani Miller theorem holds, then this is the correct objective function regardless of the preferences of $E$. But when markets are not perfect, then preferences matter for investment decisions and to use the same objective function is equivalent to impose assumptions *p1* and *p2*. Such assumptions imply that $E$ is indifferent between consuming today and consuming in the future. As a consequence in presence of financing constraints she consumes nothing and saves all her income until there is even the smallest chance of facing future financing constraints. Therefore her investment would be financially constrained only at the very beginning of her activity, because she would save and quickly accumulate financial wealth. In reality entrepreneurs must consume something to survive, and they also tend to increase consumption as their income and wealth increases. Such behaviour is represented by the following assumption:

*p3) E has to consume at least a fixed share* $\eta > 0$ *of output to be able to continue activity:*

$$x_t^* = x_t - \eta y_t \geq 0 \tag{3.1}$$

We define $(1+\eta)y_t$ as total output. This is composed by $\eta y_t$, the minimum consumption level and $y_t$, defined by (3.2), that is "financial" output that can be consumed, invested or saved. $x_t$ is total consumption, and $x_t^*$ is "voluntary" consumption above the minimum level.

Therefore assumption *p3* ensures that consumption is always positive and increasing in income[2]. In the consumption literature the same features (positive consumption increasing in wealth) are usually generated by assuming a concave utility function (see Zeldes, 1989, for a model of consumption with liquidity constraints). We will show in section 3.3.2 that the two modelling choices generate similar results in terms of the effects of future expected financing problems on current investment and saving decisions. Our approach[3] has the advantage of preserving the linearity of the preferences and of rendering the solution of the model much simpler. This feature is very important for the extended model proposed in chapter 4, which is already considerably complex in that it analyses, in a multifactor model, financing constraints together with the irreversibility of fixed capital.

Another way of justifying assumption *p3* is to interpret $E$ as a manager that receives private benefits (Jensen, 1986) proportional to the firm's output. Her objective is to maximise her intertemporal utility, and she receives from the shareholders a wage proportional to financial profits. Jensen (1986) argues that in a situation with free cash flow and private benefits a manager may become an "empire builder": she may invest in inefficient projects in order to increase her power and her privileges in the company. This overinvestment problem is a possible way of interpreting the consequences of assumption *p3* in our model. In fact in section 3.3.1 we show that the presence of such private benefits implies that the investment is inefficiently high with respect to the level that maximises financial profits. This is because while the wage of the manager is proportional to financial profits, her private benefits are proportional to the output of the firm. Therefore she wants to increase the size and the output of the firm beyond the efficient level. In this situation we show that future expected financing constraints reduce the investment of the firm because the manager values more financial profits, which increase financial wealth and reduce future

[2]Because we assume that productivity shocks are persistent (assumption *t3*), *p3* also implies that consumption is, on average, increasing in wealth.

[3]Kiyotaki and Moore (1997) use a similar assumption in a business cycle model.

financing constraints, than private benefits, which do not affect the wealth of the firm.

Therefore, although the model primarily applies to a small-medium firm owned and managed by the entrepreneur, its main results can also be extended to larger firms with separation between ownership and management.

Finally, we assume that the expected lifetime of $E$ is finite: *p4) each period $E$ may become ill with an exogenous probability $1-\gamma$, with $0 < \gamma < 1$, and must retire. Retirement is not reversible. Retired entrepreneurs sell their assets*[4] *and consume all the proceedings.*

At the aggregate level we assume that each retired entrepreneur is replaced by a newly born one, in order to keep the population of entrepreneurs constant. This assumption ensures that $E$ discounts future consumption. Later we will show that this allows for the presence of expected financing constraints in equilibrium.

- **Technology**

*t1) $E$ can invest in the risky technology of the firm she owns and manages. $k_t$ is the stock of capital, installed at time $t-1$, which will generate output at time $t$.*

The assumption that investment takes one period to become productive ensures that financing constraints are relevant when $E$ needs additional funds to exploit new investment opportunities.

*t2) Net financial output $y_t$ is produced according to a Cobb-Douglas production function:*

$$y_t = \theta_t k_t^{\alpha} \tag{3.2}$$

with $\alpha < 1$. All prices are assumed constant and normalised to 1. Capital depreciates at the rate $\delta_k$:

$$1 > \delta_k > 0 \tag{3.3}$$

*t3) $\theta_t$ is a productivity shock that follows a stationary autoregressive stochastic process.* In this and in most of the following chapter, for simplicity, we assume that $\theta_t$ is a symmetric two state Markov process, even though all the theoretical results can be easily generalised

[4] They cannot sell the "goodwill", as $E$'s work is essential to generate output using her risky technology.

for a more complicated stationary stochastic process:

$$\theta_t \in \{\theta_L, \theta_H\} \ with \ \theta_H > \theta_L; \qquad pr(\theta_{t+1} = \theta_t) = \epsilon > 0.5; \qquad pr(\theta_{t+1} \neq \theta_t) = 1 - \epsilon \tag{3.4}$$

Hence the first order autocorrelation coefficient is $\rho = 2\epsilon - 1 > 0$, and we have that:

$$E_t(\theta_{t+1} \mid \theta_t = \theta_H) > E_t(\theta_{t+1} \mid \theta_t = \theta_L) \tag{3.5}$$

- **Financial Markets**

*f1) Equity finance and risky debt are not available*; *f2) at time t E 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 $E$ is a net borrower (lender); *f3) 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:

$$b_{t+1} \leq \tau_k k_{t+1} \tag{3.6}$$

$\tau_k$ is the share of value of capital that can be used as collateral[5]:

$$\tau_k \leq (1 - \delta_k) \tag{3.7}$$

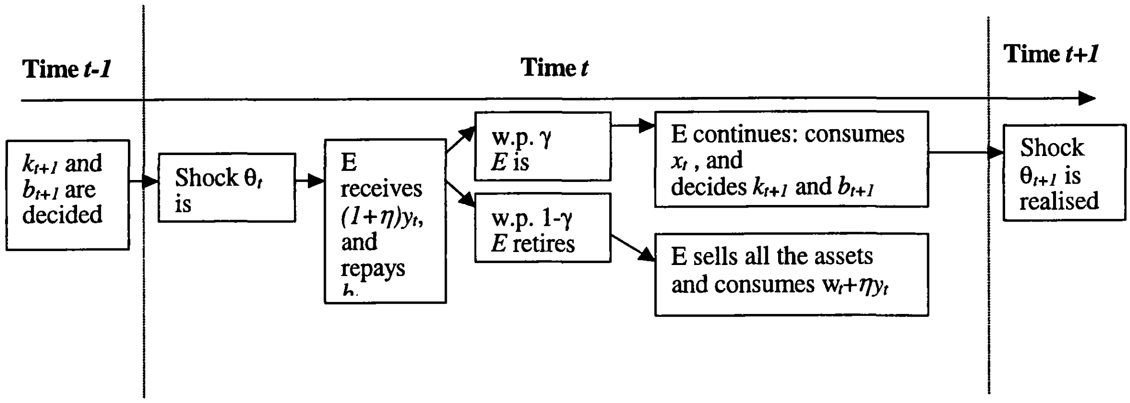
The rationale for assumptions *f1-f3* is that $E$ can hide the revenues from the production. Being unable to observe such revenues the lenders can only claim, as repayment of the debt, the value of $E$'s physical assets (Hart and Moore, 1998). Therefore $E$ can only lend or borrow one period secured debt at the market interest rate $r$ offered by the banks[6]. This collateral constraint assumption is supported by the empirical evidence illustrated in chapter 2, which shows that, in our sample of manufacturing firms, those entrepreneurs who state financing problems are much less collateralised than those who do not (see figure

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

[6] We also implicitly assume that in the case of forced liquidation to repay the debt $E$ can continue the business with her residual wealth. Otherwise it is possible that the ex post liquidation threat by the lender during the renegotiation process is such that $E$ can credibly commit ex ante to repay more than $\tau_k k_{t+1}$.

2-8). The timing of the model is represented in figure 3-1. At the beginning of time $t$ $E$

Figure 3-1: The timing of the basic model




inherits from time $t-1$ the stocks of capital $k_t$. Then $\theta_t$ is realised, $(1+\eta)\,y_t$ is produced and $b_t$ repaid. Residual wealth $w_t$, net of the minimum consumption level $\eta y_t$ is:

$$w_t = y_t + (1-\delta_k)k_t - b_t \tag{3.8}$$

After producing, with probability $1-\gamma$ $E$ becomes ill and retires. With probability $\gamma$ she instead continues activity. She borrows new one period debt with face value $b_{t+1}$, receiving the discounted value $b_{t+1}/R$. The net worth $w_t$ plus the new borrowing $b_{t+1}/R$ are allocated between consumption and investment[7]. Therefore $E$'s budget constraint is the following:

$$x_t^* + k_{t+1} = w_t + b_{t+1}/R \tag{3.9}$$

Let's define $\Theta$ as the vector of structural parameters of the problem: $\Theta' = \{\theta_L, \theta_H, \epsilon, \gamma, R, \delta_k, \tau_k, \eta, \alpha\}$. In order to analyse the problem, it is useful to consider $w^{MAX} = w(\Theta)$, which we define as the level of wealth beyond which $E$ has zero probability to face future financing problems. It is then easy to prove the following:

**Proposition 1** *(i)* $\lim_{\eta\to 0} w^{MAX} > 0$;

*(ii)* if $w_t \geq w^{MAX}$ then conditional on not retiring $pr(w_{t+j} < w^{MAX}) = 0$ for $j = 0, 1, ..., \infty$;

[7] If $b_{t+1}$ is negative, then (3.9) shows that $w_t$ is allocated between consumption ($x_t^*$), risky investment ($k_{t+1}$) and risk free lending $(-b_{t+1})/R$.

*(iii)* $\partial w^{MAX}/\partial\eta > 0$;

*(iv) there exist a finite value of* $\eta$, *called* $\eta_{\min}$, *such that* $pr(w_t > \overline{w}) = 0$ for $t = 0, 1, ..., \infty$, and $\overline{w} \leq w^{MAX}$.

**Corollary 2** *if* $\eta \geq \eta_{\min}$ *then an active* $E$ *has always some probability of facing future financing constraints, and she does not consume more than the minimum amount, i.e.* $x_t = \eta y_t$ *and* $x_t^* = 0$ *for* $t = 0, 1, ..., \infty$.

**Proof: see appendix 1.**

Intuitively when $E$ reaches $w^{MAX}$ she becomes so rich that her financial wealth can absorb any future negative productivity shock and hence she does not have any future financing problem. Given that she can costlessly postpone consumption above the minimum level, she will do so in order for her wealth to always remain above $w^{MAX}$. $\eta$ affects both the value of $w^{MAX}$ and the probability of reaching it. $\eta_{\min}$ is the minimum consumption share which ensures that, starting from $w_0 < \overline{w}$, $w^{MAX}$ is never reached, and there is always some probability of being financially constrained in the future. The intuition is that $\eta$ limits the capacity of $E$ to retain earnings and increases her incentives to expand production to boost consumption. As $w_t$ increases, $E$ becomes richer and cares less about financing constraints. Thus she increases the size of the firm to increase her consumption, even if it implies investing in less profitable projects. This means that her expected financial profits decrease, and that there is always a positive probability that, conditional on future negative productivity shocks, $w_t$ will decrease enough to push $E$ in the constrained region[8].

For the rest of this chapter and the two following we assume[9] that $\eta \geq \eta_{\min}$. Corollary 2 implies that constraint (3.1) is always binding with equality, because with future expected financing problems $E$ always prefers to postpone consumption. Therefore consumption is no longer a state variable of the problem. $x_t$ is predetermined at the beginning of time $t$, when $\eta y_t$ is realised, and $x_t^* = 0$ for any $t = 0, 1, ...\infty$. We denote the expected

[8]In section 3.3.2 we show that a bounded stochastic process for $w_t$ can be generated with a concave utility function plus an intertemporal discount factor $\beta > 1/R$.

[9]This assumption is not restrictive, given that typically, for a wide range of parameters, $\eta_{\min}$ is around 0.3-0.4, which corresponds to consumption being only 23%-28.5% of total output.

lifetime utility at time $t$ of $E$, after $\theta_t$ is realised, conditional on not becoming ill and after consuming $\eta y_t$, by $V_t(w_t, \theta_t \mid w_0, \theta_0; \Theta)$ :

$$V_0(w_0, \theta_0 \mid \Theta) = \underset{\begin{pmatrix} k_{t+1} = k(w_t, \theta_t) \\ b_{t+1} = b(w_t, \theta_t) \end{pmatrix}_{t=0,1,\ldots\infty}}{MAX} E_0 \left\{ \sum_{t=0}^{\infty} \left(\frac{\gamma}{R}\right)^t \left\{ \frac{1}{R} [\eta y_{t+1} + (1-\gamma) w_{t+1}] \right\} \right\} \tag{3.10}$$

$w_t$ and $\theta_t$ are the state variables of the problem, which is defined by (3.10) subject to (3.6) and (3.9). These constraints define a compact and convex feasibility set for $k_{t+1}$ and $b_{t+1}$, and the law of motion of $w_{t+1}$ conditional on $w_t$ and $\theta_t$ is continuous. Therefore, given assumption *t3* and the concavity of the production function, a solution to the problem exists and is unique[10]. In order to describe the optimality conditions of the model, let $\lambda_t$ and $\phi_t$ be the Lagrangean multipliers associated to constraints (3.6) and (3.9). Taking the first order conditions of (3.10), at a generic time $t$, with respect to $b_{t+1}$ and $k_{t+1}$ it is possible to show that the solution is given by the optimal sequence of $\{k_{t+1}, b_{t+1}, \lambda_t, \phi_t \mid w_t, \theta_t; \Theta\}_{t=0}^{\infty}$ that satisfies (3.11), (3.12) and (3.13) plus the standard Kuhn-Tucker complementary slackness conditions on $\lambda_t$:

$$Dk_{t+1} \leq w_t \tag{3.11}$$

$$\phi_t = 1 + RE_t \left[ \sum_{j=0}^{\infty} \gamma^j \lambda_{t+j} \right] \tag{3.12}$$

$$(1+\eta) E_t(MPK_{t+1}) = UK + E_t(\Omega_{t+1}) + DR^2 \lambda_t \tag{3.13}$$

$D = 1 - \tau_k / R$ is the downpayment required to purchase one additional unit of capital. Equation (3.11) combines together the budget constraint (3.9) and the collateral constraint (3.6) and implies that the downpayment necessary to buy $k_{t+1}$ must be lower than $E$'s net worth. Equation (3.12) is obtained by solving recursively forward the first order condition for $b_{t+1}$. Equation (3.13) is obtained using the first order condition for $b_{t+1}$ to substitute $\phi_t$ in the first order condition for $k_{t+1}$. $UK = R - (1 - \delta_k)$ is the user cost of capital,

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

and $E_t(MPK_{t+1})$=$\alpha E_t(\theta_{t+1})\, k_t^{\alpha-1}$ is the expected marginal productivity of capital in financial terms. $E_t(\Omega_{t+1})$ is the premium required by $E$ to compensate for the cost of future expected financing problems, and it is defined as follows:

$$E_t(\Omega_{t+1}) = \gamma E_t\left[(\phi_{t+1} - 1)\,(UK - MPK_{t+1})\right] \tag{3.14}$$

$\lambda_t$ is positive when the constraint (3.6) is binding, and is equal to zero otherwise. It represents the shadow cost of not being able to increase investment because of the lack of additional funds.

### 3.3.1 The precautionary saving effect

Let's first consider the solution without financing problems, by assuming that $w_0 \geq w^{MAX}$. In this case proposition 1 implies that $E$ has always enough resources to invest in her business. The collateral constraint (3.6) is never binding, and its associated Lagrange multiplier $\lambda_t$ is constantly equal to 0 for $t = 0, 1, ..., \infty$. Hence from (3.12) it follows that $\phi_t = 1$ for $t = 0, 1, ..., \infty$. Hence one additional unit of net worth increases the value function only by one. This is because, with linear utility and no financing constraints, expected marginal return from investment is exactly equal to $R$, the discount factor. Moreover this implies, from (3.14), that $E_t(\Omega_{t+1}) = 0$ for $t = 0, 1, ..., \infty$. By substituting these values in (3.13), we obtain equation (3.15):

$$(1+\eta)\, E_t\left(MPK_{t+1} \mid w_t \geq w^{MAX}\right) = UK \tag{3.15}$$

By solving (3.15) we obtain $k_{t+1}^{PM}$, the optimal investment level when $E$ has zero probability to face future financing constraints:

$$k_{t+1}^{PM} = \left[\frac{(1+\eta)\,\alpha E_t(\theta_{t+1})}{UK}\right]^{\frac{1}{1-\alpha}} \tag{3.16}$$

Equation (3.15) looks like a standard profit maximising conditions, with one important difference: $\eta E_t(MPK_{t+1})$ is the share of marginal revenues that will be consumed, while

$E_t\left(MPK_{t+1}\right)$ is marginal financial revenues. Since $\eta > 0$, it follows that:

$$E_t\left(MPK_{t+1} \mid k_{t+1} = k_{t+1}^{PM}\right) < UK \tag{3.17}$$

(3.17) implies that conditional on $k_{t+1}^{PM}$ expected marginal financial revenues are negative. We define $k_{t+1}^{*}$ as the level of investment that maximises net financial profits, and satisfies the following condition:

$$E_t\left(MPK_{t+1}\right) = UK \tag{3.18}$$

By solving (3.18) for $k_{t+1}$ we obtain the value of $k_{t+1}^{*}$ :

$$k_{t+1}^{*} = \left[\frac{\alpha E_t\left(\theta_{t+1}\right)}{UK}\right]^{\frac{1}{1-\alpha}} \tag{3.19}$$

Since $\eta > 0$ it follows that, because of the consumption constraint, $E$ overinvests with respect to the level of investment that maximises financial profits. Therefore:

$$k_{t+1}^{PM} > k_{t+1}^{*} \tag{3.20}$$

Let's consider now the case for $w_0 \in [0, \overline{w}]$. In this case $w_t$ is a stochastic process bounded above by $\overline{w}$. We define $\underline{w}_t = \underline{w}(\theta_t) < \overline{w}$ as the minimum level of net worth such that $E$ is not financially constrained at time $t$ and (3.6) is not binding. Therefore if $w_t < \underline{w}_t$ (3.11) is binding with equality and together with (3.13) determines $\lambda_t^{CB}$ and $k_{t+1}^{CB}$, where the superscript $CB$ denotes the constrained solution:

$$k_{t+1}^{CB} = \frac{w_t}{D} \tag{3.21}$$

$$\lambda_t^{CB} = \frac{(1+\eta)\, E_t\left(MPK_{t+1}\right) - UK - E_t\left(\Omega_{t+1}\right)}{DR^2} \tag{3.22}$$

In this situation $E$ borrows up to the limit without being able to exhaust all profitable investment opportunities. Therefore she invests all her resources in productive capital. (3.21) means that one additional unit of financial wealth allows $E$ to increase $k_{t+1}$ by $1/D$ units. This positive correlation between changes in wealth and investment is the key assumption of the cash flow-investment literature. However our model shows that for an entrepreneur a binding financing constraint is a very costly outcome in terms of lost investment opportunities. The cost is measured by $\lambda_t^{CB}$ as defined by (3.22), and $E$ will try to prevent it by engaging in "precautionary saving".

We explain the intuition behind the "precautionary saving effect" by noting that, from (3.12), $\phi_t$ increases in the discounted sum of the expected $\lambda_{t+j}$ for $j = 1, ..., \infty$. The intuition is that the opportunity cost of money for $E$ increases when such money is useful in reducing the chance of being financially constrained in future. Therefore future expected financing problems increase the required return on capital and reduce the optimal unconstrained investment level below $k_{t+1}^{PM}$. To show this, let's consider $k_{t+1}^{o}$, the optimal solution for $w^{MAX} > w_t > \underline{w}_t$. In this case $E$ is unconstrained today and $\lambda_t = 0$, but she will be constrained in future with some positive probability, so that $E_t \left[ \sum_{j=1}^{\infty} \gamma^j \lambda_{t+j} \right] > 0$. Therefore from (3.13) it follows that $k_{t+1}^{o}$ satisfies the following condition:

$$(1 + \eta) E_t (MPK_{t+1}) = UK + E_t (\Omega_{t+1}) \tag{3.23}$$

$E_t (\Omega_{t+1})$, as defined by (3.14), is the product between the value of money in terms of its ability to reduce future expected financing problems, $\phi_{t+1} - 1$, multiplied by the loss in monetary profit caused by the overinvestment problem, $UK - MPK_{t+1}$. $E_t (\Omega_{t+1})$ is equal to zero if there are no expected financing problems 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 problems. The existence of this well behaved solution depends on assumption *p4,* which implies that $E$ discounts future expected financing constraints at the rate $\gamma < 1$. In fact from corollary 2 it follows that $\eta \geq \eta_{\min}$ corresponds to an equilibrium where $E$ has always some future expected financing problems. Therefore $\gamma < 1$ ensures that, despite the sequence $\{E_t (\lambda_{t+j})\}_{j=0}^{\infty}$ converges to a positive constant, from (3.12) it follows that the shadow value of money is finite: $\phi_t < \infty$. The intuition is the following: $E$ knows that by not saving one additional unit of financial wealth she increases the expected financing constraints in all future periods, but she does not care about the distant future because she will probably be retired by then.

By comparing (3.15) with (3.23), it is easy to prove the following proposition:

**Proposition 3** *For* $t = 0, 1, ..., \infty$ :

i) $k^*_{t+1} < k^o_{t+1} < k^{PM}_{t+1}$ ;

ii) $\left(\partial k^o_{t+1}/\partial w_t \mid w_t < w^{MAX}\right) > 0$;

iii) $\lim_{w_t \to w^{MAX}} k^o_{t+1} = k^{PM}_{t+1}$.

**Proof: see appendix 1.**

Proposition 3.i means that the optimal unconstrained capital $k^o_{t+1}$ is lower than the optimal capital with no future expected financing problems $k^{PM}_{t+1}$. By choosing $k^o_{t+1}$ smaller than $k^{PM}_{t+1}$ $E$ chooses to precautionary reduce production and consumption, in order to increase the expected return on capital and the share of $w_t$ invested in risk free assets[11]. Proposition 3.ii means that changes in net worth affect investment choices even when the constraint is not binding, because they change the expectations about future financing problems and the amount of precautionary saving. From the above discussion it follows that $\widetilde{k}_{t+1}(w_t, \theta_t)$, the level of capital that solves the problem, is the following:

$$\widetilde{k}_{t+1}(w_t, \theta_t) = \begin{cases} k^o_{t+1}(w_t, \theta_t) & \text{if } k^o_{t+1}(w_t, \theta_t) < w_t/D \text{ (unconstrained region)} \\ k^{CB}_{t+1} & \text{otherwise (constrained region)} \end{cases} \tag{3.24}$$

### 3.3.2 Precautionary saving and accumulation of financial wealth

The above discussion implies that an entrepreneur with future expected financing constraints accumulates net financial wealth. In fact equation (3.23) can be interpreted in the following way: the higher $E_t(\Omega_{t+1})$, the more **E** reduces investment to increase marginal financial profits. Since $k^o_{t+1} > k^*_{t+1}$, it follows that this increases also total profits and the net financial wealth of $E$. Therefore (3.23) means that at the margin the advantage of reducing investment to increase financial wealth is equal to the cost of reducing output and consumption.

[11]Given that lending and borrowing rates are the same, $E$ is indifferent between maintaining some additional borrowing capacity and borrowing up to the limit to invest that additional capacity in risk free lending. Therefore to say that (3.11) is not binding is equivalent to say that $E$ is investing part of her wealth in risk free assets.

The optimal choice of $k^o_{t+1}$ defines the optimal level of net financial wealth but it does not determine how it should be allocated between debt and financial assets. In fact assumptions p1 and p2 imply that if the borrowing constraint is not binding, then $E$ is indifferent between borrowing up to the limit and investing in financial assets and keeping some spare borrowing capacity. Therefore if **E** engages in precautionary saving, the model predicts that this can happen using two equivalent mechanisms: i) accumulation of financial assets, which could be liquidated if needed to finance new investment; ii) reduction of the level of debt. The firm maintains a low level of debt in order to leave some room to increase it if future financing need will arise.

Another mechanism, not allowed by the simple model in this chapter, which can be used in reality by a firm in order to reduce future expected financing problems, is to delay the investment. If a firm has the chance to delay a profitable investment opportunity, then future expected financing problems could influence the delay decision: the firm waits to invest if the cost of delaying the new project is lower than the gain in terms of having more funds to counter future expected financing problems on the existing projects.

### 3.3.3 Solution with a concave utility function

In the previous sections we showed that the minimum consumption share (assumption *p3*) generates an equilibrium with two important properties: i) the financial wealth of $E$ follows a bounded stochastic process; ii) future expected financing problems affect the allocation of wealth between saving and investment.

Assumption *p3* is sufficient but not necessary to obtain these two results. As a robustness check of the theoretical results derived in the previous section, in this subsection we derive the optimality conditions of a version of the model where the preferences of $E$ are defined by a concave utility function instead of by a linear utility function plus a minimum consumption share $\eta > 0$. Our aim is to illustrate the similarities of the two approaches in terms of the effect of expected financing constraints on investment. More specifically, assumptions *p1, p2, p3* and *p4* are replaced by the two following assumptions:

*p1') $E$'s utility is a logaritmic function of consumption:*

$$U(x_t) = \ln x_t \tag{3.25}$$

*p2') E's subjective discount rate is equal to $\beta$, where $0 < \beta < 1/R$*

Assumption *p2'* ensures that the entrepreneur is "impatient", and prefers to consume today rather than to postpone consumption. The budget constraint is now the following:

$$x_t + k_{t+1} = w_t + b_{t+1}/R \tag{3.26}$$

Where $w_t$ is still defined by (3.8). The value function of $E$ is the following:

$$V_0(w_0, \theta_0) = \underset{\{k_{t+1}, b_{t+1}\}_{t=1}^{\infty}}{MAX} E_0 \left\{ \sum_{t=0}^{\infty} \beta^t \{\ln x_t\} \right\} \tag{3.27}$$

The problem is defined by (3.27) subject to (3.6) and (3.26). We substitute $x_t$ in (3.27) using (3.26) and we add the collateral constraint with the associated Lagrange multiplier $\lambda_t$. By differentiating (3.27) with respect to $b_{t+1}$ and $k_{t+1}$ we obtain the following:

$$\frac{1}{x_t} = R\lambda_t + R\beta E_t \left( \frac{1}{x_{t+1}} \right) \tag{3.28}$$

$$E_t(MPK_{t+1}) = \overline{UK} + \frac{\frac{D}{\beta} R\lambda_t - cov\left( \sum_{j=1}^{\infty} (R\beta)^j \lambda_{t+j}, MPK_{t+1} \right)}{\sum_{j=1}^{\infty} (R\beta)^j E_t(\lambda_{t+j})} \tag{3.29}$$

(3.28) and (3.29) are the Euler equations for consumption and investment respectively. (3.29) is obtained by taking the first order condition for $k_{t+1}$ and then by substituting $\frac{1}{x_t}$ and $E_t\left(\frac{1}{x_{t+1}}\right)$ from (3.28) solved recursively forward. The user cost of capital becomes $\overline{UK} = 1/\beta - (1 - \delta_k)$. $\overline{UK}$ is greater than $UK$ because assumption *p2'* implies that $1/\beta > R$. The intuition is that the impatient $E$ requires a higher profitability to be convinced to invest rather than consume. If $w_0 \geq w^{MAX}$ and $E$ is never constrained then $\lambda_t = 0$ for $t = 0, 1, ..., \infty$. In this case the last term at the right hand side of (3.29) is equal to zero, and (3.29) becomes equivalent to (3.18). Moreover consumption follows a martingale process:

$$x_t = \frac{1}{R\beta} E_t(x_{t+1}) \tag{3.30}$$

from (3.30) it is clear that present consumption is higher than future expected one since $\frac{1}{R\beta} < 1$. Therefore assumption *p2'* ensures that, if $\beta$ is small enough, $w_t$ follows a bounded stochastic process, because $E$ prefers to consume today rather than postpone consumption and accumulate financial wealth.

If $w_t < \underline{w}_t$ and the financing constraint is binding, then $k_{t+1}$ and $E_t(MPK_{t+1})$ are determined by (3.11) holding with equality and (3.29) determines $\lambda_t > 0$. Therefore also in this alternative model $\lambda_t$ can be interpreted as the shadow cost of a binding collateral constraint. From (3.28) it is clear that, the higher $\lambda_t$, the lower consumption $x_t$. Therefore current financing constraints affect the intertemporal allocation of consumption. This is what happens in Zeldes (1989) as well as in the model described in the previous section.

Let's consider now what happens to optimal investment choices. We consider the case with $w_0 < \overline{w}$, $\theta_t = \theta_L$ and $w_t > \underline{w}_t$, when $E$ is wealthy enough not to be financially constrained today, $\lambda_t = 0$, but faces a positive probability of being constrained in future, $\sum_{j=1}^{\infty} (R\beta)^j E_t(\lambda_{t+j}) > 0$. In this case equation (3.29) becomes the following:

$$E_t(MPK_{t+1}) = \overline{UK} - \frac{cov\left(\sum_{j=1}^{\infty} (R\beta)^j \lambda_{t+j}, MPK_{t+1}\right)}{\sum_{j=1}^{\infty} (R\beta)^j E_t(\lambda_{t+j})} \tag{3.31}$$

It is easy to see that (3.31) is similar to (3.23). In fact the term $-\frac{cov\left(\sum_{j=1}^{\infty} (R\beta)^j \lambda_{t+j}, MPK_{t+1}\right)}{\sum_{j=1}^{\infty} (R\beta)^j E_t(\lambda_{t+j})}$ represents the premium in the user cost of capital induced by future expected financing constraints. Since assumption *p2'* implies that $R\beta < 1$, this ensures that the cost of future expected financing constraints $\sum_{j=1}^{\infty} (R\beta)^j E_t(\lambda_{t+j})$ is finite[12]. The numerator $cov\left(\sum_{j=1}^{\infty} (R\beta)^j \lambda_{t+j}, MPK_{t+1}\right)$ is the covariance between future expected financing problems and marginal productivity of capital. If the technological shock is persistent, and $w_t - \underline{w}_t$ is relatively large, then this covariance term is positive, because the constraint becomes binding at time $t+1$ only conditional on a positive shock ($\theta_{t+1} = \theta_H$) that increases investment needs. The denominator $\sum_{j=1}^{\infty} E_t(\lambda_{t+j})$ is also positive and hence the

[12]The same property is ensured, in the basic model, by assumption *p4*.

term $\frac{cov\left(\sum_{j=1}^{\infty}(R\beta)^j\lambda_{t+j},MPK_{t+1}\right)}{\sum_{j=1}^{\infty}(R\beta)^j E_t(\lambda_{t+j})}$ is positive as well, and reduces the expected marginal productivity on the left hand side of (3.31). Hence it increases the optimal level of capital. This is because $E$ wants to increase production contingent to a positive shock that also increases the shadow value of retained earnings. As $w_t$ decreases towards $\underline{w}_t$, the denominator $\sum_{j=1}^{\infty} E_t(\lambda_{t+j})$ increases, because there is a higher chance of being constrained in future. Instead the numerator $cov\left(\sum_{j=1}^{\infty}\lambda_{t+j}, MPK_{t+1}\right)$ decreases, because now also a negative shock that reduces $MPK_{t+1}$ can increase future expected financing constraints. The intuition is that, if $w_t$ is very close to $\underline{w}_t$, a negative shock ($\theta_{t+1} = \theta_t = \theta_L$) does not modify expected productivity, but generates negative returns in period $t+1$ which reduce net worth and can drive $w_t$ below $\underline{w}_t$. This effect can be strong enough to make the covariance term in the numerator negative. The result is that, as $w_t$ decreases towards $\underline{w}_t$, the term $\frac{cov\left(\sum_{j=1}^{\infty}\lambda_{t+j},MPK_{t+1}\right)}{\sum_{j=1}^{\infty}E_t(\lambda_{t+j})}$ decreases. From (3.31) it follows that the required return on capital increases and the desired level of investment in the firm decreases. This mechanism is analogous to the precautionary saving effect described in the previous section.

