Structural Modelling of Financial Constraints on Investment: Where Do We Stand?

Jean-Bernard Chatelain

To cite this version:


HAL Id: halshs-00112522
https://halshs.archives-ouvertes.fr/halshs-00112522
Submitted on 8 Nov 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Structural Modelling of Financial Constraints on Investment: Where Do We Stand?

Jean-Bernard CHATELAIN

2003

Abstract

This paper surveys issues with respect to the structural modelling of econometric tests of investment facing financial constraints, to their link with firms data and asset prices, and to their consequences for macroeconomic modelling. The key issue is to provide conditions which support the interpretation of the sensitivity of investment to liquidity variables such as cash flow as a measure of financial constraints. The structural modelling of investment facing financial constraints is also limited by the structural modelling of the forces driving investment dynamics such as adjustment costs, which has not been so successful empirically.

Keywords: Investment, Financial Constraints, Structural models
JEL Classification: D92.

Résumé

Cet article présente des problèmes associés aux tests économétriques structurels des contraintes financières pesant éventuellement sur l’investissement des entreprises, à l’utilisation de données de bilans d’entreprises et à leur conséquence en terme de modélisation macroéconomique. Le problème principal est de préciser sous quelles conditions l’interprétation de la sensibilité de l’investissement à des variables financière telles que l’autofinancement est une mesure des contraintes financières. Une autre limitation vient de la qualité souvent médiocre des estimations des modèles structurels incluant des coûts d’ajustement du capital associés à des phénomènes non financiers.

Mots clés : Investissement, contraintes financières, modélisation structurale.

Laboratoire d’Économie d’Orléans (LEO), University of Orleans, Rue de Blois, BP 6739, Orleans Cedex 2. E-mail: jean-bernard.chatelain@univ-orleans.fr. I thank Catherine Fuss, Paul Butzen, Robert Chirinko and the participants to the National Bank of Belgium conference on “Firms’ Investment and Finance Decisions”, for their very helpful comments.
1. Introduction

This contribution is intended to stimulate ideas on the interplay between theory and data when estimating investment facing financial constraints, a subject of applied econometric research during more than forty years (since at least Meyer and Kuh [1957]), and probably an pervasive feature of business financing during centuries. As there are several excellent surveys dealing with these issues (e.g. Chirinko [1993], Schiantarelli [1996], Hubbard [1998], Bond and Van Reenen [1999]), I will focus on specific issues related to theoretical modelling and data and their macro-economic consequences.

Liquidity variables such as cash flow, cash stock, leverage or the coverage ratio are often found to be significant explanatory variables of investment. A recurrent debate deals with how to interpret these sensitivities: what percentage of these estimated parameters is expected to be related to financial factors and why? What is the standard error on this percentage? Conversely, what percentage of these estimated parameters is expected to be related to other other factors and why? A methodological answer is the estimation of structural parameters related to explicit characteristics of various forms of financial constraints, instead of the estimation of reduced form parameters (Lucas [1976]). The drawback of a reduced form is that estimated parameters are a mix of structural parameters and of current, past and expected variables, for example, investment liquidity sensitivities estimated in reduced form may or may not be related to financial constraints. In 1993, Chirinko’s view was that empirical evidence of financial constraints was found in many papers but was broadly inconclusive, being too distant from an explicit or ”structural” framework. Since then, some progress have been made in explaining more precisely how structural parameters are ”hidden” in those reduced form parameters. However, it may be the case that the problems in structural modelling of investment are not only due to financial constraints but also to the modelling of investment dynamics and the capital adjustment process. Nonetheless, in the recent years, the use of micro-data has greatly expanded and many papers have documented financial constraints through effects of liquidity variables on investment in several countries. It is useful to point out where caveats remains in these works.

A first section reviews three groups of testable theoretical models of investment including or not capital market imperfections. A second section details some recurrent problems in empirical tests of financial constraints on investment. A third section discusses some recent macro-economic modelling taking into account the behaviour of firms facing financial constraints. A last section concludes.
2. Testable theoretical models

2.1. Three Groups of Investment Equations

Since Jorgenson’s work [1963], traditional estimations of investment are related to the Marshallian condition equating the marginal product of capital with its marginal cost, which builds on the arbitrage between the return from investing inside the firm and the opportunity cost of investing cash outside of the firm. It is then followed by a parameterisation of the production function as a constant elasticity of substitution, which leads to the following demand for capital:

\[ \ln (K_t) = \left( \sigma + \frac{1-\sigma}{\rho} \right) \ln (Y_t) - \sigma \ln (UC_t) \]  \hspace{1cm} (2.1)

The stock of capital \( K_t \) is then a function of sales \( Y_t \) and of the cost of capital \( UC_t \), with elasticities related to structural parameters: the elasticity of substitution between capital and labour denoted \( \sigma \) and the parameter of returns to scale, \( \rho \). More precisely, the elasticity of substitution between capital and labour is exactly the opposite of the elasticity of the stock of capital with respect to the cost of capital. This intertemporal neo-classical theoretical model without adjustment costs does lead to results identical with those of a static model (Jorgenson [1963]). To compensate for this lack of dynamics, econometricians usually add auto-regressive distributed lags: this takes into account dynamics which are not directly derived from intertemporal optimisation. Lagged variable in this approach are subject to the Lucas critique in the sense that they are not exactly related to parameters of the theoretical “structural” model. These auto-regressive distributed lags models are often estimated in first differences (Chirinko, Fazzari and Meyer [1999]) or reparameterised in an error correction form (Bond et al. [1997]). Recent examples of this approach taking into account the cost of capital are found in Chirinko, Fazzari and Meyer [1999] and in the papers related to firms investment in the Monetary Transmission Network of the European System of Central Banks (for example, Butzen, Fuss and Vermeulen [2001]).

