
Author: Fabio Schiantarelli
FINANCIAL CONSTRAINTS AND INVESTMENT:

by

Fabio Schiantarelli
(Boston College)

Revised, August 1995

This paper has been prepared for the conference "Is Bank Lending Important for the Transmission of Monetary Policy?" organized by the Federal Reserve Bank of Boston, North Falmouth, MA, June 11-13, 1995. I would like to thank S. Fazzari, D. Hester and the other conference participants for their comments. I am also grateful to K. Baum, S. Bond, R. Chirinko, H. Schaller and S. Tittman for useful suggestions and discussions.
1. Introduction

Recently, there has been resurgence of interest in the determinants of firms' investment decisions. The empirical shortcomings of existing models developed mainly under the assumption of perfect capital markets and the theoretical advances in the field of information economics have stimulated an explosion of studies that focus on the effects of financial constraints on investment.

The purpose of this paper is to provide a critical assessment of the methodological issues involved in the empirical testing of the implications of capital market imperfections for investment, and to offer a critical review of the econometric evidence on this topic. In particular, I will concentrate on the empirical contributions that have used firm level panel data. It is the increased availability of panel data that has resulted in the burst of empirical work in recent years. This is because with firm level panel data a researcher can examine how the incidence and severity of information and incentive problems vary across firms and over time, and investigate what the differential effects on investment are. Finally, in this paper I will adopt an international perspective and comment on the econometric evidence on firm investment behavior available for both developed and less developed countries.

In Section 2 I begin with a brief review of the theoretical arguments that explain why information and incentive problems introduce a wedge between the cost of internal and external finance, and outline the implications for investment decisions. In Section 3 I explore the methodological issues involved in testing for the importance of financial constraints on the basis of Q models of investment. The tests for the presence of financing constraints have consisted mainly in adding proxies for the availability of internal funds and/or firms' net worth to the model derived under the assumption of perfect capital markets, and in investigating whether they are significant for the firms that are thought more likely to face information and incentive problems. I discuss the potential weaknesses of Q based models, in particular whether average
Q captures adequately future profit prospects, and review possible solutions to this problem. The most widely used alternative approach has been to estimate the Euler equation for the capital stock. Its advantages and drawbacks are reviewed in Section 4. In both the Q and the Euler equation approach it is necessary to partition the sample of firms (or firm-year observations) according to the likelihood that they will suffer from information or incentive problems. In Section 5 I investigate the conceptual and econometric problems involved in the choice of the criteria used in splitting the sample. The main issue here is how to deal with the potential endogeneity of the sample stratification criteria commonly used. Another important problem is the choice between time invariant and time varying classifications, and between criteria based on single or multiple indicators of firms financial status. Section 6 contains a critical assessment of the evidence available for several developed and developing countries. The discussion is organized around the various criteria used to classify the observations both cross sectionally and over time (dividend payout behavior, association with business groups and banks, size, concentration of ownership, etc.). I also review the evidence on variations over time in the tightness of financial constraints due to changes in business cycle conditions or in the stance of monetary policy, and also due to financial markets reforms. Finally, in Section 7 I offer some concluding remarks and suggestions for future work.

Section 2: Information and Agency Problems, Substitutability Between Internal and External Finance, and Investment

Under the Modigliani Miller Theorem (1958), a firm's capital structure is irrelevant to its value. In this case internal and external funds are perfect substitutes and firm investment decisions are independent from its financing decisions.

However, the irrelevance hypothesis fails in the presence of informational asymmetries and contract enforcement problems. These problems may give rise to agency costs. Myers and
Majluf (1984) point out the informational asymmetry problems of equity financing. They show that if outside investors are less well informed than managers about the value of the firm’s assets, then, because of adverse selection, they will demand a premium to purchase all firm’s shares in order to offset the losses incurred from financing “lemons”.

Stiglitz and Weiss (1981) show that informational asymmetries may cause credit rationing in the loans market. Since the project risk is unobservable, lenders cannot price discriminate between good and bad borrowers. When the interest rate increases relatively good borrowers drop out of the market, increasing the probability of defaults on loans made, and, possibly, decreasing lenders’ expected profits. In equilibrium, lenders may set an interest rate that leaves an excess demand for loans. The possibility of credit rationing in the context of optimally designed contracts has also been suggested by Williamson (1987), using the costly state verification model, in which profit outcomes can only be observed at a cost.

Jensen and Meckling (1976) argue that the presence of limited liability debt will give rise to moral hazard problems in the sense that a firm may have the incentive to opt for excessively risky investment projects, even if these projects are value decreasing. When debt holders anticipate this behavior, they will demand a premium on the debt they purchase or covenants that restrict the firm’s future use of debt. Moreover Myers (1977) shows that when the firm is partly debt financed, it may forgo projects with positive net present value because the returns from such investment may be captured by debt holders.

Jensen and Meckling also consider the potential conflict of interest that may arise between managers and outside shareholders. If managers have less than 100% stake in the company, they may have an incentive to use firm resources in the form of perquisites or other wasteful activities. Such activities can be monitored at a cost, and ultimately the insiders will

---

1 See also Jaffee and Russell (1976).
2 See Townsend (1979) and Gale and Hellwig (1985).
bear the cost in terms of a reduced price that prospective shareholders are willing to pay for a stake in the firm.

Informational asymmetries, costly monitoring and contract enforcement, and incentive problems outlined above lead to an imperfect substitutability between internal and external funds. The consequences of these information and incentive problems for investment have been explored in a set of more recent papers by Bernanke and Gertler (1989, 1990), Gertler and Hubbard (1988), Calomiris and Hubbard (1990), Gertler (1992), Kiyotaki and Moore (1993), Greenwald and Stiglitz (1988, 1993). Although the models differ in their details, two main results emerge from this literature. First, unless the loans are fully collateralized, external finance is more costly than internal finance. Second, the premium on external finance is an inverse function of a borrower's net worth (liquid assets plus the collateral value of illiquid assets). It follows that negative shocks to net worth lead to an increase in the premium and, therefore, to a reduction in investment and production. For this reason the initial impact of the shock will be amplified (the so called "financial accelerator" effect).

All this has important consequences for the channels of transmission of monetary policy. An increase in the interest rate will work not only through the traditional impact on the user cost of capital, but also through the adverse impact on the present value of collateralizable net worth, leading to a widening of the wedge between the cost of external and internal finance. Moreover, insofar as some borrowers are dependent upon banks because of information problems, monetary policy may restrict the supply of loans or increase their cost for this category of borrowers, inducing them to reduce their investment. Finally, the existence of information and incentive problems means that tax policy will operate both through marginal and average rates. Although

---

3 See Gertler (1988), Hubbard (1990), Bernanke, Gertler and Gilchrist (1995), and Hubbard (1995) for another perspective on the issues discussed in this paper.
4 See Bernanke (1993), Kashyap and Stein (1994), Hubbard (1994), and Cecchetti (1994) for a fuller discussion.
The marginal rates that matter in calculating the tax benefits of an additional unit of capital spending in a world of perfect capital markets, it is the average tax rate on cash flow from existing assets that determines the (post tax) availability of internal funds for investment.

Section 3: Testing for Financial Constraints Using Q Models

The basic approach to testing for the importance of financial constraints has been to assess whether there are significant departures from standard models derived under the assumption of perfect capital markets for those firms with low net worth. These firms suffer more from information and incentive problems and such models are more likely to be misspecified. At the same time their investment is likely to be more sensitive to fluctuations in proxies for internal net worth.

Many of the empirical tests of the importance of financial constraints for investment have used, as a point of departure, the standard model of investment based on the assumption of convex adjustment costs. Consider for simplicity a firm that can only finance itself either through retentions or new share issues. Under the assumptions of perfect competition, linear homogeneous technology, and capital as the only quasi fixed input, average Q is a sufficient statistic for investment. Conditional on Q no other variable should matter when the firm is either paying positive dividends or issuing new shares. The investment equation under quadratic adjustment costs can be written as:

\[
\frac{I_u}{K_u} = a + \frac{1}{b} Q_o + \epsilon_u. \tag{1}
\]

\(I_u/K_u\) denotes the investment rate. \(b\) is the multiplicative parameter in the adjustment cost function, \(a\) is the non stochastic additive parameter and \(\epsilon_u\) is the stochastic additive component.
When the tax rate on dividends exceeds the tax rate on capital gains, it is well known that the standard formulation of Q models implies that firms will not pay dividends and issue new shares at the same time. Under retention financing, the definition of tax adjusted $Q_n$ is:

$$Q_n = \frac{\beta_n (V_s - H_s)}{\gamma_s (1 - \tau_s) P_s (1 - \delta) K_n} \frac{P_n^k (1 - \xi_s)}{P_n},$$  \hspace{1cm} (2)

where $\beta_n$ is the firm discount factor, $V_s$ the market value of the equity, $H_s$ the present value of tax savings on existing capital goods, $P_s$ the price of output, $P_n^k$ the price of investment goods, $\tau_s$ the corporate tax rate, $\xi_s$ the present value of the tax savings on new investment and $\delta$ the depreciation rate. $\gamma_s$ is the tax price of retentions in terms of dividends and equals $(1 - m_s)/(1 - z_s)$, where $m_s$ is the tax rate on dividends and $z_s$ is the tax rate on capital gains.\(^5\)

When the firm finances itself through new share issues, the only difference is that $\gamma_s$ is replaced by one in the definition of $Q_n$. Let us think of the error term containing a firm specific-time invariant component, $v_s$, an idiosyncratic component, $u_s$, and a common time component, $\eta_t$, i.e. $\varepsilon_s = v_s + u_s + \eta_t$. We can eliminate the firm specific and time invariant component of the error term by appropriate transformations of the observations and include time dummies to account for time effects that are common across firms.\(^6\) Even after these formulations one should consider that, $Q_n$ is likely to be correlated with the idiosyncratic component of the error term, either because the latter is the stochastic additive component in the adjustment cost function or because of measurement error. For this reasons, an Instrumental Variable or Generalized Method

---

\(^5\) Tax parameter have been assumed constant across firms for simplicity. Moreover, it has been assumed that new investment becomes immediately productive.