## 3.4 Numerical solution and simulations

We now solve the problem and illustrate the properties of the investment policy function. We solve the dynamic nonlinear system of equations defined by (3.11), (3.12) and (3.13) using a numerical method. Adding the subscript i to indicate the *i-th* entrepreneur[13], we discretise the state space of state variables $w_{i,t}$ and $\theta_{i,t}$ and we solve the problem numerically iterating between the value function and the first order conditions, until the value function converges (Judd, 1998). First, we formulate an initial guess of the forward variables $E_t\left(\phi_{i,t+1}\right)$ and $E_t\left(\phi_{i,t+1}\theta_{i,t+1}\right)$[14]. Second, we solve the static optimisation problem conditional on this guess, for each discrete value of the state variables $w_{i,t}$ and $\theta_{i,t}$. Third, we update the guess of $E_t\left(\phi_{i,t+1}\right)$ and $E_t\left(\phi_{i,t+1}\theta_{i,t+1}\right)$. We repeat these steps until the value function converges. The numerical solution is obtained using the software package

[13]We adopt this notation from now on because we simulate an economy with many heterogeneous entrepreneurs.

[14]These are the two forward terms present in $E_t\left(\Omega_{t+1}\right)$.

Gauss. Details about the maximisation routines we developed to solve the problem are available upon request. Solving the problem for all possible values of the state variables yields the optimal policy function $k_{i,t+1}(w_{i,t},\theta_{i,t})$ and the associated Lagrange multipliers $\lambda_{i,t}(w_{i,t},\theta_{i,t})$ and $\phi_{i,t}(w_{i,t},\theta_{i,t})$, and the value function $V_{i,t}(w_{i,t},\theta_{i,t})$.

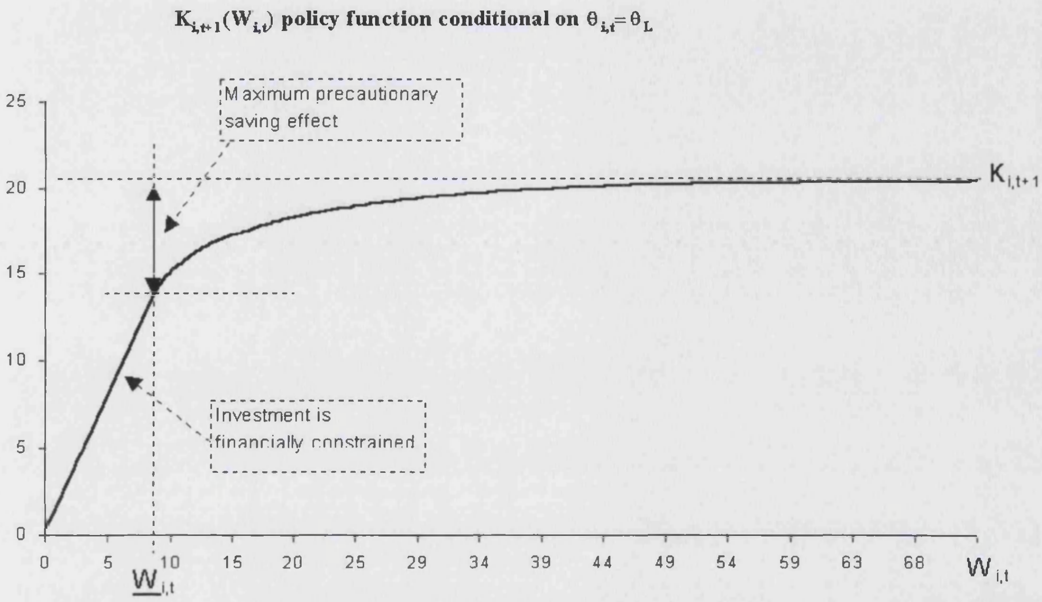
The main qualitative results of the basic model presented in this chapter hold for a very large range of parameters. The specific parameter values used in this and the next section are chosen in order to make the simulations comparable with those of the extended model presented in chapter 4, where $E$ employs a multifactor technology with fixed and variable capital. This choice is motivated by the fact that the simulations presented in the next chapter are aimed at matching more closely observed fixed capital and variable capital investment dynamics, and most of the parameters used there are directly estimated from our panel data of Italian manufacturing firms.

The parameters regarding productivity are the following: $\theta_H = 4.52$, $\theta_L = 5.71$, $\epsilon = 0.62$, $\alpha = 0.781$, $\delta_k = 0.6129$. The values of $\theta_H$ and $\theta_L$ and $\epsilon$ are chosen in order to match the level of risk faced by the entrepreneurs in the extended model (where there is also an aggregate source of uncertainty). The value of $\alpha$ ensures that the optimal level of capital in this single factor model is equal to the sum of optimal levels of fixed and variable capital in the extended model. The value of $\delta_k$ is equal to the weighted average of the depreciation factors for fixed and variable capital in the extended model.

The remaining parameters are the following: $\tau_k = 1 - \delta_k$, $R = 1.015$, $\eta = 0.4$ and $\gamma = 0.95$. The value of $\tau_k$ implies that the residual capital is fully collateralisable. The values of $R$, $\eta$ and $\gamma$ are the same ones chosen in chapter 4. The value of $\gamma$ implies that each period the exit rate is 5%. The value of $\eta$ implies that consumption is equal to a share of $\eta/(1+\eta) = 28.5\%$ of total revenues.

Figure 3-2 shows the policy function for capital $k_{t+1}(w_t)$ conditional on a negative productivity shock, $\theta_t = \theta_L$. For low values of $w_t$ the constraint is binding at time $t$, and $k_{t+1} = w_t/D$. This corresponds to the section of the policy function that is a diagonal line, with slope $1/D$. The denominator of the slope is the required down-payment for a unit of purchased capital. In the remaining part of the policy function $k_{t+1} = k^o_{t+1}(w_t;\theta_t = \theta_L) < k^{PM}_{t+1}(\theta_t = \theta_L)$. The distance between $k^o_{t+1}$ and $k^{PM}_{t+1}$ measures

Figure 3-2: Policy function of capital conditional on a low productivity shock



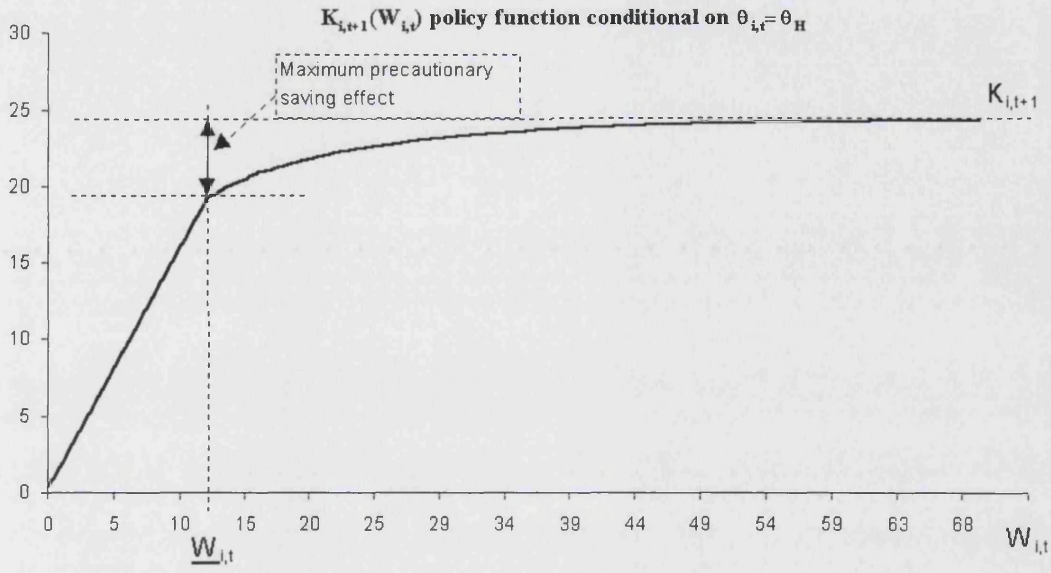

the extent to which the precautionary saving effect reduces optimal investment choices.. For $w_t = \underline{w}_t$, precautionary saving is responsible for a reduction in the optimal investment equal to $\frac{k_{t+1}^{PM} - k_{t+1}^{o}}{k_{t+1}^{PM}} = 30.7\%$ . The difference between $k_{t+1}^{o}$ and $k_{t+1}^{PM}$ converges to zero as $w_t$ increases, and the probability of facing future financing problems decreases.

Figure 3-3 shows the policy function for capital $k_{t+1}(w_t)$ conditional on a positive productivity shock, $\theta_t = \theta_H$. It shows that when $\theta_t$ increases the binding constraint region ($w_t < \underline{w}_t$) expands and the precautionary saving region ($w_t > \underline{w}_t$) shrinks. This is because $\theta_t$ is persistent. A higher $\theta_t$ means also higher productivity in future, and $E$ expects to increase financial profits and financial wealth, thereby reducing future probability of being financially constrained. The consequence is that the maximum precautionary reduction in optimal investment measured at $w_t = \underline{w}_t$ is only equal to $\frac{k_{t+1}^{PM} - k_{t+1}^{o}}{k_{t+1}^{PM}} = 20\%$ .

### 3.4.1 Simulation: single firm

Figure 3-4 shows the time path of $k_{t+1}$ for 50 periods, with a predetermined sequence of expansion and contraction phases. The firm's idiosyncratic shock is $\theta_H$ for 15 periods, then $\theta_L$ for 20 periods, and then again $\theta_H$ for the last 15 periods. Such arbitrary

Figure 3-3: Policy function of capital conditional on a high productivity shock




sequence is chosen in order to better distinguish between the three components of investment dynamics: change in expected productivity, precautionary saving and binding financing constraint. We simulate two entrepreneurs with different initial endowment $w_0^1$ and $w_0^2$, but otherwise identical. The first entrepreneur ($E^1$) has $w_0^1 > w^{MAX}$, and has a zero probability to be financially constrained at any time. Hence she changes her stock of capital only in response to changes in expected productivity: $k_{t+1}^1(w_t, \theta_t) = k_{t+1}^{PM}(\theta_t)\ \forall t$. The second entrepreneur ($E^2$) has $\underline{w} < w_0^2 < \overline{w}$. Therefore she is initially unconstrained but has a positive probability of being constrained in the future. When, after 15 periods, the productivity shock switches to $\theta_L$, she reduces her capital in a similar way with respect to the unconstrained entrepreneur. The difference is that she keeps reducing capital also in the following periods. From $t = 15$ to $t = 24$ the constraint is not binding, and $E^2$ reduces investment because the negative shocks reduce her financial wealth (see figure 3-5) and increase the chances of facing future financing problems. This is equivalent to precautionary saving, in the sense that $E^2$ borrows less and invests less, in order to keep some spare borrowing capacity for the future. From the point of view of figure 3-2, this drop in investment is equivalent to moving leftward along the $k_{i,t+1}$ policy function, while $w_{i,t}$ decreases and approaches $\underline{w}_t$. This precautionary behaviour implies that $E^2$ manages to avoid a binding constraint until $t = 24$. From $t = 25$ to $t = 34$ the subsequent contraction in investment, caused by a binding constraint, is smaller than the one caused by the precautionary saving effect in the previous periods. From the point of view of figure 3-2, this further drop in investment is equivalent to moving leftward along the diagonal section of the $k_{i,t+1}$ policy function, because $w_{i,t}$ has decreased below $\underline{w}_t$. When the productivity shock switches back to $\theta_H$, the entrepreneur is constrained for five periods. The subsequent gradual increase in capital from time $t = 40$ on is determined by the reduction in the precautionary saving effect.

Figure 3-4: Changes in capital stock for a constrained and an unconstrained entrepreneur

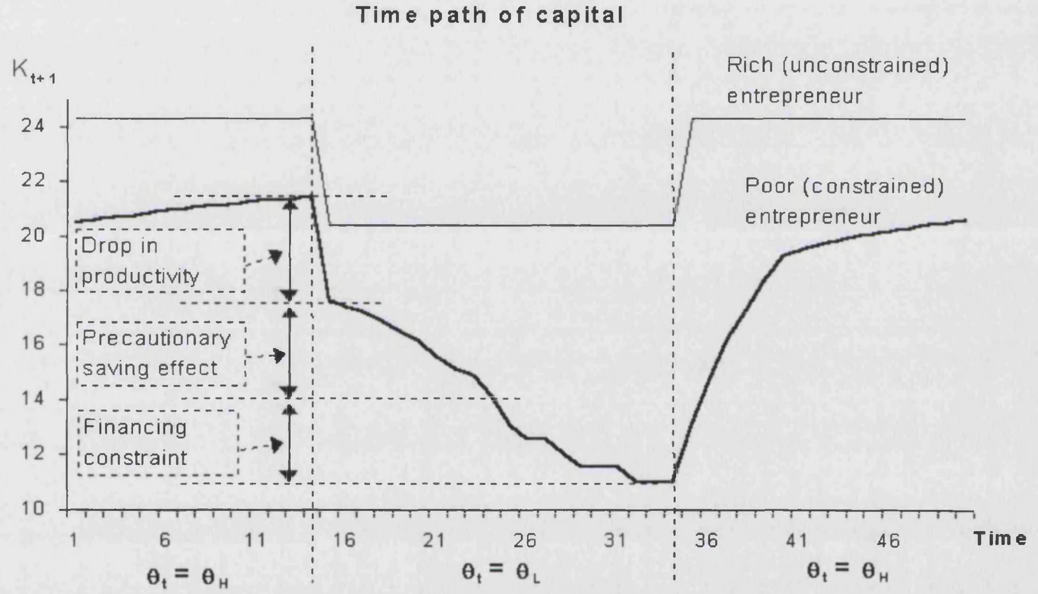
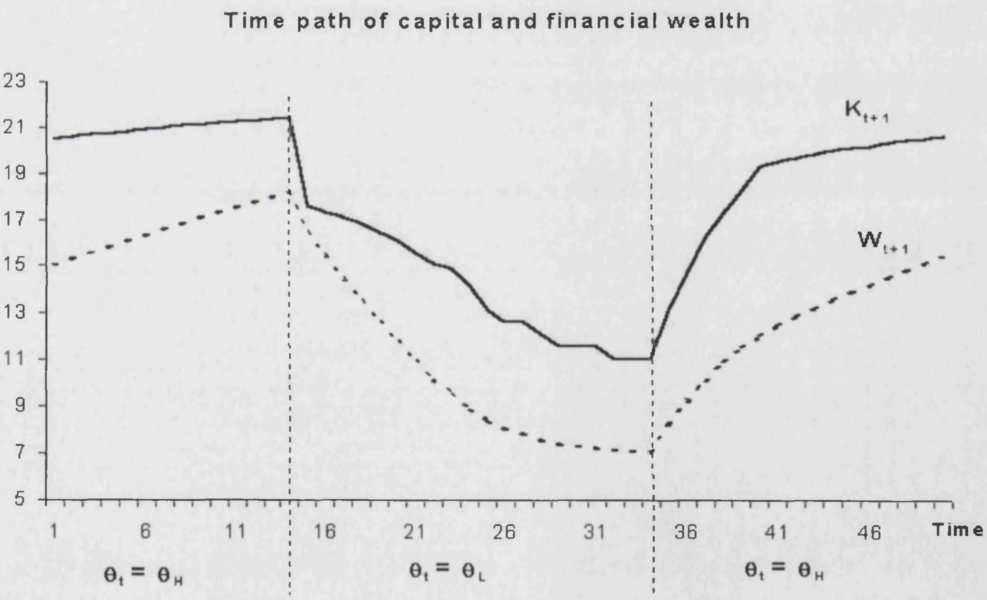



Figure 3-5 illustrates the time paths of capital ($k^2_{t+1}$) and financial wealth ($w^2_{t+1}$) for $E^2$. It clearly shows that capital has a positive correlation with expected profitability, but also with changes in financial wealth (cash flow) not related to changes in expected profitability, because of the precautionary saving effect. In the first 15 periods $E$ accumulates financial wealth. As she becomes more confident about the future, she increases the investment in the risky asset (firm's capital), reducing precautionary saving. The situation is reversed in the next 20 periods of "bad luck". Here the sensitivity of investment to cash flow becomes progressively stronger as financial wealth shirks towards the constrained region. Figure

3-5 clearly shows that it is wrong to assume that the effects of financing constraints on firm investment are confined to the periods when the financing constraints is binding. For example the sensitivity of investment to changes in financial wealth is higher between $t = 15$ and $t = 25$, when the constraint is not binding, than between $t = 26$ and $t = 31$, when the constraint is binding.

This result supports the Kaplan and Zingales (1997) critique that there is no theoretical foundation for the claim that investment-cash flow sensitivities are monotonously increasing in the intensity of financing constraints. However figure 3-5 also shows that the investment-cash flow sensitivity can be related to financing constraints also for a profitable and wealthy company. For example this is what happens between $t = 40$ and $t = 50$, when $E^2$ increases her stock of physical capital, generates profits, increases her financial wealth and reduces her debt. Also in this period the correlation between cash flow and investment is entirely due to the reduction in future expected financing constraints, and it would be absent with perfect financial markets, as is shown by the investment path of $E^1$.

Figure 3-5: Changes in capital stock and in financial wealth for a constrained entrepreneur




### 3.4.2 Simulation: aggregate

In this subsection section we use the solution of the model to simulate an artificial economy with many heterogeneous entrepreneurs. This is a simple "partial equilibrium" economy, because both interest rate and prices are exogenous. The purpose of this exercise is to show the effects of financing constraints on aggregate investment.

The procedure is simple: we simulate an economy with 10000 entrepreneurs for 200 periods. Each of the productivities of the entrepreneurs evolves accordingly to a value of $\theta_{it}$, that is independent across entrepreneurs and path dependent across time. If an entrepreneur retires and exits, a new one enters, so that the total number of them is constant. Each period the entrepreneurs are selected into small and large groups according to their total assets[15], and aggregate statistics for small and large entrepreneurs are computed to analyse the effects of an unexpected macroeconomic shock that hits after 100 periods.

It is important to note that the generated dynamics are sensitive to technology assumptions. For example one limitation of the basic model presented in this chapter is that the idiosyncratic shock is stationary and $E$'s technology is concave. Therefore she wants to expand the size of her business only up to a certain steady state that depends on the unconditional expectation $E(\theta_{t+1}) = (\theta_H + \theta_L)/2 = \overline{\theta}$. This corresponds to a situation where returns to scale are decreasing, or are constant/increasing but $E$'s "know-how" is essential for the business[16].

A consequence of this limitation is that there exists a straightforward way to generate higher volatility of small entrepreneurs. It is sufficient to impose the same $\overline{\theta}$ to all entrepreneurs and to assume that new entrepreneurs start their activity with a very small endowment. Given that there is an ongoing entry and exit, small entrepreneurs will include younger ones that are constrained because they are in the start-up phase, while large entrepreneurs will include older ones that are less constrained because they are on average closer to the steady state.

In order to avoid this effect, we assume that newly born entrepreneurs have a relatively

[15] Similar results are obtained using different selection criteria, like capital level or output.

[16] The concave technology can be interpreted as a constant/increasing return to scale technology where E supplies an essential fixed amount of "quality". That is, $\theta_t = \widehat{\theta}_t \overline{E}^{\beta}$, where $\widehat{\theta}_t$ is the stochastic component, $\overline{E}$ is $E$'s quality and $\alpha + \beta \geq 1$.

large initial endowment:

$$w_0 = k^{PM}(\theta_H) \tag{3.32}$$

(3.32) ensures that new entrepreneurs have enough resources to finance the perfect market capital level and that the constraint is not binding in the first periods of life. Assuming a smaller value of $w_0$ does not change the qualitative results of the analysis, it only increases their magnitude. Therefore this assumption has the purpose of emphasising the importance of existing firms dynamics and of the precautionary saving effect with respect to the effect of a binding constraint in the start-up phase.

Moreover, in order to further generalise the analysis, we consider another dimension of the entrepreneurs heterogeneity, regarding the value of $\overline{\theta}$. We simulate 10000 firms that belong to two types:

*Type1:* $\overline{\theta}^1 = 5;\ \theta_L^1 = 4.42;\ \theta_H^1 = 5.58$

*Type 2:* $\overline{\theta}^2 = 5.23;\ \theta_L^2 = 4.62;\ \theta_H^2 = 5.84$

Type one entrepreneurs have smaller $\overline{\theta}$, but are otherwise equal, also in terms of risk, as:

$$\frac{(\theta_H^1 - \theta_L^1)}{\overline{\theta}^1} = \frac{(\theta_H^2 - \theta_L^2)}{\overline{\theta}^2}$$

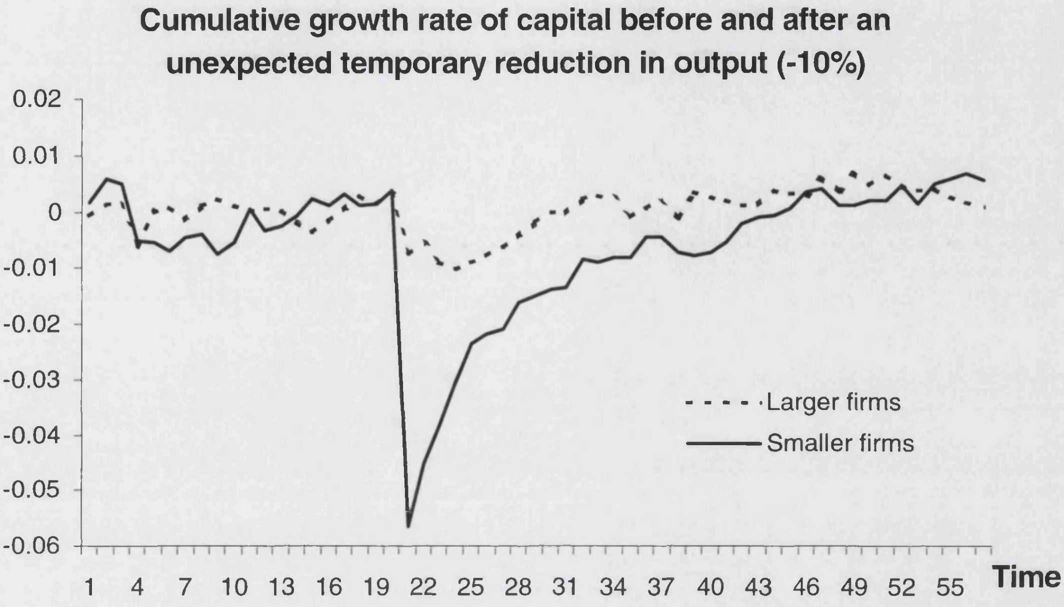
The fraction of type 1 and 2 among the 10000 entrepreneurs is such that the two types on average produce the same aggregate level of output. In this way we make the simulations more realistic, assuming that firms do not have all the same average steady state size.

Figures 3-6-3-9 show the *cumulative rate of growth of capital* for small and large firms, before and after unanticipated macroeconomic shocks. Following the empirical literature (Bernanke, Gertler and Gilchrist, 1996), we select entrepreneurs according to a fixed percentile in the cumulative size distribution function. In each period the entrepreneurs whose businesses are below the 30% percentile are selected in the "small firms" group and those above the 70% percentile in the "large firms" group.

Figure 3-6 considers a temporary reduction in output[17] of 10%. The immediate nega-

[17] Since the model does not have any aggregate uncertainty, the observed variability in the aggregate growth rates before the shock depends only on the entry/exit dynamics.

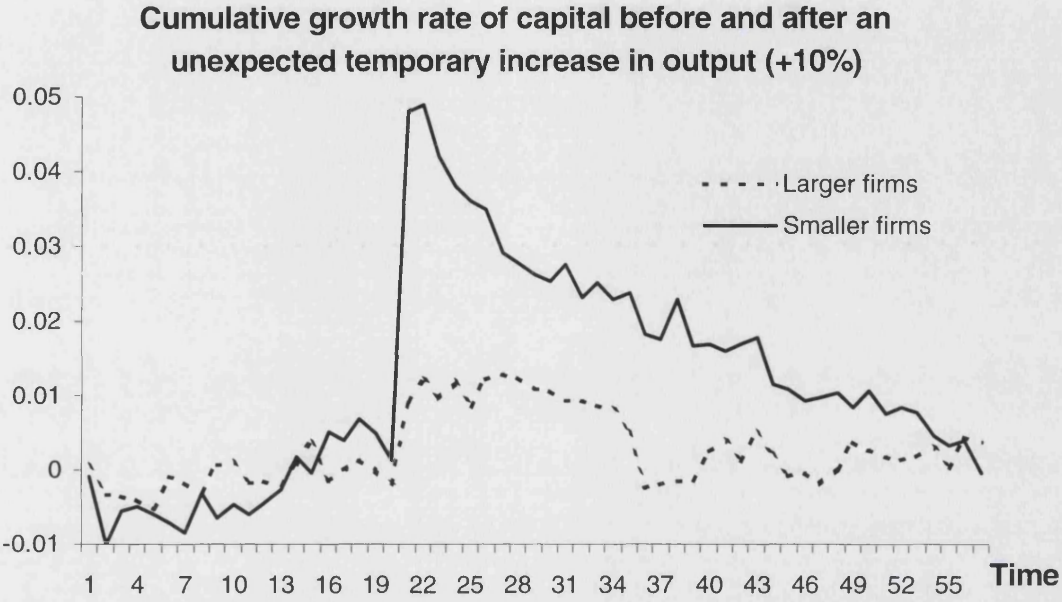
Figure 3-6: Aggregate capital response to a temporary unexpected negative income shock




tive impact on investment after the shock is more than 5% for small firms and only around 1% for large firm. The reason for the difference is that lower output reduces profits and financial wealth. Some entrepreneurs reduce investment because the wealth shock pushes them in the constrained region, while others reduce investments because of precautionary saving reasons. Both effects are stronger for small firms that are relatively less wealthy. After the first negative reaction entrepreneurs slowly return to the previous steady state. The reduction in cumulative growth rates takes around 25 periods to disappear. Such persistency depends on the precautionary saving effect. As long as entrepreneurs are on average less wealthy, with respect to the before-shock situation, they have higher expected probability to be financially constrained in future, and save more until they reach the before-shock average wealth level.

Figure 3-7 considers a symmetric positive shock. Also in this case small firms react much more than large ones, increasing capital by almost 5%. This change is relatively less, in absolute value, than after the negative shock. The reason of the asymmetry is that few entrepreneurs in the steady state experience a binding financing constraint, as argued in the previous subsection. Hence the positive shock's reaction is caused almost exclusively

Figure 3-7: Aggregate capital response to a temporary unexpected positive income shock

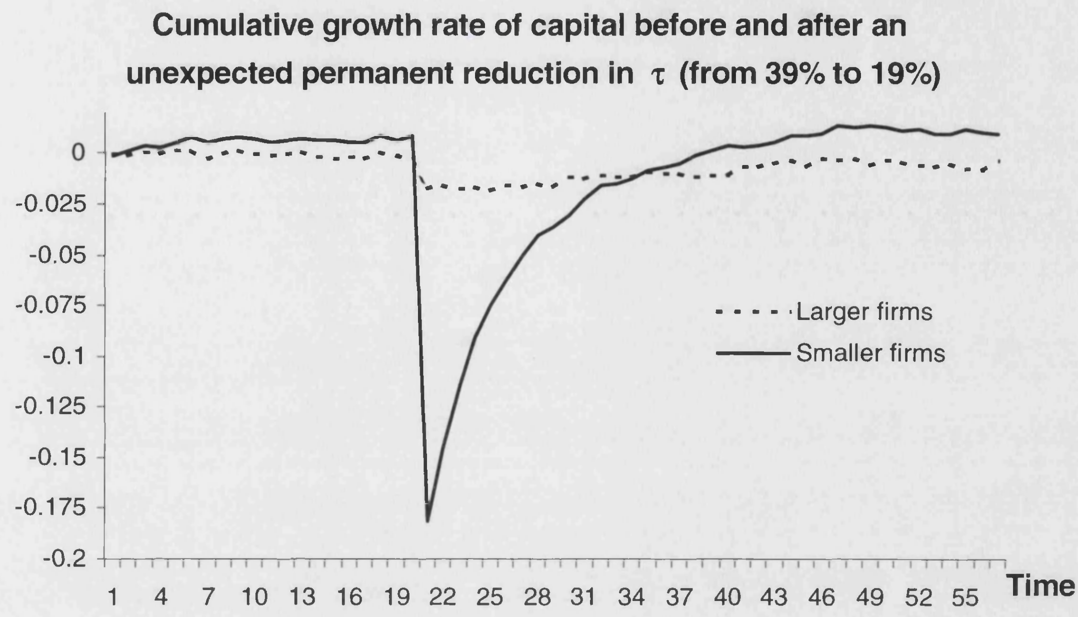



by the precautionary saving effect, while the negative shock's impact is caused by a mix of precautionary saving and of binding financing constraint.

Figure 3-8 considers a permanent reduction of $\tau$, the collateral coefficient, from 0.38 to 0.19. This reduction means that, after the shock, capital is collateralisable only up to 50% of its residual value. This shock is like a credit crunch, as it reduces the borrowing capacity of all entrepreneurs. In this case small firms' reaction is very strong, with a drop in the stock of capital of 18%, while large firms are almost unaffected. Once again this result depends both on binding constraint and precautionary saving effect.

Figure3-9 shows the effect of a permanent reduction in interest rate from 3.5% to 1.5%. This shock decreases both the user cost of capital $UK$ and the downpayment $D$. The effect on $E_t(\Omega_{t+1})$, the cost of future expected financing constraints, is instead ambiguous. This is because on the one hand the decrease in $R$ means that, conditional on having a binding financing constraint, an entrepreneur gets more financing. On the other hand it also means that the expected rate of financial wealth accumulation for an unconstrained entrepreneur is lower. Therefore she has a higher probability of being financially constrained in the future.

Figure 3-8: Aggregate capital response to a permanent unexpected increase in the collateral requirements

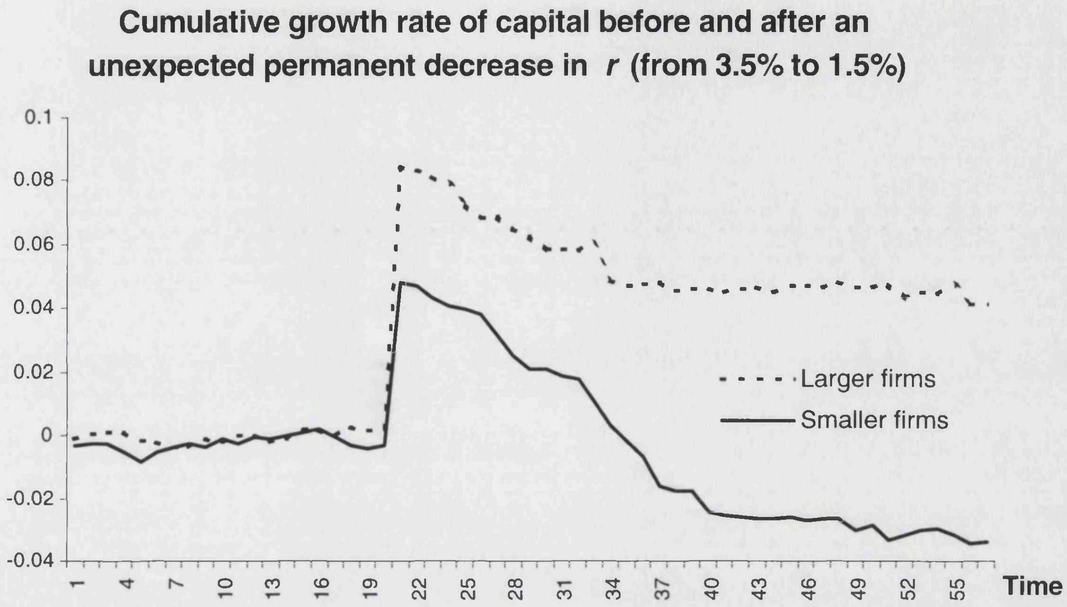



This discussion implies that $E_t(\Omega_{t+1})$ is less sensitive than $UK$ to a decrease in interest rate. It is important to note that $UK$ is the cost of capital for an unconstrained (now or in future) entrepreneur, while (3.23) implies that $UK + E_t(\Omega_{t+1})$ is the cost of capital for an unconstrained entrepreneur who engages in precautionary saving. It follows that the more *she* engages in precautionary saving, the less her investment is sensitive to interest rate changes. Therefore, since smaller entrepreneurs on average engage in more precautionary saving, they also respond less to interest rate changes.