Standard investment models basing structural dynamics on intertemporal optimisation assume convex adjustment costs (Lucas [1967]). Alternative models include fixed adjustment costs which are able to describe lumpy investment behaviour. In the convex adjustment cost case, the empirical literature splits into Q models (test on the marginal condition on investment) and Euler equation tests (test on marginal condition on the stock of capital). In the first case, the market value of the firms divided by the stock of capital (Tobin’s Q ratio, see Tobin [1969]) summarises all the expected determinants of investment under technology conditions derived by Hayashi [1982] so that:

\[ \frac{I_t}{K_{t-1}} = a + bQ_t \]  \hspace{1cm} (2.2)
where $I$ is investment, $K$ is capital, $Q$ is Tobin’s average Q ratio (market value of the firm divided by the value of its stock of capital), $a$ and $b$ are parameters.

The usual Euler equation can be written as:

$$F_K (K_t, L_t) = UC_t + J_K \left( \frac{I_t}{K_{t-1}}, \frac{I_{t-1}}{K_{t-2}} \right)$$  \hspace{1cm} (2.3)

where $F_K (K_t, L_t)$ is the marginal product of capital, with $L_t$ representing labour and $J_K$ a function representing the arbitrage between the marginal cost of investing now with respect to the marginal cost of investment in the future. The production function is then parameterized as an homogeneous function of capital and labour, that is a more general assumption than the constant elasticity of substitution production function used in error correction models, which is a particular case of an homogeneous function.

$$F_K (K_t, L_t) = \rho Y_t - F_L L_t$$  \hspace{1cm} (2.4)

$F_L$ is the marginal product of labour which is related to the marginal cost of labour. This parameterisation allows to compute the marginal product of capital in such a way that the investment flow appears explicitly only as a component of the marginal cost of investing. The higher the growth of capital of the firm, (or the investment ratio), the higher the cost of investing due to convex adjustment costs. The parameterisation of the production function as an homogeneous function of capital and labour and the fact that the investment ratio appears in the cost of capital leads to specific properties of the estimated Euler equation in the perfect capital market case. The investment ratio is a negative function of EBITDA (earnings before interest, taxes and depreciation allowances) usually revealed by the parameterisation of the marginal productivity of capital. With respect to Jorgenson’s approach with constant elasticity of substitution production function, the sensitivity of the investment ratio with respect to the interest rate is no longer constant: it is a function of parameters, such as the adjustment cost parameters, and of other variables of the Euler equation.

### 2.2. Four Major Types of Financial Constraints

The microeconomic foundations of financial constraints are found in the economics of asymmetric information. But, while there is only one way to know everything, there are many ways for outsiders to have an incomplete knowledge of, for example, the investment of a specific firm. As a consequence, there is a broad variety of auxiliary assumptions describing various types of asymmetry of information leading to a variety of micro-economic models. Getting an agreement on these auxiliary assumptions is difficult. This could require the testing of auxiliary assumptions of theoretical models and not only their consequences, which are sometimes similar among those different models. This research strategy is not often chosen, but Himmelberg, Hubbard and Love [2002] provide an example. It implies a quest for organizational data and
measures of information disclosed by the firms, in their case, measures of investors’ protection. But more generally, applied work and macroeconomic modelling have used simplified versions of financial constraints.

The literature focuses on four major "simplified" capital market imperfections. Important features are the distinction between retained earnings and new share issues, which both accumulate into equity and the distinction between various types of new debt.

- First, one can consider that external debt is related to an exogenous bankruptcy or monitoring cost (Bernanke, Gertler and Gilchrist [1998]) which implies a cost premium with respect to internal funds. This cost premium may increase with the amount of debt or of leverage (Cooley and Quadrini [2001]) and may also include a fixed cost. It may also depend on other factors such as the flow of liquidity, liquid assets, the growth of debt as well as exogenous firm characteristics (Chirinko [1997]). This can be viewed as taking into account the credit supply curve characteristics for this individual firm.

- Second, one can consider that dividends cannot be negative.

- Third, one can consider that new share issues incur a cost premium with respect to internal funds: this cost premium for new share issues may be a fixed cost (Bond and Meghir [1994]), or may increase with the amount of new share issues (Cooley and Quadrini [2001]), or it may comprise both of these costs (Gomes [2001]).

- Fourth, one can consider that the firm is facing credit rationing and new share issues rationing as in Myers and Majluf [1984], or that the fixed cost of a new share issue is sufficiently high that it is not sound to issue new shares to meet the firm’s demand for capital. An example is found in Stiglitz and Weiss [1981] and Kiyotaki and Moore [1997].

At this point, the literature on financial constraints meets the literature on financial structure (Jensen and Meckling [1976], Jensen [1986], Hart [1995], Myers [2001]). To simplify matters, three groups of results emerge: static trade-offs sometimes independent of the level of investment, financial hierarchy and credit rationing.

