Hall, Bronwyn H.

Article
The financing of innovative firms

EIB Papers

Provided in Cooperation with:
European Investment Bank (EIB), Luxembourg

Suggested Citation: Hall, Bronwyn H. (2009) : The financing of innovative firms, EIB Papers, ISSN 0257-7755, European Investment Bank (EIB), Luxembourg, Vol. 14, Iss. 2, pp. 8-28

This Version is available at:
http://hdl.handle.net/10419/44902

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
ABSTRACT

To what extent are new and/or innovative firms fundamentally different from established firms, and therefore require a different form of financing? The theoretical background for this proposition is presented, and the empirical evidence on its importance is reviewed. Owing to the intangible nature of their investment, asymmetric-information and moral-hazard, these firms are more likely to be financed by equity than debt and behave in some cases as though they are cash-constrained, especially if they are small. Recognising the role for public policy in this area, many countries have implemented specific policies to bring the cost of financing innovation more in line with the level that would prevail in the absence of market failures.

Bronwyn H. Hall (bhhall@econ.berkeley.edu) is Professor in the Graduate School, University of California at Berkeley and Professor of Technology and the Economy at the University of Maastricht. She is also a Research Associate of the National Bureau of Economic Research (NBER) in Cambridge (MA) and the Institute for Fiscal Studies, London.
The financing of innovative firms

1. Introduction

In a market-based economy, those who create and manage firms (entrepreneurs) are usually not the same individuals as those who have the means to finance this activity. This fact implies that there is likely to exist an information gap between those asking for funds and those supplying them. As will be discussed in more detail below, economists refer to the extra costs thus induced as arising from the problems of "asymmetric-information" and "moral-hazard". Both of these are expected to raise the costs of obtaining finance from sources external to the firm. And although this will be true to some extent for all firms, the problem is particularly salient in the case of new firms and firms undertaking innovative activities.

Three different levels of difficulty need to be distinguished: the problems of existing innovating firms in acquiring sufficient funds for their investments, the reluctance of non-innovators to undertake innovation due to its high cost, and the problems faced by new start-up firms. Although all of these difficulties arise from the same ultimate set of causes, the empirical analysis of each will differ substantially and the possible range of policy solutions will differ. In particular, the first case, on which most of the econometric literature has focused, is subject to marginal analysis, whereas the second two cases involve the overcoming of (often substantial) fixed costs of entry into innovation.

This paper reviews the theoretical and empirical economic literature on the financing of innovation and draws some conclusions with respect to policy designed to ameliorate some of the problems. Section 2 of the paper discusses the characteristics that make innovation investment different from ordinary investment. Then the theories of asymmetric-information and moral-hazard as applied to financing innovation are reviewed, followed by a summary of the empirical evidence on this topic. The concluding section discusses implications for public policy.

2. Why is investment in innovation different?

Investment in innovation usually consists of Research and Development spending (R&D), design and marketing expenses for bringing a new product to market, investment in the necessary new capital equipment, and investment in training. The relative importance of each of this varies with the industry and type of innovation, although the most important spending in most sectors is R&D, accounting for more than 50 percent of innovation expenditures.

From the perspective of investment theory, innovation investments have a number of characteristics that make them different from ordinary investment. First and most importantly, most of the expenditures, with the exception of those on new capital equipment, consist of worker wages and salaries. From considerable survey evidence over the past 50 years, we know that in practice 50 percent or more of the R&D portion of this investment goes toward paying scientists and engineers, who are usually highly...

1 The paper draws from – and updates – an earlier survey (Hall 2002).
2 For example, the European Community Innovation Survey Wave 4 (CIS4) asks whether the firm has the following types of innovation expenses: 1) internal R&D; 2) external R&D; 3) acquisition of new capital equipment or software; 4) acquisition of external knowledge or know-how including licenses, IP rights, etc.; 5) training; 6) marketing; and 7) other expenses associated with launching new products or processes, such as design and engineering.
Their efforts create an intangible asset, the firm’s knowledge base, from which profits in future years will be generated. Added to this knowledge base is the specific human capital created by worker training in new products and processes and the knowledge created by design and marketing investments. To the extent that all this knowledge is "tacit" rather than codified, it is embedded in the human capital of the firm’s employees, and is therefore lost if they leave or are fired.