Figure 3-9 confirms this, showing that the short term positive reaction to the interest rate shock is stronger for large firms. Figures 3-8 and 3-9 confirm the credit view (Bernanke, 1983): if financing imperfections are present in the economy, and affect firms behaviour, the availability of credit (figure 3-8) may be more important than interest rate (figure 3-9) in affecting investment for financially constrained firms.

## 3.5 Conclusions

Figure 3-9: Aggregate capital response to a permanent unexpected reduction in interest rate




In this chapter we illustrate a dynamic model of investment with financing constraints and with stochastic productivity that analyses the trade off between investing in physical capital and saving in financial assets.

The model shows that not only expected productivity but also financing problems affect investment and saving decisions of the entrepreneurs. In particular we derive the existence of a precautionary saving effect analogous to the one studied in the consumption literature: the amount of wealth allocated between risky projects and safe assets depends on future expected financing problems. We solve the model using a numerical method, and we show that the precautionary saving effect has important implications for firm level and aggregate investment: i) the cash flow-investment sensitivity is related to financing constraints but it is not necessarily monotonous in the intensity of them; ii) financing constraints have a quantitatively important effect on aggregate investment even if only a small share of entrepreneurs faces a binding financing constraint at any point in time. This is because net worth fluctuations affect the amount of precautionary saving of all entrepreneurs.

We use the model's solution to simulate a partial equilibrium economy with many heterogeneous entrepreneurs. In the steady state unexpected temporary aggregate shocks have a large and persistent impact on aggregate investment because they change the precautionary saving levels. For example a negative shock reduces the net worth of all entrepreneurs and limits their ability to absorb future idiosyncratic shocks. This makes them reduce physical investment and increase their precautionary saving of financial assets. The model therefore explains that expected financial constraints may have an important depressing effect on investment in downturns. Another interesting result is that small firms are on average more financially constrained, despite the fact that all firms are ex ante identical with respect to their ability to access external finance and as a result they are much more sensitive to aggregate output shocks than large firms. Moreover we show that a credit availability shock is more effective than an interest rate shock in affecting investment for financially constrained (small) firms, while the interest rate shock is more effective for unconstrained (large) firms.

# Chapter 4

# A dynamic multifactor model of investment with financing and irreversibility constraints

## 4.1 Introduction

In this chapter we extend the basic model developed in chapter 3 by assuming that output is produced by an entrepreneur who operates a concave risky technology 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 whole business is sold, while investment in variable capital is reversible. We analyse the interactions between financing and irreversibility constraints to show that the two problems interact and amplify each other. In particular the irreversibility of fixed capital amplifies the effect of financing constraints and of precautionary saving on variable investment dynamics. Such amplification effect has never been analysed before in investment literature, and we show that it is both quantitatively and qualitatively important in explaining certain stylised facts about inventories fluctuation: i) aggregate inventory investment is very volatile and procyclical; ii) its decline accounts for a large part of the GDP decline in recessions; iii) it is contemporaneously correlated with sales; iv) it leads the business cycle, while fixed capital lags it; v) output and inventories are more volatile and procyclical for small firms than for large ones.

We analyse the effects of financing and irreversibility constraints on the cyclical fluctuations of investment in fixed capital and variable capital by simulating an artificial economy with heterogeneous entrepreneurs and with both idiosyncratic and aggregate uncertainty. In order to understand how the interaction between irreversibility and financing constraints helps to explain the stylised facts mentioned before, let's consider the example of an entrepreneur who faces a permanent decreases in productivity, like at the beginning of an economy downturn. We distinguish two cases:

1) irreversibility of fixed capital and no financing constraints. The decrease in productivity means that the stock of fixed capital is now higher than the desired level. Because of irreversibility, the entrepreneur cannot sell fixed capital and she will take some time to reduce it to the new optimal level. During the transition period, fixed capital is inefficiently high and since the two factors of productions are complementary, variable capital is too high as well.

2) Both irreversibility of fixed capital and financing constraints. As in the previous case, it will take some time to adjust the stock of fixed capital. During the adjustment period expected profits drop and the expected rate of wealth accumulation drops as well with respect to a situation where fixed capital is reversible. As a result the entrepreneur has a higher chance to face future financing constraints: i) if the contraction period is long enough, then the drop in wealth can be so severe that she does not have enough funds to invest in variable capital; ii) if the contraction period ends, she will have not enough resources to invest in both fixed and variable capital and will be financially constrained for some time. In order to compensate these higher costs of future expected financing constraints, the entrepreneur engages in precautionary saving ex ante, more than it would have done with reversible capital. However since fixed capital cannot be reduced, such precautionary saving behaviour affects variable capital only. Therefore the combination of higher current and future expected financing constraints implies that variable capital becomes much more volatile with respect to the case without irreversibility of fixed capital, and drops more in response to the permanent decrease in productivity.

In the remainder of this chapter we analyse the implications of these effects for aggregate investment dynamics by simulating an artificial partial equilibrium economy with many heterogeneous entrepreneurs. With respect to the simulations illustrated in chapter 3 aggregate uncertainty is given by a combination of recurrent transitory and persistent aggregate shocks that generate an exogenous stochastic business cycle. This means that the cross-sectional distribution of net worth and fixed capital among entrepreneurs is determined by both idiosyncratic and aggregate uncertainty, and it affects the way aggregate output and investment react to aggregate shocks. In chapter 3 we show that the cross sectional distribution of wealth determines heterogeneous responses of small and large entrepreneurs to unexpected macroeconomic shocks. Here we also show that: i) the heterogeneity of entrepreneurs with respect to fixed capital implies that fixed investment reacts less than variable investment to transitory shocks and also that it has a lagged reaction to permanent shocks; ii) the amplification of current and future expected financing constraints implies that aggregate investment (fixed and variable) is sensitive to net worth fluctuations, that variable investment is contemporary correlated with sales and that it is much more volatile than fixed investment. All these effects become stronger the higher the volatility and the persistence of the idiosyncratic shocks, which imply a more dispersed distribution of both financial wealth and fixed capital across entrepreneurs.

Because of the time lag for the investment to produce output, changes in the level of variable capital in the model can be interpreted as investment in input inventories, such as raw materials and work in progress. Thus this extended model is consistent with several of the stylised facts mentioned in section 1.5. First, both fixed and inventory investment at firm level are found to be sensitive to the availability of internal finance. In particular US data show that "*inventory investment for small firms absorb from 15% to 40% of cash flow fluctuations*" (Carpenter, Fazzari and Petersen (1998)). Second, aggregate inventory investment is more volatile and procyclical than fixed investment. "*Changes in business inventories, which constitute but a small fraction of total GDP, account for one-fourth of the cyclical movements in GDP*" in the US (Stock and Watson (1998)). Ramey (1989), Blinder and Maccini (1991) and Ramey and West (1999) show that this is especially true during recessions, when the drop in inventory investment accounts for a large part of the GDP decline. They also provide evidence in support of our approach to model inventories as a production factor, because they emphasise that input inventories, like variable capital, are quantitatively more important and more volatile than finished goods inventories. Third, inventory investment is contemporaneously correlated with sales (Ramey and West (1999)). Fourth, inventory investment leads the business cycle, while fixed investment lags it (Stock and Watson (1998)). Fifth, small firms are more procyclical than large ones in inventories, output, and short term debt (Bernanke, Gertler and Gilchrist (1996)).

The outline of the chapter is the following: in section 4.2 we illustrate the extended model. In sections 4.3-4.5 we illustrate the qualitative features of the solution respectively with irreversibility only, with financing constraints only and with both problems. In section 4.6 we describe the numerical solution for selected parameter values. In section 4.7 we show the results of the simulation of the aggregate economy. In section 4.8 we summarise the conclusions.

## 4.2 The extended model with financing and irreversibility constraints

This model extends the one presented in chapter 3, by assuming the presence of one additional factor of production and of irreversibility of fixed capital.

- **Preferences**

Preferences are the same as in the basic model in chapter 3. Therefore see assumptions *p1-p4*.

- **Technology**

Assumption *t3,* regarding the productivity shock, is unchanged in this section. However it will be relaxed later in section 4.6, where we consider both idiosyncratic and aggregate uncertainty. Assumptions *t1* and *t2* are replaced by the following: *t1') E can invest in the risky technology of the firm she owns and manages.* $k_t$ *and* $l_t$ *are the stock of fixed and variable capital respectively, 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 assumption that investment takes one period to become productive also allows us to interpret $l_t$ as end of period $t-1$ inventories of variable capital.

*t2') Net output $y_t$ is produced according to a Cobb-Douglas production function:*

$$y_t = \theta_t k_t^{\alpha} \left(l_t + l^E\right)^{\beta} \text{ with } \alpha + \beta < 1 \tag{4.1}$$

$l^E$ is a small fixed amount of variable capital supplied[1] each period by $E$. As in the previous chapter, all prices are assumed constant and normalised to 1.

*t4) Variable capital is nondurable, while fixed capital is durable:*

$$1 = \delta_l > \delta_k \tag{4.2}$$

$\delta_l$ and $\delta_k$ are the depreciation factors of variable and fixed capital respectively.

*t5) Variable capital is reversible, while fixed capital is irreversible, and can only be disinvested if E sells her whole business. In this case she cannot start a new business and must retire.*

Therefore conditional on continuation $E$ is subject to the following constraints[2]:

$$k_{t+1} \geq (1 - \delta_k) k_t \tag{4.3}$$

$$l_{t+1} \geq 0 \tag{4.4}$$

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) (see section 1.3).

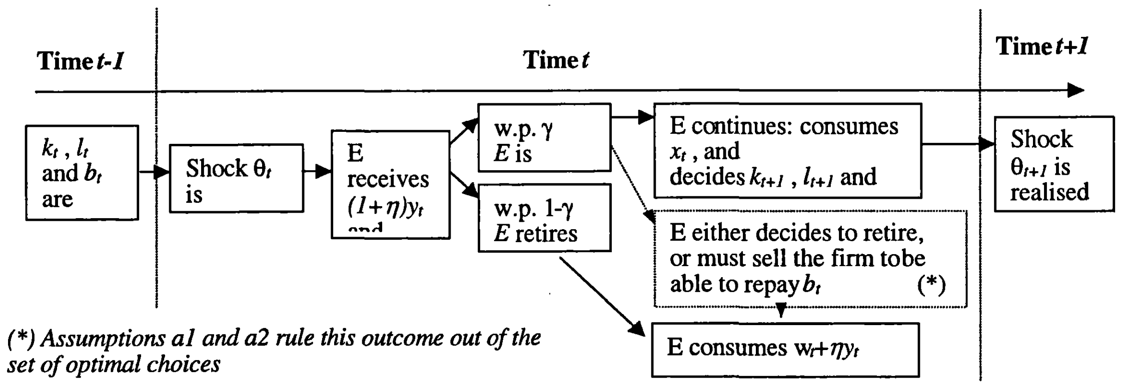
- **Financial Markets**

[1] It can be interpreted as E's effortless labour supply.

[2] Constraint (4.4) would be never binding if the Inada conditions are satisfied. However this is not true in this case, because the presence of $l^E$ implies that $\lim_{l_{t+1} \to 0} \frac{\partial y_{t+1}}{\partial l_{t+1}} < \infty$.

See assumptions[3] *f1-f3*. $\tau_k$ is the share of fixed capital value that can be used as collateral, while from (4.2) it follows that $\tau_l = 0$. Therefore variable capital has no collateral value, and (3.6) and (3.7) apply to this extended model as well. The timing of the extended model is represented in figure 4-1. At the beginning of the period $t$ $E$

Figure 4-1: The timing of the extended model




inherits from time $t-1$ the stocks of fixed and variable capital $k_t$ and $l_t$. Then $\theta_t$ is realised, $(1+\eta)\,y_t$ is produced and $b_t$ repaid. Residual wealth $w_t$, net of the minimum consumption level $\eta y_t$ is still represented by (3.8). After producing, $E$ either retires or continues activity. In the extended model retirement can happen for three different reasons: i) $E$ becomes ill; ii) $E$ must retire after having sold the business to repay $b_t$; iii) $E$ chooses to retire because the liquidation value is greater than the continuation value. We define $D_t$ as the dichotomous variable which represents the retirement choice of $E$ conditional on not becoming ill in period $t$:

$$D_t = \begin{cases} 1 \text{ if E continues activity} \\ 0 \text{ if E retires} \end{cases}$$

the choice of $D_t$ is subject to the following constraint:

$$D_t = 0 \qquad if\ y_t + \tfrac{\tau_k}{R}\,(1-\delta_k)\,k_t < b_t \tag{4.5}$$

[3]The difference with the basic model is that, in order for assumptions *f1-f3* to be consistent with the technological assumptions, we implicitly assume that in any default and renegotiation of the debt with the bank $E$ has all the bargaining power. Otherwise the bank could use the threat of liquidation of fixed capital to enforce the repayment of uncollateralised debt.

$D_t$ is constrained to be equal to 0 if the inequality (4.5) is satisfied. The left hand side of this inequality represents the maximum amount of funds available to repay $b_t$ without selling the residual value of fixed capital $(1-\delta_k)k_t$. Such funds are equal to the financial wealth plus the new borrowing available using $(1-\delta_k)k_t$ as collateral. Hence if $b_t$ is higher than the left hand side of (4.5) the only way to repay the debt is by selling all the assets and retiring. The retired entrepreneur consumes the residual value $w_t+\eta y_t$. If $E$ continues, she borrows new one period debt and allocates net worth plus the new borrowing between consumption, investment in fixed capital and investment in variable capital:

$$x_t^* + l_{t+1} + k_{t+1} = w_t + b_{t+1}/R \tag{4.6}$$

The introduction of irreversibility of fixed capital greatly increases the complexity of the solution. To show why, lets derive the Bellman equation of the dynamic optimisation problem. We denote the expected lifetime utility at time $t$ of $E$, after $\theta_t$ is realised, and conditional on not becoming ill at time $t$, by $W_t\left(w_t,\theta_t,k_t\right)$. $w_t$, $\theta_t$ and $k_t$ are the three state variables of the problem. $\overline{\Theta}$ is the new vector of parameters: $\overline{\Theta}' = \{\theta_L,\theta_H,\epsilon,\gamma,R,\delta_k,\tau_k,\eta,l^E,\alpha,\beta\}$.

$$W_t\left(w_t,\theta_t,k_t\right) = \underset{k_{t+1},l_{t+1},b_{t+1},D_t}{MAX} \mathbf{1}\left(D_t=1\right)\{x_t^* + \eta y_t + \tfrac{1}{R}E_t[\gamma W_{t+1}(w_{t+1},\theta_{t+1},k_{t+1})$$

$$+\left(1-\gamma\right)\left(w_{t+1}+\eta y_{t+1}\right)]\} + \mathbf{1}\left(D_t=0\right)\left(w_t+\eta y_t\right) \tag{4.7}$$

$E$ maximises (4.7) conditional on (3.1), (3.6), (4.3), (4.4), (4.5) and (4.6). $\mathbf{1}(.)$ is an indicator function that assumes the value of 1 if the argument is true, and 0 otherwise. $b_{t+1}$, $k_{t+1}$, $l_{t+1}$ and $D_t$ are the control variables. They also determine $x_t^*$ from (4.6).

### 4.2.1 Endogenous retiring

The solution of (4.7) is more complicated than the one of the basic problem analysed in chapter 3, because the contemporaneous presence of financing and irreversibility constraints implies that $D_t$, the retirement decision, becomes endogenous. Conditional on not being forced to retire $E$ chooses voluntary retirement at time $t$ if her utility conditional on

$D_t = 0$ (liquidation value) is greater than her utility conditional on $D_t = 1$ (continuation value). Without irreversibility of fixed capital $E$ never wants to retire, because she can always choose a combination of $k_{t+1}$ and $l_{t+1}$ that yields an unconditional expected return greater than $R$. Moreover she is never forced to retire as well, because she is always able to liquidate the fraction of assets needed to repay the debt without having to sell the entire business.

Instead, if $k_{t+1}$ is irreversible, forced retirement happens at time $t$ if revenues are so low that $E$ must sell the fixed capital to repay the debt, as shown by condition (4.5). If revenues allow $E$ to repay the debt and to continue, she may still prefer to retire if expected short term return is so low as to offset long term gains from continuing activity. Short term return can be very low because of the combined effects of financing and irreversibility constraints. In order to explain how this can happen, lets consider the case with $l^E = 0$. Moreover suppose that at time $t$ $\theta_t = \theta_H$ and $E$ borrows up to the limit to invest in fixed and variable capital, while at time $t+1$ $\theta_{t+1} = \theta_L$. This means that output is low, and we assume that it is just enough to repay all the debt, so that no funds are left for investing in $l_{t+2}$. Therefore at time $t+1$ E has now two possibilities: 1) sell the business now, and obtain the residual value of her assets $w_{t+1}$, which is by definition positive. 2) Continue.

It is easy to prove that in this case continuation is not optimal. In fact since $E$ has no funds to invest in variable capital, $l_{t+2} = 0$. By substituting this value and $l^E = 0$ in (4.1) we get that $y_{t+2} = 0$. Hence at time $t+2$ $E$ cannot repay her debt and is forced to sell the business and retire. Since the residual wealth $w_{t+2}$ is smaller than $w_{t+1}$, it would have been better to sell and retire at time $t+1$ instead.

The above discussion clarifies that the interaction between financing and irreversibility constraints can influence the retirement decisions of firms. While this is an interesting intuition to explore in future research, it goes beyond the scope of this chapter. Here we want to note that the presence of forced and voluntary retirement implies that the discount factor of the problem is time variant, and that $D_t$ is itself a control variable: $D_t = D(\theta_t, k_t, w_t)$. This makes the dynamic maximisation problem extremely difficult to solve, even numerically. Therefore in order to simplify the analysis we make two additional assumptions to rule out endogenous retiring. It is important to note that such assumptions do not affect the qualitative results of the model. In fact they rule out the extreme outcomes that would increase the expected cost of irreversibility and financing constraints, and hence would strengthen rather than weaken the model's results. The assumptions are the following:

*a1)* $l^E \geq l^E_{\min}\left(\overline{\Theta}\right)$;

*a2)* $w_o, \theta_L, \theta_H$ *and* $\epsilon$ *are such that:*

*i)* $prob(w_t \leq w^{\min}(k_t, \theta_t)) = 0$ *for* $t = 0, 1, ..., \infty$;

*ii)* $\lim\limits_{w_t \to w^{\min}(k_t,\theta_t)} \left[(1+\eta) E_t(\theta_{t+1}) k^{\alpha}_{t+1} \left(l^E\right)^{\beta} + (1-\delta_k) k_{t+1}\right] \geq R$ *for* $t = 0, 1, ..., \infty$.

$l^E_{\min}$ is defined in equation (6.22) of appendix 1. It represents the minimum amount of $l^E$ that allows $E$ to generate enough revenues to repay the debt without liquidating $(1-\delta_k)k_t$. Therefore she is never forced to retire[4]. *a2* ensures that an active $E$ never voluntarily retires. $w^{\min}(k_t, \theta_t)$ is the level of net worth such that $\lim\limits_{w_t \to w^{\min}(k_t,\theta_t)} l_{t+1}(w_t, k_t, \theta_t) = 0$. *a2* ensures that $w_t$ is always greater than the minimum level $w^{\min}(k_t, \theta_t)$. This implies that in the worst case scenario, when $w_t \to w^{\min}(k_t, \theta_t)$ and $l_{t+1} \to 0$ because $E$ does not have funds to invest in variable capital, expected return on the business is at least equal to $R$. Assumptions *a1* and *a2* allow us to state the following proposition:

**Proposition 4** *Under assumptions a1 and a2 E* does not retire voluntarily, and is never forced to retire, i.e. $D_t = 1$ for $t = 0, 1, ..., \infty$.

**Proof: see appendix 1.**

Proposition 4 implies that $D_t$ is no longer a state variable of the problem and that constraints (4.4) and (4.5) are never binding[5]. Therefore we can substitute $D_t = 1$ for $t = 0, 1, ..., \infty$, in (4.7). Applying also proposition 1 and corollary 2 (see appendix 1), we denote the expected lifetime utility at time $t$ of an active $E$, after $\theta_t$ is realised, conditional on not becoming ill and after consuming $\eta y_t$, by $W'_t(w_t, \theta_t, k_t)$ :

---

[4] In the simulations presented in the next section, $l^E_{\min}$ is never greater than 3% of optimal unconstrained $l_t$.

[5] Given the concavity of the production function, *a2* implies that $l_{t+1}(w_t, k_t, \theta_t) > 0$ for $t = 0, 1, ..., \infty$.

$$W_0'(w_0,\theta_0,k_0) = \underset{\begin{pmatrix} k_{t+1} = k(w_t,\theta_t,k_t) \\ l_{t+1} = l(w_t,\theta_t,k_t) \\ b_{t+1} = b(w_t,\theta_t,k_t) \end{pmatrix}_{t=0,1,\ldots\infty}}{MAX} E_0\left\{\sum_{t=0}^{\infty}\left(\frac{\gamma}{R}\right)^t\left\{\frac{1}{R}\left[\eta y_{t+1}+(1-\gamma)w_{t+1}\right]\right\}\right\} \tag{4.8}$$

The problem is now defined by (4.8) subject to (3.6), (4.3) and (4.6). Also in this case it is easy to show that a solution to the problem exists and is unique. In order to describe the optimality conditions of the extended model, let $\lambda_t$, $\mu_t$ and $\phi_t$ be the Lagrangean multipliers associated respectively to constraints (3.6), (4.3) and (4.6). Taking the first order conditions of (4.8) with respect to $b_{t+1}$, $l_{t+1}$ and $k_{t+1}$ it is possible to how that the solution is given by the optimal sequence of $\left\{b_{t+1}, k_{t+1}, l_{t+1}, \lambda_t, \mu_t, \phi_t \mid k_t, w_t, \theta_t; \overline{\Theta}\right\}_{t=1}^{\infty}$ that satisfies (4.3), (4.9), (4.10), (4.11) and (4.12) plus the standard Kuhn-Tucker complementary slackness conditions on $\lambda_t$ and $\mu_t$:

$$Dk_{t+1} + l_{t+1} \leq w_t \tag{4.9}$$

$$\phi_t = 1 + RE_t\left[\sum_{j=0}^{\infty}\gamma^j\lambda_{t+j}\right] \tag{4.10}$$

$$(1+\eta)\,E_t\left(MPL_{t+1}\right) = UL + E_t\left(\Omega_{l,t+1}\right) + R^2\lambda_t \tag{4.11}$$

$$(1+\eta)\,E_t\left(MPK_{t+1}\right) = UK + E_t\left(\Omega_{k,t+1}\right) + (1-\delta)\,\gamma E_t\left(\mu_{t+1}\right) + DR^2\lambda_t - R\mu_t \tag{4.12}$$

$E_t\left(MPL_{t+1}\right) = \beta E_t\left(\theta_{t+1}\right) k_{t+1}^{\alpha}\left(l_{t+1}+l^E\right)^{\beta-1}$ is the marginal productivity of variable capital and $E_t\left(MPK_{t+1}\right) = \alpha E_t\left(\theta_{t+1}\right) k_{t+1}^{\alpha-1}\left(l_{t+1}+l^E\right)^{\beta}$ is the marginal productivity of fixed capital. $UL = R$ is the user cost of variable capital. Equations (4.9), (4.10) and (4.11) are the equivalent of (3.11), (3.12) and (3.13) in the basic model. $E_t\left(\Omega_{k,t+1}\right)$ is identical to $E_t\left(\Omega_{t+1}\right)$ defined in (3.14), and $E_t\left(\Omega_{l,t+1}\right)$ is the equivalent term for variable capital:

$$E_t\left(\Omega_{l,t+1}\right) = \gamma E_t\left[\left(\phi_{t+1}-1\right)\left(UL - MPL_{t+1}\right)\right] \tag{4.13}$$

The first order condition for fixed capital, (4.12), is the most obvious difference with respect to the solution of the basic model. $(1-\delta)\gamma E_t(\mu_{t+1})$ is the cost of future expected irreversibility problems. $\mu_t$ is positive when the irreversibility constraint is binding, and is equal to zero otherwise. Since it is not possible to obtain a solution to this problem analytically, in section 4.6 we will solve it using a numerical method. Instead in the following three sections we wish to describe the main qualitative features of the model. We first analyse the solution without financing problems, then we analyse the solution without irreversibility problems, and finally we explain how the two problems interact together.

## 4.3 The two factors model with irreversibility of fixed capital.

In this subsection we rule out current and future expected financing constraints by imposing $w_0 = w^{MAX}$. Together with proposition 1 this implies that $E$ will never be financially constrained during her active life, and that $\lambda_t = 0$, $\phi_t = 1$ and $E_t(\Omega_{k,t+1}) = E_t(\Omega_{l,t+1}) = 0$ for $t = 0, 1, ..., \infty$. Furthermore in this subsection we also restrict the analysis to the set of parameters $\overline{\Theta}^1 \subset \overline{\Theta}$ such that the irreversibility constraint is binding, at time $t$, after a negative productivity shock (i.e. when $\theta_{t-1} = \theta_H$ but $\theta_t = \theta_L$). First order conditions (4.11) and (4.12) simplify to:

$$(1+\eta)E_t(MPL_{t+1}) = UL \tag{4.14}$$

$$(1+\eta)E_t(MPK_{t+1}) = UK + (1-\delta)\gamma E_t(\mu_{t+1}) - R\mu_t \tag{4.15}$$

since in this case financial wealth is irrelevant, the problem is now defined by equations (4.14) and (4.15) and constraint (4.3), which jointly determine $\mu_t(k_t, \theta_t)$, $k_{t+1}(k_t, \theta_t)$ and $l_{t+1}(k_t, \theta_t)$. The term $(1-\delta)\gamma E_t(\mu_{t+1})$ represents the cautious investment effect analysed by the irreversibility literature. In fact (4.3), (4.14) and (4.15) describe the solution to a simplified version of the problem analysed by Bertola and Caballero (1994). The main difference is that we allow for a reversible factor of production to be used in conjunction with the irreversible one. The intuitive consequence is that $l_{t+1}$, the reversible factor, is more volatile than $k_{t+1}$, the irreversible one, both after a positive and a negative shock.

This is clear from the comparison of (4.14) and (4.15). After a negative productivity shock at time $t$ the marginal productivity of both $k_{t+1}$ and $l_{t+1}$ decreases. $E$ reduces the investment in variable capital $l_{t+1}$, while $k_{t+1}$ cannot be reduced below $(1-\delta_k)k_t$ because of constraint (4.3). Following Caballero (1997) we use the notation of the "the desired level of capital", which in this case is $k_{t+1}^{PM}$. Therefore a binding irreversibility constraint implies that $(1-\delta_k)k_t = k_{t+1} > k_{t+1}^{PM}$. This corner solution in the fixed capital is reflected by a positive $\mu_t$ on the right hand side of (4.15), which indicates that the marginal productivity is lower than the user cost of fixed capital. It is interesting to note that while a positive $\lambda_t$, in case of a binding collateral constraint, indicates that investment is inefficiently low, a positive $\mu_t$ signals that investment in fixed capital is inefficiently high. In this sense financing and irreversibility problems have opposite effects.

Instead after a positive productivity shock $E$ wants to invest more in both factors. Therefore (4.3) is not binding and $\mu_t = 0$, while $E_t\left(\mu_{t+1}\right) > 0$ because, applying the same reasoning made before, (4.3) can be binding at time $t+1$ conditional on a future negative shock. The positive $E_t\left(\mu_{t+1}\right)$ represents the cost associated to future expected irreversibility. Such cost increases the required marginal productivity of fixed capital $E_t\left(MPK_{t+1}\right)$, thereby reducing $k_{t+1}$. Therefore $k_{t+1}$ increases less than $l_{t+1}$ after a positive shock.

On the one hand this result is consistent with the empirical evidence that input inventories are relatively more volatile than fixed capital. On the other hand it implies that they are not very volatile in absolute terms, because irreversibility indirectly affects them trough the complementarity of the two factors of production. Hence irreversibility alone does not explain the high volatility of inventories and the other stylised facts cited in section 1.5. Therefore we now analyse the effects of financing constraints.

## 4.4 The two factors model with financing constraints

In this section we rule out current and future expected irreversibility constraints by considering the parameter set $\overline{\Theta}^2 \subset \overline{\Theta}$ such that the irreversibility constraint (4.3) is never binding[6]. Hence $\mu_t = E_t(\mu_{t+1}) = 0$ at $t = 0, 1, ..., \infty$. In this case the problem defined by (4.9), (4.10), (4.11) and (4.12) is equivalent to the basic problem analysed in the previous chapter, but with two factors of production. In appendix 1 we show that proposition 3 applies to both fixed and variable capital. Therefore all the results derived in chapter 3 regarding future expected financing constraints and the precautionary saving effect are extended to the two factor model. In particular, let's consider (4.11) and (4.12) evaluated at $\mu_t = E_t(\mu_{t+1}) = 0$ :

$$(1+\eta) E_t(MPL_{t+1}) = UL + E_t(\Omega_{l,t+1}) + R^2\lambda_t \tag{4.16}$$

$$(1+\eta) E_t(MPK_{t+1}) = UK + E_t(\Omega_{k,t+1}) + DR^2\lambda_t \tag{4.17}$$

(4.16) and (4.17) are the counterparts of (3.23). If $w_t < \underline{w}_t$ then the collateral constraint (3.6) is binding, and $\lambda_t > 0$. In this case $\lambda_t, k_{t+1}$ and $l_{t+1}$ are jointly determined by (4.9), (4.16) and (4.17). The three nonlinear equations cannot be solved analytically, but the intuition of the result is straightforward: the amount of wealth available for investing is shared between fixed and variable capital according to their respective productivities and costs. Lets consider the ratio between (4.16) and (4.17):

$$\frac{E_t(MPL_{t+1})}{E_t(MPK_{t+1})} = \frac{UL + E_t(\Omega_{l,t+1}) + R^2\lambda_t}{UK + E_t(\Omega_{k,t+1}) + DR^2\lambda_t} \tag{4.18}$$

(4.18) shows that current and future expected financing constraints affect the optimal mix between $l_{t+1}$ and $k_{t+1}$. If financing constraints do not matter (i.e. $w_0 = w^{MAX}$), then $\lambda_t = E_t(\Omega_{l,t+1}) = E_t(\Omega_{k,t+1}) = 0$ and (4.18) becomes the standard condition

$$\frac{E_t(MPL_{t+1})}{E_t(MPK_{t+1})} = \frac{UL}{UK} \tag{4.19}$$

Alternatively if the financing constraint is binding, for $w_t$ that decreases towards $w^{\min}(\theta_t, k_t)$ the intensity of the financing constraint $\lambda_t$ monotonously increases, and from (4.18) it follows that the optimal mix converges to the following:

[6]This happens when $\delta_k$ is large relative to the volatility of the productivity shock.

$$\lim_{\lambda_t \to \infty} \frac{E_t\left(MPL_{t+1}\right)}{E_t\left(MPK_{t+1}\right)} = \frac{R}{RD} \tag{4.20}$$

(4.20) means that when $E$ is constrained the relevant cost is not the user cost $UL$ and $UK$, but rather the capitalised value of the downpayment necessary to buy one additional unit of capital, which is equal to R and R$D$ for $l_{t+1}$ and $k_{t+1}$ respectively.

This however implies that (4.20) converges to (4.19) when the residual value of capital is fully collateralisable ($\tau_k = 1 - \delta_k$). In fact in this case $\frac{R}{RD} = \frac{UL}{UK}$ and financing constraints do not alter the optimal mix between $k$ and $l$. This result is counterintuitive. Given that fixed capital can be used as collateral while variable capital cannot, one would expect that, ceteris paribus, a constrained **E** uses more fixed capital than an unconstrained one. In order to explain why this does not happen, we note that in this model what distinguishes fixed capital from variable capital is not the collateral capacity. In fact strictly speaking both factors have the same collateral capacity, in the sense for both of them only the next period residual value can be used as collateral. Such value is always zero for variable capital.

Therefore the real difference between the factors is the fact that fixed capital does not fully depreciates after one period. This implies that, since the prices of fixed and variable capital are identical and equal to one, fixed capital has a lower user cost than variable capital. Therefore an unconstrained firm, which compares the user costs of the two factors in order to determine the optimal mix between fixed and variable capital, will choose an equilibrium marginal return on fixed capital lower than the one of variable capital.