Let us consider first a well known static trade-off between bankruptcy costs and the tax advantage of debt with respect to equity. Taxation is crucial for means of finance, as it creates large cost distortions between equity and debt, since debt service is deducted from corporate income tax; it also creates large cost distortions between dividends and capital gains, as well as between new share issues and retained earnings in most tax systems (Auerbach [1983], King and Fullerton [1984], Sinn [1990]). A well known example, the weighted average cost of capital, is presented in corporate finance textbooks such as Brealey and Myers [1999], where the cost of capital is a weighted mean of the opportunity cost of equity and the marginal cost of debt. It can
be derived from the tax distortion between debt and equity and the capital market imperfection described in the “first bullet point” alone, that is a rising marginal cost of debt as leverage increases. The arbitrage between the taxation gain and the costs of bankruptcy leads to an optimal debt/equity ratio which provides the optimal weights in the weighted average cost of capital (Auerbach [1983]). In this case, this provides an optimal debt/equity ratio which can always be reached instantaneously following shocks affecting capital demand or the costs of the means of financing capital, with the help of negative dividends or new share issues who do not attract tax and/or asymmetric information distortions of their cost with respect to the costs of retained earnings. As soon as one assumes that negative dividends are prohibited and that new share issues are more costly than new debt, there is no longer an optimal capital structure derived from a static trade-off, but a hierarchy of means of finance (Myers [2001]). This example suggests that assuming only one of the four major capital market imperfections alone lead to focus on one among several financial regimes that may confront firms in the real world.

Credit rationing differs from an increasing cost of debt as leverage increases due to a Lagrange multiplier related to credit rationing. This Lagrange multiplier measures the gap between desired investment and realised investment, but it alters marginal conditions (an explicit expression of this Lagrange multiplier can be found in Chate-lain [2000]). Jaramillo, Schiantarelli and Weiss [1996] tested the opposition between an increasing cost of debt as leverage increases versus credit rationing in an Euler equation context, where negative dividends remain a possibility. More precisely, a credit rationing regime implies that investment and leverage are always complements: they both increases or both decreases. Conversely, the regime where the cost of debt rises as leverage increases allows the possibility that investment and leverage can be substitutes or complements. It depends on the relative shifts from one period to the next in the capital demand curve (due to productivity shocks or demand shocks, for example) with respect to the capital supply curves (retained earnings, credit and new share issues curves).\footnote{The capital demand curve may be inelastic to the cost of capital in the constant returns to scale case (e.g. Kiyotaki and Moore [1997] and Bernanke, Gertler and Gilchrist [1999]) or decreasing with respect to the cost of capital in the decreasing returns to scale case (Kaplan and Zingales [1997], Gomes [2001], Cooley and Quadrini [2001]). In both cases, this curve shifts upwards following a rise in productivity.}

Consider the case where the capital demand curve is unchanged (for example, no productivity shocks from this year to next year) and the firm accumulates retained earnings, this causes the credit supply curve, which rise with the interest rate, to shift to the right. It turns out that the share of debt in the means of finance decreases whereas capital increases. Marginal increase of leverage and marginal increase of capital are “substitutes”. In the case of constant returns to scale, capital increases and the amount of debt remains constant and equity rises so that leverage decreases. In the case of decreasing returns to scale, capital increases and the amount of debt decreases
(and not only leverage): it proves rational to decrease debt and, as a consequence, the cost of capital in order to match a lower marginal return on capital related to a larger size of capital. Then, marginal increase of capital and marginal increase of debt are "substitutes".

Conversely, consider now that a rise of productivity shifts the capital demand curve upwards relatively more than the accumulation of internal equity shifts the credit supply to the right. In this alternative case, investment and leverage both increase. They are "complements" as in the credit rationing regime. Chatelain [1999] proposed a theoretical investigation of this phenomenon, where the hypothesis of a positivity constraint on dividends proved crucial, and a preliminary empirical test. This example shows that in these investment models of financial constraints, the question of the adjustment of the debt or dynamic capital structure is related to investment demand. It then interacts with the theoretical and applied econometrics literature explaining leverage and the variation of debt (e.g. Anderson and Nyborg [2001]).

2.3. Financial Constraints and Other Features of Firms’ Investment

The interaction between financial constraints and other specific characteristics of business investment offers a fruitful way to improve the accuracy of our modelling of investment behaviour of firms. I have already mentioned taxation. But financial constraints can also interact with irreversibility and uncertainty effects (Arrow and Kurz [1970], Dixit and Pyndick [1994], Trigeorgis [2002]). As mentioned above, financial constraints can increase the fixed costs sometimes related to investment irreversibility. Outside investors may also take into account the cost of the real option involved in investing now in a specific project, that is the loss incurred by losing the option to invest later on in other projects (Dixit and Pyndick [1994]). Similarly, a high uncertainty over demand or costs may foster asymmetric information, which is related to moral hazard. The higher the economic uncertainty, the more difficult it is to assess ex post what is due to management failures related to moral hazard or to economic uncertainty. Financial constraints and uncertainty effects are likely to be simultaneous. Investigating such a joint effect is proposed by Lensink and Sterken [2001] and Gérard and Vershueren [2002].

3. Recurrent Problems in Empirical Testing

3.1. Investment Cash-Flow Excess Sensitivity

The literature on financial constraints puts forward two different interpretations of investment cash flow sensitivities in the literature of financial constraints for any of the testable model currently used (non structural reduced form, Q model, Euler equation).