\(^6\) Taking first differences, deviation from firm means or orthogonal deviations would accomplish the desired effect.
of Moments procedure is appropriate, although many empirical contributions rely on the Least Squares Dummy Variables (or within) estimator.\footnote{See Arellano and Bond (1992) for a discussion of the GMM estimator in the context of panel data. See also Hayashi and Inoue (1991), and Blundell, Bond, Devereux and Schiantarelli (1992) for a discussion in the context of Q models. If a first difference transformation is used, and the error term in the level equation is white noise, the investment rate or $Q$ lagged twice would be legitimate instruments.}

If dividends have been exhausted and yet it is not profitable to issue new shares today or this is expected to be the case in the future, marginal $Q$ and average $Q$ no longer are equal to each other, and it is not possible to find a relationship between average $Q$ and the investment rate that does not involve present or future values of the unobservable non-negativity multiplier for dividends.\footnote{We are assuming for simplicity that the minimum dividend payment is zero.} In this financing regime investment simply equals cash flow. If a researcher estimates equation (1) using a definition of $Q$, derived under the assumption that the firm either has not exhausted retentions or is issuing new shares, this will lead to mispecification. This model can be enriched by specifying the kind of capital market imperfection that firms may face. For instance, Fazzari, Hubbard and Petersen (1988) in their seminal paper suggest that firms have to pay a lemon premium $s_n$ for issuing new shares, as suggested by Myers and Majluf(1984). In this case $1 + s_n$ should replace $\gamma_n$ in the definition of $Q_n$. The existence of a premium on new equity issues increases the range of values of $Q_n$ for which dividends have been exhausted and yet it is not profitable to issue new shares.

Debt can also be introduced in the problem. In order to make firm's financial policy determinate, one may want to assume that the interest rate is an increasing function of the debt to capital ratio. The basic idea is that incentive problems are more severe when the amount of debt is large relative to the value of collateral. If the premium above the safe rate increases linearly in leverage, the only change in the model is that the value of debt must be added to the market value of shares in the numerator of $Q_n$ so that this form of imperfection \textit{per se} does not call into
question the validity of Q models. Obviously, also in this case the Q model is misspecified if the firm pays zero dividends and issue no new shares. Another form of misspecification can also be generated if a ceiling is introduced in the amount of debt a firm can issue, and such a ceiling is binding.\footnote{We will discuss this issue at length below in the context of the Euler equation approach.} Even if firm pays and is expected to pay dividends in the future, it is easy to show that additional linear and quadratic terms in the debt to capital ratio should appear in equation (1).

The implementation of the test for the presence of financial constraints has consisted, following Fazzari, Hubbard and Petersen, of adding proxies for the availability of internal funds and/or net worth to the equation and checking whether they are significant for the firms that a priori are thought more likely to face information and incentive problems. The measurement of net worth (liquid assets plus the collateralizable value of illiquid assets) is a very difficult problem in an intertemporal context since it is related to the expectations of future returns. Typically cash flow is used as a proxy for internal net worth in empirical work. Sometimes stock measures of liquidity are also included. Both cash flow and liquid assets act not only as a proxies for net worth (which is inversely related to the premium to be paid for external finance), but also convey information about what proportion of investment spending can be internally financed. All the theories surveyed above suggest that internal funds are less costly than external finance, so that an increase in liquidity is likely to lead to greater investment.

The cross sectional criteria most commonly used to identify firms for whom information and agency problems are more severe are the dividend payout ratio (Fazzari, Hubbard and Petersen), the affiliation to industrial groups and to banks (Hoshi, Kashyap and Scharfstein (1991)), size and age (Devereux and Schiantarelli (1990)), the presence of bond ratings (Whited (1992)), degree of shareholder concentration, and one or more of the above (Oliner and Rudebush (1991), Schaller (1993)). In Section 4 I will discuss the issues involved in choosing
the criterion for sample separation, while in Section 5 I will review in detail the international evidence. On the whole the evidence from both developed and developing countries suggest that for a subset of firms internal and external finance are not perfect substitutes and that for these firms investment decisions display excess sensitivity to the availability of internal resources. There is also evidence that cash flow is significantly related to investment also for the group of firms that are thought \textit{a priori} to be less likely to face financial constraints (although not as strongly as for constrained firms).

The basic problem with testing for financial constraints in the context of Q models is that average Q may be a very imprecise proxy for the shadow value of an additional unit of new capital. The model can be extended to allow for imperfect competition in output markets, and for the presence of more than one quasi fixed factor. This introduces a wedge between marginal and average Q that is a function of observable variables and Q models can be reformulated to account for all this. However, when stock markets are not efficient and stock prices are driven by fads and fashions, or when market expectations and insider expectations diverge this problem is not easily fixed. When Q does a bad job in measuring investment opportunities, the significance of

\footnote{Chirinko (1994) argues that care must be taken in interpreting the difference in the cash flow coefficients as a sign that firms are differentially constrained. He produces a model based on the presence of flotation costs in which the size of the latter depends upon the ratio of the cash flow and Q coefficients. It is debatable, however, if one would want to summarize the degree of financial constraints faced by firms on the basis of the parameters of the flotation cost function. Nevertheless, there is a genuine difficulty in giving a "structural" interpretation to the cash flow coefficient, since one is forced to specify the precise form of the capital market imperfection to be included in the firm's optimization exercise. This problem had been noted by Devereux and Schiantarelli (1990), who had assumed that the interest rate paid by firms was a function of the cash flow rate, as well as leverage.}

\footnote{If there is imperfect competition in the output market, the shadow value of the capital stock also depends upon present and future values of the capital output ratio. See Schiantarelli and Georgousos (1990) on estimating Q models under imperfect competition and Galeotti and Schiantarelli (1991) for a model that allows also for adjustment costs for labor.}

\footnote{Under the assumption of perfect and efficient markets a relationship exists between the quasi difference in investment and dividends. Galeotti and Schiantarelli (1994) show that when a proxy for stock market fads is added to this model, it is a significant determinant of investment. On this issue see also Blanchard, Rhee and Summers (1993) and Mork, Shleifer and Vishny (1990). Both papers find that cash flow based proxies for fundamentals play a bigger role than Q in explaining investment. Again, however, the problem remains of sorting out the informational and liquidity role of cash flow.
cash flow may simply reflect the fact that it contains information about future profitability. This may be particularly true for firms that are classified a priori as more likely to suffer from information problems, so that differences in cash flow coefficients across firms cannot be interpreted as representing only the incidence and severity of such problems.

One way to address this issue is to estimate the Euler equation for the capital stock derived from the same underlying model. Although this is the prevalent solution found in the literature (see Section 4), I will first discuss other approaches that have been used to isolate the role of cash flow as a proxy for a firm's net worth. An attempt to separate the liquidity and informational content of cash flow is contained in Gilchrist and Himmelberg (1994). Following Abel and Blanchard (1986), they use a simple VAR in the profit rate and sales to capital ratio (in addition to aggregate variables) to calculate an estimate of the present value of profits resulting from an additional unit of capital today (its shadow value). This proxy is used in place of average Q in an investment equation that also contains cash flow. If the information set used in generating this proxy adequately represents the one used by the agents, the cash flow coefficient in the investment equation should only reflect its role as a source of internal liquidity or as a proxy for net worth. The evidence suggests that, even controlling for future profits, the previous conclusion on the relative magnitude of sensitivities between constrained and unconstrained firms is not affected.

Another approach is to identify changes in cash flow that represent variations in internal net worth or liquidity, and at the same time are not correlated with investment opportunities. Lamont (1993) analyzes the investment behavior of US companies that operate both in oil related and non-oil related lines of business. He finds that variation in the oil related cash flow has an effect on the investment in the non-oil related business. This is likely to reflect the fact that cash flow plays a role that goes beyond providing information about future profitability. Calomiris and Hubbard (1993) and Cummins, Hassett and Hubbard (1994) use, instead, changes in tax policies to identify changes in cash flow not related to future profitability. In particular
they analyze how investment reacts to changes in the relative taxation of retentions relative to
 dividends. If internal and external funds are perfect substitutes, one would anticipate that such
tax changes should affect payout behavior but not necessarily investment. Conversely, firms
should respond to a reduction, for instance, in the tax rates on retained earnings by increasing
investment only if they face financing constraints. The evidence here is somewhat mixed. Tax
related fluctuations in cash flow have an effect on investment in some US firms in the 1930's, but
not in Germany, France, and Japan in the 1980's and 1990's.