At the same time, since fixed capital does not fully depreciates, it has some residual value that can be used as collateral. For this reason the constrained **E** values fixed capital more than variable capital. When $\tau_k = 1 - \delta_k$ then the advantage of fixed capital over variable capital for an unconstrained **E** (lower user cost) is identical of the advantage for a constrained **E** (collateral value). As a consequence both the constrained and the unconstrained **E** choose the same optimal mix of factors. In order for fixed capital to be used more intensely by the constrained **E** it must be that fixed capital is "more collateralisable" than variable capital. Since $\tau_l = 1 - \delta_l$ , then this happens when $\tau_k > 1 - \delta_k$. In this case (which is not considered in the paper since it would be inconsistent with the implicit assumptions about the enforceability problem) we would have that the constrained **E** would choose an higher proportion of fixed capital with respect to the unconstrained **E**.

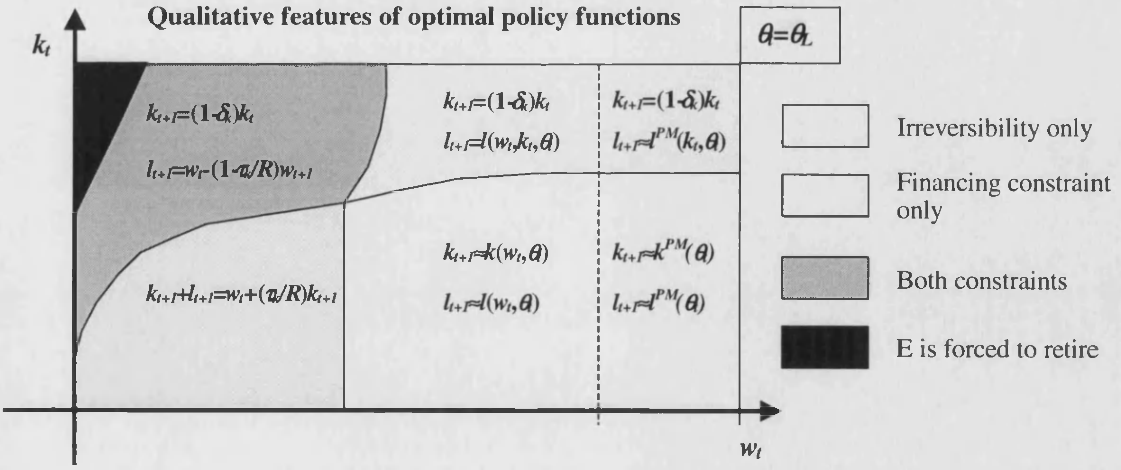
Now let's consider the case with $w_0 < \overline{w}$ and $\overline{w} > w_t \geq \underline{w}_t$. In this case $\lambda_t = 0$ but $E_t(\Omega_{l,t+1}) > 0$ and $E_t(\Omega_{k,t+1}) > 0$. These two terms represent the precautionary saving effect on variable capital and fixed capital respectively. In this case $k^o_{t+1}$ and $l^o_{t+1}$ represent the optimal levels of fixed and variable capital, and we can rearrange (4.16) and (4.17) to show that the optimal mix between $k^o_{t+1}$ and $l^o_{t+1}$ is determined by:

$$\frac{E_t(MPL_{t+1})}{E_t(MPK_{t+1})} = \frac{UL\left[1-\gamma+\gamma E_t(\phi_{t+1})\right] - \gamma cov\left[\phi_{t+1}MPL_{t+1}\right]}{UK\left[1-\gamma+\gamma E_t(\phi_{t+1})\right] - \gamma cov\left[\phi_{t+1}MPK_{t+1}\right]} \tag{4.21}$$

Assuming that $cov\left[\phi_{t+1}MPL_{t+1}\right]$ and $cov\left[\phi_{t+1}MPK_{t+1}\right]$ have a marginal effect, and substituting them with 0 in (4.21), we note that (4.21) simplifies again to (4.19). The conclusion of this section is that, with financing constraints only, the two factor model simply replicates the results of the single factor one. In fact future expected financing constraints affect the allocation between saving and investment, but do not alter the optimal mix between fixed and variable capital.

## 4.5 Interaction between financing and irreversibility constraints

Figure 4-2: Policy functions conditional on low productivity shock




We consider now the solution of the problem with both constraints. Figure 4-2 summarises the different types of optimal policy functions $k_{t+1}(w_t, \theta_t, k_t)$ and $l_{t+1}(w_t, \theta_t, k_t)$ in the $\{k_t, w_t\}$ space, conditional on $\theta_t = \theta_L$. Instead of describing in detail such solutions, we focus only on the most interesting feature: the fact that irreversibility and financing constraints interact and amplify each other. We firstly note that, when both constraints are binding, $\mu_t$ is determined by equation (4.12). By substituting recursively we obtain the following:

$$R\mu_t = E_t\left[\sum_{j=0}^{\infty} \nu^j \left(\Gamma_{k,t+j+1} + \Omega_{k,t+j+1} + D_k R \lambda_{t+j}\right)\right] \tag{4.22}$$

where $\Gamma_{k,t+1} = U_k - (1+\eta) E_t(MPK_{t+1})$ and $\nu = (1-\delta_k)\gamma$. Equation (4.22) shows that the cost of irreversibility increases in the present and expected costs of financing constraints. More importantly, the reverse is also true: the irreversibility constraint increases the chances of facing financing constraints now or in the future. The intuition is that, when fixed capital is irreversible $E$ knows that, in the case of a negative persistent productivity (or demand) shock, like at the beginning of an economic downturn, it will take some time to reduce the stock of capital to the new optimal level. During the adjustment period expected profits drop, and the expected rate of wealth accumulation drops as well relatively to a situation where fixed capital is reversible. As a result $E$ has a higher chance of facing future financing constraints: i) if the contraction period is long enough, then the drop in wealth can be so severe that $E$ does not have enough funds to invest in variable capital; ii) if the contraction period ends, she has not enough resources to invest in both fixed and variable capital, and will be financially constrained for some time. In order to compensate these higher costs of future financing problems $E$ engages in precautionary saving ex ante, more than she would have done with reversible fixed capital. These considerations can be summarised in the following proposition:

**Proposition 5** *for any set of state variables* $\{w_t, k_t, \theta_t\}$ *such that* $\lambda_t = 0$, $\mu_t = 0$, $E_t\left(\sum_{j=1}^{\infty} \lambda_{t+j}\right) > 0$ *and* $E_t(\mu_{t+1}) > 0$, *we have that* $\phi_t(w_t, k_t, \theta_t \mid k_{t+1} \text{ is irreversible}) > \phi_t(w_t, k_t, \theta_t \mid k_{t+1} \text{ is reversible})$.

**Proof: see appendix 1.**

Proposition 5 shows that the shadow value of money, which reflects the cost of future expected financing constraints, is higher when fixed capital is irreversible. This implies that the precautionary saving effect, which is monotonously increasing in $\phi_t$, is stronger when fixed capital is irreversible. A quantification of this amplification effect is computed in the next section in figure 4-6.

## 4.6 Model's numerical solution

We solve the dynamic nonlinear system of equations defined by (4.3), (4.9), (4.10), (4.11) and (4.12) using the same numerical method adopted to solve the basic model in chapter 3. First, we formulate an initial guess of the forward variables $E_t\left(\phi_{i,t+1}\right)$, $E_t\left(\phi_{i,t+1}\theta_{i,t+1}\right)$ and $E_t\left(\mu_{i,t+1}\right)$. Second, we solve the static optimisation problem conditional on this guess, for each discrete value of the state variables $w_{i,t}, k_{i,t}$ and $\theta_{i,t}$. This static optimisation problem is nonlinear as well. Conditional on the value of $\theta_{i,t}$ the solution falls in four possible categories, depending on the values of the couple $\{w_{i,t}, k_{i,t}\}$. These categories correspond to the four areas in figure 4-2. Third, we update the guess of $E_t\left(\phi_{i,t+1}\right)$, $E_t\left(\phi_{i,t+1}\theta_{i,t+1}\right)$ and $E_t\left(\mu_{i,t+1}\right)$. We repeat these steps until the value function converges. Solving the problem for all possible values of the state variables yields the optimal policy functions $k_{i,t+1}\left(w_{i,t}, \theta_{i,t}, k_{i,t}\right)$ and $l_{i,t+1}\left(w_{i,t}, \theta_{i,t}, k_{i,t}\right)$, the associated Lagrange multipliers $\lambda_{i,t}\left(w_{i,t}, \theta_{i,t}, k_{i,t}\right)$, $\mu_{i,t}\left(w_{i,t}, \theta_{i,t}, k_{i,t}\right)$ and $\phi_{i,t}\left(w_{i,t}, \theta_{i,t}, k_{i,t}\right)$, and the value function $\mathrm{W}'_{i,t}\left(w_{i,t}, \theta_{i,t}, k_{i,t}\right)$. From now on we relax assumption *t3* and we model uncertainty in the following way:

$$\theta_{i,t} = \theta^f_{i,t}\varepsilon_t \tag{4.23}$$

$\theta^f_{i,t}$ is the idiosyncratic productivity shock, and has the same structure for all entrepreneurs:

$$\theta^f_{i,t} \in \{\theta_L, \theta_H\} \ with \ \theta_H > \theta_L \qquad pr(\theta^f_{i,t+1} = \theta^f_{i,t}) = \epsilon > 0.5 \qquad pr(\theta^f_{i,t+1} \neq \theta^f_{i,t}) = 1 - \epsilon$$

$\varepsilon_t$ can be interpreted as an economy-wide demand shock. Each realisation of $\varepsilon_t$ is a transitory shock that affects revenues and net worth but does not affect future expected profitability. The probability distribution of $\varepsilon_t$ depends on the "state of the world" $S_t$:

| | Expansion phase ($S_t = S_G$) | Contraction phase ($S_t = S_B$) |
|---|---|---|
| $pr(\varepsilon_t = \varepsilon_H)$ | $\nu$ | $1-\nu$ |
| $pr(\varepsilon_t = \varepsilon_L)$ | $1-\nu$ | $\nu$ |

$$\nu > 0.5$$

$$pr(S_{t+1} = S_t) = \xi > 0.5 \qquad pr(S_{t+1} \neq S_t) = 1 - \xi$$

A change in $S_t$ can be interpreted as an unexpected change in monetary policy, or another unexpected event that modifies expectations about economic activity. Since both $\nu$ and $\xi$ are greater than 0.5, it follows that a change in $S_t$ represents a persistent macroeconomic shock that changes future expected profitability, and shifts the economy from a contraction phase to an expansion phase or vice-versa. This modelling choice of aggregate uncertainty is useful because it allows us to better distinguish the financing constraints economy, in which any transitory shock that changes financial wealth affects investment choices, from the perfect market economy, in which only persistent shocks that change expected productivity affect investment choices[7].

The parameter values used for the solution are selected according to the following criteria: first, the parameters regarding the idiosyncratic productivity shock $\theta_{i,t}^{f}$ are estimated from our panel-data of Italian manufacturing firms. We estimate[8] the idiosyncratic productivity shock $\widehat{\theta}_{i,t}^{f}$ for each firm $i$ and for each year $t$. We then discretise $\widehat{\theta}_{i,t}^{f}$ in two intervals and estimate a two state transition matrix. $\theta_L$ and $\theta_H$ are the average value of the lower and the upper interval and $\epsilon$ is the estimated probability of remaining in the same state[9]: $\theta_L = 4.86$, $\theta_H = 5.37$ and $\epsilon = 0.66$. The parameters calibrated for the aggregate shock are the following[10]: $\xi = 0.85$, $\nu = 0.6$, $\varepsilon_L = 0.93$ and $\varepsilon_H = 1.07$. Another set of pa-

[7]It is important to note that such assumption is not essential for the aggregate results derived in the next section. All the results would hold with a simplier two state aggregate shock with the same structure than the idiosyncratic shock $\theta_{i,t}^{f}$. Only the result illustrated in figure 4-7 would be quantitatively less important and in general more sentive to the choice of the other parameters.

[8]The estimation procedure employed is described in detail in the next chapter.

[9]The estimated transition probabilities to change state, $pr(\theta_{t+1} = \theta_H \mid \theta_t = \theta_L)$ and $pr(\theta_{t+1} = \theta_L \mid \theta_t = \theta_H)$ are almost identical, allowing the use of a symmetric stochastic process. We chose the value of $\epsilon = 0.66$ because the estimated $\widehat{\epsilon}$ ranges between 0.65 and 0.68 in five out of the 7 industrial sectors considered and it is slightly smaller in the other two groups (see figure 5-5).

[10]Their implications for aggregate fluctuations are discussed in the next section.

rameters is "indirectly" estimated: $\alpha = 0.039$, $\beta = 0.783$, $\delta_k = 0.06$, $\delta_l = 1$, $\tau_k = 0.94$ and $R = 1.015$. The chosen values of $\alpha$ and $\beta$[11] imply an average optimal fixed and variable capital in the steady state equivalent to the average fixed and variable capital observed in our dataset, for a 6 months productive period. The other parameters are estimated from Italian macroeconomic data for the same period covered by our dataset, 1982-1992.

The last three parameters are calibrated as follows: $l^E = 0.027 * l_{t+1}^{PM} \left(\theta_t = \theta_H; S_t = S_G\right)$, $\eta = 0.4$ and $\gamma = 0.95$. $l^E$ is the minimum value that satisfies assumption *a1*, and is equal to 2.7% of the maximum unconstrained variable capital. The value of $\gamma$ implies that each period the exit rate is 5%. The value of $\eta$ implies that consumption is equal to a share of $\eta / (1+\eta) = 28.5\%$ of total revenues. In figures 4-3, 4-4 and 4-5 we show the optimal policy functions of $k_{i,t+1}$ and $l_{i,t+1}$ as a function of $w_{i,t}$, for given values of $\{k_{i,t}, \theta_{i,t}\}$[12]. In figures 4-3 and 4-4 $k_{i,t} = 0$ and the irreversibility constraint is not binding for any value of $w_{i,t}$, while $\theta_t$ is equal to $\theta_L$ and $\theta_H$ and $S_t$ to $S_B$ and $S_G$ respectively.

These figures confirm the qualitative analysis presented in chapter 3. Both $k_{i,t+1}$ and $l_{i,t+1}$ increase linearly[13] in the net worth in the interval $w_{i,t} \in (w_{i,t}^{\min}, \underline{w}_{i,t})$ where the borrowing constraint is binding. For $w_{i,t} \geq \underline{w}_{i,t}$ the borrowing constraint is not binding any more, and $k_{i,t+1}^o$ and $l_{i,t+1}^o$ gradually converge towards the perfect market levels $k_{i,t+1}^{PM}$ and $l_{i,t+1}^{PM}$, as $w_{i,t}$ increases and expected financing constraints decrease. Figure 4-3 shows that in the left end of the unconstrained region, for $w_{i,t} = \underline{w}_{i,t}$, the precautionary saving effect on variable capital is equal to $\frac{l_{i,t+1}^{PM} - l_{i,t+1}^o}{l_{i,t+1}^{PM}} = 32.4\%$, while the same effect for fixed capital is equal to $\frac{k_{i,t+1}^{PM} - k_{i,t+1}^o}{k_{i,t+1}^{PM}} = 31.1\%$. Given that the parameters used for these solutions are comparable to those used in chapter 3 for the solution of the basic model, it is possible to compare these values with the one showed in figure 3-2. The maximum precautionary saving effect conditional on $w_{i,t} = \underline{w}_{i,t}$, $k_t = 0$ and $\theta_t = \theta_L$ is a bit lower in the basic model, being equal to 30.7% . The difference is minimal because figure 3-2 is relative to a situation in which there are very low future expected irreversibility problems. In fact $k_t = 0$, which

[11] $\alpha$ and $\beta$ are not directly estimated because the production function we estimate in the next section also includes labour cost. The values above are in line with previous estimations we conducted computing output as added value net of labour cost.

[12] The values are equivalent to billion of Italian lira at 1982 prices. 1 billion liras was equivalent to 0.71 million US$ at the 1982 exchange rate.

[13] $k_{t+1}$ is slightly bented because the relative importance of $l_m^E$ decreases as $l_{t+1}$ increases.

Figure 4-3: Policy functions of fixed and variable capital conditional on low productivity shock and low stock of fixed capital

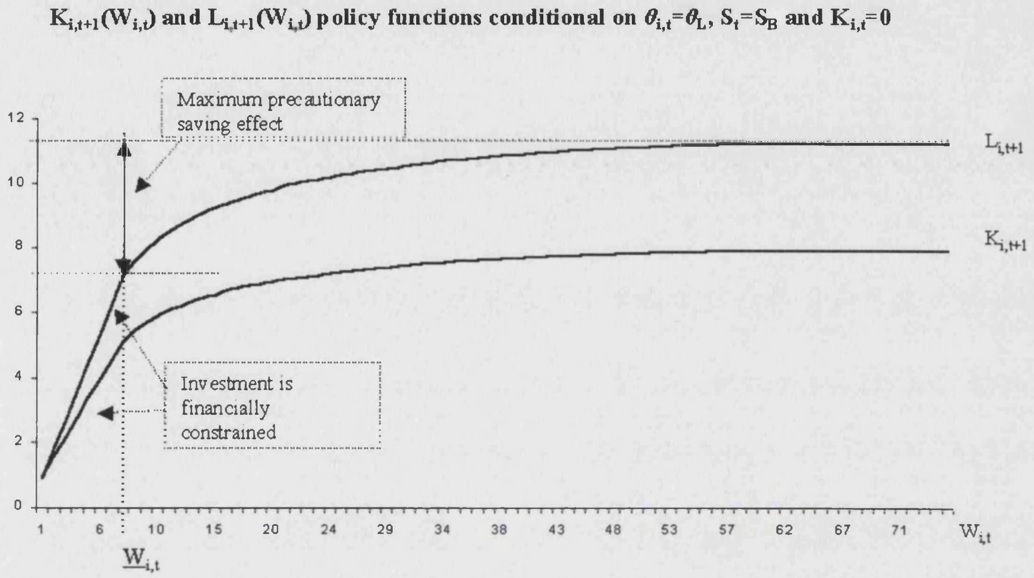
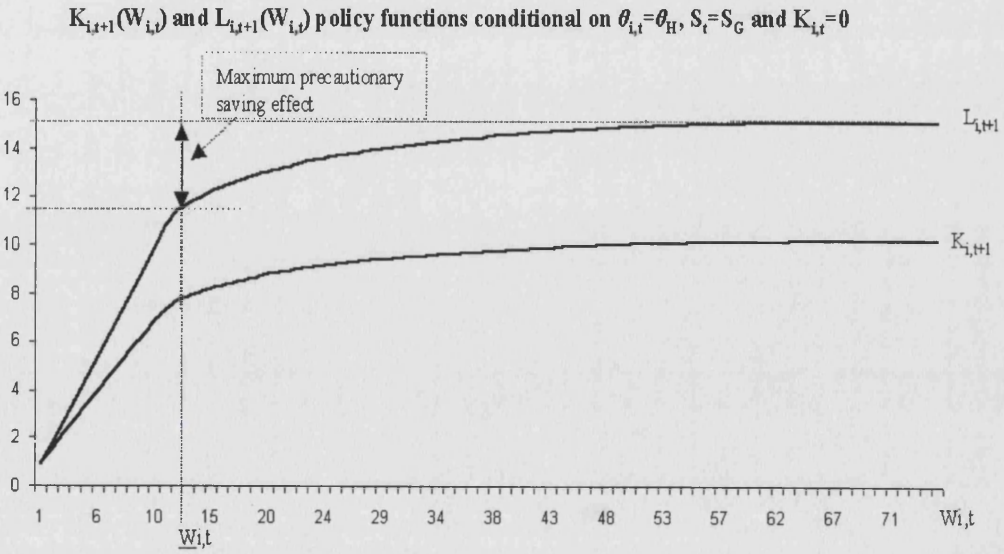
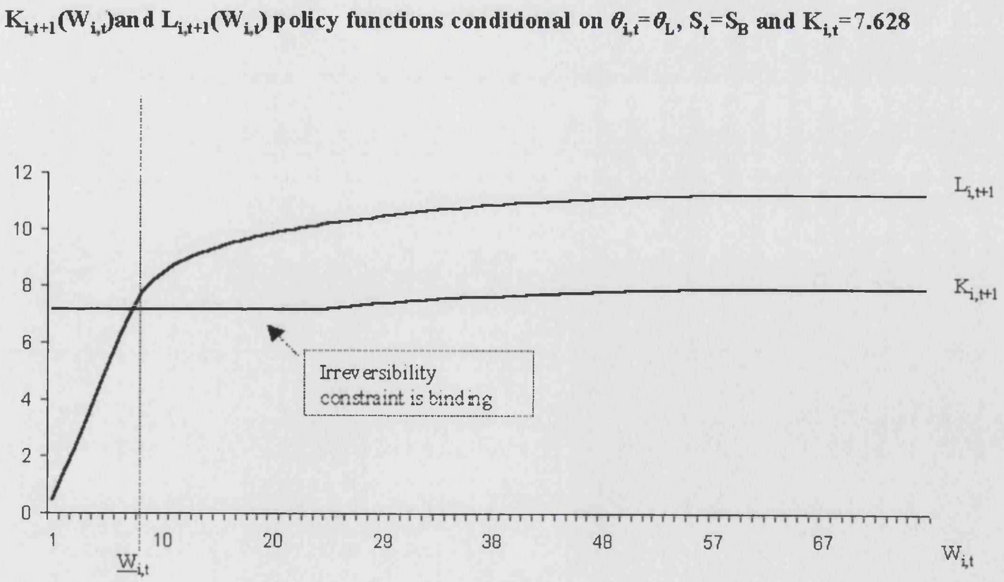



Figure 4-4: Policy functions of fixed and variable capital conditional on high productivity shock and low stock of fixed capital



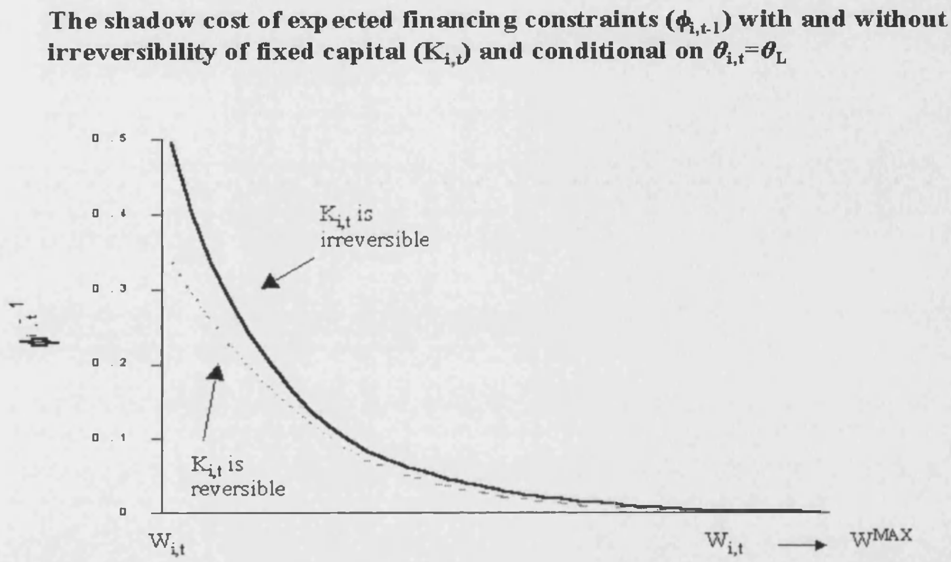
means that $E$ does not have an excess stock of fixed capital at time $t$, but also $\theta_t = \theta_L$ and $S_t = S_B$, which means that things "can only get better" and $E$ is expecting to increase the level of fixed capital in the immediate future. As a consequence, without current and (immediate) future expected irreversibility problems the amplification effect defined by proposition 5 is almost absent. Figure 4-4 shows that conditional on good news ($\theta_t = \theta_H$ and $S_t = S_G$) the binding constraint region expands, because $E$ wants to invest more today, while the precautionary saving region shrinks, because $E$ is also more optimistic about the future. The maximum precautionary saving effect conditional on $w_{i,t} = \underline{w}_{i,t}$ is equal to 24% for both $l_{t+1}$ and $k_{t+1}$, while it is only equal to 20% in the comparable policy function in the basic model (see figure 3-3). This larger difference is due to the fact that in figure 4-4, for $\theta_t = \theta_H$ and $S_t = S_G$, there are some future expected irreversibility problems. This is because productivity is at its top, and conditional on future negative persistent shocks ($\theta_{t+1} = \theta_L$ and/or $S_{t+1} = S_B$) fixed capital will be inefficiently high and profits very low for some time. Hence the precautionary saving effect is amplified by future expected irreversibility problems.

Figure 4-5: Policy functions of fixed and variable capital conditional on low productivity shock and high stock of fixed capital




In figure 4-5 $\theta_t$=$\theta_L$ and $S_t = S_B$, but $k_{i,t} = 7.628$ and $k_{i,t+1}$ is flat in correspondence

Figure 4-6: The amplification of future expected financing constraints due to the irreversibility of fixed capital




with the lower values of $w_{i,t}$ because the irreversibility constraint is binding. Therefore in this region the changes in net worth $w_{i,t}$ affect $l_{i,t+1}$ only. When the irreversibility constraint is binding then the amplification effect is stronger, but the increase in future expected financing constraints does not translate in a larger difference between $l_{t+1}^{PM}$ and $l_{t+1}^{o}$. This is because of two opposite effects: on the one hand $E$ would like to precautionarily reduce $l_{t+1}$, on the other hand with $k_{t+1}$ inefficiently high it is also convenient to maintain a higher level of $l_{t+1}$, because the two factors are complementary. It is important to clarify that this still implies that variable capital is more volatile as a consequence of the amplification effect. This is because both $l_{t+1}$ and $k_{t+1}$ are inefficiently high and this means that profits and financial wealth will drop more if the productivity shock is low in the future. Therefore if there is a sequence of negative shocks financial wealth deteriorates more rapidly towards the binding constraints region, and in figure 4-5 variable capital declines faster as well, moving leftward along the policy function schedule.

This consequence of the amplification effect, a sharp drop in variable capital during downturns, is quantified in the aggregate simulations presented in the next section. An alternative way to illustrate graphically the amplification effect derived in proposition 6 is shown in figure 4-6. In the previous section we explained that $\phi_{i,t} - 1$ can be interpreted as the premium in the shadow value of financial resources induced by future expected financing problems. In figure 4-6 we compare the value of $\phi_{i,t} - 1$ for $w_{i,t} \in [\underline{w}_{i,t}, \infty)$ ("the precautionary saving region") for two identical simulations, which differ only regarding the reversibility of fixed capital. The figure shows that, when $k_{i,t+1}$ is irreversible, the value of $\phi_{i,t} - 1$ is up to 50% higher in comparison with when $k_{i,t+1}$ is reversible. It is important to note that this difference further increases in the region (not showed in figure 4-6) with $w_{i,t} \in (w_{i,t}^{\min}, \underline{w}_{i,t})$, where the borrowing constraint is binding.

## 4.7 Dynamics of aggregate output and investment

In this section we use the model's solution to simulate the investment and production path of many heterogeneous entrepreneurs. The aim is to show how the combination of irreversibility and financing constraints generates a behaviour of aggregate variables consistent with the stylised facts outlined in the introduction. In the simulated economy, the behaviour of aggregate investment and production depends on the heterogeneity of entrepreneurs in terms of the state variables. All entrepreneurs are identical ex ante, but each of them is subject to different realisation of the idiosyncratic productivity shock $\theta_{i,t}^f$, which is uncorrelated across entrepreneurs and serially correlated for each entrepreneur. Therefore at time $t$ different entrepreneurs have different values of $w_{i,t}$ and $k_{i,t}$, depending on $\left\{\theta_{i,j}^f\right\}_{j=0}^t$. The steady state distributions of $\{w_{i,t}, k_{i,t}\}$ at time $t$ depends on the parameters set $\overline{\Theta}$ and on the rule regarding the endowment level of newborn entrepreneurs in the economy. In each period the number of new entrepreneurs is equal to the number of retired ones, to keep the active population constant. For the simulations presented in this section we keep the same parameters used in the previous section, and we assume that the initial endowment of new entrepreneurs at time $t$ is equal to $w_{i,t}^{new} = 2 * k_{t+1}^{PM}(\theta_t = \theta_H,$ $S_t = S_G)$. Such endowment implies that a new entrepreneur is not initially constrained but must borrow, to reach the optimal investment level, more than an average existing entrepreneur. This implies that new entrepreneurs engage in more precautionary saving until they build the desired level of net wealth. A lower value of $w_{i,t}^{new}$ would imply that new entrepreneurs are closer to the binding financing constraint region. This would even increase the quantitative results of this section. By choosing a higher value of $w_{i,t}^{new}$ we want to show that expected financing constraints are sufficient to generate the results regarding the difference between small and large firms behaviour. As mentioned in the introduction, the shape of the distribution of $\{w_{i,t}, k_{i,t}\}$ influences the way aggregate investment responds to the recurrent transitory ($\varepsilon_t$) and permanent ($S_t$) shocks. This feature is essential to explain why variable investment leads the business cycle, while fixed investment lags it and why small firms are more procyclical than large ones in output and variable investment.

The statistics illustrated in the remainder of this section are generated by the simulation of an economy of 3000 firms, for 3000 periods. The chosen values of the aggregate shock $\xi = 0.85$, $\nu = 0.6$, $\varepsilon_L = 0.93$ and $\varepsilon_H = 1.07$ generate stochastic business cycles which vary for length and amplitude, depending on the realisations of $S_t$ and $\varepsilon_t$. We can have periods with no cycles, when $S_t$ changes frequently, periods with "V" shaped cycles when a change from $S_G$ to $S_B$ is followed by a sequence of negative transitory shocks in $\varepsilon$ (and vice-versa), and periods with more gradual contractions and expansions in the simulated economy, when $S_t$ is constant for many periods but $\varepsilon$ has both positive and negative realisations. To give an idea of the amplitude of the cycles generated in the simulated industrial sector, we compute the cumulative growth of aggregate output and calculate the average difference between the 300 observations with higher cumulative growth and the 300 observations with lower cumulative growth. The average difference is equal to 30%, which is comparable with the empirical evidence of the cumulative (de-seasoned) real growth rate of the Italian manufacturing sector in the 90s (see figure 2-18).

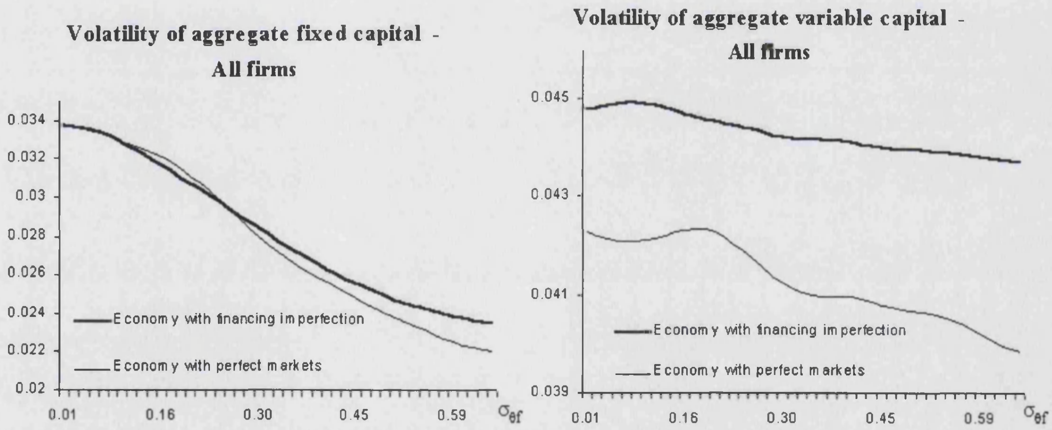
Figure 4-7 shows aggregate statistics regarding the volatility of fixed and variable capital, for the economies with and without financing constraints. The economy with financing constraints is the one where all the assumptions of the extended model hold, while the economy with perfect markets has the two following changes:

*PM1)* There is no minimum consumption constraints. All output $(1+\eta)\, y_t$ is financial and can be consumed, invested or saved.

*PM2)* The entrepreneurs can use future expected output as collateral.