The first one is as follows. Since at least Meyer and Kuh [1957] did their empirical work on investment, it has been debated whether the investment cash-flow sensitivity is a signal of financial constraints or merely a signal of expect profit. In this former
case, cash flow is correlated with components of Tobin’s Q numerator, which is the discounted sum of future cash flows of the firm. Cash flow may prove to be a significant explanatory variable of investment due to the omission of the Q ratio in non structural models or due to measurement errors in Q and Cash Flow model of investment (Bond and Cummins [2001], Whited [2000], Gomes [2001]). In fact, alternative specifications of the neo-classical model such as the Euler equations, as well as common sense, suggests that both effects play a role in panel data due to the heterogeneity of the finance regime for firms. An attempt to isolate these two components inside observed cash flow has been made by Gilchrist and Himmelberg [1998], for example, using vector auto-regressive techniques on panel data. This general problem is widely acknowledged and the isolation of these two effects inside the investment cash-flow sensitivity remains a difficult task.

The second point has been put forward by Fazzari, Hubbard and Petersen [1988] and combines two features. First, they estimated that, even if investment cash-flow sensitivity does not necessarily reflect financial constraints, the excess sensitivity of investment cash flow for some firms with respect to a benchmark group is more likely to reflect financial constraints. Second, these groups of financially constrained firms have to be found using sample separation criteria which measure the extent of asymmetric information problems or the extent of difficulties in obtaining external finance (size of the firm, size of intangibles, long term relationship with banks, high trade credit, and so on). In more general terms, the higher the financial constraints and/or the asymmetric information problems, the higher the investment cash-flow sensitivity.

The hypothesis that the excess sensitivity of investment to cash flow is rising the greater the financial constraint has been refined by Kaplan and Zingales [1997]. They proposed an inverted U-shape curve (rising then decreasing) for the investment cash flow sensitivity as a function of the extent of financial constraints. When firms are facing financial distress and a serious threat of bankruptcy, then the investment cash-flow sensitivity will fall. A likely explanation is that the firm probably cuts investment by much more than the fall in its cash flow, because it needs to use this cash for debt repayments. Kaplan and Zingales’s critique can be viewed as a useful extension of Fazzari, Hubbard and Petersen [1988], adding a new financial regime labelled ”financial distress”, measured by a specific sample split indicator. They emphasize that it is difficult to get good sample selection criteria in practice because of the diversity of financial constraints regimes (see also Kaplan and Zingales [2000], Fazzari Hubbard and Petersen [1996 and 2000]).

There is an important econometric problem confronting these sample separation criteria. Most of the ones used in the recent literature are likely to be endogenous. This would imply a specific econometric treatment for this endogenous selection problem. For example, the sample separation criteria can be explicitly estimated using Probit or Tobit estimation. Then switching regression techniques can be used (Hu and Schiantarelli [1998]). Besides the general endogenous selection problem (which may or may not cause significant bias in the estimated coefficients), further analysis
of sample separation criteria may be very useful to assess the first questions outlined in our introduction: what percentage of these estimated investment liquidity variables excess sensitivities is expected to be related to financial factors and why? What is the standard error on this percentage? To assess the quality of sample separation criteria as a signal of financial constraints, probit estimations can also provide this additional information. If other variables signalling financial fragility explains ”more” a sample separation criterion with respect to another sample separation criterion, and if both of them lead to investment liquidity excess sensitivity, then one is able to rank the quality of sample separation criteria as signalling more financial constraints than other characteristics. For example, size is a sample separation criterion which is intuitively perceived as a less precise signal of financial constraint than the fact that the firm is financed by specific institutions (for example, ”keiretsu” in Japan as in Hoshi, Kashyap and Scharstein [1991]) or the firm’s bankruptcy risk rating measured by scoring based on several financial ratios (Whited [1992], von Kalckreuth [2001] and Chatelain and Tiomo [2001]). These Probit explorations would complement intuition on the accuracy of sample selection criteria with respect to excess sensitivity tests of financial constraints.

3.2. Q Model and Financial Constraints

3.2.1. Average Q ratio capitalizes some of the financial constraint effects but not all of them.

The link between investment, marginal Q (ratio of a marginal change of the value of the firm divided by a marginal change of capital) and average Q has been stated by Hayashi [1982] under the assumption of convex costs of adjustment for investment. Marginal Q and average Q are equal for a competitive firm with a constant return to scale production provided that the adjustment cost function is linearly homogeneous in the rate of investment and the level of the capital stock. A widely estimated equation is the following:

\[ \frac{I_t}{K_{t-1}} = a + bQ_t + cLIQ_t + \varepsilon_t \]  (3.1)

where \( I \) is investment, \( K \) is capital, \( Q \) is Tobin’s average Q ratio (market value of the firm divided by the value of its stock of capital), \( LIQ \) is a liquidity variable, \( a, b \) and \( c \) are parameters and \( \varepsilon_t \) is an error term. When the liquidity variable is cash flow, the model is called the Q cash flow model of investment and \( c \) is the investment cash flow sensitivity. With respect to structural modelling, depending on circumstances, average Q capitalizes the impact of some or all finance constraints so that the investment liquidity sensitivity variable may reflect a specific financial constraint effect (Chirinko and Schaller [1995], Chirinko [1997]) or prove to be significant due to measurement errors in Q (Gomes [2001]). In particular, Chirinko and Schaller [1995] point out that the liquidity variable may be related to the hypothesis of a convex dependence of the
cost of external finance on the growth rate of debt (related to flotations costs) in much the same way as convex dependence of adjustment cost on the growth rate of capital affects the technology of the firm. A constraint such as the dependence of the cost of external finance on leverage or on the stock of debt (not on the growth rate of debt) is likely to be taken into account in the Q ratio (Chirinko [1997] and Gomes [2001]). Testing this specific hypothesis amounts to considering that the relevant ”investment financial constraint sensitivity” is \(\frac{c}{c + b}\) under certain conditions instead of the investment liquidity sensitivity \(c\). This leads Chirinko [1997] to reinterpret the effects of former studies. However, a competing structural interpretation of the Q cash flow model excluding financial constraints was proposed recently.