Fazzari and Petersen (1993) sidestep the multiple roles played by cash flow by analyzing
the relationship between investment and the variation in (end of period) working capital. Under
the assumption that working capital is less costly to adjust than fixed investment, one would
expect a negative relationship between the latter and the former in the presence of capital market
imperfections, because working capital is used as a buffer to avoid changing investment when
external funds are more expensive than internal resources or impossible to obtain. Since changes
in working capital are likely to be positively related to profit expectations, their expectational
role would instead generate a positive correlation with fixed investment. The fact that working
capital is significantly and negatively related to fixed investment for low dividend paying US
firms is suggestive of the importance of capital market imperfections.

Section 4: Controlling for Profit Opportunities Using The Euler Equation Approach

The main alternative to using augmented Q models of investment has consisted in
directly estimating the Euler equation for the capital stock. The advantage of the Euler equation
approach is that it avoids relying on measures of profitability based on firms’ market value. The
Euler equation is a different way to rearrange the first order conditions from the same
maximization problem used to derive Q equations.\textsuperscript{13} It states that the value of the marginal product of capital today, net of adjustment costs, must equal the cost of a new machine minus the cost savings due to the fact that the firm can invest less tomorrow and still maintain the capital stock on its optimal path. More precisely, allowing for imperfect competition in the output market:

\[
\frac{1}{1 + \mu} \left[ F_k(K_n, L_n) - G_k(I_a, K_n) - G_l(I_a, K_n) \right] = \frac{(1 - \xi_t)P^t}{(1 - \tau_t)P^t} \\
- E \left\{ \psi_{t+1} \beta_{t+1} \left( 1 - \tau_{t+1} \right) P_{t+1} \left( 1 - \delta \right) \left[ \frac{(1 - \xi_{t+1})P^t_{t+1}}{(1 - \tau_{t+1})P^t_{t+1}} + G_l(I_{t+1}, K_{t+1}) \right] \right\},
\]

where output, \( y_n \), equals \( F(K_n, L_n) - G(I_a, K_n) \). \( \psi_{t+1} \) represents \( \frac{\gamma_{t+1} + \lambda^0_{t+1}}{\gamma_t + \lambda^0_{t}} \), where \( \lambda^0_t \) is the non-negativity multiplier for dividends. \( \mu \) denotes the markup up of prices over marginal costs assumed to be constant through time:\textsuperscript{14}

For estimation purposes, under quadratic adjustment costs and linear homogeneity the equation can be written (omitting the constant term) as:

\[\text{\textsuperscript{13} It should be clear that neither the Q nor the Euler equation approaches yields an investment rule, in which investment is written as a function of predetermined variables and present and expected values of exogenous variables.}\]

\[\text{\textsuperscript{14} See the Appendix for details.}\]
\[
\frac{I_{t}}{K_{t}} = \left( \frac{I_{t}}{K_{t}} \right)^{2} + \psi_{t+1}\beta_{t+1} (1 - \delta) \frac{I_{t+1}}{K_{t+1}} + \left( 1 + \mu \right) \frac{\pi_{t}}{b} \frac{(1 - \xi_{t}) P^{k}_{t}}{(1 - \tau_{t}) P_{t}} + \\
\psi_{t+1}\beta_{t+1} (1 - \delta) \frac{(1 - \xi_{t}) P^{k}_{t+1}}{(1 - \tau_{t+1}) P_{t+1}} - \frac{\mu}{b} \left( \frac{Y_{t}}{K_{t}} \right) + \nu_{t+1}
\] (4)

\(\pi_{t}\) is net revenue minus variable costs and \(\nu_{t+1}\) now also includes the error generated because expected future variables have been replaced by their realizations. \(^{15}\)

Again if the firm pays dividends in both periods, both \(\lambda_{t}^{D}\) and \(\lambda_{t+1}^{D}\) will be zero. In this case \(\psi_{t}\) will equal \(\frac{Y_{t}}{Y_{t+1}}\) and, conditional on defining a proxy for \(\beta_{t+1}\), the Euler equation can be consistently estimated by IV or GMM techniques using lagged values of the included variables as instruments. If there is no stochastic component in the adjustment cost function and there are no measurement error problems, \(\nu_{t+1}\) is only an expectational error and variables dated \(t-1\) are potentially legitimate instruments after differencing to eliminate firm specific-time invariant components. Otherwise variables lagged at least twice should be used as instruments. \(^{16}\) The test of the validity of the orthogonality conditions proposed by Hansen (1982) can be used as a general mis specification test. If the firm faces the zero dividend constraint in either of the two periods, the instruments will be invalid and the test of over identifying restrictions should lead to a rejection of the model.

When debt is introduced in the model, one has to make a choice on the source and form of the capital market imperfection. One possibility is to assume that there is an exogenous limit

---

\(^{15}\) Note that in equation (4), the term in square brackets is operating revenue minus Jorgenson's user cost of capital.

\(^{16}\) If the equation is first differenced to eliminate firm specific-time invariant effects, endogenous variables must be lagged at least twice.
on the amount of debt the firm can issue (Whited (1992), Hubbard and Kashyap (1992),
Hubbard, Kasyap and Whited (1995)). The Euler equation for capital is still equation (4). Using
the first order condition for debt one can show that:

\[ \beta_{t+1} = \frac{1 - \lambda_n^B + \omega_{t+1}}{\psi_{t+1} (1 + (1 - \tau)) \gamma_{t+1}}, \]  

(5)

where \( \lambda_n^B \) is the multiplier associated to the debt ceiling and \( \omega_{t+1} \) is the error in forecasting
future variables in the first order condition for debt. Substituting out \( \psi_{t+1} \) in the Euler equation
for capital using (5), and forgetting about \( \omega_{t+1} \) for the moment, one can see that the firm
discount rate equals the interest rate only when the firm is at an interior solution for debt. When
the firm is at a debt ceiling, \( \lambda_n^B \) will differ from zero and this will invalidate the orthogonality
conditions used in estimation and, hopefully, this will be detected by the test of over identifying
restrictions. Notice that in order to implement this approach we need to make the somewhat
unpalatable assumption that the conditional covariance between \( \omega_n \) and the future variables in
the Euler equation for capital is constant. The restrictiveness of this assumption must be traded
off against the necessity to choose, again somewhat arbitrarily, a proxy for \( \beta_{t+1} \), when the latter
is not substituted out of the estimating equation. Since the power properties of the test of over
identifying restrictions may be poor in some circumstances, in order to sharpen the test for
financial constraints, the three papers mentioned above allow the multiplier to depend in an ad
hoc fashion on variables that capture firms' internal net worth, like cash flow or general
macroeconomic conditions. The coefficients on these variables measure the response of the
firm's discount rate that occur because of micro or macro factors when financial constraints are
binding.
Another option in modeling the nature of the financial constraints is to assume that the premium paid over the safe rate is a function of the debt to capital ratio. If this premium is linear in the degree of leverage and equals \( \frac{c}{2} \frac{B_{t+1}}{P_{t+1}^u K_{t+1}} \), then we simply need to add the following term to the right hand side of equation (4) (see Bond and Megir (1994)):

\[
\frac{c(1 + \mu) \psi_{t+1} \beta_{t+1}(1 - \tau_{t+1}) B_{t+1}^2 P_{t+1}^u}{2b} \frac{1}{(1 - \tau_{t})(P_{t+1}^u K_{t+1})^2 P_{t+1}}. \tag{6}
\]

This term contains basically the squared value of leverage and reflects the fact that an increase in capital lowers the premium for debt finance. Its significance is suggestive of the existence on a premium on debt. The augmented Euler equation will still be misspecified if the dividend constraint is binding in any period. Note that the sign of the leverage term should be positive, which means that there should be a negative partial correlation between leverage (squared) at the beginning of the period and investment during that period.\(^{17}\)

A combination of the two approaches illustrated so far is to allow for a premium over the safe rate and to use the first order condition for debt in order to substitute out \( \psi_{t+1} \) from equation (4). If the solution for debt is an interior one then:

\[
\beta_{t+1} = \frac{1 + \omega_{t+1}}{\psi_{t+1}(1 + (1 - \tau_{t+1}) \lambda_{t+1} + \frac{b}{2} \frac{P_{t+1}^u}{P_{t+1}^u K_{t+1}})}. \tag{7}
\]

\(^{17}\) Both capital and debt are defined as end of period quantities, so that equation (4) implies that leverage at the end of period \( t \) is negatively related to investment in \( t+1 \).
This introduces additional non-linear interaction terms between leverage and future variables in the model (see Johansen (1994) for a linearized version of this model). 18

The assumption of an exogenous ceiling on debt is rather unsatisfactory. The firm’s accumulation of collateralizable assets is likely to affect the maximum amount that firms are allowed to borrow. A simple way to capture this is to assume that there is a ceiling on the debt to capital ratio, implying that the maximum amount of debt is proportional to the capital stock, i.e.

\[ \frac{B_n}{K_n} \leq M_n. \]

Assume, moreover, that the firm has to pay a premium for debt that is linear in leverage. Than the following term should be added to the right hand side of the Euler equation (Jaramillo, Schiantarelli and Weiss (1994)).

\[ \frac{c}{2} \frac{(1+\mu)\psi_{t+1}\beta_{t+1}(1-\tau_{t+1})B_n^2P_n^4}{(1-\tau_t)P_n^2K_n^2} + \frac{\lambda_n^gB_n}{\gamma_t(1-\tau_t)P_n^3P_n^4K_n^2}. \tag{8} \]

The term containing the multiplier associated with the ceiling reflects the fact that additional units of capital are beneficial because they relax the borrowing constraint. Even if dividends are strictly positive in both periods, the unobservable multiplier associated with the debt ceiling appears in the equation when the ceiling is binding, and this again would invalidate the orthogonality conditions. However, if dividends are strictly positive, one can use the first order condition for debt to substitute \( \lambda_n^g \) out in the Euler equation. This leads to the inclusion in equation (4) of the terms:

18 For evidence of the effect of leverage on investment in the context of a more ad hoc specification of the investment equation see also Harris, Schiantarelli and Siregar (1994) for Indonesia, Calomiris, Orphanides and Sharpe (1994), and Lang, Ofek, and Stulz (1995) for the US, and Schiantarelli and Sembenelli (1995) for Italy. For evidence of the impact of the degree of indebtedness on the response of employment to demand shocks in US firms, see Sharpe (1994) and Calomiris et al (1994). Nickell and Nicolitas (1994) analyze the effect of leverage on employment, productivity and wages in UK companies.
\[
- \frac{c(1+\mu)}{2b} \left[ \gamma_{t+1} (1-\tau_{t+1}) \beta_{t+1} B_{t+1}^2 P_{t+1}^u \right] + \frac{(1+\mu)}{b} \left\{ \frac{\gamma_t - \beta_{t-1} \gamma_{t-1} (1-\tau_t) \beta_{t-1} B_{t-1}^2 P_{t-1}^u}{\gamma_t (1-\tau_t) P_t^u K_u P_u} \right\}.
\] (9)

The presence of a term linear in leverage and the fact that the sign of the quadratic term has changed relative to the case of a non-binding ceiling (see equation (6)) allows one to assess which form, if any, of the imperfection is consistent with the data.

While I will provide a detailed critical summary of the results in Section 6, the overall evidence suggests that there are significant departures from the perfect capital market paradigm. Tests of the over identifying restrictions tend to be rejected for the sub-sample of firms thought a priori to face more severe information and agency problems. For those firms leverage terms also tend to be significant, indicating the existence of a premium on external finance and sometimes the existence of binding credit constraints. In some cases, there are also signs of mispecification for the firms for whom the perfect capital markets assumption is thought to be more reasonable.

The main advantage of the Euler equation approach is that it does not rely on average Q to measure expected profitability. The market value of the firm (relative to the replacement value the capital stock) may be a poor proxy for investment opportunities and, moreover, it precludes an investigation of those firms that are not quoted on the stock market, and it is likely that information problems are particularly severe for this kind of firms. Notice that in many countries, a large fraction of production takes place in private companies. This is certainly true for developing countries, but it also applies to many developed countries.

What are the drawbacks of the Euler equation approach? A first potential problem has been outlined by Zeldes (1989) in the context of liquidity constraints on consumption. The Euler equation approach may fail to detect the presence of financial constraints if the tightness of such constraints is approximately constant over time. This can be easily seen by focusing on the non
negativity multipliers for dividends. If $\lambda^D$ and $\lambda^O$ are close in value, then the evolution of $\psi_{t,t+1}$ in equation (4) will be dominated by the changes in the tax parameters. In this case tests of over identifying restrictions may not be able to detect departures from the null hypothesis of no constraints. Although this is a risk in very short panels, it seems less of a problem when we have data over period of time long enough to record changes in individual firms' financial strength and overall macroeconomic conditions. Moreover, we have seen that if one is willing to formulate the nature of the alternative hypothesis to the one of perfect capital markets, this may introduce additional financial variables (like leverage or cash flow) into the investment equation. The significance of their coefficients may provide a sharper test of the financial constraints hypothesis.

Moreover, parameter estimates in Euler equations are often sensitive to the normalization rule (Mairesse (1994)). Although the overall conclusions on the importance of capital market imperfections tend not to be affected, the change in parameter estimates across normalizations is somewhat worrisome. Although it may be simply the result of the poor small sample properties of the GMM estimators used, it may, instead, be suggestive of some general form of misspecification that goes beyond capital market imperfections.\(^\text{19}\) There is also evidence in some studies of instability over time in the underlying adjustment costs parameters for both Euler and Q models.\(^\text{20}\) Obviously parameter instability in models derived under the assumption of perfect capital markets may be the result of the existence of financing constraints. For instance, changes in the tightness of the non-negativity constraints for dividends leads to variations in $\psi_{t,t+1}$, while changes in the tightness of the exogenous ceiling on debt leads to a non stable relationship between the interest rate and the firm’s discount rate $\beta_{t,t+1}$ in equation (4). However, parameter

\(^\text{19}\) In order to sort out the origin of the problem, it would be useful to estimate the Euler equation with a method that is not sensitive to the choice of normalization, like LIML.

Instability might also have a different origin. Ideally what is needed are tests of parameter stability for different categories of firms. Evidence of instability for firms that are not likely to suffer from financial constraints would be suggestive of the existence of additional specification problems.

Both the Euler equation and Q type of investment equations share the same underlying model based on the assumption of convex adjustment costs. If there are fixed or linear components to adjustment costs, or if there are irreversibility constraints on investment or other forms of asymmetries in adjustment costs, both models would be misspecified in a fundamental way for both groups of firms. Still, it is comforting that the model tends to be rejected more often for firms classified as constrained. The possibility remains that one may be also picking up differences in adjustment technology. This topic certainly deserves further investigation.

A final issue with the specification of the standard model of investment is the choice of the maximand itself. The underlying assumption is either that ownership and control coincides or that managers' objective is to maximize the market value of shares of existing shareholders. However, managers may have incentives to make the firm expand beyond its optimal size because this increases their power by increasing the resources under their control. Moreover, their compensation may be directly tied to growth, or their chances of promotion may be de facto related to an increase in the size of the organization. In this situation Jensen (1986) suggests that the availability of "free cash flow" (cash flow in excess of that required to fund net present value projects) will lead to an increase in investment spending.\textsuperscript{21} For this reason the association between cash flow and investment may not reflect the information and agency problems associated to new share issues or debt. It may instead be a sign of the non value maximizing behavior of management. This issue of interpretation affects the tests of the imperfect

\textsuperscript{21} See Grossman and Hart (1982), Stulz (1990), and Hart and Moore (1990) for formal models of financial structure based on the disciplinary role of debt.
Reinvestitability of internal and external funds conducted using either the Q or Euler equation approach. Both models, in fact include cash flow type variables. The main problem with the "free cash flow" hypothesis is that it difficult, if not impossible, to test, since the variable central to the hypothesis is essentially unmeasurable. However, the merit of the "free cash flow" hypothesis is to re-emphasize the importance of agency problems between management and outside shareholders described originally in Jensen and Meckling (1976) and to focus on managers incentives and behavior as a potential source of the correlation between investment and liquidity. The actions taken to control management behavior (audits, budgetary restrictions, design of compensation systems) are costly and generate a cost premium for outside equity finance. It may be difficult to distinguish this cost from the information costs due to adverse selection problems, described by Myers and Majluf (1984). More generally, it is possible that the desire by managers not to be subject to the close scrutiny that may occur when they resort to external finance, or the fear of being replaced in case of bankruptcy or when there are changes in the structure of share ownership, may lead them to rely primarily on internal funds in order to finance investment spending. These are certainly open and difficult questions for which no definitive answers are available and that deserve further investigation. I will review the empirical results that bear on some of these issues in Section 6.

Section 5: Sample Separation Criteria

The common feature in almost all of the tests of the consequences for investment of informational asymmetries and agency costs is that they are based on the identification of a subset of firms (or firm-year observations) for whom financial constraints are likely to be more important. In this section I want to discuss some of the general issues and problems involved in choosing how to partition the sample.
First, in some papers, whether a firm belongs to the financially constrained or unconstrained group is fixed over the entire sample period. However, it is possible for firms to face financial constraints of varying intensity at different points in time. For instance, if average firm characteristics (like dividend behavior or size) over the sample or pre-sample characteristics are used, one is neglecting the information that the financial constraints may be binding for the same firm in some years but not in others. It would be more advisable in these cases to allow firms to transit between different financial states.

A second observation concerns the endogeneity of the sample splitting criteria. Some, if not most, of the criteria used to split the sample are likely to be correlated with both the firm specific and time invariant component of the error term, as well as with the idiosyncratic component. This is certainly true when one uses contemporaneous or average dividend payout behavior or firm size. Correlation with the time invariant component can be easily eliminated by appropriate transformations of the variables used in the model (taking deviations from the firm's mean, first differencing, etc.). Correlation with the idiosyncratic component can also be addressed in most, but not all, cases. Probably the simplest strategy is to use contemporaneous information in partitioning the observations in the context of a single equation, and use lagged information as instruments in the context of IV or GMM procedures. For instance, one could interact the cash flow coefficient with a dividend (size) dummy depending upon whether dividends are high or low (the firm is above or below a certain size). Alternatively, if we think that the severity of financial constraints varies continuously with certain characteristics like size, we may simply want to interact a measure of size with cash flow. In any case, consistent

\footnote{For instance in the paper by Fazzari, Hubbard and Petersen (1988) firms are classified as low or high paying firms using the prevalent payout ratio over the entire period used for estimation (1970-1984). Whited (1992) uses the pre-sample existence of a bond rating to classify firms. Hubbard, Kashyap, and Whited (1995) split the sample on the basis of dividend behavior in the two years preceding the estimation period.}
estimates can then be obtained using appropriately lagged values of these interaction terms. If the model is first differenced and the idiosyncratic component of the error term in the level equation is white noise, endogenous variables lagged twice would be legitimate instruments.