The fact that all the output is financial in the perfect market economy implies that $k_{t+1}^{PM}$ and $l_{t+1}^{PM}$ are the level of fixed and variable capital that maximise net financial profits. Moreover *PM2* implies that banks are always willing to finance all positive net present value projects, and hence entrepreneurs are never financially constrained in their investment decisions. These assumptions imply that optimal investment in the perfect market economy is independent from the financial wealth of the entrepreneurs and is determined by the solution of (4.3), (4.14) and (4.15). In figure 4-7 aggregate variables are plotted as a function of the level of idiosyncratic risk measured by $\sigma_{\theta f}$. We simulate the economies for different values of $\sigma_{\theta f}$, ranging from 0 (no idiosyncratic uncertainty) to 0.65. This interval includes the vale of $\sigma_{\theta f}$ estimated from our panel data of Italian firms. In fact the estimated $\widehat{\theta}_L = 4.86$ and $\widehat{\theta}_H = 5.36$ correspond to $\widehat{\sigma}_{\theta f} = 0.36$.

Figure 4-7: Volatility of aggregate fixed and variable capital - economies with and without financing constraints (measured as the standard deviation of percentage changes in the stock of capital)



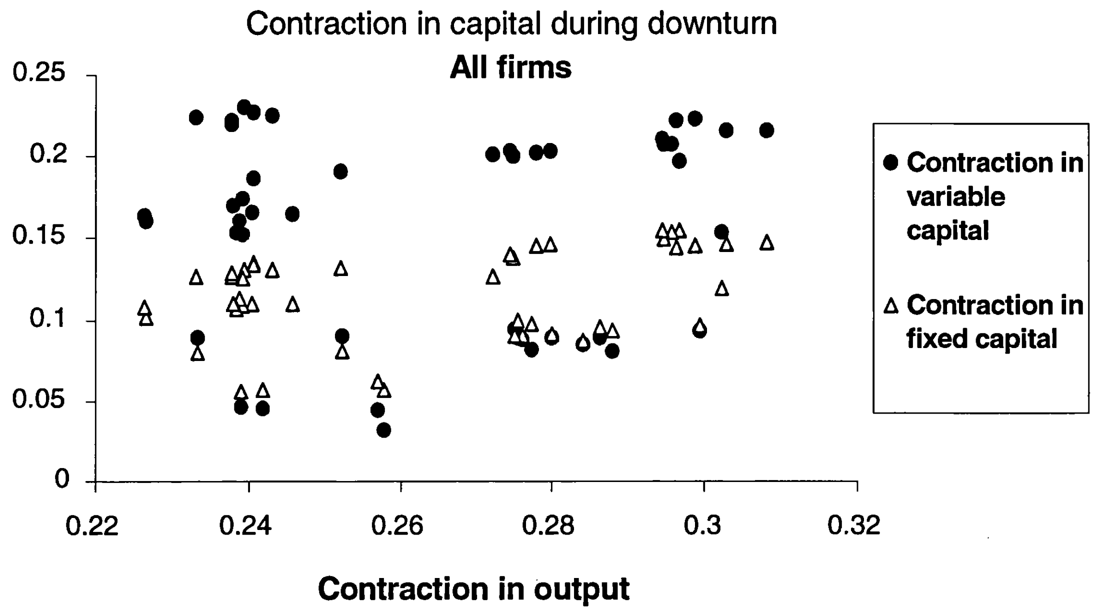

On the left side of figure 4-7 the volatility of aggregate fixed capital decreases in $\sigma_{\theta f}$, because higher idiosyncratic risk implies that the steady state distribution of fixed capital becomes more dispersed and hence a bigger fraction of entrepreneurs experiences in each period a binding irreversibility constraint. The two economies show similar levels of fixed capital volatility, as two opposite effects almost exactly compensate: on the one hand investment in the economy with financing constraints reacts to transitory shocks, which affect aggregate wealth, while transitory shocks are irrelevant in the economy with perfect markets. On the other hand investment in the perfect markets economy reacts more to permanent shocks than investment in the financially constrained economy.

On the right side of figure 4-7 we show the volatility of aggregate variable capital. Due to its reversibility variable capital is more volatile than fixed capital in both economies. However variable capital also decreases in $\sigma_{\theta f}$, because the irreversibility has an indirect effect through the complementarity of the two factors.

Despite such complementarity, while fixed capital has similar volatility in the two economies, variable capital is more volatile in the economy with financing constraints. This is because the irreversibility of fixed capital amplifies the effects of financing constraints on variable capital. In fact after a negative persistent shock entrepreneurs can only adjust variable capital, fixed capital being irreversible. But the amplification effect implies that after such shock current and future expected financing problems also increase, and the sensitivity of variable capital to changes in net worth becomes stronger as well. As a consequence variable capital reacts more to such negative permanent shocks in the constrained economy. Moreover the difference between the volatility of variable capital in the two economies increases in $\sigma_{\theta f}$, because such amplification effect increases in the amount of firms with a binding irreversibility constraint. This result implies that the empirical findings of Ramey (1989), Blinder and Maccini (1991) and Ramey and West (1999) regarding the high volatility of input inventories can be explained by the presence of financing imperfections. In fact the same authors emphasise the fact that input inventories are more volatile than finished goods inventories, and $l_{t+1}$ can be interpreted as the end of the period $t$ stock of input inventories, like raw materials and work in progress. They also show that the high volatility of inventories is important because it implies that the decline in inventories accounts for a large part of the decline in GDP in both US and G7 countries recessions. The amplification effect helps to explain this stylised fact. To show it more clearly, we illustrate in figure 4-8-4-10 the relation between episodes of sharp decline in industrial output in our simulated economy and the correspondent decline in fixed and variable capital. We define as a "sharp decline" a sequence of 5 consecutive periods of decline in aggregate industrial output that cause an overall decline of at least

20% in its cumulative growth rate. We find 46 of such sequences in our simulation. Figure

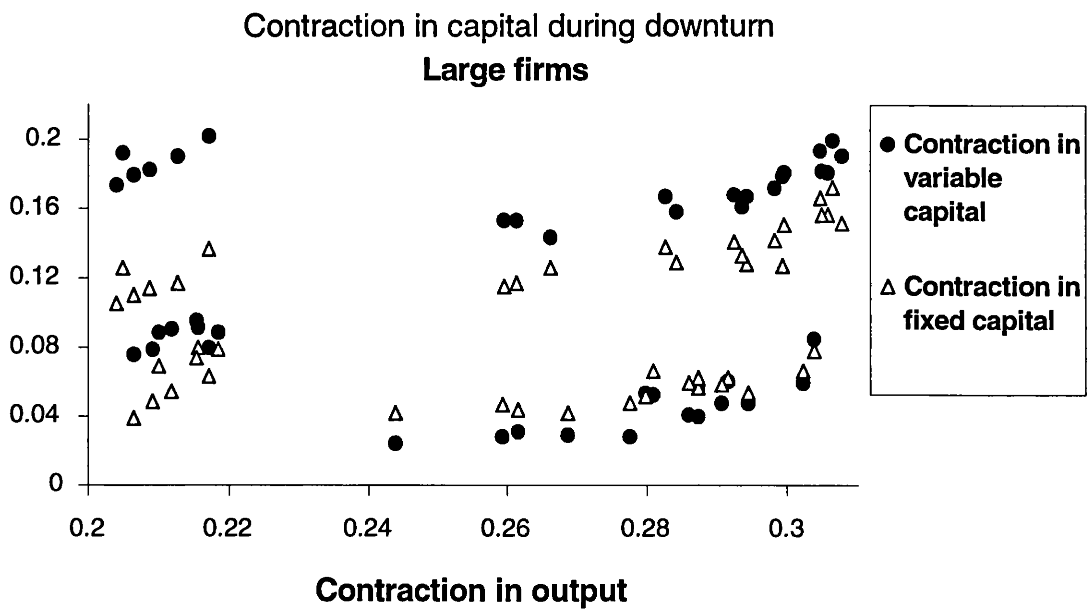
Figure 4-8: Contraction in fixed and variable capital during downturns - all entrepreneurs




4-8 shows that there is a considerable heterogeneity in the response of variable and fixed capital to the output decline in each episode. This is not only because the nature of the cycles depends a lot on the combination of permanent and aggregate shocks, but also because aggregate responses to the same aggregate shock change a lot depending on the distribution of wealth and fixed capital among entrepreneurs. In chapter 1 we mention that Caballero and Engel (1999) derive the aggregate investment function in presence of non convex adjustment costs, to show that in general aggregate investment depends on the cross sectional density of plants capital imbalances, so that a sequence of positive aggregate shocks not only causes some plants to invest, but also changes the cross sectional distribution of imbalances, meaning that more and more plants are likely to adjust at each subsequent aggregate shock. Our theory implies similar nonlinear effects generated by the distribution of financial wealth and fixed capital. If a sequence of negative shocks reduces the wealth for all firms, each subsequent shock has a larger negative effect on variable capital, because entrepreneurs engage in more precautionary saving. This effect is more or less amplified by the irreversibility of fixed capital depending on the cross sectional distribution of fixed capital among entrepreneurs.

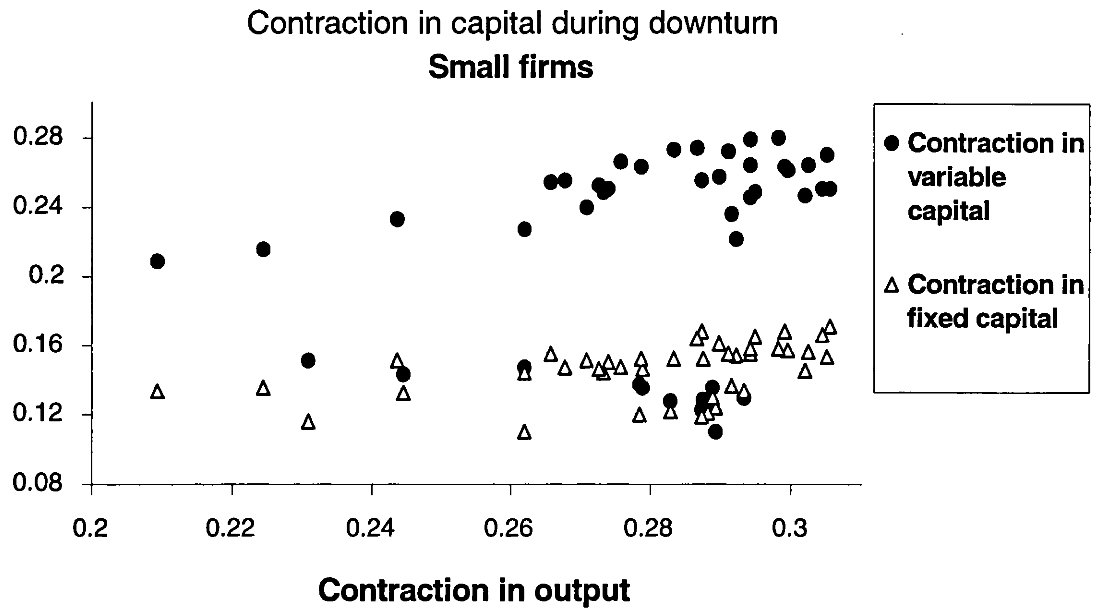
These considerations are illustrated in figure 4-8. The contraction in variable capital is much bigger than the one in fixed capital, and both are roughly increasing in the magnitude of the drop in output. How much current and future expected financing constraints are responsible for the drop in variable capital? The easiest way to quantify it is to compare the same statistic for small and large firms. Smaller firms in the sample are less wealthy, more financially constrained and more sensitive to aggregate shocks than large firms. This is due to two effects: i) "age effect": small entrepreneurs are younger ones, which are accumulating wealth to reach their optimal steady state size; ii) "precautionary saving effect": small entrepreneurs are the ones that experienced a period of low revenues in the recent past, and as a consequence increased their precautionary saving and reduced the investment in the risky technology to balance future expected financing constraints. On the contrary large firms are more wealthy and behave more like unconstrained firms. Therefore if financing constraints are important in affecting the drop in variable capital, we expect such drop to be more pronounced for smaller firms. This conjecture is confirmed

Figure 4-9: Contraction in fixed and variable capital during downturns - large entrepreneurs




by figures 4-9 and 4-10. As in chapter 3, each period the small and large firms group are

Figure 4-10: Contraction in fixed and variable capital during downturns - small entrepreneurs




composed by the smallest and largest 30% of businesses in the sample. The size is measured as the amount of total assets (real plus financial). In the same contraction episodes analysed before small entrepreneurs are more volatile (figure 4-10), especially in variable capital. The sharp contraction in variable capital for smaller entrepreneurs is caused by the amplification effect. It corresponds to a situation in which small entrepreneurs face a policy function of the type illustrated in figure 4-5. In the previous section we argue that in such situation, when fixed capital is irreversible, a sequence of negative shocks reduces financial wealth rapidly towards the binding constraint region and variable capital drops rapidly as well. This result can explain the stylised fact that the decline in inventories accounts for a large part of the GDP decline in recessions (see stylised fact n.2 in section 2.5). Moreover it is consistent with Bernanke, Gertler and Gilchrist's (1996) computations, which show that most of the fluctuations in the US manufacturing sector can be accounted for by small firms.

The precautionary saving effect also implies that the procyclicality of variable investment is not confined only to these extreme contraction episodes. Figure 4-11 shows in general a positive correlation between inventory investment and sales in the economy with borrowing constraints, while the correlation is negative in the perfect markets economy. We measure inventory investment as the difference between the end of period $t$ ($l_{t+1}$) and end of period $t-1$ ($l_t$) stock of variable capital. Sales are measured as beginning of period $t$ total output ($(1+\eta)y_t$). The correlation is negative in the perfect market economy because the productivity shock is mean reverting and hence it affects current output more than inventory investment, which is a forward looking variable. The correlation is instead positive in the financially constrained economy because of an accelerator effect: any change in output, even when it is caused by a transitory shock, changes financial wealth and affects also investment. Moreover such positive correlation increases in $\sigma_{\theta f}$, again because of the amplification effect. This result for the constrained economy can explain the stylised fact regarding the contemporaneous correlation between inventory investment and sales (Blinder and Maccini (1991) and Ramey and West (1999)).

Figure 4-11: Commovements between inventory investment and sales - economies with and without financing constraints

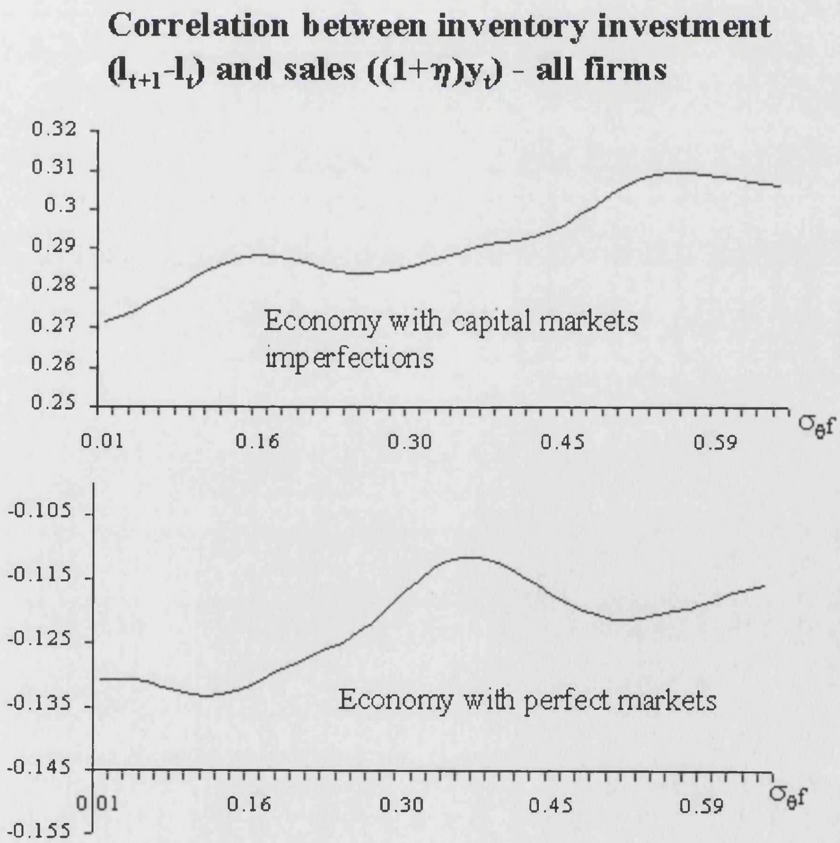
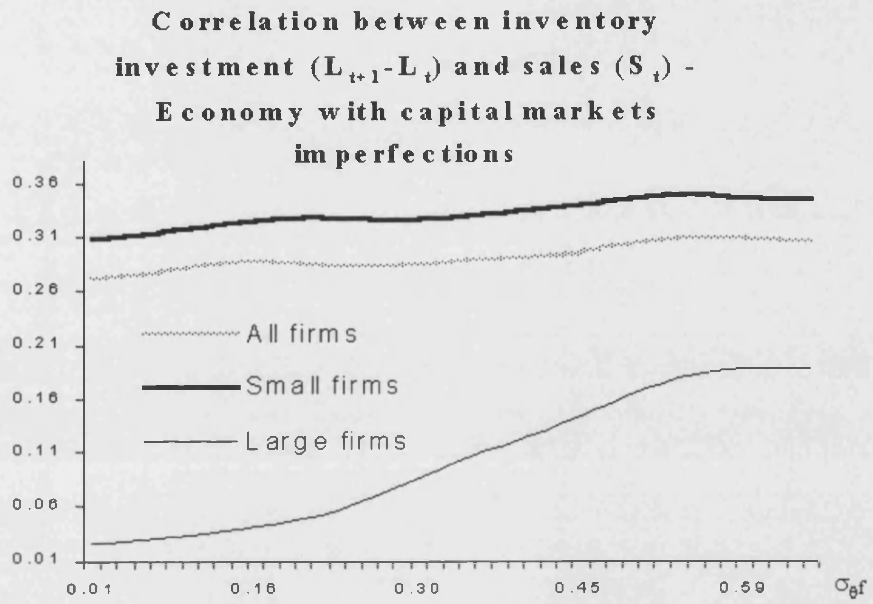


Figure 4-12 shows the same statistic for the whole sample and for small and large firms. It shows that the procyclicality of inventory investment is mainly driven by the small firms subgroup. This result is consistent with the findings of Bernanke, Gertler

Figure 4-12: Commovements between inventory investment and sales - small and large firms




and Gilchrist (1996), regarding the higher procyclicality of small as opposed to large firms inventory investment and output. We illustrate it more clearly in figure 4-13, which shows that the difference in the cumulative growth rates of output and of variable capital between small and large entrepreneurs are both highly procyclical. In the left graphic of figure 4-13 we plot the difference between small and large entrepreneurs cumulative growth rate of output. The other line plotted is an indicator of the stochastic business cycle, the cumulative growth rate of total output. The fact that the two lines move together indicates that small entrepreneurs are more procyclical than large ones. The right graphic of figure 4-13 shows the same evidence regarding the difference in the cumulative growth rates of variable capital. These aggregate dynamics of our simulated economy are very similar to the observed dynamics of inventories and output of small and large US manufacturing firms during the 1978-1992 period, computed by Bernanke, Gertler and Gilchrist (1996) on quarterly US data.

Figure 4-13: Procyclicality of variable capital and output - small versus large firms

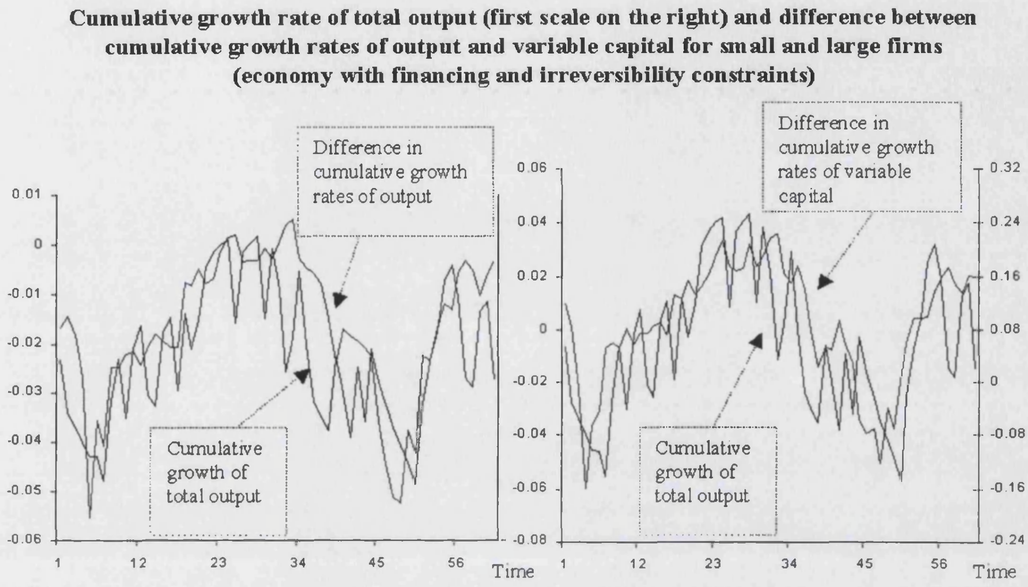
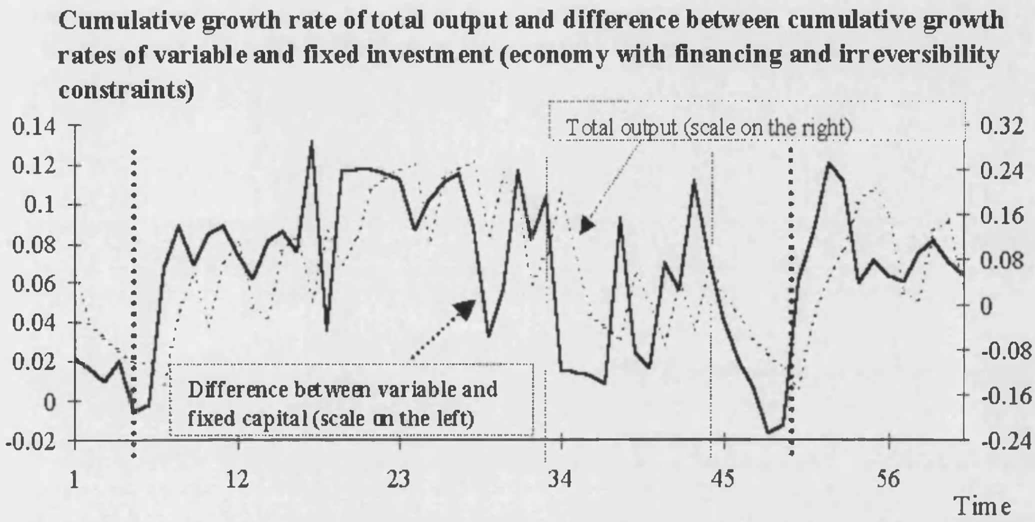



Figure 4-14 addresses the empirical fact that variable capital leads the business cycle, while fixed capital lags it (Stock and Watson, 1998). As in figure 4-13, the "business cycle" is represented by the cumulative growth rate of total output. In figure 4-14 we also add vertical lines that mark changes in the state of the economy. A thick vertical line marks a switch from $S_B$ to $S_G$, which implies the beginning of an expansion phase. A thin vertical line marks a switch from $S_G$ to $S_B$, which implies the beginning of a downturn. Figure 4-14 shows that variable capital is more procyclical than fixed capital, because the difference in the cumulative growth rates is itself procyclical. The fact that variable capital leads the cyclical fluctuations, while fixed capital has a lagged reaction to them, is shown by the fact that the difference between variable and fixed capital increases especially at the beginning of upturns (thick vertical line), and decreases at the beginning of downturns (thin vertical line). This is to say that variable capital reacts faster than fixed capital to turnings points. The reason is a combination of two effects: i) every period some entrepreneurs have a binding irreversibility constraint and do not adjust fixed capital

Figure 4-14: Procyclicality of investment - variable versus fixed capital




to changes in expected profitability; ii) all entrepreneurs adjust variable capital more than they would have done without financing constraints, because at turning points the change in expected profitability also changes the levels of precautionary saving.

## 4.8 Conclusions

This chapter extends the structural model developed in chapter 3 by assuming two factors of production, one of which is irreversible. It solves the model and analyses the interactions between financing and irreversibility constraints. The two constraints may seem unrelated, because the financing constraint limits the ability to increase capital, while the irreversibility constraint limits the ability to reduce it. Indeed the existing investment literature, reviewed in chapter 1, always analyses them separately. Instead in this chapter we show that the two problems interact with each other, and that current and future expected financing problems are amplified by the presence of irreversibility of fixed capital. This has important consequences for investment dynamics, and we illustrate them by simulating an artificial economy with both idiosyncratic and aggregate uncertainty and examining the implications of the model for the cyclical behaviour of fixed investment, working capital investment and output. We show that the interaction between financing and irreversibility constraints explains a number of stylised facts about investment dynamics: i) aggregate inventory investment is very volatile; ii) its decline accounts for a large part of the GDP decline in recessions; iii) it is contemporaneously correlated with sales; iv) it leads the business cycle, while fixed capital lags it; v) output and inventories are more volatile and procyclical for small firms than for large ones.

# Chapter 5

# A new empirical method to test for the presence of financing constraints on firm investment

## 5.1 Introduction

The structural model developed in chapter 4 shows that the interactions between financing constraints and irreversibility of fixed capital help to explain a number of stylised facts about aggregate investment dynamics. One essential condition for these aggregate results is that the optimal investment and saving choices of the entrepreneurs are affected by current and future expected financing problems. In this chapter we test this financing constraints hypothesis at the microeconomic level, using a sample of Italian small and medium-sized manufacturing firms. The contribution of this chapter is not only in providing new evidence about the effects of financing imperfections on firm investment, but also in developing an original test able to discriminate between the financing imperfections and the perfect markets hypothesis.

We formulate the test by considering one implication of the model developed in chapter 4: because variable capital investment is reversible, the "premium" of expected marginal productivity over user cost of variable capital reflects the tightness of current and future expected financing constraints. We call this premium the "excess" expected marginal productivity of variable capital, and we use it as the indicator of the intensity of financing constraints. More specifically the financing constraints hypothesis implies that the value of the indicator is monotonously decreasing in the financial wealth of the entrepreneurs, conditional on fixed capital stock and productivity shock. We test this hypothesis by estimating empirical measures of the productivity shocks, the user cost and the excess expected marginal productivity of variable capital for our sample of Italian manufacturing firms[1].

This test is more efficient than the test based on the cash flow-investment correlation in discriminating between the financing constraints and the perfect markets hypothesis, because of two reasons: i) it is based on a structural model that identifies an indicator that is monotonously increasing in the intensity of financing constraints. Thus it is robust to the Kaplan and Zingales (1997) critique (see chapters 1 and 3) to the cash flow-investment type of test; ii) it maintains its power of discriminating the financing constraints hypothesis from the perfect markets hypothesis in presence of two potential misspecification problems: the presence of convex adjustment costs and the misspecification of the stochastic process for the productivity shock. That is, if our model is misspecified in these ways, then this is likely to increase the chances of "rejecting the financing constraints hypothesis when it is true".

This is a considerable advantage with respect to the cash flow-investment type of test, which is biased towards "accepting the financing constraints hypothesis when it is false" if future expected profitability is not properly estimated. Consider for example the reduced form investment estimation performed by Fazzari, Hubbard and Petersen (1988), where investment is a function of cash flow and of average Tobin's Q. If average Q does not properly takes into account future expected profitability, this is likely to increase the correlation between investment and cash flow. This is because cash flow is very sensitive to current profitability, which in turn is correlated to future expected profitability. Therefore the cash flow investment correlation could simply be caused by the fact that the former absorbs the positive effect of the unobserved future expected profitability on the latter. This empirical problem can explain why many authors (including Fazzari, Hubbard and

[1] A similar procedure has been recently 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*".

Petersen (1988) and Kaplan and Zingales (1997)) estimate positive cash flow coefficients for very large firms. Such firms have direct access to equity and bond markets, and it is very difficult to argue that they can be financially constrained, in the sense of being unable to raise funds to finance profitable investment opportunities.

For the empirical analysis of this chapter we use our subsample of 897 small and medium-sized Italian manufacturing firms, for which we have 11 years of balance sheet data available from 1982 to 1992, plus the qualitative information from the First Mediocredito Centrale Survey on Small and Medium Italian Manufacturing Firms. Out of the initial sample of 897 firms we restrict our analysis to the subset of 561 firms that reported richer information about their fixed assets, namely the details about their stocks of land & building and plant & equipment.

Among the information in the survey we have the statements from the entrepreneurs about the financing problems they faced in the 1989-1991 period (see appendix 2). Entrepreneurs were asked whether they had any of the following problems regarding investment financing: Q1) lack of collateral; Q2) lack of medium-long term financing; Q3) too high cost of banking debt (see figure 2-2). This qualitative information is not used to directly test the financing constraint hypothesis, but rather to perform an independent robustness check which strongly supports the validity of our indicator of the intensity of financing constraints: entrepreneurs with an high value of the indicator are three times more likely to state financing problems in funding investment than entrepreneurs with a low value.

This chapter is organised as follows: section 5.2 defines the financing constraints test; section 5.3 defines the estimation strategy; section 5.4 illustrates the results of the production function estimation; section 5.5 verifies the validity of our indicator of the intensity of financing constraints using the qualitative data from Mediocredito Centrale; section 5.6 tests the financing constraints hypothesis; section 5.7 summarises the conclusions.

## 5.2 A new test of financing constraints on firm investment

In chapter 4 we derive the optimality conditions of the extended model as defined by (4.3), (4.9), (4.10), (4.11) and (4.12). We consider now equation (4.11). $UL$ is the exogenous user cost of variable capital, while both $E_t\left(\Omega_{l,t+1}\right)$ and $\lambda_t$ are endogenous and not directly observable. We rewrite equation (4.11) adding the subscript $i$ for the *i-th* entrepreneur:

$$E_t\left(\Psi_{i,t+1}\right) = (1+\eta)\, E_t\left(MPL_{i,t+1}\right) - UL_{i,t} \tag{5.1}$$

Where $E_t\left(\Psi_{i,t+1}\right)$ is the following:

$$E_t\left(\Psi_{i,t+1}\right) \equiv E_t\left(\Omega_{l,i,t+1}\right) + R^2\lambda_{i,t} \tag{5.2}$$

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

**Proposition 6** *For any $t \geq 0$:*

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

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

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

**Proof: see Appendix 1**

Proposition 6.iii applied to (5.1) implies that variable investment is sensitive to net worth fluctuations because current and future expected financing problems, which determine $E_t\left(\Psi_{i,t+1}\right)$, are monotonously decreasing in net worth, conditional on $\theta_{i,t}$ and $k_{i,t}$. Proposition 6 proves that $E_t\left(\Psi_{i,t+1}\right)$, the excess expected marginal productivity of variable capital, has the important property of being monotonously increasing in the intensity of financing constraints. This means that $E_t\left(\Psi_{i,t+1}\right)$ is a theoretically consistent indicator of financing constraints and is robust to the Kaplan and Zingales critique to the cash flow-investment literature mentioned in chapters 1 and 3. 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. Therefore we can 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 $E_t\left(\Psi_{i,t+1}\right)$ and $w_{i,t}$, conditional on $\theta_{i,t}$*

*and* $k_{i,t}$.

*H1) If financing constraints are irrelevant then we expect no monotonously decreasing 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, which is equivalent to testing proposition 6.iii, 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 $\theta_{i,t}$: i) we have positive productivity shocks which increase $\theta_{i,t}$, but they are transitory and hence do not affect the investment in variable capital $l_{t+1}$ of unconstrained firms; ii) we have positive persistent productivity shocks which increase $\theta_{i,t}$, but they do not immediately increase the investment in variable capital $l_{t+1}$ of financially unconstrained firms because of the presence of adjustment costs. In both cases we would observe high excess marginal productivity of variable capital, because $E_t(\Psi_{i,t+1})$ increases in $\theta_{i,t}$, but such high value would not be related to financing constraints. However such high value of the indicator would also be accompanied by positive cash flows or increases 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.