3.2.2. Imperfect Competition and Decreasing Returns to Scale May Blur Financial Constraints Signals

Recent theoretical papers departed from Hayashi’s (1982) set of hypothesis and considered firms facing imperfect competition and/or decreasing returns to scale to challenge the interpretation of the investment cash flow sensitivity as a signal of financial constraints. They focus on the applied econometric model where investment is estimated as a function of average Q and of cash flow. It turns out that decreasing returns to scale for a firm facing uncertainty imply that the investment ratio is a reduced form linear function of average Tobin’s Q and of the ratio of cash flow over capital, even without financial constraints. These demonstrations complement the initial Lucas’ critique that the parameter related to cash flow in the Q cash flow reduced form model of investment is not derived from a structural model of financial constraints.\(^2\) Abel and Eberly [2002] remove the assumption of adjustment costs and do not consider financial constraints, but assume uncertainty. To introduce time series variations of the investment ratio, they had to introduce uncertainty on the growth rate of the productivity factor, which implies uncertainty on average Q. To introduce time series variation of the cash flow/capital ratio (cash flow is the revenue function before paying the cost of capital), that is variation of average productivity, itself related to marginal productivity, they assumed uncertainty on depreciation, which implies uncertainty on the marginal cost of capital (the marginal product of capital is equal to the marginal cost of capital). The existence of decreasing return to scale and/or monopoly power in an uncertain framework leads to a non-linear dependence of the expected investment ratio with respect to average Q and to cash-flow. Both linearised coefficients are a function of the cash flow/capital ratio (average productivity). Under some uncertainty parameter conditions, firms with high average productivity (combined with a high cost of capital, hence small firms with respect to capital, as they face decreasing returns to scale or high market power) exhibit a higher investment

\(^2\)Note that these demonstrations do not invalidate the existence of a cash flow channel of monetary policy. But they found its source in uncertainty, imperfect competition and/or decreasing returns to scale instead of financial constraints.
cash flow sensitivity and a lower investment Q sensitivity than other firms. Hence, sample separation criteria such as average productivity, the cost of capital, the size of capital or market power indicators lead to excess sensitivity of investment to cash flow.\footnote{Their model does not take into account an industry equilibrium of firms with large or small capital as in Gomes [2001] but deals with a monopoly in partial equilibrium.} Finally, in that context, the uncertain growth rate of productivity is exogenous and is not necessarily connected to size. Abel and Eberly also show that fast-growing firms exhibit a higher investment cash-flow sensitivity and a lower investment Q sensitivity than other firms. As a consequence, uncertainty and decreasing returns to scale are able to generate excess investment cash flow sensitivity when growth is used as a sample separation criterion without assuming financial constraints.

However, if the sample separation criteria are financial variables such as dividend payouts, leverage, interest coverage, bond ratings (Whited [1992], Hubbard, Kashyap and Whited [1995], Gilchrist and Himmelberg [1995]) or the composition of external finance (Kashyap, Stein and Wilcox [1993]), or the Bank affiliation (van Ees et al. [1998]) or 'keiretsu' affiliation in Japan (Hoshi, Kashyap and Scharstein [1991]) which are unrelated to the real characteristics of the firm (size, expected growth rate), then evidence using these variables to identify financial constraints is not subject to the confounding of financial effects and firm real characteristics in the Q-cash flow model of investment. What is more, a higher cost of capital may be due not so much to a high depreciation rate, leading to a high investment ratio, but rather to a risk premium related to financial constraints included in the financial cost of capital. If we add some of the four major financial constraints in an Abel and Eberly setting, uncertainty and decreasing returns are likely to sharpen the effects of financial constraints on investment with respect to the certainty model with constant returns to scale.

Cooper and Ejarque [2001] produce similar findings using another research avenue. Using simulations of a model with structural parameters and indirect inference based on the method of moments, they are able to replicate the reduced form results of a Q cash-flow model of investment estimated on a panel data set by Gilchrist and Himmelberg [1995]. They found that firms with identical market power or identical (decreasing) returns to scale but with a higher adjustment cost parameter (using quadratic adjustment costs) and a higher autocorrelation of productivity shocks exhibit a higher investment cash-flow sensitivity and a higher investment Q ratio sensitivity. They indirectly infer that small firms are likely to exhibit a higher adjustment costs parameter and higher persistence of productivity shocks.