This fact has an important implication for the conduct of R&D investment and to a lesser extent, innovation investment more broadly. The focus in the discussion below is on R&D as it is the measure on which empirical research has usually been based. Because part of the resource base of the firm itself disappears when knowledge workers leave or are fired, firms tend to smooth their R&D spending over time, in order to avoid having to lay them off. This implies that R&D spending at the firm level will behave as though it has high adjustment costs (confirmed empirically for the US by Hall et al. 1986; Lach and Schankerman 1988), with two consequences, one substantive and one that affects empirical work in this area. First, the equilibrium required rate of return to R&D may be quite high simply to cover the adjustment costs. Second, and related to the first, is that it will be difficult for empirical studies to measure the impact of changes in the costs of capital on such investment, because such effects can be weak in the short run due to the sluggish response of R&D to any changes in its cost.

The above conclusion needs to be tempered somewhat by the observation that in the recent past, the variance of R&D spending growth in publicly traded US firms has increased somewhat, for at least two reasons. The first is the increased importance of the ICT sector, where there is fairly rapid obsolescence of R&D outputs, and a consequent reduction in the incentives to safeguard human capital (Hall 2006). The second is that markets for technology have become somewhat more important, which reduces the need to keep firm’s entire knowledge in-house (Arora et al. 2001). Nonetheless, it remains true that the variance in R&D investment growth rates is about one quarter to one fifth that for ordinary investment.

A second important feature of R&D investment is the degree of uncertainty associated with its output. This uncertainty tends to be greatest at the beginning of a research program or project, which implies that an optimal R&D strategy has an options-like character and should not really be analyzed in a static framework. R&D projects with small probabilities of great success in the future may be worth continuing even if they do not pass an expected rate of return test. The uncertainty here can be extreme and not a simple matter of a well-specified distribution with a mean and variance. There is evidence such as that in Scherer (1998) that the distribution of profits from innovation sometimes takes the form of a Pareto distribution where the variance does not exist. When this is the case, standard risk-adjustment methods will not work well. Looked at from the perspective of standard finance theory, the variance of a portfolio constructed from such assets is unbounded so the usual diversification analysis does not apply. High uncertainty of returns to innovation has been exacerbated in the recent past by the rise of network, or “winner-take-all” industries, such as software or Web-based services. Examples such as Microsoft, Goggle, Yahoo, eBay, and now Facebook are familiar – these are characterized by very high returns to the initial investment, but there are a number of similar entrants who either fail or never reach critical mass and settle for a small niche of the market. In fact, the high returns experienced by a few firms do succeed in attracting a number of entrants, but there is no guarantee that it is the “right” number.

An important characteristic of uncertainty for the financing of investment in innovation is the fact that as investments are made over time, new information arrives, which reduces or changes the uncertainty.

---

3 The simplest Pareto distribution is a one-parameter probability distribution with the property that under certain values of the parameter the variance of the outcomes drawn according to such a distribution is infinitely large.
The consequence of this fact is that the decision to invest in any particular project is not a once and for all decision, but has to be reassessed throughout the life of the project. In addition to making such investment a real option, the sequence of decisions complicates the analysis by introducing dynamic elements into the interaction of the financier (either within or without the firm) and the innovator.

The final characteristic of R&D as investment that has implications for financing is that the “capital” thus created is intangible. Not only is it intangible, but as discussed above, much of it is in the form of human capital embedded in the heads of the employees. Such capital typically has relatively low salvage value because it is also idiosyncratic – for example, the fact that the firm owning the capital goes out of business is a signal that its value turned out to be low. Except for the type of effort now underway to harvest patents from such firms (e.g., Ocean Tomo or Intellectual Ventures), there is little market for distressed intangible assets. The human capital involved goes with the employee, and usually he or she will capture any residual value from that in the form of wages in future employment. Thus debt instruments that are secured by the value of the capital asset are not likely to provide a useful source of funding for R&D. See Harhoff (2009) for a description of recent efforts to change this situation via the creation of patent funds for the commercial exploitation of unused inventions.

Summarizing, R&D and other innovation investments have the following characteristics: (i) they need to be smoothed in order to retain valuable employees and their knowledge; (ii) they are highly uncertain and information about success or failure is revealed over time; and (iii) they create an idiosyncratic intangible capital with a limited resale market. The next section of the paper reviews the theory of investment in the light of these characteristics and discusses the implications of the theory for the financing of innovation.

3. Theoretical framework: Is finance different for innovative firms?

What follows focuses on the analysis of ongoing innovative firms rather than new start-ups. The greater part of empirical work has been performed using data on these kinds of firms, partly because they perform the great bulk of R&D and partly due to data availability. The start-up problem will be discussed later in the paper.