Thus it is not necessary to split the sample on the basis of predetermined criteria in order to obtain consistent estimates of the parameters. For instance, using pre sample information is certainly legitimate, but may lead to a misclassification of firms in the later years of the panel. However it is much more difficult to obtain consistent estimates when past, present, and future values of endogenous variables are used in defining the dummy used to partition the sample (unless truly exogenous instruments are available that are reasonably correlated with the endogenous variables). This is the case when average (or prevalent) endogenous characteristics are using as sample separation criteria because even lagged values of the interaction terms between the dummy and other regressors are correlated with the error term.

One may think that the issue of getting consistent coefficient estimates may not be that important. Even if the estimates are biased, it could be argued, the estimated difference is not, provided that the bias is the same for the two sets of firms. This is a potential rationale for using the Least Squares Dummy Variable estimator when estimating Q equations. However, even abstracting from the issue of measurement errors that would invalidate this procedure, not properly accounting for the endogeneity of the selection criteria is likely to generate different biases for the two sets of firms. For instance, firms classified as constrained may be those with a higher correlation between cash flow and the unobservable component of investment opportunities, which is likely to lead to a larger upward bias on the cash flow coefficient.24

24 This problem is related to the one that occurs in the Q formulation of the investment equation, when the variable that is assumed to capture investment opportunities (Q and possibly sales) does a good job for unconstrained firms but not for the others.
There is another issue that should be discussed in relation to sample separation. A common characteristic to most of the work on financial constraints is that firms or observations are partitioned into groups on the basis of a single indicator that may or may not be a sufficient statistic for the existence of liquidity constraints. In some cases two indicators are interacted, typically when a cross sectional classification criteria is used in conjunction with period dummies that capture changes in macroeconomic conditions or structural characteristics of the financial system at different points in time. In principle there is no reason not to use more than one cross sectional characteristic in order to partition the sample. It is obvious, however, that the interaction terms and, consequently, the number of parameters to be estimated increases rapidly and this may lead to imprecise inferences.\textsuperscript{25} One possible way to address this issue, and at the same time allow the data to speak as to which firm-year observations belong to constrained or unconstrained regimes, is to use endogenous switching regressions methods with sample separation unknown (Hu and Schiantarelli (1994), using panel data for US firms). In this case the probability of being constrained or unconstrained is determined by a switching function that is written as a function of a vector of firm characteristics and macroeconomic conditions. Depending upon the switching function, the firm can be in either of two regimes ("constrained" and "unconstrained"), each characterized by different values of the coefficients on Q and cash flow in the investment function.\textsuperscript{26} The model can be estimated by Maximum Likelihood. Using the data to endogenously determine which set of multiple characteristics determine the likelihood of financial constraints and how it evolves over time is obviously attractive, but this gain comes at the cost of having to make precise assumptions about the distribution of the error term. This has to be contrasted with the absence of such needs when one uses IV or GMM procedures.

\textsuperscript{25} Faroque and Ton-That (1995) suggest the use of non-nested tests in order to select the "best" among different stratification criteria. Although the idea is interesting, it as well relies on the belief that a single criteria is adequate to partition the sample.

\textsuperscript{26} Notice that the researcher does not observe in which regime each firm is in a given year.
Another example of the use of switching regression models to assess the importance of financial constraints is Nabi (1989) who uses cross sectional data for 119 firms in Pakistan to estimate an accelerator model of investment. In this case the sample separation criterion is known (whether the firms have access or not to the formal credit market) and the estimation is carried out using standard two step methods.


In this section I will review in detail the international evidence on the impact of capital market imperfections on investment decisions. I organize my discussion around the more commonly used criteria that have been employed to identify firms that are more likely to suffer from financial constraints. Most of these criteria emphasize the cross sectional differences that exist across firms. However, the importance of financial constraints is likely to vary over the business cycle and with the stance of monetary policy. Moreover, structural changes in financial markets can potentially affect the degree of substitutability between internal and external finance. For these reasons I will discuss the evidence concerning the variation of the severity of financial constraints over time. Most of the empirical contributions I will survey are based on individual firm level panel data. In some cases the individual firm data are aggregated into size classes. Unless otherwise stated the results are based on individual firm level data. In addition to evidence based on Q and Euler equations, I will also discuss results based on variations on the

27 An interesting way used to partition US companies has been the presence/absence or quality of a firm bond rating. However, this information is not generally available for other countries. Whited (1992) finds that firms with a bond rating display less sensitivity of fixed investment to cash flow. Similarly, Calomiris, Himmelberg and Wachter (1995) find that inventory investment of firms with a commercial paper rating is less sensitive to cash flow fluctuations.
flexible accelerator model. In this case future profit prospects are summarized by changes in sales. These models can be rationalized as being derived from the standard neoclassical model of investment without adjustment costs (Jorgenson (1963)) when the real user cost of capital is (relatively) constant, or from a putty clay model when the cost of labor relative to the purchase price of a machine does not change significantly (Nickell (1978), ch. 11).

6.1: Dividend Payout Behavior

The original contribution by Fazzari, Hubbard and Petersen (1988) classified US firms according to their prevalent payout behavior over the period used for estimation, and showed that firms with a low dividend payout ratio were more sensitive to cash flow, in the context of Q models of investment. The use of payout behavior tries to identify the group of firms that have exhausted their retentions and are forced to rely on external financing that is an imperfect substitute for internal finance. As we have argued above, the use of prevalent (or average) payout behavior does not take into account that firms may transit between states in which they face binding constraints and states in which they do not, and it is likely to make it virtually impossible to obtain consistent parameter estimates.\textsuperscript{28} Hubbard, Kashyap and Whited (1995) produce evidence that the test of over identifying restrictions in an Euler equation model points to a rejection for low dividend payment firms, but not for high dividend paying firms. Firms are sorted on the basis of average payout behavior in the two pre-sample years. This addresses the econometric issues of endogenous sample selection, but firms are still not allowed to transit between different financial states. Moreover, the classification criteria is less accurate for the later years, compared to the earlier ones. Bond and Megir (1994) allow firms to transit between

\textsuperscript{28} The econometric results in the paper are mostly obtained using the Least Square Dummy Variable (or Within) estimator.
constrained and unconstrained states by defining a dummy variable that equals zero when dividends are positive in both adjacent periods and one otherwise. They then interact this variable with all the regressors in the Euler equation for capital. Obviously the dummy variable is endogenous but appropriately lagged values of the interaction terms provide valid instruments. The results indicate that the cash flow coefficient is wrongly signed (negative, instead of positive, in the context of equation (4)) and significant for the constrained firms, while it is not significantly different from zero for the unconstrained firms. This result is not as clearly supportive of the importance of financial constraints as the ones obtained for the US. In fact, in terms of the sign and significance of the cash flow coefficient, the Euler equation for UK firms is not satisfactory for either group of firm-year observations, even though it is less satisfactory for those in which the dividend constraint binds.

Alonso Borrego (1994) follows Bond and Megir using data for Spanish firms and he also finds that the standard Euler equation model is rejected by the test of over identifying restrictions and that the coefficient of cash flow is wrongly signed when estimated over the entire sample, while it performs somewhat better for firms that are paying dividends.

In a recent paper, Kaplan and Zingales (1995), undertake a closer analysis of the forty-nine low dividends firms identified by Fazzari et al. as financially constrained. Using qualitative information in the companies financial statements, including statement by managers, they suggest that only in 15% of the firm year observations can be classified in the constrained group. They then show that the sensitivity of investment to cash flow is greater for the unconstrained group, contrary to the implications of information based stories. However, the Kaplan-Zingales classification is open to criticism because of its subjective nature. Moreover, it is likely to

---

29 Basically, in equation (4), the data demand a positive correlation between the investment rate at time $t$ and cash flow at $t-1$. A negative or, at best, non significant coefficient for cash flow is also obtained by Rondi, Sembenelli and Zanetti (1994), using a panel of large Italian companies.
identify financially distressed firms, which had been excluded by design from the sample used by Fazzari et al., who had chosen firms with positive real sales growth during the sample period. It should not, therefore, be a surprise that only a small number of firm-year observations are included in the constrained group. As we have already argued, the original choice by Fazzari et al. of classifying firms on the basis of their prevalent dividend payout behavior has serious potential drawbacks. Indeed, many of the contribution that have followed have addressed these problems and the other ones that have been mentioned in the previous sections. Although for different reasons, the Kaplan and Zingales finer classification within the group of constrained firms is also open to criticism, and it is not clear what general conclusion can be derived from the econometric results they obtain.³⁰

6.2: Association with Business Groups and with Banks

Business groups are a pervasive form of organization in several countries. Although this is certainly not the only way to look at them, business groups can be seen as an organizational form that helps to cope with information and contract enforcement problems in the capital markets. The knowledge by financial intermediaries or individual investors that in case of financial distress individual firms may also rely on the financial resources of the group is likely to improve their access to external financial resources. The diversification of the group's activities is an added bonus in this respect. Moreover, even in the absence of financial distress, business groups allow the formation of an internal capital market that supplements the capital allocation function of the external market. Finally, in some countries groups are organically

³⁰See Fazzari, Hubbard and Petersen (1995) for a detailed discussion of the Kaplan and Zingales' paper, including the econometric reasons that may explain the differences observed in the estimated value of the cash flow coefficient.
linked with banks. In Japan banks provide a large proportion of a firm's financing, own shares, and sit on the board of directors of industrial firms. In Germany there are also close relationships between banks and firms through board representation and the control of voting rights for their own shares and for the shares left in custody with the banks. However, contrary to common belief, bank financing does not represent a large share of corporate financing in the post war era (see Mayer (1990)).