This result is slightly different from Abel and Eberly [2002], where an increase of the investment cash-flow sensitivity goes hand in hand with a decrease of investment Q ratio sensitivity. Both approaches indicate that the investment Q ratio sensitivity is no longer exactly the inverse of the quadratic adjustment cost parameter, so that the low values of this sensitivity found in applied studies do not necessarily reflect extremely high adjustment costs in the context of decreasing returns to scale. With respect to financial constraints, the observed persistence of productivity shocks may
not necessarily be a purely technological phenomenon. Financial constraints may create persistence in output due to internal equity accumulation or due to the persistence of informational asymmetry characteristics. They are likely to decrease steadily over time when outside investors learn more about the firm.

3.2.3. Measurement errors and capital markets short run valuation errors

Several papers investigate how measurement errors in marginal Q lead to significant cash flow effect. Gomes [2001] explores measurement errors on the price of capital goods (often taken at a sectoral level and not at the individual level), which values the stock of capital at the denominator of the Q ratio as well as measurement error in average Q. Simulating his industry model by removing financial constraints, he finds that significant cash-flow effects can be found due to measurement error. Moving on real data and a moment estimator robust to measurement errors, Erickson and Whited [2000] empirically confirm the findings that Q can explain investment once measurement errors are taken into account.

Plots of time series of variations of Q with respect to variations of investment may show that the changes of Q are often much larger than the changes of investment, so that the investment Q ratio sensitivity has to be small. Depending on circumstances, it seems that the capital market may present short-run valuation errors. In Cummins, Hassett and Oliner [1999], the response of investment rates to variation in average Q are quite small and cash flow is a significant regressor. However, when they replace average Q with their measure of Q based upon earnings expectations, financial variables are no long significant. Bond and Cummins [2001] produce a similar finding. They consider to what extent the empirical failings of the Q model of investment can be attributed to the use of share prices to measure average Q. They show that the usual empirical formulation may fail to identify the Q model when stock market valuation deviates from the present value of expected net distributions in ways that are consistent with weak and semi-strong forms of the Efficient Markets Hypothesis. They show that the structural parameters of the Q model can still be identified in this case using a direct estimate of the firm’s fundamental value, and implement this using data on securities analysts’ earnings forecasts for a large sample of publicly traded US firms in the 1990s. Their empirical results suggest that stock market valuations deviate significantly from fundamental values. Controlling for this, they find no evidence that the Q model of investment is seriously mis-specified.

The recent renewal of interest in Q model should not disguise the fact that more and more data set are available for firms which are not traded on the equity market. For those firms, which are more likely to be financially constrained, the market value is not available so that the traditional estimation of the market value Q model simply cannot be run. Applied econometrician would have to find analysts’ earnings forecast for these non traded firms or build their own market value indicator. Firms can be valued with the help of a few retrospective balance sheets and income statements.
However, this requires assumptions regarding demand forecasts for the firm. But the valuation proposed by analysts can be approximated in order to extend Q testing to non-traded firms.

Due to difficulties in tracking financial constraints precisely in the Q-Cash flow model, the Euler equation was devised as a new method of structural estimation in the early 1990s (Gertler, Hubbard and Kashyap [1991], Hubbard and Kashyap [1992]). Let us now investigate why applied econometricians are not so enthusiastic about this alternative approach.

### 3.3. Euler Equations and Financial Constraints

On one hand, Q has the virtue to capture some, if not all, of the profit expectations. On the other hand, as seen in the preceding section, the Q and Cash Flow model of investment faces difficulties due to non-linear or non-structural parameter in the estimated reduced form where both sensitivities may or may not depend on the extent of financial constraints, measurement errors, short run valuation error on the equity market, and the lack of data for the value of non traded firms at the microeconomic level. Use of the Euler equation has been seen as an alternative to structural tests of investment which avoids several of the problems faced by the Q model. It estimates the equality between the marginal product of capital and the cost of capital including marginal adjustment costs of investment now and marginal costs of investing next period. The marginal productivity of capital is usually computed under the assumptions of homogeneity of capital and labour and of imperfect competition or decreasing return to scale.

This marginal condition has been estimated taking into account the financial constraint designed as an increasing cost of debt as leverage increases (e.g. Bond and Meghir [1994], Estrada and Valles [1998], Jaramillo, Schiantarelli and Weiss [1996], van Ees et al. [1998], Gilchrist and Himmelberg [1998], Chatelain and Teurlai [2003]). For example, Kaplan and Zingales [1997] proposed a similar theoretical model without adjustment costs as an alternative to the Q cash-flow reduced form model, so that these Euler equation estimations avoids some of their criticisms. The parameter related to the cost of debt can be directly estimated. It provides a measure of the financial constraints related to leverage. It has the same status than a structural parameter.

However, if Euler equation estimations were relatively successful to integrating leverage into investment equation, they prove disappointing with respect to the role of cash-flow. Cash flow are often introduced in Euler equations which take account of credit rationing:

\[
L (LIQ_t) F_K (K_t, L_t) = L (LIQ_t) \left[ UC_t + J_K \left( \frac{I_t}{K_{t-1}}, \frac{I_{t-1}}{K_{t-2}} \right) \right]
\]

(3.2)

The Lagrange multiplier related \( L \) to the credit rationing constraint is parameterised as a linear function of cash-flow and other liquidity variables denoted \( LIQ_t \).
But the new parameter related to cash flow is not a structural one. It is a combination of individual variables and of other parameters of the model, namely the adjustment cost parameter and the mark-up parameter (Chatelain [2000]). On the other hand, omitting cash flow in the Lagrange multiplier regularly leads to a mis-specification of the Euler equation (Whited [1998], Chatelain and Teurlai [2003]). The need to alter the specification of cash flow in the neo-classical Euler equation without financial constraints may be related to the fact that this equation constrains the investment ratio to be a negative function of EBITDA, a feature that the Lagrange multiplier add-on corrects (this problem was noticed by Bond and Meghir [1994]). If EBITDA was to be correlated with future cash-flow, then the Q result would favour an opposite sign (positive) for the relationship between the investment ratio and cash-flow instead of the one found in the neo-classical Euler equation.