The combination of qualitative and quantitative information in our dataset allows not only to test the financing constraints hypothesis, but also to directly verify that $E_t(\Psi_{i,t+1})$ is a reliable indicator of the intensity of financing constraints. Therefore, after estimating the empirical counterpart of $E_t(\Psi_{i,t+1})$, called ${}_t\widehat{\Psi}_{i,t+1}$, we proceed in two steps: (i) we check the validity of ${}_t\widehat{\Psi}_{i,t+1}$ as a financing constraints indicator using the direct qualitative information about financing problems. (ii) We test *H0*: we discretise the steady state 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.

We believe that this strategy is feasible because the discretisation of $\theta_{i,t}$ and $k_{i,t}$ can be relatively simple. In fact $w_{i,t}$ is the only variable that directly affects $E_t(\Psi_{i,t+1})$ by determining the probability of present and future financing constraints. If $w_{i,t}$ is low, then $E_t(\Psi_{i,t+1})$ is expected to be greater than zero, regardless of $\theta_{i,t}$ being low or high, unless the persistency of $\theta_{i,t}$ is very high. Since the estimated persistency of the idiosyncratic productivity shock is quite low, 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 it is very high, and $E_t(\mu_{i,t+1}) > 0$, and/or $\mu_{i,t} > 0$. In this case $k_{i,t}$ amplifies the sensitivity 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 $E_t(\mu_{i,t+1})$ is very small. Hence we can eliminate the distortion effect caused by $E_t(\mu_{i,t+1}) > 0$ and/or $\mu_{i,t} > 0$ by focusing on firm year observations with relatively smaller fixed capital/variable capital ratio[2].

## 5.3 Estimation strategy

We define $(1+\eta)\, y_{i,t+1} \equiv Y_{i,t+1}^T$ as total revenues of firm $i$ at time $t+1$. This implies that $\left(\frac{\eta}{1+\eta}\right) Y_{i,t+1}^T$ is the unobservable share of revenues that is consumed[3]. Hence we can also define $(1+\eta)\, E_t(MPL_{i,t+1}) = E_t\left(MPL_{i,t+1}^T\right)$ as "total" expected marginal productivity of variable capital. Therefore, given (5.1), in order to test proposition 6.iii we need empirical estimates of the expected marginal productivity of variable capital $E_t\left(MPL_{i,t+1}^T\right)$ and of the user cost of variable capital $UL_{i,t}$. The production function considered in this section is the following:

$$Y_{i,t}^T = cA_i\theta_{i,t}K_{i,t}^{\alpha}L_{i,t}^{\beta}N_{i,t}^{\gamma} \tag{5.3}$$

with $\alpha > 0$; $\beta > 0$; $\gamma > 0$ and $\alpha +\ \beta + \gamma < 1$. With respect to the theoretical section, we maintain the Cobb-Douglas production function[4], but we include labour $(N_{i,t})$. The

[2] A more efficient method would have been to use the observed fixed capital/variable capital ratios to estimate, for each firm year observation, the probability to face a binding irreversibility constraint, and then to use the estimated probability to correct for the above distortion. This procedure is currently work in progress.

[3] We tried an alternative approach to estimate directly the share $\left(1-\frac{1}{\eta}\right)$ of output that is consumed using panel data information about the number of entrepreneurs working in the firm together with the total labour cost. Results from this alternative approach do not differ substantially from the ones presented here.

[4] Given that $l^E$, which in the model represents the fixed supply of inputs from the entrepreneur, is assumed to be very small, we do not attempt to quantify it, and assume that it is included in $L_{i,t}$.

inclusion of an additional factor of production does not modify the theoretical results showed in the previous section[5]. $c$ is the constant common to the whole sample. $A_i$ includes all assets that are fixed in the time period used for the estimation[6]. The unobservable productivity shock $\theta_{i,t}$ is the product of three components:

$$\theta_{i,t} = \varepsilon_t \chi_{s,t} \theta^f_{i,t} \tag{5.4}$$

$\varepsilon_t$ is an exogenous market wide shock, $\chi_{s,t}$ is an exogenous sector specific shock, and $\theta^f_{i,t}$ is a firm specific idiosyncratic shock. The subscript $s$ refers to the *s-th* industrial sector. Following the specification of the theoretical model, $\ln \theta^f_{i,t}$ is a first order stochastic process: $\ln \theta^f_{i,t+1} = \rho \ln \theta^f_{i,t} + v_{i,t}$ 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, $\alpha$, $\beta$ and $\gamma$ can be estimated as the factors shares of output. 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. In fact in the context of our model the regularity conditions mentioned above would be met if $\eta = 0$ and if there are no financing constraints ($w_{0i} = w^{MAX}_i \ \forall i$). In this case the optimal solution is $E_t\left(MPK_{i,t+1}\right) = UK_{i,t}$ and $E_t\left(MPL_{i,t+1}\right) = UL_{i,t}$. Since $E_t\left(MPK_{i,t+1}\right) \equiv \alpha E_t\left(\frac{y_{i,t+1}}{k_{i,t+1}}\right)$ and $E_t\left(MPL_{i,t+1}\right) \equiv \beta E_t\left(\frac{y_{i,t+1}}{l_{i,t+1}}\right)$ , and since the rational expectation hypothesis implies that $\left(\sum\limits_{i=1}^{N} \frac{y_{i,t+1}}{l_{i,t+1}}\right)/N$ is a consistent estimator for $E_t\left(\frac{y_{i,t+1}}{l_{i,t+1}}\right)$ when $N$ is large, it follows that: $\widehat{\beta} = \frac{UL_t l_{t+1}}{y_{t+1}}$, and $\widehat{\alpha} = \frac{UK_t k_{t+1}}{y_{t+1}}$. Hence it is possible to directly estimate $\widehat{\alpha}$ and $\widehat{\beta}$, as average factor's shares of output. This is not possible if financing constraints affect $E$'s optimal choices. For example in our model, with $\eta > 0$ and $w_t < w^{MAX}$, $\widehat{\beta}$ would not be a consistent estimator of $\beta$ :

[5] 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 during the sample period Italian firms were unable to freely reduce employment, and therefore labour is closer to an irreversible than to a reversible factor of production.

[6] 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 $i-th$ entrepreneur. This formulation is consistent with the assumption that the "know-how" of the entrepreneur is essential for her business (See footnote n.16 in chapter 3).

$$P\lim\left(\widehat{\beta}\right)=\beta\frac{\left[1+\eta+\gamma E_t\left(\phi_{t+1}\right)\right]UL}{\left[1+\gamma E_t\left(\phi_{t+1}\right)\right]UL+R^2\lambda_t-cov(MPL_{t+1},\phi_{t+1})}\neq\beta \tag{5.5}$$

This issue is important, because we estimate $\alpha$, $\beta$ and $\gamma$ separately for firms in different industrial sectors, and hence not only we would have biased estimates, but also the bias would differ among sectors if the intensity of financing and irreversibility constraints differs among sectors as well. Hence we choose to directly estimate these parameters from the production function, using an instrumental variable estimation technique.

## 5.4 Estimation results

We estimate equation (5.3) using the following data: $p_t^y Y_{i,t}^T$ is total revenues in monetary terms. $p_t^k K_{i,t}$ is the replacement value of plant, equipment and other intangible fixed assets. $p_t^l L_{i,t}$ is the nominal value of working capital. $p_t^n N_{i,t}$ is labour cost in monetary terms. Detailed information about these variables is provided in appendix 3. Given that land and building are not included elsewhere in the production function, $A_i$ also proxies for the size of these assets[7]. In order to transform the variables in real terms, we divide each variable at time $t$ by the ratio $\frac{p_1^z}{p_t^z}$, and we redefine $p_1^y Y_{i,t}^T = y_{i,t}$ and $p_1^z Z_{i,t} = z_{i,t}$, with $z \in \{k,l,n\}$ .Variables $y,k,l$ and $n$ are therefore in real terms (valued at constant 1982 prices). Figure 5-1 reports summary statistics of $y_{i,t}$, $k_{i,t}$, $l_{i,t}$ and $n_{i,t}$. By taking logs,

Figure 5-1: 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

[7]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 $K_{i,t}$.

we have the following linearised version of equation (5.3):

$$\ln y_{i,t} = c + \ln A_i + \ln \varepsilon_t + \ln \chi_{s,t} + \alpha \ln k_{i,t} + \beta \ln l_{i,t} + \gamma \ln n_{i,t} + \ln \theta^f_{i,t} \tag{5.6}$$

In order to allow for heterogeneity in the technology employed by firms in different sectors, equation (5.6) is separately estimated for seven groups of firms. Each group is composed of firms with as homogeneous as possible production activities. Figure 5-2 shows their composition. 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

Figure 5-2: 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

firm idiosyncratic shock $\ln \theta^f_{i,t}$ can neither be estimated as a fixed effect, nor eliminated through a transformation of the data. In the theoretical model we assume that $\ln \theta^f_{i,t}$ 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 $cov(\ln \theta^f_{i,t}, \ln z_{i,t}) = 0$ for $z \in \{k, l, n\}$. Unfortunately this is not necessarily true in reality. Even assuming that $\rho = 0$, we can still expect $\ln \theta^f_{i,t}$ 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. Therefore in order to correct this problem we need to 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 $\ln \theta^f_{i,t}$. 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 $\ln \theta^f_{i,t}$ between firms are captured by the fixed effect $A_i$. Therefore the residual persistency of $\ln \theta^f_{i,t}$ 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 (5.6) 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[8].

In figure 5-3 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 $\theta^f_{i,t}$, but the persistency of $\ln \theta^f_{i,t}$ is not very high, then we expect only the 0 and -1 instruments to be rejected. Figure 5-3

Figure 5-3: 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)<$\chi^2$(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)<$\chi^2$(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

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.

The discussion above and figure 5-3 justify the choice of the Generalised Method of Moments (GMM) as the estimation method[9]. We first eliminate the firm specific effect $A_i$. The within-firm transformation would be the obvious choice to do it, but unfortunately it

[8] 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.

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

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[10]. 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 (5.7), where the symbol * denotes the transformed variables, imposing the equality of parameters across equations:

$$\begin{aligned}
\ln y^*_{i,92} &= c_{92} + d_{ts} + \alpha \ln k^*_{i,92} + \beta \ln l^*_{i,92} + \gamma \ln n^*_{i,92} + \ln \theta_{i,92} \\
\ln y^*_{i,91} &= c_{91} + d_{ts} + \alpha \ln k^*_{i,91} + \beta \ln l^*_{i,91} + \gamma \ln n^*_{i,91} + \ln \theta_{i,91} \\
&\;\;. \\
\ln y^*_{i,85} &= c_{85} + d_{ts} + \alpha \ln k^*_{i,85} + \beta \ln l_{i,85} + \gamma \ln n_{i,85} + \ln \theta_{i,85}
\end{aligned} \tag{5.7}$$

$d$ 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 5-4 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[11]. The first column in figure 5-4 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 for all groups except group 7. Therefore 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 output elasticity of labour $\widehat{\gamma}$. These high estimates of $\beta$ 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 0.04 and 0.11. 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

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

[11] 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.

Figure 5-4: 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 |
|---|---|---|---|---|---|---|---|---|
| $\alpha$* | 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) |
| $\beta$* | 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) |
| $\gamma$* | 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*** | 37 | 37 | 36** | 36** | 36** |
| p-value | 0.00 | 0.38 | 0.53 | 0.34 | 0.35 | 0.37 | 0.14 | 0.58 |
| Chi square**** | | 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 |

* Standard deviation in parenthesis; **One coefficient relative to a two-digit sector dummy variable is estimated here; ***Only 6 years used for the estimation of this group; **** Wald test of the following restriction: $\alpha+\beta+\gamma=1$**

(see figure 5-1), and the difference in the user costs of fixed and variable capital generated by the difference in depreciation factors (see appendix 3). Overidentifying restrictions are rejected for the estimation of the whole sample, but not for single groups estimations. This confirms the validity of our strategy to separately estimate technology parameters for each group.

Before proceeding to estimate $E_t\left(MPL_{t+1}^T\right)$, it is important to emphasise some problems which could have affected the estimates obtained above: i) the choice of the functional form for the production function. We chose the Cobb-Douglas production function in (5.3) in order to maintain the same functional form used in the theoretical chapters 3 and 4. This allows us to use the estimated $\widehat{\theta}_{i,t}, \widehat{\alpha}$ and $\widehat{\beta}$ to calibrate the parameters for the simulations presented in those chapters. However we do not test this functional form against other possible ones, which may fit better the data. ii) 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 series as instruments for the level equations.

On the one hand our estimates of the production function coefficients (5.7) 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. On the other hand we believe that further work is needed to improve our estimations, following the results of Blundell and Bond (2000), especially to check whether the bias caused by weak instruments is responsible for the rejection of the constant return to scale assumption in 6 groups out of 7. iii) Since we assume $\theta_{i,t}$ to be an AR(1) process, we could also substitute $\theta_{i,t} = \rho\theta_{i,t-1} + v_{i,t}$ in (5.6) and transform the model to include $\ln y_{i,t-1}$, $\ln k_{i,t-1}$, $\ln l_{i,t-1}$ and $\ln n_{i,t-1}$ as regressors. This estimation and the testing of the AR(1) process against some alternative functional forms is currently work in progress. iv) Following the theoretical model in chapter 4, we define the productivity shock in this empirical section as $\varepsilon_t$ (equation (5.4)). Since in the simulations in chapter 4 all entrepreneurs are ex ante identical, their output is affected in the same way by such aggregate shock. If this is true also for the firms in our sample, then the time dummy should perfectly identify the effect of such aggregate shock on the level of output. In reality firms are different, also in the way a common aggregate shock affects their output. Therefore the time dummies only partially control for the effect for such shocks. This problem is however mitigated by the fact that we estimate the production function separately for firms in different industrial sectors.

In order to obtain an empirical estimate of $E_t\left(MPL^T_{i,t+1}\right)$, we proceed in the following way: i) we 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 $\widehat{TFT}_{i,t} \equiv \ln c + \ln A_i + \ln\varepsilon_t + \ln\chi_{s,t} + \ln\theta^f_{i,t}$; ii) 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 $\widehat{\rho}$[12] separately for the seven groups of firms, as shown in figure 5-5. The $\widehat{\rho}$ estimates

[12] 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 $t-5$

Figure 5-5: 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 |
|---|---|---|---|---|---|---|---|
| $\rho$^^ | 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 |
| $\rho$~ | 0.22 | 0.36 | 0.37 | 0.30 | 0.25 | 0.35 | 0.32 |

*Alternative estimator of $\rho$ based on the estimated transition probabilities of $\theta_{i,t}$ discretised in a two states transition matrix.

are positive and significant, but relatively 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 (5.3) we define the expected marginal productivity of variable capital as the following:

$$E_t\left(MPL^T_{i,t+1}\right) = E_t\left(\frac{\partial Y^T_{i,t+1}}{\partial L_{i,t+1}}\right) = \beta E_t\left(cA_i\theta^f_{i,t+1}\varepsilon_{t+1}\chi_{s,t+1}K^{\alpha}_{i,t+1}L^{\beta-1}_{i,t+1}N^{\gamma}_{i,t+1}\right) \quad (5.8)$$

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

$$E_t\left(MPL^T_{i,t+1}\right)^W = \beta E_t\left(\theta^f_{i,t+1}\right) E_t\left(c\varepsilon_{t+1}\chi_{s,t+1}k^{\alpha}_{i,t+1}l^{\beta-1}_{i,t+1}n^{\gamma}_{i,t+1}\right) \quad (5.9)$$

Equation 5.9 also assumes[13] that $cov\left(\theta^f_{i,t+1}, k^{\alpha}_{i,t+1}l^{\beta-1}_{i,t+1}n^{\gamma}_{i,t+1}\right) = 0$. Hence our estimator of $E_t\left(MPL^T_{t+1}\right)^W$, called ${}_t\widehat{MPL}^W_{t+1}$, is the following:

$$\widehat{MPL}^w_{i,t+1} = \widehat{c\beta}\left({}_t\widehat{\theta}_{i,t+1}k^{\widehat{\alpha}}_{i,t+1}l^{\widehat{\beta}-1}_{i,t+1}n^{\widehat{\gamma}}_{i,t+1}\right) \quad (5.10)$$

where ${}_t\widehat{\theta}_{i,t+1} \equiv \exp\left(\widehat{\rho}\ln\widehat{\theta}^f_{i,t} + \ln\widehat{\varepsilon}_t + \ln\widehat{\chi}_{s,t}\right)$.

lags as instruments.

[13] This assumption is not likely to cause a relevant bias in the estimates of $E_t\left(MPL^T_{i,t+1}\right)^W$, because the estimated correlation between the empirical counterparts $\widehat{\rho}\widehat{\theta}^f_{i,t}$ and $k^{\widehat{\alpha}}_{i,t+1}l^{\widehat{\beta}-1}_{i,t+1}n^{\widehat{\gamma}}_{i,t+1}$ 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.

In order to derive (5.10) from (5.9) 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, $\widehat{\rho}\ln\widehat{\theta}^f_{i,t}$ is the estimate 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[14]:

$$UL_{i,t} = \frac{p^l_t}{p^l_1}(1+r_t) - \frac{p^l_{t+1}}{p^l_1}(1-\delta_l) \tag{5.11}$$

we apply to it the same normalization[15] 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 $\delta_l = 0$ by construction, because we include in $l_{i,t+1}$ only variable capital consumed during time $t+1$ (see appendix 3 for details). Hence the user cost simplifies to:

$$ul_{i,t} = 1 + r_t \tag{5.12}$$

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 of period $t-1$ and the fourth quarter of period $t$). In order to use (5.12) as the equation that defines the user cost of variable capital, we adopt a series of simplifying assumptions[16] which include the fact that we maintain assumptions *p1* and *p2* regarding the linearity of preferences. This implies that the user cost of capital is independent on the risk faced by the firms' projects. in reality agents are risk averse, and in equilibrium the user cost of

[14] 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 $\varepsilon_t$. 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 $\varepsilon_t$ does not affect the results in any relevant way.

[15] Another simplification of this formulation is that we do not consider relative prices differences. In order to consider relative prices changes between variable capital and output we should multiply both terms on the right hand side of (5.11) by $p^y_1/p^y_t$. This distortion is likely to have no effects on the results, for two reasons: i) the two indexes we use to normalise output and variable capital (see appendix 3) are almost identical in all the years of the sample; ii) since the distortion is constant across firms, it is captured by yearly dummies $\varepsilon_t$.

[16] See footnotes 14 and 15.

capital should include a component above the riskless interest rate representing the price of risk. While this implementation is currently work in progress, it is important to note that the main aim of this section is to test wether the marginal productivity of variable capital is abnormally high for financially constrained firms. Recall that, from (5.2), the expected marginal productivity of variable capital is higher for financially constrained firms because of the two components $E_t(\Omega_{l,i,t+1})$ and $\lambda_{i,t}$. $E_t(\Omega_{l,i,t+1})$ is the cost of future expected financing constraints. Simulations presented in the previous chapter (see figure 4-6) show that this component can be up to 50%. $\lambda_{i,t}$ is the cost of currently binding financing constraints, and its magnitude goes from the maximum value of $E_t(\Omega_{l,i,t+1})$ upwards. Figure (5-7) below and the estimations performed in the next subsection confirm this orders of magnitude, showing that for firms with low financial wealth the premium in the expected marginal productivity of variable capital is around 60-80%. Therefore we believe that the bias in $ul_{i,t}$, even though it could seriously influence our results, it is probably not enough to explain the magnitude of them.

Given (5.12), the estimator for $E_t(\Psi_{i,t+1})^W$ is the following[17]:

$$_t\widehat{\Psi}^w_{i,t+1} = \left( \widehat{_tMPL}^w_{i,t+1} - ul_{i,t+1} \right) \tag{5.13}$$

Figures 5-6 and 5-7 show basic statistics[18] and the kernel estimation of the distribution function of $_t\widehat{\Psi}^w_{i,t+1}$ . Figure 5-7 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.

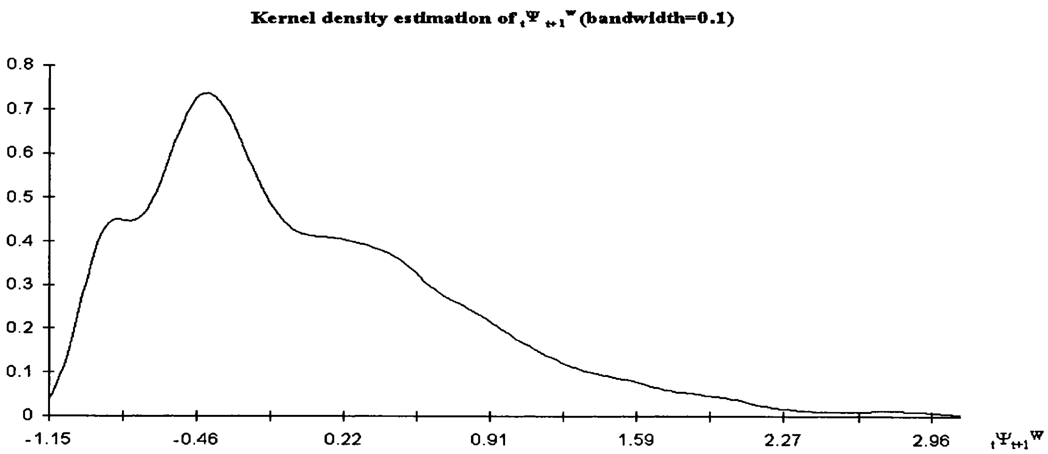
[17]The estimator of $E_t(\Psi_{i,t+1})^W$ does not include fixed effects $A_i$ as well. Since by construction $\sum_{i=1}^{N} \ln \widehat{A}_i = 0$, and since the exponential is a convex function, it follows that $\sum_{i=1}^{N} \widehat{A}_i > \exp\left(\sum_{i=1}^{N} \ln \widehat{A}_i\right) = 1$,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

[18]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 $\widehat{_t\Psi^w_{i,t+1}}$.

Figure 5-6: Estimated excess expected marginal productivity of variable capital ( ${}_t\widehat{\Psi}^{w}_{i,t+1}$)

| Summary statistics for ${}_t\Psi^{w}{}_{i,t+1}$ ,t | | | |
|---|---|---|---|
| Mean | St. dev. | Min. | Max. |
| 0.026 | 0.727 | -1.05 | 3.013 |

Figure 5-7: Density estimation of the excess expected marginal productivity of variable capital



## 5.5 Empirical evidence of financing constraints on investment

Before testing the financing constraints hypothesis, we can verify the validity of the theoretical model and of our estimator ${}_t\widehat{\Psi}^w_{i,t+1}$ using the direct information about financing problems available in the Mediocredito Centrale survey. We consider the entrepreneurs that stated problems in accessing external finance in the 1989-91 period (questions Q1, Q2 and Q3). Such problems are directly related to the ${}_t\widehat{\Psi}^w_{i,t+1}$ variable. The bigger ${}_t\widehat{\Psi}^w_{i,t+1}$, the higher the shadow value of additional funding for the *i-th* entrepreneur and the higher the probability that she answers positively to one of the questions regarding financing constraints. Among the 561 firms considered, 21.6% of their entrepreneurs 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 $i-th$ entrepreneur does not state any financing problem, 1 if she 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 ${}_t\widehat{\Psi}^w_{i,t+1}$ as a valid indicator of the intensity of financing constraints by regressing $ration_i^j$ on $\overline{\Psi}_i^w$, which is the average value of ${}_t\widehat{\Psi}^w_{i,t+1}$ in the period covered by the Mediocredito Centrale survey:

$$\mathbf{ration}_i^j = \alpha_0 + \alpha_1 \overline{\Psi}_i^w + \alpha_2 \dim_i \tag{5.14}$$

$\overline{\Psi}_i^w = \sum\limits_{t=1989}^{1992} \left({}_t\widehat{\Psi}^w_{i,t+1}\right)$. 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 ith firm in number of employees, included to control for size effects. Figure 5-8 shows estimation results. The first column is relative to the whole sample and to $ration_i^4$ as dependent variable. The coefficient relative to $\overline{\Psi}_i^w$ 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 large firms is suggested by figure 5-9, which shows the tri-dimensional smoothing of $ration^4$ with respect to both $\overline{\Psi}_i^w$ and $\dim_i$. Figure 5-9 shows that the positive correlation between the probability of

Figure 5-8: Relation between stated financing problems and the financing constraints indicator

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

| | 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) |
| $\alpha_0$ | -0.64*** (0.09) | -0.69 (0.42) | -0.78*** (0.16) | -2.05*** (0.31) | -1.48*** (0.20) | -1.04*** (0.17) |
| $\alpha_1$ | 0.24** (0.11) | -0.16 (0.36) | 0.30*** (0.12) | 0.36* (0.20) | 0.28** (0.14) | 0.29** (0.13) |
| $\alpha_2$ | -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; $\Psi^w_i$ = average value of the premium in the expected productivity of variable capital; dim$_i$ = dimension in number of employees;

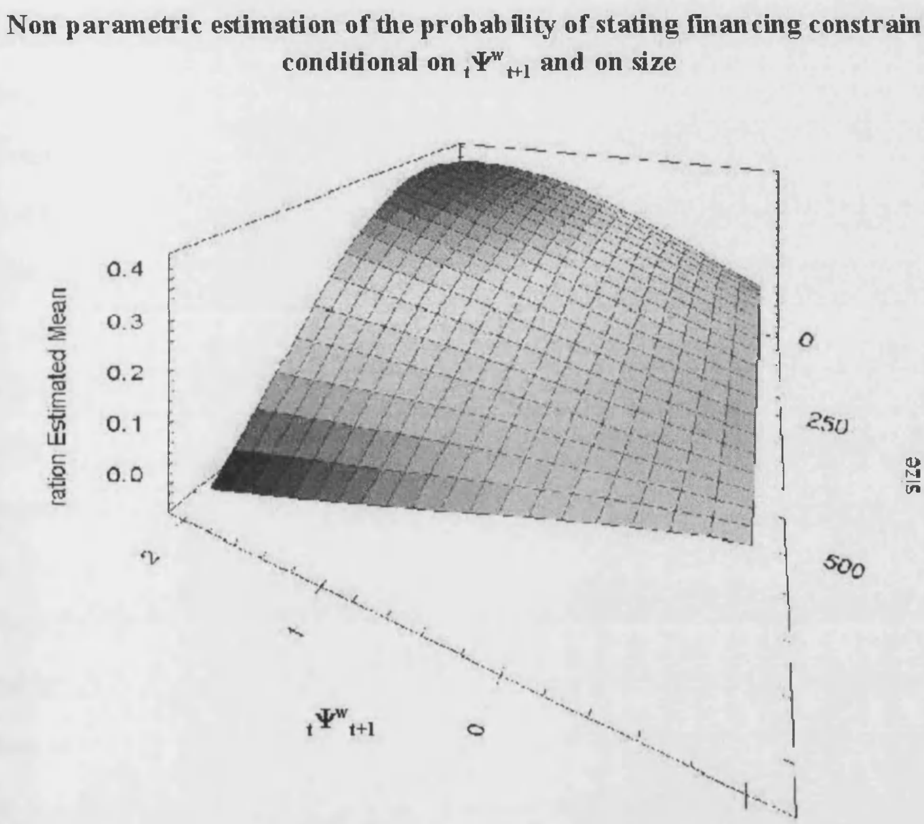
stating financing problems and $\overline{\Psi}^w_i$ is strong for all the firms except the larger ones, so that on average "small-medium" firms with a high value of $\overline{\Psi}^w_i$ are three times more likely to declare financing constraints than entrepreneurs with a low value of it. Such relation 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 $\theta^f_{i,t}$ 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 such steady state according to the realisations[19] of the idiosyncratic shock $\theta^f_{i,t}$. Therefore the result illustrated in figures 5-8 and 5-9 is consistent with the assumption that the higher the average size of firms, the less likely they are to face the informational or contractual problem which causes the financing constraint (3.6). This assumption is realistic as 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 $ration^j_i$ and $\overline{\Psi}^w_i$ for small and medium firms below 300 employees is confirmed by the probit regression results in figure 5-8. The last four columns show that the $\overline{\Psi}^w_i$ coefficient is positive and strongly significant, especially for the

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

Figure 5-9: Probability of stating financing problems as a function of size and financing constraints indicator



specification ($j = 4$) that pools together the three different questions. This result shows that ${}_t\widehat{\Psi}^w_{i,t+1}$ is a valid indicator of the intensity of financing constraints, supporting the validity of our theoretical model and our empirical approach, and rejecting the view of efficient financial markets. This result is robust because of the following considerations: i) the qualitative and quantitative information come from different sources (see appendix 2). This reduces the probability that those entrepreneurs that declare financing constraints also manipulate their balance sheets data to show that their investment is inefficiently low; ii) we condition for firms size, thus ruling out the possibility that $\overline{\Psi}^w_i$ is on average higher for small firms, which are also more likely to state financing constraints; iii) the result is not driven by sectorial differences: table 2-14 shows that financing constraints are equally distributed in the different industrial sectors.

Given that ${}_t\widehat{\Psi}^w_{i,t+1}$ 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. This result is also not contradicted by the observation that firms that declare financing problems have lower net income (see figure 2-15). In fact ${}_t\widehat{\Psi}^w_{i,t+1}$ is the estimated marginal productivity of variable capital, not the total marginal productivity of the projects in the firm. That is, proposition 6 implies that financially constrained firms have high marginal productivity of variable capital, being unable to invest efficiently. At the same time they may have overinvested in the past in fixed capital, and may have an irreversibility problem. The discussion in chapter 4, and in particular in section 4.5, emphasises the fact that financing constraints are amplified by the presence of a binding irreversibility constraint, which implies that investment in variable capital is inefficiently low while the one in fixed capital is inefficiently high. If the firm borrowed in the past to finance the current level of fixed capital, it can have a situation with high marginal productivity of variable capital, due to the binding borrowing constraint, and low total profits, due to the low total productivity because of the unbalanced use of factors, and the high cost to pay the interest on the stock of existing debt.

This interpretation is supported by our analysis in chapter 2. When we select out firms more likely to have a binding irreversibility constraint (firms with negative sales growth) and firms that are more likely to have high interest payments (firms that declare the problem of the high cost of debt), then the average net income of the group of firms with financing problems increases considerably (see figure 2-16).

## 5.6 A formal test of the financing constraints hypothesis

After verifying that ${}_t\widehat{\Psi}^w_{i,t+1}$ is indeed positively related to directly revealed financing constraints, we test the financing constraints hypothesis by performing a nonparametric estimation of ${}_t\widehat{\Psi}^w_{i,t+1}$ with respect to $w^w_{i,t}$, conditional on $E_t\left(\theta^f_{i,t+1}\right)$ and $k_{i,t}$ :

$$ {}_t\widehat{\Psi}^w_{i,t+1} = g\left(w^w_{i,t} \mid \widehat{\theta}^f_{i,t+1}, k^w_{i,t}\right) \tag{5.15} $$

$g(.)$ is a nonlinear function estimated using a nonparametric estimation method[20]. The variables considered are the following:

Financial wealth: we consider two alternative variables:

i) $w^1_{i,t}$ = net financial wealth at the end of year $t-1$ (liquidity plus short term financial assets[21] 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 end of year $t-1$.

$w^1_{i,t}$ 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 not always the case, as time $t$ cash flow is partly generated by time $t-1$ and partly by time $t$ investment. This implies that our nonlinear estimation of (5.15) could be biased. Therefore $w^2_{i,t}$ is a less precise but more consistent estimator, because it excludes by construction the time $t$ cash flow generated by time $t$ investment. 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 effect $A_i$, we apply the same procedure to $w^1_{i,t}$ and $w^2_{i,t}$ , and

[20] The estimation of (5.15) 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 available at www.quanttools.com.

[21] 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.

consider the values $w_{i,t}^{1w} = w_{i,t}^{1} - \sum_{t=1}^{T} w_{i,t}^{1}$ and $w_{i,t}^{2w} = w_{i,t}^{2} - \sum_{t=1}^{T} w_{i,t}^{2}$ that are the deviations form firm specific means.

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

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

In order to estimate (5.15) we condition by $E_t\left(\theta_{i,t+1}^{f}\right)$ by discretising the state space of its estimator ${}_t\widehat{\theta}_{i,t+1}^{f}$ in 3 equally spaced intervals and by estimating (5.15) for each interval. Moreover we condition by $k_{i,t}$ by excluding from our analysis the observations in the fourth quartile of $k_{i,t}^{w}$[22].