Even though formal ties between banks and firms are absent for Italy, banks represent the dominant source of outside finance. Moreover, the dominant large business groups have special informal relationships with national financial institutions. Some of the latter play an important role not only in the financing of enterprises, but also acting as exclusive clubs where mutual share holdings are organized and strategic decisions on corporate control are taken. Notice that Italian business groups are often organized around a family nucleus and in most case the controlling group owns a large stake of total equity. Business groups also play an important role in developing countries, like Korea and Indonesia.

Whatever the form, strong ties between banks and certain firms represent a way to reduce information costs.\textsuperscript{31} In this sense we would expect that firms affiliated to a business group will be less sensitive to cash flow, both because of the mitigation of information problems in accessing external finance (especially, if there are bank links), or because of the creation of an internal capital market. The use of affiliation to industrial groups, particularly in situations in which such affiliation is a stable characteristic, is probably less subject to the problems generated by endogeneity of the sorting criterion.

Evidence on this issue is available for Japan (Hoshi et al. (1991)), Italy (Schiantarelli and Serbenelli (1995)), Korea (Cho (1995)) and Germany (Elston (1993)). In all three cases the test

\textsuperscript{31} See Diamond (1984) for a formal role that emphasizes the role and advantages of financial intermediaries as delegated monitors.
are of the excess sensitivity type, using the Q approach for Japan and Germany and an accelerator type of model for Italy.\textsuperscript{32} For Japan, Hoshi, Kashyap and Scharfstein find that firms that are members of a business group (Keiretsu) are less sensitive to cash flow compared to the independent firms. Schiantarelli and Sembenelli partition firms into three groups: members of the eighteen major Italian groups, subsidiaries of multinational corporations, and the other non affiliated firms. The coefficient of cash flow is between ten and fifteen times larger for non affiliated firms compared to affiliates of Italian groups or subsidiaries of multinational corporations. Cho finds that cash flow effects are greater for Korean firms affiliated to a business group (Chaebol) in a variety of specifications of the investment equation. However, the results from Elston suggest that the coefficient of cash flow is twice as large for German firms with bank ties (.84 versus .48), although it is less significant.\textsuperscript{33} Notice that the sample used in estimation is rather small (approximately a hundred firms). Preliminary work by Chirinko and Elston (1995) also suggest that bank-influenced firms do not enjoy lower cost in terms of bank debt or total debt. They also suggest that firms with bank links have a more dispersed share ownership structure. Since a more dispersed ownership is, everything else equal, associated with greater agency problems between management and outside investors, it would be interesting to use the two characteristics simultaneously in partitioning the sample. Finally, Schaller (1993) and Chirinko and Schaller (1995) provide evidence that members of major Canadian conglomerates do not display excess sensitivity to cash flow. Canadian conglomerates often contain distinct enterprises with their own publicly traded shares and have points of similarity with business groups in Japan or Italy. However, there is no suggestion in the two papers mentioned above that conglomerates have a special relationship with banks.

\textsuperscript{32} In a similar vein, Calem and Rizzo (1994) provide evidence that chain-affiliated hospitals in the US are less sensitive to cash flow than independent hospitals.

\textsuperscript{33} The category with closer bank ties contains those firms for whom bank ownership of stocks exceeds 10%.
On balance all of these results are consistent with the idea that group membership relaxes financial constraints. How much this is due to the role of banks' ties and how much is due to the creation of an internal capital market is a matter of conjecture, and the answer is likely to differ across countries. Detailed information on both consolidated and unconsolidated balance sheets, and on intra group loans and equity issues, in principle, could help in assessing the relative importance of these two effects. While data on these financial flows may be available, it is likely to be difficult to assess intra group flow of funds achieved through transfer pricing.

So far I have focused on the differences between different types of firms within each country in order to draw inferences on the importance of bank affiliation. Another possible way to assess the importance of financial intermediaries in minimizing the adverse consequences of informational asymmetries can be obtained by analyzing the cross country differences in the excess sensitivity to cash flow. The empirical study on financing patterns in developed countries by Mayer (1990) based on flow of funds data suggests that in all countries retentions are the dominant source of finance, while banks in general are more important than market sources of external finance. However, bank finance is particularly important in France, Italy, and Japan, while it is relatively less important in the UK and the US. Surprisingly, the proportion of total finance provided by German banks is more similar to the UK and US figures. Bond, Elston, Mairesse, and Mulkay (1995) estimate various versions of the investment equation (in its Euler equation form, flexible accelerator, etc.) on panel data for the UK, France, Belgium, and Germany and conclude that the sensitivity to cash flow is greater for the UK than for all the other countries. This is suggestive that the availability of internal finance may be more important in

---

34 All international comparisons are fraught with difficulties and conclusions reached must be treated with care. The Bond, Elston, Mairesse and Mulkay paper contains a detailed discussion of these issues and of the efforts made to render the international comparisons as meaningful as possible.
financial systems that are more market based. However caution is needed before jumping to this conclusion because it is also possible that the different role of cash flow reflects differences in the nature of the data for each country. In particular, while the UK data are consolidated accounts, the main data available for the other countries are not. Although the purpose of their study was not an inter country comparison of cash flow sensitivity, Cummins, Hassett and Hubbard (1994) find that, out of a set of fourteen countries included in the Global Vantage database, the cash flow coefficient is significant only for Japan, Norway, the UK and the US in Q equations. Obviously these four countries differ greatly in terms of the market or bank orientation of the system of external finance. Again, the results may be driven by the vastly different composition (and numbers) of the firms included in the data base for each country.

6.3: Size

One criterion that is frequently used to identify firms that are more likely to be financially constrained has been size, on the presumption that size is highly correlated with the fundamental factors that determine the probability of being constrained. Smaller firms are more likely to suffer from idiosyncratic risk and, insofar as size is positively correlated with age, are less likely to have developed a track record that helps investors to distinguish good from bad firms. Moreover, small firms may have lower collateral relative to their liabilities and unit bankruptcy costs are likely to decrease with size. Finally, it is likely that transaction costs for new share issues decrease with size. However, size may also be inversely related to concentration of ownership, and concentrated share ownership is likely to mitigate agency problems between

---

35 Notice that Q is likely to be more informative in countries in which the provision of external finance is (relatively) more market based.
managers and outside investors. This last consideration is probably more important when dealing with samples of relatively large quoted companies.

The evidence is indeed mixed. When the size criterion is applied to large data sets that include quoted and unquoted companies and cover a broad spectrum of the size distribution, then the results tend to suggest that smaller firms face significantly higher hurdles in accessing external funds. This is true both for developed countries (see Galeotti, Schiantarelli, and Jaramillo (1994) for Italy and Johansen (1994) for Norway) and for developing countries (see Jaramillo, Schiantarelli, and Weiss (1993), using data on Ecuadorian firms and Harris, Schiantarelli, and Siregar (1994) using panel data for Indonesia). Carpenter, Fazzari, and Petersen (1994) find that the impact of internal finance on inventory investment is greater for small US firms relative to large firms, although internal finance is economically important for large firms. Time series data disaggregated by firm size confirm the greater sensitivity to cash flow (relative to interest payments) of inventory investment in the US (see Gertler and Gilchrist (1994), using Quarterly Financial Report data), of fixed investment and inventory investment in Italy (see Rondi, Schiantarelli and Sembenelli (1993)), and of investment in Columbia (Tybout (1983)). However, Devereux and Schiantarelli (1990) find that large firms are more sensitive than small firms to cash flow fluctuations using a sample of relatively large quoted firms. The fact that a firm must be quoted to be included in the sample means that there probably is a selection bias in favor of picking only the best of the small firms. However, it may also be the case that larger firms have more disperse share ownership (see below). Unfortunately, the UK panel does not contain enough information to assess whether this explanation is correct. Results

36 Hall (1992) finds that R&D expenditure by US firms respond significantly to cash flow. Himmelberg and Petersen (1994) also provide similar evidence for a panel of smaller firms. The cash flow effect is stronger for their sample than for the sample of larger firms used in Hall’s paper.

37 Athey and Laumas (1994) find that large Indian firms are more sensitive to cash flow than small firms and explain their result as a reflection of the Indian government credit policies for promoting small enterprises.
on the role of size for a small sample of US firms listed on the New York Stock Exchange or traded in the Over the Counter market, suggest that there are no significant differences between size classes (see Oliner and Rudebush (1991)). Hu and Schiantarelli (1994) find that, everything else equal, size is positively related to the probability of being financially constrained for quoted companies present continuously between 1978 and 1987 in the Compustat Annual Industrial File and in the Over-the-Counter File. The probability of being constrained increases with stock and flow measures of leverage and decreases with the stock of liquid assets. Summarizing, size seems to be a useful criteria to identify firms that are more likely to be financially constrained but only when the sample used for estimation includes at least a portion of the lower tail of the size distribution.

6.4: Agency problems and Concentration of Ownership

The interests of inside shareholders are likely to be more aligned to those of outside shareholders when the former have a large equity stake in the company. Moreover, more efficient monitoring of management will occur when outside share holding is highly concentrated. In this case the agency cost premium for equity finance should be smaller. Oliner and Rudebush do not find evidence that the structure of share holding for a small sample of US firms affects the sensitivity of cash flow in Q type of equations. Results for Canada suggest, instead, that cash flow is less important for companies with more concentrated share ownership and are on average smaller (see for example Schaller (1993), Chirinko and Schaller (1995) using Q models, and Ng and Schaller (1991) using the Euler equation approach).