Another major problem with Euler equations is related to the quadratic adjustment costs assumption. Estimates of the adjustment cost parameter are sometimes very small and not significant (whereas estimates of the adjustment cost parameter are usually too large in the Q model, assuming Hayashi’s [1982] conditions). An example using Belgium data is given by Barran and Peeters [1998]. As the production function is parameterised assuming homogeneity with respect to capital and labour, the marginal productivity of capital is replaced by a function of average productivity where capital appears only at the denominator. Investment then appears only in marginal adjustment costs. If the marginal adjustment cost is not significant, investment behaviour is no longer explained by the Euler equation. There have been proposals for removing the assumption of quadratic adjustment cost to a polynomial specification (Whited [1998], Chatelain and Teurlai [2003]) or to another specification which allows a higher number of lags of the investment ratio (Gerard and Verschueren [2002]) or by assuming non-convex costs of adjustment (Cooper and Haltiwanger [1999])

It turns out that those problems have maintained alive and well the alternative traditional auto-regressive distributed lags (ADL) specification of the neo-classical Jorgensonian model omitting proxies for the cost of capital (Hall, Mairesse and Mulkay [2000]) or taking into account the cost of capital (Chirinko, Fazzari and Meyer [1999]), because of the flexibility of the distributed lag structure. What is more, the effect of the Jorgenson’s cost of capital on capital is not easy to isolate properly in the Euler equation, because an homogeneous production function does not necessarily exhibit a constant elasticity of substitution (hence a constant elasticity of the stock of capital with respect to its cost) and because marginal adjustment costs, which are a function of the investment ratio, are added besides the user cost of capital. On the other hand, the ADL structure allows attempts to isolate the cost of capital channel of monetary policy from the broad credit channel or the cash flow channel. This specification has been chosen in the monetary transmission network of the European System of Central Banks (for example, Chatelain et al. [2003]). As seen above, structural models of investment (Q model and Euler equation) stop half way in the structural modelling of financial constraints. The interpretation of the parameters of cash flow or of other
liquidity variables faces the Lucas critique in these models as much as in the traditional auto-regressive distributed lags models.

3.4. Private Accounting Data

The fact that these data are not aggregated should not conceal the fact that accounting data are also fragile. Accountants can choose optimistic or pessimistic valuations for several items related to future cash inflows or cash outflows. However, it is possible to improve our use of accounting data for understanding investment. Off balance sheet information such as leasing and discount could be taken into account in some countries.

Discount is a short-run means of obtaining cash from banks by firms where they give property rights to some of their trade receivables to banks. As a counterpart, the banks provide the term with liquidity on which interest is charged. For financially distressed firms, this can be a mean of finance which is accepted by banks because they have the collateral in the form of the claims on other firms. By taking into account off balance sheet discount, we increase the amount of trade credit on the asset side and increase the amount of liquid debt on the liabilities side. Trade credit (trade payables less trade receivables) can be a complement to external finance (Ramey [1998]).

A similar collateral argument holds for leasing. Leased capital remains the property of the "lender" even when the firm goes bankrupt. The lender receives rents from the firm. The collateral consists of the leased capital itself, which mostly incurs the risk of being depreciated at accelerated rate. In this context, leasing may be an opportunity for some firms which are financially constrained. However, while it is easy to add leased capital on the gross asset side, there are several ways of splitting rental contract flows between depreciation flows and debt flows, as to whether this capital good had been bought instead of rented. On the one hand, one can find the equivalent interest rate for a debt contract (International Accounting Standards Committee recommendation) and consider that the remaining value of the good is depreciated, but the depreciation scheme does not fit any standard accounting depreciation rule. On the other hand, one can apply a standard depreciation accounting rule and consider that the remaining flows are related to a debt contract. In this case, the interest payments do not fit existing debt contracts (this rule is used in Chatelain and Teurlai [2003], see also De Bodt et al. [1996] and Sharpe and Nguyen [1995]).

3.5. Asset Prices

Trying to isolate the cost of capital channel of monetary policy from the broad credit channel or the cash flow channel has been one of the goals of the firm studies in the monetary transmission network of the European System of Central Banks [2003]. The credit channel emphasises that credit availability is a function of collateral. The cyclical movement of asset prices may modify the value of collateral. It is then useful to investigate the asset price channel as a particular channel of the broad credit channel.
Before going further, it is worth mentioning that monetary policy faces more difficulties in influencing asset prices than the consumer price index. Asset price expectations driving speculative bubbles are not systematically affected by monetary policy shocks or central bankers’ declarations. Unfortunately, when speculative bubbles burst, central banks sometimes have to provide liquidity for some financial intermediaries. Then, after over-investment, the decline in asset prices may be long lasting.