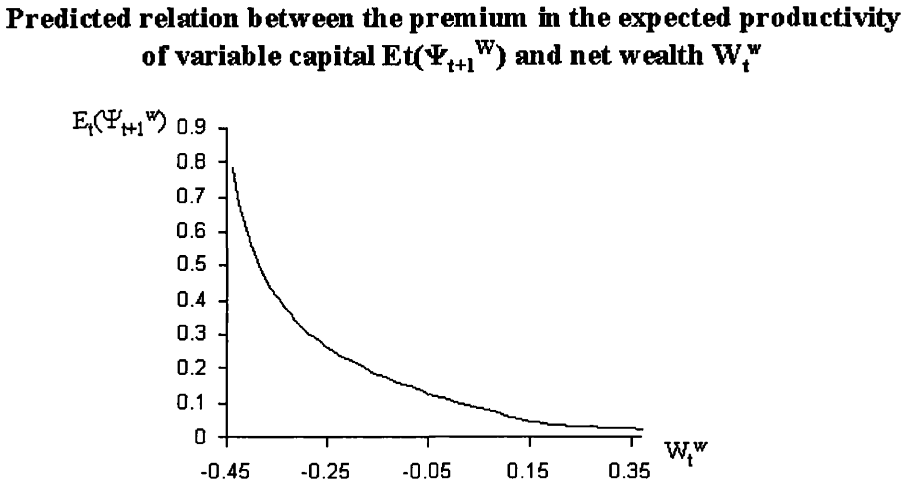
If the financing constraints hypothesis is not rejected by the data, we expect to find a negative slope of the conditional mean of ${}_t\widehat{\Psi}_{i,t+1}^{w}$ with respect to $w_{i,t}^{w}$. Such a slope should be convex: steeper when $w_{i,t}^{w}$ is very low, then gradually flatter as $w_{i,t}^{w}$ increases. This is illustrated in figure 5-10, which shows the predicted relation between $E_t\left(\Psi_{i,t+1}^{w}\right)$ and $w_{i,t}^{w}$ for a given $E_t\left(\theta_{i,t+1}\right)$ and $k_{i,t}$, simulated using the same parameters used in the calibration in the previous section and applying to the simulated data the same within-transformation described above. The highest values of $E_t\left(\Psi_{i,t+1}^{w}\right)$, in correspondence with the lowest values of $w_{i,t}^{w}$, 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, $E_t\left(\Psi_{i,t+1}^{w}\right)$ decreases, at a gradually slower pace.

Figures 5-11, 5-12 and 5-13 show the estimation of (5.15) for smaller firms (less than 300 employees) using both $w_{i,t}^{1w}$ and $w_{i,t}^{2w}$. 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[23] for both $w_{i,t}^{1w}$ and $w_{i,t}^{2w}$ (Figures 5-11 and 5-12). For $w_{i,t}^{1w}$ the model predictions are not confirmed for high values of financial wealth, for which we observe an upward

[22] See the discussion in section 5.2.

[23] The minimum of ${}_t\widehat{\Psi}_{i,t+1}^{w}$ is lower than zero, and this could be due to one of the reasons mentioned before: i) the bias induced by the elimination of $A_i$ (see footnote n.17); ii) the overestimation of the user cost of capital, given that we do not explicitly measure tax differentials (see footnote n.14).

Figure 5-10: Expected relation between the intensity of financing constraints and financial wealth




sloping relation. This could be due to one of the two misspecification problems mentioned in section 5.2. Suppose a firm receives a positive shock between time $t-1$ and time $t$, which increases financial wealth at the end of time $t-1$ and also cash flow at the beginning of time $t$, but does not increase investment during time $t$ (which determines ${}_t\widehat{\Psi}^w_{i,t+1}$), either because the shock is transitory, or because of the presence of convex adjustment costs. For this firm we should observe high ${}_t\widehat{\Psi}^w_{i,t+1}$ but also high net worth. The effect should be stronger on $w^{1w}_{i,t}$ than on $w^{2w}_{i,t}$, which does not include time $t$ cash flow. This can explain why the positive relation between ${}_t\widehat{\Psi}^w_{i,t+1}$ and net worth for the high values of the latter is stronger when using $w^{1w}_{i,t}$ than 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. In this respect this test is more efficient than the test based on the cash flow-investment correlation. This is because, as we mentioned before, an high cash flow-investment correlation is likely to be caused by unobserved productivity shocks, while it is more difficult to argue that the relation showed in figures 5-11 and 5-12 is due to something else than the presence of financing constraints.

One possible alternative explanation for such downward sloping relation between ${}_t\widehat{\Psi}^w_{i,t+1}$

and $w_{i,t}^{1w}$ is the bias in the user cost of capital induced by the absence of the risk premium. Since, as mentioned above, the user cost of capital as defined in (5.12) does not include a term to reflect the risk premium, it is possible that high ${}_t\widehat{\Psi}_{i,t+1}^w$ observations are relative to firms with risky projects, which are on average smaller and hence less wealthy. At the same time such riskier firms may also find more difficult to obtain bank financing and may be more likely to complain about the lack of banking debt or the too high cost of debt. We think that this alternative explanation may partly explain our results, and further work needs to be done to assess the influence of the risk factor both on ${}_t\widehat{\Psi}_{i,t+1}^w$ and on our qualitative information. At the same time, as argued in the previous subsection, we think that the risk factor is not enough to explain the magnitude of our findings. In fact figure (5-10) 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 (5-11) and (5-12) not only confirm the downward sloping relationship between ${}_t\widehat{\Psi}_{i,t+1}^w$ and $w_{i,t}^w$, but also they show a similar magnitude of ${}_t\widehat{\Psi}_{i,t+1}^w$ , 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.

Observations with high productivity shock in figure 5-13 instead exhibit a slightly upward sloping relation between $w_{i,t}^w$ and ${}_t\widehat{\Psi}_{i,t+1}^w$ also in this case more pronounced for $w_{i,t}^{1w}$. This is probably caused by the fact that by considering all the firm-year observations with higher productivity shocks, we select those observations for which the two misspecification problems mentioned in section 5.2 are more severe, and as a consequence we reject the financing constraints hypothesis. This interpretation is supported by the qualitative information in figure 5-14: in the subset of observations of small firms with high productivity shock (third column) fewer entrepreneurs state financing constraints, and the probability of stating them is not correlated to the value of $\overline{\Psi}_i^w$.

The qualitative information in figure 5-8 also suggests that the financing constraints hypothesis should be rejected for larger firms, for which $\overline{\Psi}_i^w$ is not related to financing constraints. Figure 5-15 confirms this. The estimation of (5.15) rejects model's prediction,

Figure 5-11: Estimated relation between the intensity of financing constraints and financial wealth - small firms - low productivity

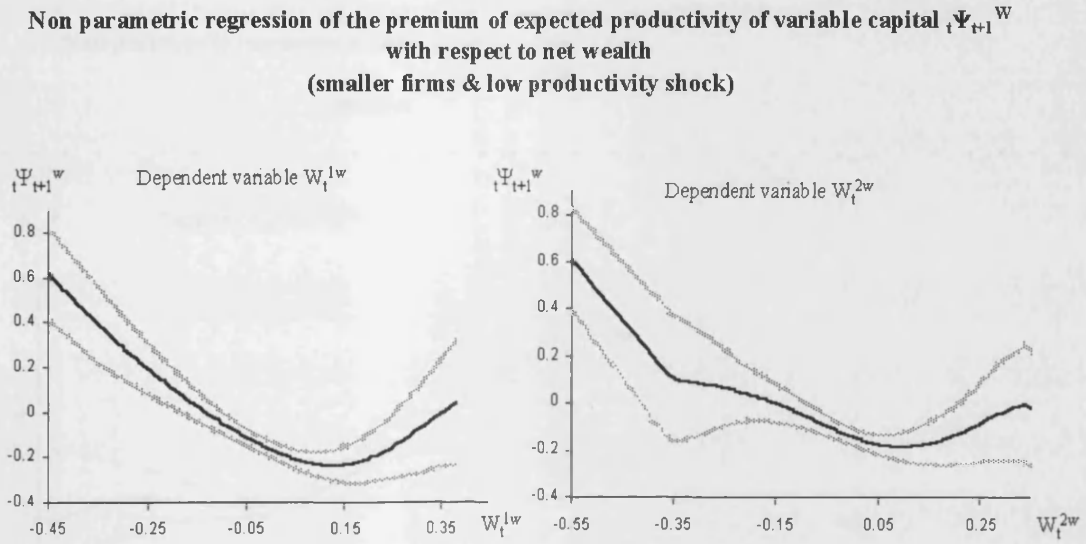



Figure 5-12: Estimated relation between the intensity of financing constraints and financial wealth - small firms - medium productivity

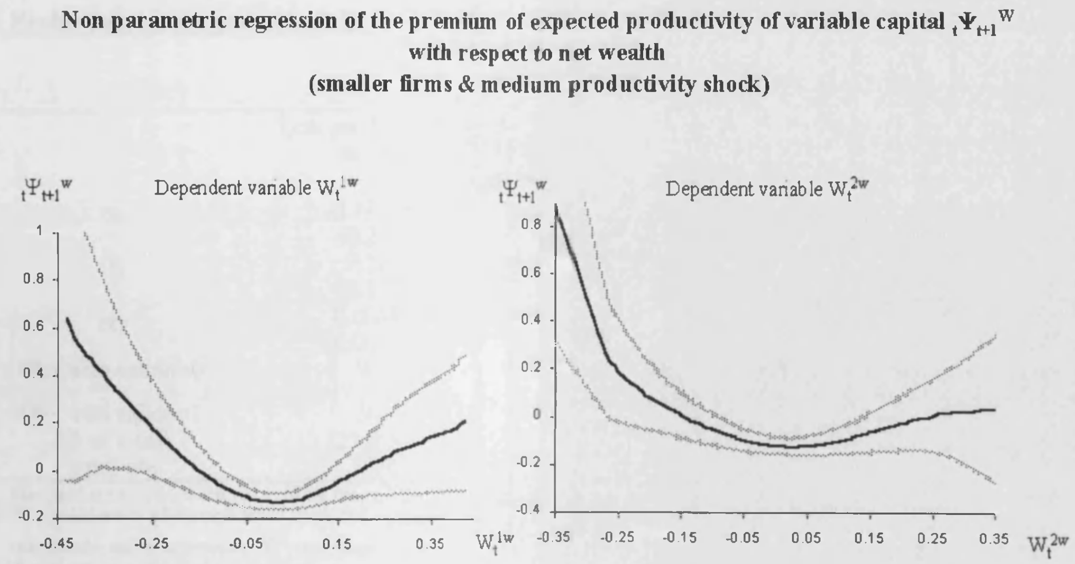


Figure 5-13: Estimated relation between the intensity of financing constraints and financial wealth - small firms - high productivity

**Non parametric regression of the premium of expected productivity of variable capital ${}_t\Psi_{t+1}^{W}$ with respect to net wealth (smaller firms & high productivity shock)**

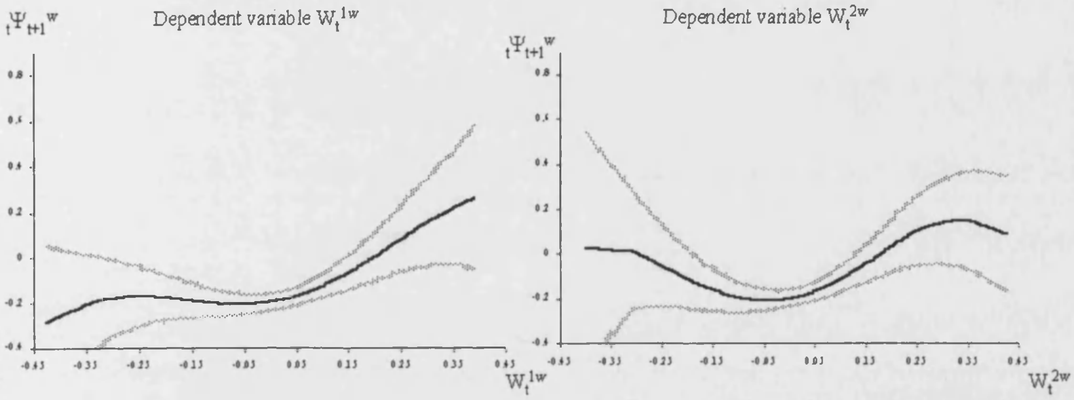



Figure 5-14: Relation between stated financing problems and the financing constraints indicator - small firms

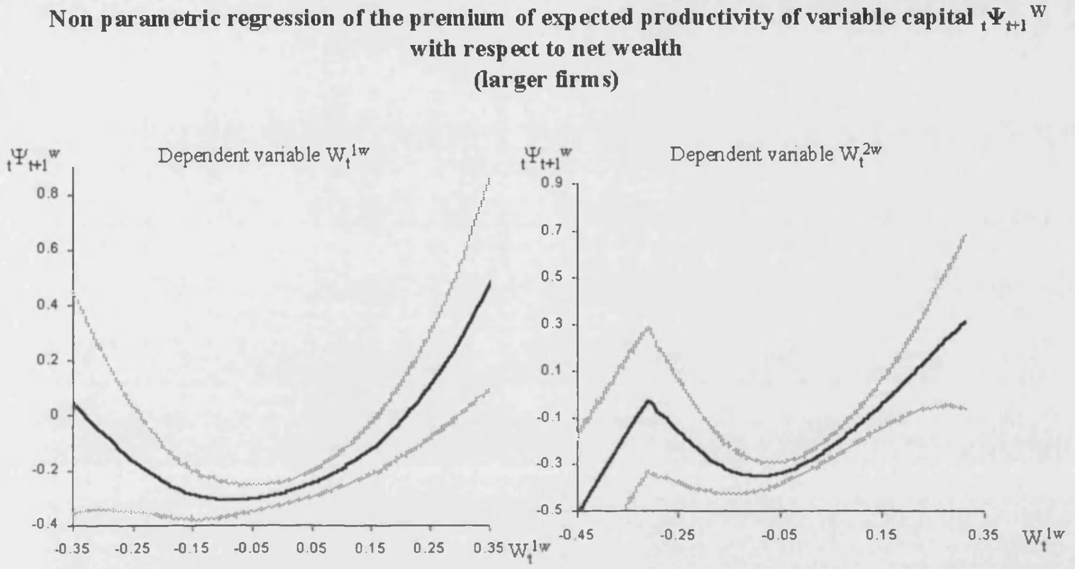
**Probit regression and productivity shock levels:** $ration^j_i = a_0 + a_1\,\Psi^w_i + a_2 dim_i$

| | Smaller firms[2]<br>Dependent variable: Ration[4]<br>(all problems) | | |
|---|---|---|---|
| | Low productivity shock | Medium productivity shock | High productivity shock |
| $\alpha_0$ | -0.77*** | -0.58*** | -1.09*** |
| | (0.27) | (0.24) | (0.32) |
| $\alpha_1$ | 0.28 | 0.43** | 0.09 |
| | (0.19) | (0.21) | (0.27) |
| $\alpha_2$ | 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; $\Psi^w_i$ = average value of the premium in the expected productivity of variable capital; $dim_i$ = dimension in number of employees;

showing instead an upward sloping relationship between ${}_t\widehat{\Psi}^w_{i,t+1}$ and $w^w_{i,t}$. As before this upward sloping relation is steeper for $w^{1w}_{i,t}$ than for $w^{2w}_{i,t}$ and it could again be caused by the two misspecification problems mentioned before.

Figure 5-15: Estimated relation between the intensity of financing constraints and financial wealth - large firms




## 5.7 Conclusions

In this chapter we use the structural model developed in chapter 4 to formulate an original test for the presence of financing constraints on firm level investment. After estimating the excess marginal productivity of variable capital, we verify that it is monotonously decreasing in firms financial wealth for all the firms in the sample except the larger ones. This result allows us to reject the perfect market hypothesis in favour of the financing constraints one. In the chapter we explain why this test is more efficient than the tests based on the cash flow-investment correlation. This is because of three main reasons: i) our theory shows that, after conditioning for fixed capital stock and productivity shock, the excess marginal productivity of variable capital is monotonously increasing in the intensity of financing constraints, and therefore it is a theoretically consistent indicator for them; ii) the validity of the indicator is strongly supported by our qualitative information: entrepreneurs with a high value of it are three times more likely to state financing constraints than entrepreneurs with a low value, even after controlling for the size of their companies; iii) our test is efficient in discriminating the financial imperfections hypothesis from the perfect market hypothesis. We argue that for at least two forms of misspecification of the model the test is likely to be biased towards rejecting the financing constraints hypothesis when it is true rather than accepting it when it is false.

# Bibliography


[1] Abel, A.B. (1979), *Investment and the Value of Capital*, Garland, New York.

[2] Abel, A.B. and O.J. Blanchard (1986), "The Present Value of Profits and Cyclical Movements in Investment", *Econometrica*, 54, 249-273.

[3] Abel, A.B. and J. Eberly (1994), "A Unified Model of Investment under Uncertainty", *American Economic Review,* 84, 1369-1384.

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

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

[6] Bagella, M., Becchetti, L. and A. Caggese (2001), "Financial Constraints on Investments: a Three Pillar Approach", *Research in Economics*, 55, 219-254.

[7] Bernanke, B. (1983), "Nonmonetary Effects of the Financial Crisis in Propagation of the Great Depression", *American Economic Review,* 73, 257-76.

[8] Bernanke, B. and M. Gertler (1989), "Agency Costs, Net Worth, and Business Fluctuations", *American Economic Review*; 79, 14-31.

[9] Bernanke, B. and M. Gertler (1990), "Financial Fragility and Economic Performance", *Quarterly Journal of Economics,* 105, 87-114.

[10] Bernanke, B., Gertler, M. and S. Gilchrist (1996), "The Financial Accelerator and the Flight to Quality", *The Review of Economics and Statistics*, 78, 1-15.

[11] Bernanke, B., Gertler, M. and S. Gilchrist (1998) "Credit Market Frictions and Cyclical Fluctuations", *Handbook of Macroeconomics*, North-Holland, Amsterdam.

[12] Bertola, G. and R. Caballero (1994), "Irreversibility and Aggregate Investment", *Review of Economics Studies*, 61, 223-246.

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

[14] Bils, M. and J.A. Kahn (2000), "What Inventory Behaviour Tells Us about Business Cycles" American Economic Review, 90, 458-81.

[15] Blanchard, O.J. (1983), "The Production and Inventory Behaviour of the American Automobile Industry", *Journal of Political Economy*, 91, 365-400.

[16] Blanchard, O.J. and C. Wyplosz (1981), "An Empirical Structural Model of Aggregate Demand", *Journal of Monetary Economics*, 7, 1-28.

[17] Blinder, A.S. (1986), "Can the Production Smoothing Model of Inventory Behaviour Be Saved?", *Quarterly Journal of Economics,* 101, 431-453.

[18] Blinder, A.S. and L. Maccini (1991), "Taking Stock: A Critical Assessment of Recent Research on Inventories", *Journal of Economic Perspectives*, 5, 73-96.

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

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

[21] Bond, S. and C. Meghir (1994), "Dynamic Investment Models and the Firm's Financial Policy", *Review of Economic Studies*, 61, 197-222.

[22] Caballero, R.J. (1991), "Earnings Uncertainty and Aggregate Wealth Accumulation", *American Economic Review*, 81, 859-871.

[23] Caballero, R.J. (1997), "Aggregate Investment", NBER WP #6264

[24] Caballero, R.J. and E. Engel (1999), "Explaining Investment Dynamics in the U.S. Manufacturing: A Generalised (S,s) Approach", *Econometrica*, 1999, 783-826.

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

[26] Caballero, R.J. and J. Leahy (1996), "Fixed Costs: The Demise of Marginal q" NBER WP #5508.

[27] Carpenter, R.E., Fazzari, S.M. and B. Petersen (1994), "Inventory Investment, Internal-Finance Fluctuations, and the Business Cycle" *Brookings Papers on Economic Activity*, 75-122.

[28] Carpenter, R.E., Fazzari, S.M. and B. Petersen (1998), "Financing Constraints and Inventory Investment: A Comparative Study with High-Frequency Panel Data", *Review of Economics and Statistics*, 80, 513-519.

[29] Carrol, C.D. (1994), "How Does Future Income Affect Current Consumption?", *Quarterly Journal of Economics,* 109, 111-147.

[30] Carrol, C.D. and A.A. Samwick (1998), "How Important is Precautionary Saving", Review of Economics and Statistics, 80, 410-419.

[31] Chow, G.C. (1997), *Dynamics Economics*, Oxford University Press, Oxford.

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

[33] Cooley, T., Marimon, R. and V. Quadrini (2000), "Aggregate Consequences of Limited Contract Enforceability", mimeo, New York University.

[34] Devereaux, M. and F. Schiantarelli (1989), "Investment, Financial Factors, and Cash Flow: Evidence from UK Panel Data", in *Asymmetric Information, Corporate Finance and Investment,* Ed. R.Glenn Hubbard. U. of Chicago Press, Chicago.

[35] Dixit, A. and R.S. Pindyck (1994), *Investment Under Uncertainty*, Princeton University Press, Princeton.

[36] Doms, M. and T. Dunne (1998), "Capital Adjustment Patterns in Manufacturing Plants", *Review of Economic Dynamics*, 2, 409-29.

[37] Dixit, A.K. and R.S. Pindyck (1994), *Investment under uncertainty*, Princeton University Press, Princeton.

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

[39] Fazzari S.M., Hubbard G.R. and B.C. Petersen (2000), "Investment-Cash Flow Sensitivities are Useful: A Comment", *Quarterly Journal of Economics,* 115, 695-705.

[40] Fisher, I. (1933), "The Debt-Deflation Theory of Great Depression", *Econometrica*, 1, 337-357.

[41] Gentry, W.M. and R.G. Hubbard (2000), "Entrepreneurship and Household Saving", Mimeo, Columbia University.

[42] Gertler, M. and S. Gilchrist (1994), "Monetary Policy, Business Cycle, and the Behaviour of Small Manufacturing Firms", *Quarterly Journal of Economics,* 109, 309-340.

[43] Gilchrist S., and C.P. Himmelberg (1995), "Evidence on the Role of Cash Flow for Investment", *Journal of Monetary Economics*, 36, 541-572.

[44] Gilchrist, S. and C.P. Himmelberg (1998), "Investment, Fundamentals and Finance", NBER WP #6652.

[45] Gomes, J. (1998), "Financing Investment", mimeo, University of Pennsylvania.

[46] Greenwald, B.C. and J.E. Stiglitz (1993), "Financial Market Imperfections and Business Cycles", *Quarterly Journal of Economics,* 108, 77-114.

[47] Griliches, Z. and J.A. Hausman (1986), "Errors in Variables and Panel Data", *Journal of Econometrics*, 31, 93-118.

[48] Gross, D.B. (1994), "The Investment and Financing Decisions of Liquidity Constrained Firms", mimeo, MIT.

[49] Guiso, L., Jappelli, T. and D. Terilzzese (1992), "Earnings Uncertainty and Precautionary Saving", *Journal of Monetary Economics*, 30, 307-332.

[50] Gurley, J.G. and E.S. Shaw (1955), "Financial Aspects of Economic Development", *American Economic Review*, 45, 515-538.

[51] Gurley, J.G. and E.S. Shaw (1960), *Money in a Theory of Finance*, Brookings Institutions, Washington.

[52] Hall, B.H. (1987), "The Relationship Between Firm Size and Firm Growth", *Journal of Industrial Economics,* 35, 583-606.

[53] 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.

[54] Hayashi, F. (1982), "Tobin's Marginal Q. and Average Q.: A Neoclassical Interpretation", *Econometrica*, 1982, 213-224.

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

[56] Hart, P.E. and N. Oulton (1996), "Growth and Size of Firms", *Economic Journal,* 106, 1242-52.

[57] Hoshi T., Kashyap A. and D. Scharfstein (1992), "Corporate Structure, Liquidity and Investment: Evidence from Japanese Industrial Groups", *Quarterly Journal of Economics* 90, 33-61.

[58] Hu, X. and F. Schiantarelli (1998), "Investment and Capital Market Imperfections: A Switching Regression Approach Using U.S. Firm Panel Data", *Review of Economics and Statistics*, 80, 466-479.

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

[60] Hubbard, G.R., A.K. Kashyap and T.Withed (1995), "Internal Finance and Firm Investment", *Journal of Money Credit and Banking*, 27, 683-701.

[61] Hubbard, G.R., Skinner, J.S. and S.P. Zeldes (1994), "The Importance of Precautionary Motives for Explaining Individual and Aggregate Saving", *Carnegie-Rochester Conference Series on Public Policy, 40, 59-126.*

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

[63] Jorgenson, D.W. (1963), "Capital Theory and Investment Behaviour", *American Economic Review,* 53, 247-259.

[64] Judd, K.L. (1998), *Numerical Methods in Economics*, MIT Press, Boston.

[65] 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.

[66] Kaplan, S.N. and L. Zingales (2000), "Investment-Cash Flow Sensitivities are Not Valid Measures of Financing Constraints", *Quarterly Journal of Economics,* 115, 707-712.

[67] Kashap, A., Lamont, O. and J. Stein (1994), "Credit Conditions and the Cyclical Behaviour of Inventories", *Quarterly Journal of Economics*, 109, 565-592.

[68] Kashap, A., Stein, J. and D. Wilcox (1993), "The Monetary Transmission Mechanism: Evidence From the Composition of External Finance", *American Economic review,* 83, 78-98.

[69] Kashap, A. and D. Wilcox (1993), "Production and Inventory Control at the General Motors Corporation During the 1920s and 1930s", *American Economic Review*, 83, 383-401.

[70] Kiyotaki, N. and J. Moore (1997), "Credit Cycles", *Journal of Political Economy,* 105, 211-248.

[71] Koopmans, R. and A. Lamo (1995), "Cross-Sectional Firm Dynamics: Theory and Empirical Results From the Chemical Sector", CEP DP #229.

[72] Krane, S.D. and S.N. Braun (1991), "Production Smoothing Evidence from Physical Product Data", *Journal of Political Economy*, 99, 558-581.

[73] Leland, H.E. (1968), "Saving and Uncertainty: The Precautionary Demand for Saving", *Quarterly Journal of Economics*, 82, 465-473.

[74] Lucas, R.E. and N. Stokey (1989), *Recursive Methods in Economic Dynamics*, Harvard University Press, Harvard.

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

[76] Modigliani, F. and M. Miller (1958), "The Cost of Capital, Corporation Finance, and the Theory of Investment", *American Economic Review*, 48, 261-297.

[77] Monge, A. (1998), "Long Term Bank-Entrepreneur Relations", mimeo, Nortwestern.

[78] Myers, S.C. and N.S. Majluf (1984), "Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have", *Journal of Financial Economics,* 13, 187-221.

[79] Oliner, S.D. and G.D. Rudebusch (1996), "Is There a Broad Credit Channel for Monetary Policy?", *Federal Reserve Bank of San Francisco Economic Review,* 1, 3-13.

[80] Quah, D. (1993), "Empirical Cross-Section Dynamics in Economic Growth", *European Economic Review*, 37, 426-434.

[81] Quah, D. (1996), "Empirics for Economic Growth and Convergence", *European Economic Review*, 40, 1353-1375.

[82] Quah, D. (1997), "Empirics for Growth and Distribution: Polarization, Stratification, and Convergence Clubs", *Journal of Economic Growth*, 2, 27-59.

[83] Ramey, V.A. (1989), "Inventories as Factors of Production and Economic Fluctuations", *American Economic Review*, 79, 338-354.

[84] Ramey, V. and K. West (1997), "Inventories", NBER WP #6315.

[85] Redding, S. (1999), "The Dynamics of International Specialisation", mimeo, London School of Economics.

[86] Rondi, L., Sack, B., Schiantarelli, F and A. Sembenelli (1998), "Firms' Real and Financial Responses to Business Cycle Shocks and to Monetary Tightening: Evidence for Large and Small Italian Companies.", *Giornale Degli Economisti e Annali di Economia,* 57, 35-64.

[87] Schuh, S. (1996), "Evidence on the Link Between Firm-Level and Aggregate Inventory Behaviour", Federal Reserve Board of Governors Finance and Economics DP #46.

[88] Stanca, L. (1999) "The Dynamic Relation Between Financial Conditions and Investment: Evidence from Panel Data", *Industrial and Corporate Change*, 8, 551-572.

[89] Stiglitz, J. and A.Weiss (1981), "Credit Rationing in Markets with Imperfect Information", *American Economic Review* 71, 912-927.

[90] Stock, J. H. and M. Watson (1998), "Business Cycle Fluctuations in U.S. Macroeconomic Time Series", NBER WP #6528.

[91] Summers, L.H. (1981), "Taxation and Corporate Investment: a $q$-Theory Approach", *Brookings Papers on Economic Activity*, 67-127.

[92] Tobin, J. (1969), "A General Equilibrium Approach to Monetary Theory", *Journal of Money Credit and Banking,* 1969, 15-29.

[93] Von Furstenberg, G.M. (1977), Corporate Investment: Does Market Valuation Matter in the Aggregate?", *Brookings Papers on Economic Activity*, 347-397.

[94] Whited, T.M. (1992), "Debt, Liquidity Constraints, and Corporate Investment: Evidence from Panel Data", *Journal of Finance,* 47, 1425-60.

[95] Zeldes, S.P. (1989), "Consumption and Liquidity Constraints: An Empirical Investigation", *Journal of Political Economy*, 97, 305-46

# Chapter 6

# Appendices

## Appendix 1

**Proof of proposition 1.**

Following the notation in paragraph 3.3.1, we consider $k_{t+1}^{PM}$, which is the level of capital that maximises expected profits $E_t(\Pi_{t+1})$:

$$E_t\left(\Pi_{t+1} \mid k_{t+1}^{PM}\right) = (1+\eta)\, y_{t+1}^{PM} - UKk_{t+1}^{PM} \tag{6.1}$$

where $y_{t+1}^{PM}$ is the level of output generated using $k_{t+1}^{PM}$. $E_t(\pi_{t+1})$ instead denotes the expected profits *net* of the minimum consumption share $\eta y_t$:

$$E_t\left(\pi_{t+1} \mid k_{t+1}^{PM}\right) = E_t\left(\Pi_{t+1} \mid k_{t+1}^{PM}\right) - \eta y_{t+1}^{PM} \tag{6.2}$$

From equation (3.18) it follows that $E_t(\pi_{t+1})$ is maximised by the level of capital $k_{t+1}^{*} < k_{t+1}^{PM}$. Let's derive from (3.8), (3.9) and (6.2) the law of motion of $w_t$ :

$$\Delta w_{t+1} = r(w_t - x_t^{*}) + \pi_{t+1} \tag{6.3}$$

The proof of point (i) of proposition 1 is very simple. From (6.2) it follows that $\lim_{\eta \to 0} E_t\left(\pi_{t+1}^{PM}\right) = E_t\left(\Pi_{t+1}^{PM}\right)$, and from (3.15) and (3.18) it follows that $\lim_{\eta \to 0} k_{t+1}^{*} = k_{t+1}^{PM}$. Moreover from (3.6) and (3.9) it follows that $\lim_{w_t \to 0} k_{t+1} = 0$ and $\lim_{w_t \to 0} \pi_{t+1} = 0$ . Therefore for $w_t = 0$ it follows from (6.3) that $w_{t+1} = 0$, and $E$ is still financially constrained at time $t+1$. Therefore it must be that $w^{MAX}$ is greater than 0.

Regarding point (ii) of proposition 1, we first claim that, as long as there are some future expected financial constraints, it is optimal to set $x_t^* = 0$. We now show that $w^{MAX}$ is the level beyond which financial wealth becomes a monotonously increasing stochastic process. Given (6.3) $w^{MAX}$ is the following:

$$w^{MAX} = -\frac{\left(\pi_{t+1} \mid k_{t+1} = k_{t+1}^{PM}; \theta_{t+1} = \theta_L\right)}{r} \tag{6.4}$$

Intuitively, $w^{MAX}$ ensures that even in the worst case scenario, with $\theta_{t+1} = \theta_L$, the return from existing capital $rw^{MAX}$ is equal to the loss in profits. Therefore from (6.3) it follows that $\triangle w_{t+1} = 0$ conditional on $\theta_{t+1} = \theta_L$ and $\triangle w_{t+1} > 0$ conditional on $\theta_{t+1} = \theta_H$. Therefore when $E$ reaches a level of wealth equal to $w^{MAX}$ she becomes so rich that she has a zero probability of being financially constrained in future, and always chooses $k_{t+j}^{PM}$, for $j = 1, ..., \infty$, as optimal level of capital. Moreover she becomes indifferent between consuming today or postponing consumption. That is, $x_t^*$ can be greater than 0. The claim that it is optimal to set $x_t^*$ equal to 0 if there are some future financing constraints implies that with $w_t > w^{MAX}$ $E$ is indifferent in choosing any $x_t^*$ between 0 and $w_t - w^{MAX}$. This together with (6.4) proves point (ii).