Additional evidence on the source of the premium for external finance is contained in Hubbard, Kashyap and Whited (1995) for the US. Within the group of low dividend payout firms, they separate those in mature sectors. These are the firms for which the problems outlined by Jensen and Meckling (1976) and Jensen (1986) should be the most important, yet the test of
over identifying restrictions does not suggest a rejection of the model for this subgroup, while it
does for the other low dividend paying firms. However, using a similar data set, Vogt (1994)
divides the low dividend paying firms into four size classes and shows that the cash flow
coefficient is greater for larger firms that presumably are the more likely to suffer from this type
of agency problem. Summarizing, the evidence is mixed and more research work is needed if we
want to identify the relative importance of the various sources of the discrepancy between the
cost of internal and external finance.

6.5: Variations over Time in the Tightness of Financial Constraints: Asymmetric Cash Flow
Effects, Business Cycle Conditions, and Monetary Policy

One of the implications of the information based models of investment is that the severity
of financial constraints is likely to vary with overall macroeconomic conditions and with the
stance of monetary policy because they influence the value of firms' net worth. We would
expect, therefore, that during recessions or after a monetary tightening the cost of external
finance increases and/or the access to it decreases.

The evidence for the US that there is time variation in the severity of financial constraints
is quite robust. Gertler and Hubbard (1988) provide empirical evidence for the US that the cash
flow coefficient for firms with low payout ratios in a Q type of investment equation is greater in
recessions. Kashyap, Lammont, and Stein (1992), using panel data, find that inventories for
firms without a bond rating are sensitive to measures of the stock of liquidity during years of
recession, but not during the subsequent boom years. They do not detect any excess sensitivity
in any period for firms with a bond rating. Similar results for investment are obtained using the
QFR data set by Oliner and Rudebush (1994). The interesting twist in their paper is that the cash
flow coefficient increases in the four quarters following a monetary contraction, defined either on
the basis of the Romer dates (see Romer and Romer (1989, 1990)) or on the basis of the behavior of the spread between the Federal Funds rate and a long term government bond rate.

All the contributions mentioned above are based on classifying firms cross sectionally and temporally, prior to estimation. In the endogenous switching regression approach of Hu and Schiantarelli (1994) macroeconomic conditions affect the probability of a firm being constrained or unconstrained both through balance sheet variables (stock and flow measures of indebtedness, stock of liquid assets, size) and the year dummies included in the switching function. This allows the data to speak about the determinants of the probability of facing constraints and the evolution of such a probability. As a summary measure of the effect of macroeconomic conditions they use the parameter estimates to calculate the average probability (across firms) of being constrained in each year. This probability varies substantially over time: it reaches its highest value in the recession of 1982 and in its aftermath, and its movements closely follows (with a lag of approximately two years) the behavior of the Federal Funds rate.\footnote{When the Federal Funds rate is included directly in the switching function in place of the year dummies, it has a positive and significant impact on the probability of being financially constrained.}

Gross (1994) provides both a theoretical and empirical analysis of the dynamics of US firms' investment and financing decisions. In his paper firms decide about fixed and liquid assets in order to protect themselves against bankruptcy, while insuring themselves of the availability of resources to undertake profitable investment. Rather than assuming exogenously that some firms are constrained, he shows that the tightness of financial constraints varies over time, depending upon the amount of internal financial resources. Kernel regression estimates of the policy function for capital and debt suggests that the capital stock is not sensitive to the amount of internal financial resources when the latter are large. When the firm is somewhat constrained, a large portion of each extra dollar of internal funds is invested. When firms are very constrained, they resort to borrowing in order to prevent the capital stock from falling further.
The international evidence on this issue is not as rich. Rondi, Sack, Schiantarelli, and Sembenelli (1993), using annual firm data for Italy, aggregated into two size classes (large and small), also find that both fixed and inventory investment respond more to changes in cash flow relative to interest payments in periods following monetary tightening. The same occurs for large firms, although their sensitivity is less than that of smaller firms in all sub periods. Schiantarelli, and Sembenelli (1995), using Italian panel data, obtain the result that the effect of cash flow is asymmetric particularly for firms that are not associated with business groups. They allow the cash flow coefficient to differ depending whether cash flow increases or decreases, and find that it is greater when cash flow decreases. This means that lack of availability of internal resources causes a decrease in investment while an increase in such availability has a weaker positive effect. Preliminary results by Guariglia (1994) using UK panel data to estimate finished goods inventory equations also suggest greater sensitivity to cash flow during recessions.


The tightness of financial constraints over time may vary, not only following changes in business cycle conditions and monetary policy, but also because of structural changes in financial markets. During the 1980's several developing countries introduced financial reforms to facilitate capital accumulation and growth. These reforms consisted mainly of the removal of administrative controls on the interest rate and in the elimination or scaling down of directed credit programs. Barriers to entry in the banking sector were also lowered and the development of securities markets was stimulated. The main objective of the banking deregulation was to provide higher returns to depositors and to increase the supply of funds for investment, although whether this happens at the economy wide level is a matter of controversy. It is likely, however,
that the amount of saving intermediated by the banking system will increase. To the extent that there are economies of scale in information gathering and monitoring, it is possible that banking intermediaries may have an advantage over the curb (informal) market in allocating investment funds and this may lead to a reduction in the premium of external finance over internal finance. On the other hand, the elimination of subsidized credit programs will increase the financing constraints on those firms that previously benefited from the system of administrative allocation of credit. This means that there are distributional consequences to programs of financial liberalization and whether they relax financing constraints for different categories of firms it is ultimately an empirical question.

Evidence about the effects of financial liberalization is provided by Harris, Schiantarelli, and Siregar (1994) for Indonesia, and by Jaramillo, Schiantarelli and Weiss (1992) for Ecuador. Harris et al. find that cash flow is large and significant in an accelerator type of equation for small firms, but not for large firms. However the cash flow coefficient decreases dramatically in the second half of the 1980's in the post liberalization period. Moreover, the coefficient on the beginning of period degree of leverage, meant to capture the marginal cost of debt, also becomes less negative, suggesting that the cost of funds schedule increases less rapidly as a function of leverage. It is interesting that these results still hold if next period's profits are included in the equation in order to control for the informational role of cash flow. One can therefore conclude that banking deregulation has indeed relaxed financial constraints for small firms in Indonesia. No significant changes are detected for large firms. There is evidence that large firms that were largely the beneficiary of subsidized credit have been able to replace directed credit with

\[ \text{Note that the decrease in the value of the cash flow coefficient is not likely to be explained by the fact that the economy was more buoyant after banking deregulation. GDP fell steadily in fact until it reached the trough in 1987.} \]
borrowing on the foreign markets. Notice that many large firms are member of conglomerates or owned by ethnic Chinese with connection to the financial markets in Hong Kong and Singapore.

The results for Ecuador by Jaramillo, Schiantarelli and Weiss, based on estimating Euler equations for capital that allow both for an interest rate that increase with the degree of leverage and a ceiling on leverage, suggest that small firms face constraints, while large firms do not. However, there are no changes in the structural coefficients over time. This may be the result of the fact that in Ecuador financial liberalization was less profound that in Indonesia, or of the fact that some of the subsidized credit programs benefited small firms in the pre-reform period. Moreover, while financial liberalization is a process that may take time before its effects can be felt, the panels used for estimation are rather short. Additional years of data will be necessary to pass final judgment, particularly on the effect of the introduction of securities markets that occurred only at the end of the 1980's and the beginning of the 1990's in Ecuador, Indonesia and other countries.

Financial deregulation is not a phenomenon limited to developing countries and it has taken place also in a set of developed countries. The paper by Hoshi, Kashyap and Scharfstein (1990) provides some evidence on the consequences of financial reforms that increased the financing options for Japanese corporations. Reforms basically involved the repeal of regulations that hampered the issues of bonds in the domestic and international markets and eliminated interest ceilings that reduced the demand for bonds. They focus their analysis on firms that had strong bank ties during the 1977-1982 period (the focus of their previous work), and show that firms that reduced those ties after 1982 and show much stronger sensitivity to cash flow than firms that maintained bank ties in the latter period. The decrease of the proportion of borrowing from banks within the group relative to total borrowing between 1977 and 1986 is used as a sample separation criterion. This result is consistent with the idea that there are benefits from intermediation, but at the same time raises the issue why would a firm choose to weaken its bank ties. Presumably if a firm decides to do so, it is because there were net benefits
from emancipating itself from the group main bank. Moreover, there is the possibility that there may be a greater correlation between cash flow and unobserved investment opportunities for firms that have decided to weaken their bank ties. In this case there would be a greater upward bias on the cash flow coefficient for such firms. Finally, given the nature of the sample split criterion which uses future information, an instrumental variable procedure based on lagged values of the regressors would not lead to consistent estimates of the cash flow coefficients. Notice that sorting by bank association is probably less of an issue for the estimation period preceding financial deregulation, a period characterized by stable and long lasting group links. Moreover, while the growth opportunities for group and independent firms in the 1977-82 period do not greatly differ, the group firms that weakened their ties after 1982 are characterized by better investment opportunities. In conclusion, the evidence on the benefit of bank ties derived from documenting the consequences of financial deregulation in Japan is less convincing. More work is needed in order to assess the consequences for financial constraints of moving to a more market oriented (or less bank oriented) financial system, including the analysis of deregulation episodes in other developed countries.