A related problem concerns how to adapt asset prices indexes, where they exist, to microeconomic data. Asset prices matter on the asset side in the case of buildings, for example, which are partly taken into account by the sectoral investment price index for computing real investment. Asset prices also matter for the valuation of the financial assets owned by the firm and for the valuation of the firm’s equity liabilities. As financial assets are valued at acquisition cost, one may revalue them by using the equity price index. In principle, such revaluation may be anticipated by accountants and put on the liability side under provisions, if the market value of the assets is decreasing. Then, the share of financial assets may reflect the fact that a firm is more sensitive to asset price fluctuations. It may also reflect the fact that a firm can obtain finance from a group. It is also possible to revalue equity by taking account of changes in asset prices. Finally, it leads to investigation of the trade-off between financial investment and real investment, taking into account the risks of asset price fluctuations in the case of financial investment.


Once some of the four major financial constraints are taken into account, specific intertemporal dynamics emerge as accumulation into equity of retained earnings, obtained from profits. These dynamics modify the reactions of firms to shocks with respect to the standard convex adjustment cost models. Without asset price fluctuations, the accumulation of internal funds of credit rationed firms smooths capital accumulation after a rise of productivity (or a fall in the cost of capital) with respect to the neo-classical model with convex adjustment costs (Chatelain [1998]). As a consequence, additional features are usually taken into account for explaining a financial accelerator, where financial constraints are supposed to amplify productivity or demand shocks and exacerbate fluctuations. These additional features may be asset prices fluctuations, the removal of the hypothesis of adjustment costs or of its convexity, shifts in the allocation of savings towards private productive investment along the business cycle, and so on. These features are more developed in recent business cycles theory based on models of intertemporal investment facing financial constraints. Hence, econometric evidence of financial constraints on investment supports some of the four simplified assumptions of financial constraints affecting investment behaviour used in those models.

In the recent years, the accumulation of internal funds has been investigated in
business cycle theory. This economic literature is highly indebted to the ideas developed in details by Irving Fisher's [1933] seminal article on the debt/deflation episode of the 1930s. Heterogeneity of the current year equity or current year debt/equity ratio leads to different accumulation path for firms in the certainty case (Kiyotaki and Moore [1997]). This heterogeneity can be increased taking into account uncertainty on profits next period due to productivity or demand shocks, which alters the amount of retained earnings, and hence equity, in the next period. Research started with constant returns to scale technology. This assumption allows capital aggregation so that the distribution of equity and of capital is not necessary for studying the cycle (Kiyotaki and Moore [1997], Bernanke, Gertler and Gilchrist [1998], Carlstrom and Fuerst [1997], Aghion, Banerjee and Piketty [1998]). In Kiyotaki and Moore [1997], asset price fluctuations are endogenous but there is no focus on monetary policy. In Bernanke, Gertler and Gilchrist [1998], asset price fluctuations or bubbles are not endogenous. The cycle is driven by exogenous productivity shocks. But they consider a monetary policy rule, which alters the economic cycle.

Further research considered decreasing returns to scale technology, once one removes the fixed costs in some models (Gomes [2001], Cooley and Quadrini [2001], Cooley and Quadrini [1999], Barlevy, G. [1998], Caballero, R. and M. Hammour [1998], Den Haan, W. G. Ramey and J. Watson (1999)). Those recent papers have examined the general equilibrium compositional effects of shocks if financing constraints are present. It is necessary to know the distribution of equity (the firm's net worth) to compute or simulate the aggregate dynamics of capital. In those models, the heterogeneity of equity is fully taken into account but the investigation of the heterogeneity of other characteristics of financial constraints is not yet so well developed. An attempt to deal with the consequences of the heterogeneity of financial constraints for monetary policy has been proposed recently by Bean, Larsen and Nikolov [2001]. A key question is how to adapt monetary policy rules to the heterogeneity of firms reactions to monetary policy shocks.

Finally, the accumulation of retained earnings in the equity dynamics can also affects long-term growth, as well as the business cycle. This can happen under three conditions: if equity growth limits the tangible and intangible capital growth of individual firms, if the share of firms facing this growth limit is large in the economy, and if the growth of intangible capital (for example research and development) and of intangible capital is one of the driving forces of the economic growth of GDP (see Romer [1986], and the subsequent literature assuming constant returns to scale for the aggregate of all types of capital in the economy). A recent theoretical model is proposed by Amable, Chatelain and Ralf [2002] and recent empirical evidence on the relationship between financial structures and growth can be found in Demirguc-Kunt and Levine [2001].
5. Conclusion

This paper reviewed problems related to the estimation of investment facing financial constraints. A central question is the interpretation in applied work of the sensitivities of investment with respect to cash flow and other liquidity variables, which still continues to face the Lucas critique. It is not easy to isolate the component of these sensitivities related to financial constraints from components related to other features of investment, such as technology, imperfect competition, expectations, and so on. Much progress has been made in using new data sets and obtaining more and more useful information out of them. Extensive use has been made of sophisticated panel data econometric estimates dealing with endogeneity and endogenous selection problems. On the one hand, theory has raised more and more precise objections against direct interpretations of these sensitivities as measures of financial constraints. On the other hand, theory has produced new results on financial constraints, capable, for example, of handling the heterogeneity of firms. However, there are still no answers to the Lucas critique. One of the reasons is that some characteristics of investment other than financial constraints also present testable difficulties (for example, adjustment costs, see Hamermesh and Pfann [1996]). Another reason may be the diversity of financial regimes faced by firms.

References