The usual starting point for the analysis of any type of investment financing is the “neo-classical” marginal profit condition, suitably modified to take the special features of R&D into account. This condition sets the marginal product of capital equal to the rate of return on investment in that capital. Ignoring adjustment costs and uncertainty for the moment, the discrete time cost-of-capital condition is the following:

\[ MPK = c = p \frac{(1-\delta)}{(1-\rho)} p_{t+1} \]

That is, the marginal product of capital (MPK) that is used during period \( t \) is equal to its price less what the firm would receive from selling the capital at \( t+1 \) (which is \( (1-\delta) p_{t+1} \)), discounted by the required rate of return \( \rho \). \( \rho \) is the return received by the investor, which is after corporate taxes are paid but before any personal or capital gains taxes are paid.

The above formulation implicitly contains three factors that will affect R&D financing: the rate at which the knowledge capital thus created depreciates or becomes obsolescent, the required rate of return, and rate of change of the real R&D price (the price of R&D inputs relative to the price of the firm’s output).

---

4 Firms with more than 250 employees account for 90 percent of R&D worldwide (OECD 2007).
It is obvious that higher depreciation rates increase the cost of capital, while lower post-corporate-tax required returns, as in the case of an R&D tax credit, would lower it. In addition, if R&D is expected to become relatively more expensive tomorrow \( (\rho_{t+1} > \rho_t) \), this lowers the cost of capital today but this effect is relatively small in practice.

The required rate of return in the equation above is the return after corporate taxes have been paid. However, as Auerbach (1984), among others, has shown, the marginal source of financing used by the firm will also impact the required rate of return that a prospective investor perceives. He analyzed the US case, where interest expense is deductible to the corporation, and the tax on capital gains has generally been lower than that on dividends. The same type of analysis would apply whenever profits from longer holding periods are taxed at a lower rate than those from shorter holding periods. The table below shows the cost of financing based on source of finance that he derived; in general, tax considerations suggest that debt finance will be cheapest, followed by retained earnings, and lastly by new share issues.

### Table 1. Tax-adjusted financing cost

<table>
<thead>
<tr>
<th>Source of finance</th>
<th>Cost of finance</th>
<th>Assumptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Debt</td>
<td>( \rho (1-\tau) )</td>
<td>Interest is deductible at the corporate level</td>
</tr>
<tr>
<td>Retained earnings</td>
<td>( \rho (1-\tau_p)/(1-\tau_c) )</td>
<td>Avoids personal tax on dividends in favour of eventual capital-gains tax</td>
</tr>
<tr>
<td>New share issues</td>
<td>( \rho/(1-\tau_c) )</td>
<td>Eventual capital-gains tax paid</td>
</tr>
</tbody>
</table>

\( \rho = \text{required return}; \tau = \text{corporate tax rate}; \tau_p = \text{personal tax rate}; \tau_c = \text{capital-gains tax rate} \)

Source: Auerbach (1984)

Obviously the story does not end here. Implicitly in constructing this table I have assumed that there is a single risk-adjusted rate of return available, but we have already seen that this is unlikely to be true, given the degree of uncertainty faced by innovating firms. That is, some innovation investments are so risky that a simple risk-adjustment based on the variance of returns is not available. In addition, different types of investors are likely to prefer different risk profiles. Bondholders will care more about salvage value (which also favours tangible over intangible assets), whereas equity holders may see an unbounded upside to the returns and may therefore prefer risk.

Two other major reasons for differences in the required rate of return across financing source have been the subject of considerable theoretical and empirical interest, on the part of both industrial organization and corporate finance economists: asymmetric-information and moral-hazard. These factors are widely viewed as driving a wedge between the cost of internal and external sources of finance, a wedge that is likely to be larger in the case of innovation investment than it is for ordinary investment.

One of the implications of the well-known Modigliani-Miller theorem (1958, 1961) was that a firm choosing the optimal levels of investment should be indifferent to its capital structure, and should face the same price for all types of investment (including investments in creating new products and processes) on the margin. The last dollar spent on each type of investment should yield the same
expected rate of return (after adjustment for non-diversifiable risk). A large literature, both theoretical and empirical, has questioned the bases for this theorem, but it remains a useful starting point.