Section 7: Conclusions

The weight of the evidence I have reviewed suggests that for a substantial subset of firms informational asymmetries and incentive problems generate significant departures from the model derived under the assumption of perfect capital markets. The investment decisions of these firms are quite sensitive to the availability of internal funds.

---

40 The econometric results in the paper are obtained using OLS in differences.
This conclusion is derived from both Q models and Euler equations for capital, it holds independently from the specific cross sectional criteria used in classifying firms, and it is supported by most of the empirical evidence for a number of countries. Moreover, there is also substantial support for the proposition that the severity of financial constraints varies over the business cycle and with the stance of monetary policy. There is also evidence that for some developing countries financial liberalization and the ensuing process of financial re-intermediation has led to a relaxation of constraints for those firms that had restricted access to finance in the pre-reform period. More research is needed on the comparative performance of market based and bank based financial systems as well as on the consequences of those forms of financial deregulation that have led to an increased role for security markets, *vis a vis* banks. There is some panel data evidence within individual countries that bank association is beneficial, but the cross country evidence is still too weak to draw definitive conclusions.

Several other problems remain open at this stage, some of them quite general in nature. Quite a few of the results reviewed here suggest that there is evidence of excess sensitivity to internal funds or of misspecification of the estimated equations even for firms that are thought a priori not to suffer from severe information problems (for instance, large mature companies). Moreover, a simple look at the data reveals that retentions are their prevalent source of finance (just like for most other companies). This may be because even these firms cannot costlessly and credibly communicate their real investment opportunities to lenders and investors and consequently suffer from adverse selection problems. Alternatively, the agency problems between managers and the providers of external finance (both outside shareholders and suppliers of loan capital) may be severe. The research agenda for the future should include efforts to identify more carefully the nature of the information and agency problems that make external finance more expensive than internal finance. More in general, it would be useful to investigate more in depth how managerial preferences and incentives may generate a close association between firms' investment and the availability of internal resources.
Another direction for future research is provided by the desirability to move away from the standard assumption of convex adjustment costs underlying the model used so far for econometric testing. The evidence of mis specification, including the change in parameters across normalizations and their instability over time (in some studies), may not be wholly explained by capital market imperfections. There may also exist other forms of mis specification related, perhaps, to the irreversibility of investment and non-convexities in adjustment costs. The simultaneous treatment of both capital market imperfections and more complex forms of adjustment costs is likely to be very fruitful.
REFERENCES


Kyiotaki, N., and J. Moore (1993), "Credit Cycles", mimeo, University of Minnesota and LSE.


APPENDIX

The Firm's Optimization Problem

Denote by $R_u$ the required rate of return, the following standard arbitrage condition must hold for a firm's shareholder:

$$R_u = \frac{(1 - m_u)D_t + (1 - \zeta)E_t(V_{t+1} - V_t - S_{n}^u)}{V_t},$$

(A1)

where $D_t$ denotes dividends, $V_u$ the value of the firm, $S_{n}^u$ the nominal value of new shares, $m_u$ the personal tax rate, $\zeta$ the tax rate on capital gains, and $E_t$ the conditional expectations operator. Solving (A1) recursively gives us:

$$V_t = E_t\sum_{j=0}^{\infty} \beta_t^j \left[ \gamma_{t+j} D_{t+j} - S_{n}^{t+j} \right].$$

(A2)

$V_t$ is the value of the firm for existing shareholders. We assume this is the objective function that is maximized subject to the following constraints:

$$D_{t+j} = \left(1 - \tau_{t+j}\right) \left[p_{t+j}\left(F(K_{t+j}, L_{t+j}) - G(I_{t+j}, K_{t+j})\right) - w_{t+j}L_{t+j} - \left(i_{t+j} + A(B_{t+j}K_{t+j}^{-1}/B_{t+j-1})\right)B_{t+j-1}\right] + \left(B_{t+i} - B_{t+j-1}\right) - p_{t+i}^k I_{t+i} + S_{n}^{t+j} + C_{t+j},$$

(A3)

$$K_{t+j} = (1 - \delta) K_{t+j-1} + I_{t+j},$$

(A4)

$$D_{t+j} \geq 0,$$

(A5)

$$B_{t+j} \geq 0,$$

(A6)

$$S_{n}^{t+j} \geq 0,$$

(A7)

$$M_{t+i} \geq \frac{B_{t+i}}{p_{t+i}^k K_{t+i}},$$

(A8)
where:

\[ \beta_{t}^{i} = \prod_{i=0}^{i} \left(1 + R_{t}^{r} \right) \]

\[ R_{t}^{r} = \frac{R_{t}^{r+j}}{1 - z_{t+j}} \]

\[ \gamma_{t}^{r+j} = \frac{1 - m_{t+j}}{1 - z_{t+j}} \]

\[ \tau_{t}^{r+j} = \text{corporate tax rate}, \]

\[ p_{t}^{r+j} = \text{output price}, \]

\[ K_{t}^{r+j} = \text{capital stock}, \]

\[ L_{t}^{r+j} = \text{labor}, \]

\[ I_{t}^{r+j} = \text{investment}, \]

\[ w_{t}^{r+j} = \text{wage rate}, \]

\[ i_{t}^{r+j} = \text{riskless interest rate}, \]

\[ B_{t}^{r+j} = \text{stock of debt}, \]

\[ P_{t}^{r+j} = \text{price of investment goods}, \]

\[ C_{t}^{r+j} = \text{tax savings associated with depreciation allowances on existing capital goods}, \]

\[ \delta = \text{depreciation rate}. \]

Assume that the firm faces a downward sloping demand function. Denote with \( e_{t} \) the price elasticity of demand and define with \( \mu_{t} \) the markup of prices over marginal cost, where

\[ 1 + \mu_{t} = \frac{1}{1 + e_{t}}. \]  

To simplify notation set \( \beta_{t}^{0} = \beta_{t}^{r+j} \). Assume that the firm always issues a positive amount of debt. The first order conditions are:
\[(
\gamma_r + \lambda^D_r)(1 - \tau_r)(1 + \mu_r) - \gamma_r + \lambda^D_r(1 - \tau_r)\beta_{t+1}A_r(1 - \delta) = 0,
\]
\[(\gamma_r + \lambda^D_r)(1 - \tau_r)(1 + \mu_r) - \gamma_r + \lambda^D_r(1 - \tau_r)\beta_{t+1}(1 - \delta) = 0, \tag{A9}\]
\[(\gamma_r + \lambda^D_r)(1 - \tau_r)(1 + \mu_r) - \gamma_r + \lambda^D_r(1 - \tau_r)\beta_{t+1}(1 - \delta) = 0, \tag{A10}\]
\[(\gamma_r + \lambda^D_r)(1 - \tau_r)(1 + \mu_r) - \gamma_r + \lambda^D_r(1 - \tau_r)\beta_{t+1}(1 - \delta) = 0, \tag{A11}\]
\[(\gamma_r + \lambda^D_r)(1 - \tau_r)(1 + \mu_r) - \gamma_r + \lambda^D_r(1 - \tau_r)\beta_{t+1}(1 - \delta) = 0, \tag{A12}\]
\[(\gamma_r + \lambda^D_r)(1 - \tau_r)(1 + \mu_r) - \gamma_r + \lambda^D_r(1 - \tau_r)\beta_{t+1}(1 - \delta) = 0, \tag{A13}\]

where \(\xi_r\) is the present value of tax savings associated with depreciation allowances on investment, and \(\lambda^k_r, \lambda^D_r, \lambda^c_r\) and \(\lambda^p_r\) are the Lagrange multipliers associated with the capital accumulation equation, with the non-negativity constraint on dividends, new share issues, and the ceiling on the debt to capital ratio. \(A/(B_{t+1}, P^{t+1}_{t+1} B_{t+1})\) denotes the premium that must be paid over and above the safe interest rate. Equations (A9) through (A13), in addition to the complementary slackness condition (not reported here for brevity's sake) define the firm's optimal plan.

Assume that the gross production and the adjustment cost function are linear homogenous. Assume, moreover, that adjustment costs are quadratic.

\[
G(I_r, K_r) = \frac{b}{2} \left( \frac{I_r}{K_r} - a - c_r \right) K_r. \tag{A14}\]

When we omit debt entirely from the problem, and we assume perfect competition \((\mu_r = 0)\), then it is easy to show that (A9), (A10), (A11), and (A12) and the complementary slackness conditions imply:
\[ \lambda^k = \frac{\beta_b (V_a - H_a)}{P_a (1 - \tau_a) (1 - \delta) K_{t,t-1}} \]  

(A15)

In case when dividend payments are strictly positive. \( H_a \) is the present value of tax savings associated to the depreciation allowances on past investment. Equations (1) and (2) in the main text and variations thereof follow immediately from (A10) and (A15). To derive the basic Euler equation for the case of no debt, simply omit the \( \lambda^k \frac{B_a}{p^k_a K_a^2} \) term from (A9). Using (A10) to substitute out \( \lambda^k \) and \( \lambda^k_{i,t+1} \) from (9), one obtains equation (3) in the main text. The extensions due to the inclusion of debt can also be easily derived. Note that when the ceiling on debt is exogenous, i.e. \( B_a \leq \bar{B}_a \), \( \lambda^g \frac{B_a}{p^k_a K_a^2} \) should be omitted from (A9) and \( \frac{\lambda^g}{p^k_a K_a} \) is replaced by \( \lambda^g \) in (A12).