In order to prove point (iii) of proposition 1, let's note that from equation (3.15) and from the concavity of the production function it follows that $dk_{t+1}^{PM}/d\eta > 0$. Since $k_{t+1}^*$ is the level of capital that maximises $E_t(\pi_{t+1})$, condition (3.20) and the concavity of the production function imply that $dE_t(\pi_{t+1})/d\eta < 0$. Applying this result to (6.4) proves point (iii) of proposition 1. Finally, in order to prove point (iv) of proposition 1, we define $\eta_{\min}(\Theta)$ as the minimum level of $\eta$ that satisfies the following condition:

$$\left(\pi_{t+1} \mid k_{t+1} = k_{t+1}^{PM};\ \theta_{t+1} = \theta_H\right) = -r\overline{w} \geq -r\underline{\underline{w}}(\theta_t = \theta_H) \tag{6.5}$$

where $\underline{\underline{w}}(\theta_t = \theta_H)$ is the minimum level of financial wealth that supports an investment level of $k_{t+1}^{PM}$ conditional on $\theta_t = \theta_H$. (6.5) implies, for $w_t = \overline{w}$, the following:

$$\triangle w_{t+1} < 0 \text{ conditional on } k_{t+1}^{PM} \text{ and } \theta_{t+1} = \theta_L \tag{6.6}$$

$$\triangle w_{t+1} = 0 \text{ conditional on } k_{t+1}^{PM} \text{ and } \theta_{t+1} = \theta_H \tag{6.7}$$

(6.6) and (6.7) imply that $w_t$ never grows above $\overline{w}$, which by definition is smaller than $w^{MAX}$.■

**Proof of proposition 1 for the extended model.**

We redefine the following:

$$E_t\left(\Pi'_{t+1} \mid k^{PM}_{t+1}, l^{PM}_{t+1}\right) = (1+\eta)\, y^{PM}_{t+1} - UKk^{PM}_{t+1} - ULl^{PM}_{t+1} \tag{6.8}$$

$$E_t\left(\pi'_{t+1}\right) = E_t\left(\Pi'_{t+1}\right) - \eta y_{t+1} \tag{6.9}$$

From (4.16), (4.17) and (4.19) it follows that $k^*_{t+1} < k^{PM}_{t+1}$ and $l^*_{t+1} < l^{PM}_{t+1}$. The proof of point (ii) is the same as before, with (6.4) redefined as follows:

$$w'^{MAX} = -\frac{\left(\pi_{t+1} \mid k_{t+1} = k^{PM}_{t+1}; l_{t+1} = l^{PM}_{t+1}; \theta_{t+1} = \theta_L\right)}{r} \tag{6.10}$$

The proof of point (i) follows the proof for the basic model, by noting that assumptions *a1* and *a2* imply that $\lim\limits_{w_t \to w^{\min}(k_t,\theta_t)} l_{t+1}(w_t, k_t, \theta_t) = 0$ and that in this case $l^E_{\min}$ is such that $E$ has just the necessary resources to repay the debt without liquidating her business. Therefore it follows that for $w_t \to w^{\min}(k_t, \theta_t)$ $E$ is constrained today and has a positive probability to be constrained tomorrow, and therefore $w^{MAX} > w^{\min}(k_t, \theta_t) > 0$ for any $(k_t, \theta_t)$. The proof of point (iii) follows the proof of the basic model, because from (4.16), (4.17) and (4.19) it follows that $dk^{PM}_{t+1}/d\eta > 0$ and $dl^{PM}_{t+1}/d\eta > 0$. The proof of point (iv) is also analogous to the one of the basic model, using (6.9) instead of (6.2).■

**Proof of corollary 2.**

In order to prove corollary 2, lets consider the expected lifetime utility at time $t$ of an active entrepreneur, after $\theta_t$ is realised, conditional on not becoming ill and *before* consuming $x_t$, denoted by $V'_t\left(w_t, \theta_t, k_t \mid \Theta, w_0, \theta_0\right)$. We write it as a dynamic Lagrangean, including the constraints (3.1) and (3.6) and the associated Lagrangean multipliers $\vartheta_t$ and $\lambda_t$ respectively:

$$V_0(w_0,\theta_0) = \underset{\begin{pmatrix} k_{t+1} = k(w_t,\theta_t,k_t) \\ b_{t+1} = b(w_t,\theta_t,k_t) \end{pmatrix}_{t=0,1,\ldots\infty}}{MAX} E_0\left\{\sum_{t=0}^{\infty}\left(\frac{\gamma}{R}\right)^t \left\{\begin{array}{c} (1+\vartheta_t)\,x_t^* + \eta y_t \\ +\frac{(1-\gamma)}{R}(w_{t+1}+\eta y_{t+1}) \\ +\lambda_t\left[\tau_k k_{t+1} - b_{t+1}\right] \end{array}\right\}\right\} \tag{6.11}$$

By using (3.9) to substitute $x_t^*$ in (6.11), and by taking the first order condition with respect to $b_{t+1}$, we have the following:

$$\vartheta_t = R\lambda_t + \gamma E_t(\vartheta_{t+1}) \tag{6.12}$$

we can solve equation (6.12) forward to get the following:

$$\vartheta_t = RE_t\left(\sum_{j=0}^{\infty}\gamma^j\lambda_{t+j}\right) \tag{6.13}$$

$\lambda_t$ is positive when the borrowing constraint is binding, and zero otherwise. Proposition 1.iv implies that $E_t\left(\sum_{j=0}^{\infty}\gamma^j\lambda_{t+j}\right) > 0$ and therefore $\vartheta_t > 0$ for any $t \geq 0$, and hence the constraint (3.1) is always binding with equality and $x_t^* = 0$, which proves corollary 2.■

**Proof of corollary 2 for the extended model.**

The proof is analogous to the one for the basic model, using (4.7) instead of (6.11).■

**Proof of proposition 3.**

We first analyse the properties of the term $E_t(\Omega_{t+1})$ defined by (3.14). For $w_0 > 0$ and given the absence of fixed costs in the model it follows that conditional on not retiring $w_t > 0$ for $t = 1, ..., \infty$. Therefore $E_t(MPK_{t+1}) < \infty$, $\lambda_t < \infty$, and $E_t(\lambda_{t+j}) < \infty$ for $t = 0, 1, ..., \infty$ and $j = 0, 1, ..., \infty$. As $j$ increases, the future expected productivity shock $E_t(\theta_{t+j} \mid \theta_t)$ converges to the unconditional expectation $E(\theta_{t+j})$, and proposition 1.iv implies that $E_t(\lambda_{i,t+j})$ converges to a constant greater than zero. Therefore from (3.12) it follows that $1 < \phi_t < \infty$ for $t = 0, 1, ..., \infty$. Now suppose that $E$ sets $k_{t+1} = k_{t+1}^*$. It follows that $E_t(MPK_{t+1}) = UK$, and from (3.14) that $E_t(\Omega_{t+1})$ is very close to 0. This and (3.23) imply[1] that $k_{t+1}^o > \dot{k}_{t+1}^*$. Therefore from (3.23) and (3.14) it follows that, for

[1]Because of the positive correlation between $\phi_{t+1}$ and $MPK_{t+1}$ in (3.14) it follows that $E_t(\Omega_{t+1})$

$t = 1, ..., \infty$:

$$E_t(\Omega_{t+1}) < \infty \tag{6.14}$$

as a consequence from the comparison of (3.15) and (3.23) it follows that $k_{t+1}^{PM} > k_{t+1}^{o}$. This proves proposition 3.i.

We now note that, as $w_t$ increases, the probability of having a sequence of negative shocks long enough to drive $w_t$ below $\underline{w}_t$ decreases. Such probability becomes zero when $w_t \geq w^{MAX}$. Hence, for $t = 0, 1, ..., \infty$ and for $j = 0, 1, ..., \infty$:

$$\frac{\partial \lambda_{t+j}}{\partial w_t} < 0 \tag{6.15}$$

$$\lim_{w_t \to w^{MAX}} pr(\lambda_{t+j} > 0) = 0 \tag{6.16}$$

Therefore from (3.12) it follows that $\lim_{w_t \to w^{MAX}} \phi_t = 1$ for $t = 1, ..., \infty$. Applying these results to (3.14) we have the following:

$$\lim_{w_t \to w^{MAX}} E_t(\Omega_{t+1}) = 0 \tag{6.17}$$

Finally, we note that from (6.15) and (6.16) it follows that $\partial E_t(\phi_{t+j}) / \partial w_t < 0$ for $j = 0, 1, ..., \infty$ and for $t = 0, 1, ..., \infty$. Applying this result to (3.14) and (3.23) we have the following:

$$\left( \frac{\partial E_t(\Omega_{t+1})}{\partial w_t} \mid w_t < w^{MAX} \right) < 0 \tag{6.18}$$

Proposition 3.ii and 3.iii follow immediately from (6.17) and (6.18) applied to (3.23).■

**Proof of proposition 3 for the extended model.**

Proposition 3 can be proved for $l_{t+1}$ without any additional assumption. Instead to prove it also for $k_{t+1}$, we need to impose the following:

*a3) $\theta_L, \theta_H$ and $\epsilon$ are such that, for $\eta \geq \eta_{\min}$:*

$$\eta E_t(MPK_t) - \gamma(1 - \delta_k) E_t(\mu_{t+1}) > 0 \quad \forall \ (k_t, w_t, \theta_t) \tag{6.19}$$

evaluated at $k_{t+1}^{*}$ is slightly smaller than zero. Therefore for $\eta$ very small it is possible that $k_{t+1}^{o} < k_{t+1}^{*}$. numerical solutions of the problem show that such value of $\eta$ would be very close to zero, and indeed much smaller than $\eta_{\min}$.

The smaller is the expression in the left hand side, the less the precautionary saving effect tends to influence $k_{t+1}$. This assumption is not essential for model's main results and for the empirical test in chapter 5, which rely on proposition 3 to be satisfied for $l_{t+1}$. The proof is then analogous to the one for the basic model. ■

**Proof of proposition 4.**

We first prove that assumption *a1* implies that $E$ is never forced to retire. We substitute (3.6) in (4.6) obtaining the following:

$$k_{t+1} + x_t^* + l_{t+1} \leq w_t + \frac{\tau_k}{R} k_{t+1} \tag{6.20}$$

the left hand side of (6.20) is constrained downwards by constraints (3.1), (4.3) and (4.4). Therefore, if we substitute $w_t$ using (4.6), and $x_t^*$, $k_{t+1}$ and $l_{t+1}$ from constraints (3.1), (4.3) and (4.4) holding with equality in (6.20), we get the following condition:

$$y_t + \frac{\tau_k}{R} (1 - \delta_k) k_t \geq b_t \tag{6.21}$$

condition (6.21) is symmetric to (4.5) and is necessary to ensure that $E$ is not forced to violate (4.3) to repay the debt. We determine $l_{\min}^E$, the constant level of variable capital supplied by $E$, as the level of $l^E$ such that (6.21) is always satisfied for all the possible levels of state variables $w_t, k_t$ and $\theta_t$. The left hand side of (6.21) is monotonously decreasing in $\theta_t$ and $l_t$, therefore the worst possible situation is the one in which $l_t \to 0$ and $\theta_t = \theta_L$. We substitute these two limit values in (6.21), and we substitute $y_t$ using (4.1). We also define $\overline{k}$ as the maximum feasible level of $k_t$ compatible with $\theta_t = \theta_L$. Solving (6.21) in terms of $l^E$ yields the following:

$$l^E \geq \left[ \frac{\tau_k \left(1 - \frac{1-\delta_k}{R}\right) \overline{k}^{1-\alpha}}{\underline{\theta} A} \right]^{\frac{1}{\beta}} \equiv l_{\min}^E \tag{6.22}$$

Where $\underline{\theta} = E_t(\theta_{t+1} \mid \theta_t = \theta_L)$. We then prove that assumption *a2* implies that $E$ never voluntary retires. Since $E$ is risk neutral, in order to prove that continuation is always optimal it is sufficient to prove that expected return from the investment in the firm is always greater than $R$. If irreversibility is not binding, $E$ is always able to choose a combination of $k_{t+1}$ and $l_{t+1}$ such that this condition is satisfied. We consider instead the case when irreversibility is binding. The case in which continuation is less likely to be optimal is the one in which $E$ is most constrained. At the limit: $w_t = w_t^{\min}, x_t^* = 0$ and $l_{t+1} = 0$, because all the resources are used to repay $b_t$, and there is no money to invest in variable capital. $E$ compares two choices: i) if she continues, she borrows up to the maximum. By substituting the limit values $x_t^* = 0$, $l_{t+1} = 0$ and $b_{t+1} = \tau_k k_{t+1}$ in (4.6) we get the following:

$$k_{t+1} = w_t / \left(1 - \frac{\tau_k}{R}\right) \tag{6.23}$$

therefore $E$ invests $k_{t+1}$, and receives expected total revenues, net of debt repayment, equal to:

$$E_t\left(\pi_{t+1} \mid D_t = 1\right) = (1+\eta)\, E_t\left(\theta_{t+1}\right) k_{t+1}^{\alpha} l^E + \left(1 - \delta_k - \tau_k\right) k_{t+1} \tag{6.24}$$

ii) if $E$ retires, she is indifferent between consuming and investing $w_t$ in the risk free asset. Therefore substituting $w_t$ from (6.23), her expected return is:

$$E_t\left(\pi_{t+1} \mid D_t = 0\right) = R w_t = R\left(1 - \frac{\tau_k}{R}\right) k_{t+1} \tag{6.25}$$

therefore $E$ always prefers to continue if $E_t\left(\pi_{t+1} \mid D_t = 1\right) > E_t\left(\pi_{t+1} \mid D_t = 0\right)$ for $t \geq 0$, which is exactly the condition imposed by *a2.* ■

**Proof of proposition 5.**

We start by noting that from (6.15) it follows that, for $j = 0, 1, ..., \infty$ and for $s \geq j$:

$$E_t\left(\frac{\partial \lambda_{t+s}}{\partial w_{t+j}}\right) |_{w_{t+j} < w^{MAX}} < 0 \tag{6.26}$$

from (4.10) and (6.26) it follows that $\phi_t$ decreases if the expected rate of growth of financial wealth increases. Therefore in order to prove proposition 5 it is sufficient to prove that the expected rate of accumulation of financial wealth is lower when irreversibility is binding. In order to prove it, let's consider (6.3) evaluated at $x_t^* = 0$ :

$$\triangle w_{t+1} = r w_t + \pi_{t+1}^{'} \tag{6.27}$$

by taking expectations:

$$E_t\left(\triangle w_{t+1}\right) = rw_t + E_t\left(\pi'_{t+1}\right) \tag{6.28}$$

Now recall that by definition a binding irreversibility constraint implies that fixed capital is inefficiently high: $(1-\delta_k)k_t = k_{t+1} > k^o_{t+1}$. $k^o_{t+1}$ is the optimal level of capital when fixed capital is reversible. Also proposition 3 implies that $k^o_{t+1} > k^*_{t+1}$, which is the level of capital that maximises $E_t\left(\pi'_{t+1}\right)$. Therefore, since the concavity of the production function implies that $E_t\left(\pi'_{t+1}\right)$ monotonously decreases in the level of capital for $k_{t+1} \geq k^*_{t+1}$, it follows that:

$$E_t\left(\triangle w_{t+1} \mid \ (1-\delta_k)k_t = k_{t+1} > k^o_{t+1}\right) < E_t\left(\triangle w_{t+1} \mid k^o_{t+1}\right) \tag{6.29}$$

proposition 5 is proved by generalising (6.29) for any future $E_t\left(\triangle w_{t+j}\right)$ with $j = 1, 2, ...\infty$.■

**Proof of proposition 6.**

Proposition 6 follow directly from (6.16), (6.17) and (6.18), and from the fact that, when $w_t$ is smaller than $\underline{w}_t$, the concavity of the production function implies that $\lambda_{i,t}$ monotonously decreases in $w_{i,t}$.■

# Appendix 2: description of the data

The empirical analysis illustrated in chapters 2 and 5 of this thesis is based on two datasets of Italian manufacturing firms:

I) A panel of balance sheet data for 5289 Italian manufacturing firms. This is a subset of the Centrale dei Bilanci dataset (64.463 firms at 1992 according to the Cerved database) which includes only firms with complete balance sheet data from 1982 to 1992[2]. The advantage of this dataset is the amount of quantitative information available on the firms: the panel data includes assets and liabilities, profit and losses, and detailed information on the financial flows. The disadvantage is the absence of entry and exit during the sample period and the fact that the sample is not representative of the population of Italian manufacturing firms, because it is mainly composed of small and medium firms below 500 employees.

II) The three Mediocredito Centrale surveys on small and medium Italian manufacturing firms. The research department of Mediocredito Centrale, the Italian largest investment bank, compiled the three surveys in 1992, 1995 and 1998. Each survey includes a large questionnaire on firms activity, ranging from the form of ownership, to the investment and financial policy, to other organisational aspects. Each questionnaire covers the three years before its compilation (1989-91, 1992-94 and 1995-97 respectively) and includes balance sheet data for the same three years for around 4000 firms. The unique advantage of this dataset is the direct information about firms financial policy. In particular in this thesis we use the information about the financing problems the entrepreneurs faced in funding new investment projects and the information about the financing sources mostly used to fund them. The other advantage is that the surveyed firms are representative of the Italian population: in each survey the sample is randomly stratified (it reflects the sector's geographical and dimensional distribution of Italian firms) for firms below 500 employees, while it is by census for firms with more than 500 employees. The main disadvantage of this dataset is the fact that the sample composition changes across surveys,

[2] Actually the sample ranges from 1992 to 1994, but we do not use in the thesis the information from 1993 and 1994 balance sheets, except than in figures 6-2 and 6-3, because there are inconsistencies in some assets and liabilities time series between 1992 and 1993, probably due to the fact that the accounting rules changed in 1993 as a result of the E.C. armonisation of accounting procedures.

and as a consequence only 347 firms have both qualitative and quantitative data from all the three surveys.

The most interesting aspect of these datasets is that they have a fairly large number of firms in common. In particular in section 2.2 and in chapter 5 we consider the 891 firms present in both the Centrale dei Bilanci dataset and in the first Mediocredito Centrale survey. Such firms have both complete balance sheet data, from 1982 to 1992, and the detailed qualitative information from the first Mediocredito Survey, which refers to the 1989-1991 period. In particular the survey asked if the firms had one of the following problems in financing new investment projects in the 1989-1991 period: i) "lack of medium-long term loans"; ii) "too high cost of debt" or iii) "lack of collateral"[3]. Also the second and third Mediocredito surveys asked direct questions about financing problems. Unfortunately each survey asked substantially different questions to reveal financing constraints, and hence it is not possible to pool such information across surveys.

Such pooling is instead very useful when we describe the more general link between financial policy and growth in section 2.3. This is because all the three surveys asked identical questions regarding the composition of the new investment financing.

### Basic descriptive statistics

Figure 6-1 shows descriptive statistics about the size of firms. The first column refers to the Centrale dei Bilanci sample of 5289 firms; the second column to the 891 firms with also qualitative information from the First Mediocredito Survey, and the third column the 347 firms with full qualitative data from the three Mediocredito surveys. The first two samples are clearly more homogeneous in terms of size, while the third one is the more heterogeneous, ranging from very small (the 10% percentile average size of total assets is of 6.3 billion lira, that is around 5.4 million us$) to relatively large (the 90% percentile is of 260 billion lira).

The sample 2 is the one more extensively used in the thesis, in that it is the basis for the empirical analysis in chapter 5. In this sample, as in the samples 1 and 3, the balance

[3]In addition each firm could optionally assess a degree of intensity ranging from 1 to 3. Since the large majority of firms simply chose one problem without selecting intensity, we decide not to use this information here, because previous work on this dataset (see Bagella Becchetti and Caggese, 2001) found little relevance of this additional information.

Figure 6-1: Summary statistics

**Basic statistics of the 3 datasets of italian manufacturing firms (size measured as the value of total assets in millions* of italian lira, end of 1992)**

| | Sample 1: balance sheet data for 5289 firms from 1982 to 1992 | Sample 2: balance sheet data for 891 firms from 1982 to 1992 plus qualitative information for the 1989 -1991 period | Sample 3: complete information (balance sheet and qualitative) for 347 firms for the 1990 -1997 period. |
|---|---|---|---|
| N of firms | 5289 | 891 | 347 |
| Mean | 29626.79 | 54744.59 | 126687.3 |
| Median | 13800 | 2578 1 | 35026.5 |
| Std. Deviation | 71303.88 | 118057.3 | 316731.6 |
| Skewness | 14.12254 | 8.744808 | 5.832289 |
| Kurtosis | 294.189 | 105.1583 | 41.79036 |
| Minimum | 745 | 2401 | 1385 |
| Maximum | 1955000 | 1955000 | 3293412 |
| 10% percentile | 4828 | 7352.7 | 6371.4 |
| 20% p. | 6572 | 11436.6 | 11378 |
| 30% p. | 8579 | 15477 | 17258.6 |
| 40% p. | 10837 | 19929.2 | 22984.2 |
| 50% p. | 13800 | 25781 | 35026.5 |
| 60% p. | 17551 | 34102.8 | 49804.8 |
| 70% p. | 23324 | 44712.8 | 80022.9 |
| 80% p. | 34810 | 69093 | 136765.8 |
| 90% p. | 58709 | 110167.2 | 260007.9 |

*** One million of italian lira equals to 900 US$ at the averag e 1992 exchange rate.**

sheets are at company level instead that consolidated at group level.

This could be a problem for two obvious reason: i) our theory refers to a owner/manager that decides about investment and allocates funds, while if the firms analysed belong to a group then it means that production and financial decisions are taken at different levels. ii) If the firms belong to a group, then their financial wealth as accounted in the balance sheets may not be related at all to their borrowing capacity.

Our qualitative information allows us to determine that this is a relatively small problem for sample 2. In fact 65% of all the firms and 71% of the firms with less than 300 employees are independent firms. The qualitative information from Mediocredito Centrale also reports information on mergers, acquisitions and splitting of companies. In the 1982-1992 period 7.3% of all the companies in sample 2 split in two or more companies, while 27.8% of all the companies merged with or acquired another company. These episodes do not constitute a problem because they are taken into account by the Centrale dei Bilanci in forming the panel of balance sheets: for companies merged-acquired during 1982-1992 data are aggregated between the merged companies before the merging date, so that data are relative to a uniform productive unit.

The following descriptive statistics, illustrated in figures from 6-2 to 6-5, are computed using the Centrale dei Bilanci Sample of 5289 firms. We first show that the characteristics of the Italian sample are comparable to those of the most widely used UK and US samples of manufacturing firms. For example figure 6-2 compares the logaritmic distribution of the sample of 5289 Italian firms with respect to the distribution of a more comprehensive sample of UK manufacturing firms analysed by Hart and Oulon (1996). The distribution of our sample is less dispersed with respect to the UK sample, given that large firms are not included in it. Nonetheless deviations from normality are similar in both samples. Firms distribution is more concentrated in 1994 than in 1982, because of the absence of new entries in the period. Figure 6-3 shows the results of a simple regression that relates

Figure 6-2: Summary statistics - comparison between Italian and UK data about manufacturing firms

**Measures of size - summary statistics**

| | UK Firms (Hart & Oulton 1996)* | | Italian Firms - 1982 ** | | Italian Firms - 1994 ** | |
|---|---|---|---|---|---|---|
| | Sales*** | Net Assets*** | Sales**** | Total Assets**** | Sales**** | Total Assets**** |
| Mean | 7.20 | 5.55 | 8.74 | 8.48 | 9.87 | 9.78 |
| Standard Dev. | 1.66 | 1.96 | 1.03 | 1.04 | 1.01 | 1.01 |
| Skewness | 0.19 | 0.43 | 0.28 | 0.44 | 0.47 | 0.61 |
| Kurtosis | 3.18 | 1.83 | 1.50 | 1.44 | 0.58 | 0.690 |
| N. of firms | 57812 | 79491 | 5272 | 5289 | 5289 | 5289 |

* 87109 firms, 1989-1993
** 5289 firms, 1982-1994
*** Log of Br. Pounds **** Log of Italian liras

initial size to the rate of growth and the variability of growth. The results confirm the stylised fact that growth rate and variability of growth are decreasing in size (Hall, 1987; Hart & Oulton, 1996). Another issue we briefly want to explore is the degree of lumpiness of investment. This is done in figures 6-4 and 6-5. Figure 6-4 shows that, at firm level, investment is quite lumpy if measured with respect to total investment in years 1982-1994. On average the year with higher investment accounts for more than 25% and 38% of total investment in equipment and structures respectively. This investment is however small compared to total assets (first two columns of figure 6-4), as it accounts for less than 10

Figure 6-3: Summary statistics - conditional growth

**Growth rates and variability of Growth as a function of initial dimension**

Dependent Variable: rate of growth of total assets

| 1982-1985 | | | 1986-1989 | | | 1990-1994 | | |
|---|---|---|---|---|---|---|---|---|
| Variab. | Coeff | t-stat | Variab. | Coeff. | t-stat | Variab. | Coeff | t |
| (Const.) | 55.8 | 33.5 | (Const.) | 1.426 | 6.358 | (Const.) | 2.671 | 7.409 |
| Ln8201 | -7.5 | -32.4 | Ln8601 | -0.093 | -3.274 | Ln9001 | -.253 | -5.806 |
| Ln8202 | -6.9 | -32.9 | Ln8602 | -0.099 | -3.758 | Ln9002 | -.250 | -6.152 |
| Ln8203 | -6.5 | -32.9 | Ln8603 | -0.103 | -4.100 | Ln9003 | -.240 | -6.238 |
| Ln8204 | -6.1 | -33.0 | Ln8604 | -0.098 | -4.146 | Ln9004 | -.231 | -6.347 |
| Ln8205 | -5.4 | -33.0 | Ln8605 | -0.092 | -4.324 | Ln9005 | -.212 | -6.483 |

Dependent Variable: st. dev. of the rate of rowth of total assets: 1982-1994

| Variab. | Coefficient | t |
|---|---|---|
| (Constant) | 5.601 | 26.954 |
| Ln8201 | -.734 | -25.358 |
| Ln8202 | -.684 | -25.852 |
| Ln8203 | -.646 | -25.891 |
| Ln8204 | -.607 | -25.959 |
| Ln8205 | -.539 | -25.967 |

Each variable is Log (Total Assets) * Dummy. A dummy is equal to one for the corresponding dimensional class (5 equal dimensional classes in 1982), zero otherwise. A variable Ln *yycl* is Log of total assets computed at year *yy* for dimensional class *cl*. Smallest class is 01 and largest 05

% of total assets for more than half of the firms in the sample. Figure 6-5 considers the lumpiness for different groups of firms. Small firms make more concentrated investment in structures than large firms do while firms with financing problems do not exhibit a different behaviour with respect to firms without financing problems. These simple statistics are broadly comparable to the results of Doms and Dunne (1998), cited in chapter 1[4].

[4]Even though Doms and Dunne use plant level data while we use firm level data, we expect this not to cause too much difference as most of our firms are small and hence are most likely located in one single plant.

Figure 6-4: Summary statistics - investment

**Lumpiness of investment**

| Deciles | Maxinveq1 | Maxinvstr1 | Maxinveq2 | Maxinvstr2 |
|---|---|---|---|---|
| 10 | 0.024 | 0.009 | 0.146 | 0.213 |
| 20 | 0.042 | 0.033 | 0.172 | 0.254 |
| 30 | 0.061 | 0.050 | 0.197 | 0.295 |
| 40 | 0.082 | 0.067 | 0.222 | 0.333 |
| 50 | 0.103 | 0.084 | 0.251 | 0.382 |
| 60 | 0.130 | 0.104 | 0.285 | 0.441 |
| 70 | 0.163 | 0.129 | 0.335 | 0.506 |
| 80 | 0.210 | 0.160 | 0.410 | 0.610 |
| 90 | 0.299 | 0.227 | 0.522 | 0.801 |
| Mean | 0.143 | 0.107 | 0.304 | 0.4431 |

Maxinveq= maximum investment in equipment in one year, in the period 1982-1994
Maxinvstr= maximum investment in structures in one year, in the period 1982-1994
Variables 1: investments over average total assets
Variables 2: investments over total investment in the 1982-1994 period

Figure 6-5: Summary statistics - investment for subclasses of firms

**Maximum investment in one year over total investment* in the period 1982-1994**

| | Investment in structures | Investment in equipment | | Investment in structures | Investment in equipment |
|---|---|---|---|---|---|
| Small Firms | 0.54 | 0.33 | Firms without financial problems | 0.39 | 0.27 |
| Std. Dev. | 0.24 | 0.19 | Std. Deviation | 0.19 | 0.15 |
| Medium Firms | 0.45 | 0.31 | Firms with financial problems | 0.39 | 0.27 |
| Std. Dev. | 0.22 | 0.17 | Std. Deviation | 0.21 | 0.16 |
| Large Firms | 0.37 | 0.28 | | | |
| Std. Dev. | 0.17 | 0.16 | | | |
| Mean | 0.44 | 0.30 | | | |
| Std. Dev. | 0.22 | 0.17 | | | |

* total investment is the sum of all positive investments in the years 1982-1994

# Appendix 3: description of the variables used in the estimation of the production function.

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

$p^k_t K_{i,t}$: we compute $p^k_t K_{i,t}$ 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 in chapter 4 we assume that it takes one period for fixed and variable capital to become productive. The assumption is realistic, but 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 $p^k_t K_{i,t}$ 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:

$$p^{kj}_{t+1} K^j_{i,t+1} = p^{k^j}_t K^j_{i,t}(1 + \pi^j_t)(1 - \delta^j) + p^{kj}_{t+1} I^j_{i,t+1}$$

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). $\delta^j$ 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: $\delta^1$ ranges from 9.3% to 10.7%, and $\delta^2$ from 8.4% to 10.6%.

$p^l_t L_{i,t}$ : 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 4. transition dynamics theory

To illustrate how the estimation of transition matrices works, lets' call the variable of interest for the researcher at time $t$ $X_t$ (with $t$ an integer) and assume it can take values in a certain set $E$. Let $F_t$ be the distribution of that variable at time $t$: its law of motion is described by a first order autoregressive process (Quah, 1997):

$$F_{t+1} = T^* (F_t)$$

The operator $T^*$ maps the distribution from period $t$ to period $t+1$. More precisely, the operator $T^*$ can be interpreted as a *transition function* or *stochastic kernel* $P(x, \cdot)$.[5] Let $A$ be any subset of $E$; then the distribution at time $t+1$ is defined by:

$$F_{t+1}(A) = \int P(x, A) F_t(dx)$$

Thus the transition function maps the distribution $F_t$ from one period to the other. Although it assumes a markovian structure of the underlying process, the approach of distribution dynamics is different from the traditional Markov process theory. In the latter, the emphasis is on a scalar process, from which an unobservable sequence of probability distributions is usually inferred. Distribution dynamics shows his originality in the fact that a sequence of entire (empirical) cross section distributions is actually observed, while the (*dual*) scalar process is implied but never observed (Quah, 1996). Estimation of the kernel is carried out by first estimating non-parametrically the joint density function of the process at times $t$ and $t+1$ and then normalizing it by the marginal in $t$ (see, for example Quah, 1996). The estimated transition probability density is independent of the time period $t$ (it is a *stationary* transition density): this is a common assumption in Markov chain theory.

[5] The stochastic transition function, or stochastic kernel, $P(x, A)$, describes the probability that the next step will take us in a certain set $A$, given that we are currently in state $x$:

$$P(x, A) = \Pr(X_{t+1} \in A \mid X_t = x)$$

for all values $x$ in $E$ and all the subsets $A$. This framework is appropriate in this context, as the variables of interest in this work are continuous variables in the sense that they can assume any value on the real line or a subset of it.