Of course, there are a number of reasons why the theorem might fail to hold in practice. First, uncertain returns in combination with incomplete markets imply that insurance over all outcomes is not available. Second, the cost of capital may vary across the source of the funds either for non-tax reasons or for tax reasons. Finally, the cost of capital may also vary across type of investment (tangible and intangible) both for tax and for other reasons.

Summing up, with respect to innovation investment, economic theory advances a plethora of reasons why there might be a gap between the external and internal costs of capital; these can be divided into three main types, of which the first two arise from market failures: (i) asymmetric-information between innovator and investor; (ii) moral-hazard arising from the separation of ownership and management; and (iii) tax considerations that drive a wedge between external and internal finance. The following three sections treat each of these in turn.

3.1 Asymmetric-information problems

In the R&D setting, the asymmetric-information problem refers to the fact that an inventor or entrepreneur frequently has better information about the nature of the contemplated innovation project and the likelihood of its success than potential investors. Therefore the marketplace for financing the development of innovative ideas looks like the “lemons” market modelled by Akerlof (1970). In his model, the good (used) cars sells for a lower price in order to compensate the buyer for the possibility that the car is a lemon. In this setting, the seller of potential returns to R&D or innovation offers a higher return (lower price) to compensate the buyer for the possibility that the project is not as good as is claimed. The lemons’ premium for R&D or innovation will be higher than that for ordinary investment because investors have more difficulty distinguishing good projects from bad when the projects are long-term R&D investments than when they are short-term or low-risk projects (Leland and Pyle 1977).

In the most extreme version of the lemons model, the market for R&D projects may disappear entirely if the asymmetric-information problem is too great. Informal evidence suggests that some potential innovators believe this to be the case in practice. Reducing information asymmetry via fuller disclosure is of limited effectiveness in this arena, due to the ease of imitation of inventive ideas. Firms are reluctant to reveal their innovative ideas to the marketplace and the fact that there could be a substantial cost to revealing information to their competitors reduces the quality of the signal they can make about a potential project (Bhattacharya and Ritter, 1983; Anton and Yao, 1998). Thus the implication of asymmetric-information coupled with the costliness of mitigating the problem is that firms and inventors will face a higher cost of external than internal capital for R&D due to the lemons’ premium. When the level of R&D expenditure is an observable signal subject to external audit, as it is under current accounting rules in several countries, we might expect that the lemons’ problem is somewhat mitigated, but certainly not eliminated.

Asymmetric-information problems can sometimes be mitigated by reputations developed through repeated interactions and this setting is no exception. There are several forms of reputation-building observed. One of the important roles played by specialized venture capital (VC) funds is precisely to supply informed monitoring of early stage technology start-ups, but experienced venture capitalists will also have developed a reputation for honouring nondisclosure agreements that will enable them to gather better information about projects being proposed. On the other side of the transaction, serial entrepreneurs often face less difficulty in obtaining financing for new ventures, presumably because they have developed a reputation in prior start-ups.
3.2 Moral-hazard problems

Moral-hazard in R&D investing arises in the usual way: Modern industrial firms normally have separation of ownership and management. This leads to a principal-agent problem when the goals of the two conflict, which can result in investment strategies that are not share value maximizing, and hence can lead to under or over-investment. Two possible scenarios may co-exist: one is the usual tendency of managers to spend on activities that benefit them but not necessarily their firm (growing the firm beyond efficient scale, nicer offices, etc.) and the second is a reluctance of risk-averse managers to invest in uncertain R&D projects.

Agency costs of the first type may be avoided by reducing the amount of free cash flow available to the managers by leveraging the firm, but this in turn forces them to use the higher-cost external funds to finance R&D (Jensen and Meckling, 1976). Empirically, there seem to be limits to the use of the leveraging strategy in R&D-intensive sectors. See Hall (1990, 1994) for evidence that the leveraged-buyout (LBO) and restructuring wave of the 1980s, viewed by most researchers as driven by the need to reduce free cash flow in sectors where investment opportunities were poor, was almost entirely confined to industries and firms where R&D was of no consequence. One reason for this fact may be that over-investment in R&D and innovation is not usually a major problem for ongoing managerial firms and therefore the discipline of cash-flow limiting leverage is unnecessary for restraining these kinds of investment.

According to the second type of principal-agent conflict, managers are more risk averse than shareholders and avoid innovation projects that will increase the riskiness of the firm. If bankruptcy is a possibility, managers whose next best opportunity is a job with lower compensation than their present job and potential bondholders may both wish to avoid variance-increasing projects that shareholders would like to undertake. The argument of the theory is that long-term investments can suffer in this case. The optimal solution to this type of agency cost would be to increase the long-term incentives faced by the manager rather than reducing free cash flow. Many innovative firms make heavy use of stock option compensation for this reason, although this solution comes with its own incentive problems.

In the case of start-up firms, there is often a third type of principal-agent conflict, involving overconfidence on the part of the entrepreneur. When there is uncertainty and the probability of innovation success is revealed only gradually over time, the possibility of asymmetric-information and moral-hazard in the investor-innovator relationship creates further problems for achieving the optimal contract. For example, it is often observed that entrepreneurs or R&D managers wish to continue projects that investors would like to terminate (Cornelli and Yoshia 2003), presumably because the possibility of an ultimate benefit to the entrepreneurs looms large and they do not face the investment cost in the case of failure. If they are also over-confident (as is often the case), they will be even more biased toward continuation. Asymmetric-information about the project will imply that the investor has relatively more difficulty than the innovator even in determining the probability of success. The combination of information rents and agency costs will lead to inefficient funding of projects over time (waiting too long to cancel them or cancelling them too soon) as well as to inefficiently low levels of funding.

In a recent paper, Bergemann and Hege (2005) have analyzed these tradeoffs in a multi-stage investment financing decision under changing uncertainty, with renegotiation allowed. They look at the choice between relationship financing (where the investor is able to monitor the progress of the project accurately) and arm’s length financing (where the investor must rely on the innovator for information). Investors are able to speed up or slow down the rate of financing, depending on the progress of the project and their expectations of success. In general, Bergemann and Hege find that agency costs will...
lead to non-optimal stopping rules for projects, stopping them too soon on average. Surprisingly, arm’s length contracts can lead to higher project values, because in these the investor can commit in advance to a stopping rule, which eliminates any benefit to the entrepreneur from attempts to prolong the project.

3.3 Taxes and the source of funds

As discussed earlier, tax considerations that yield variations in the cost of capital across source of finance have been articulated by Auerbach (1984) among others. In many tax systems including that in the US, interest is deductible against earnings and the proceeds from equity held for longer periods is taxed at a lower rate. Therefore the cost of financing new investment by debt has been effectively lower than that of financing it by retained earnings, which is in turn less than that of issuing new shares. However, based on the discussion above, this ranking of the financing costs is unlikely to hold for the intangible assets created by innovation investment. Low salvage values relative to the original investment makes these assets unsuitable for debt finance in spite of the tax advantage, so that firms whose investments are mostly intangible will rely more heavily on retained earnings and equity.

Differential tax treatment not only obtains for the sources of funds but also for their use. The specific tax treatment of R&D and other innovation investments in most OECD economies is very different from that of other kinds of investment. Because R&D, marketing costs, training costs, etc. are expensed as they are incurred, the effective tax rate on the corresponding assets is lower than that on either plant or equipment, with or without various tax credits in place. The economic depreciation of innovation assets is considerably less than the depreciation allowed for tax purposes - which is 100 percent - so that the required pre-tax rate of return for such investment would be lower relative to ordinary investment. In addition, some countries offer a tax credit or subsidy to R&D spending and occasionally to other activities such as employee training, which can reduce the after tax cost of capital even further.

Thus the conclusions from the theoretical analysis of firm-level investment in innovation is that the presence of either asymmetric-information or a principal-agent conflict or both implies that new debt or equity finance from sources external to the firm will be relatively more expensive for R&D and innovation than for ordinary investment, and that considerations such as the lack of collateral further reduce the possibility of debt finance. Together, these arguments suggest an important role for retained earnings in the R&D investment decision. In addition, retained earnings may serve as a signal of future profitability. Hall (1992) and Himmelberg and Petersen (1994) have argued that positive cash flow may be more important for R&D than for ordinary investment, a proposition that has since received empirical support (see Section 4).

A final important conclusion from the theory above is that whatever problems are created for the financing of innovative firms by asymmetric-information and moral-hazard are likely to be exacerbated in the case of young firms and start-ups. These firms are often the most reliant on external sources of finance and of course they have not yet developed reputations that would allow them to signal their quality adequately. Evidence that this is true in practice as well as theory is provided in Section 4.5.

5 A detailed discussion of tax regimes in different countries is beyond the scope of this survey, but it is quite a common feature in several countries for long term capital gains on funds that remain with a firm for more than one year to be taxed at a lower rate than ordinary income. Yet even if the tax rates on the two kinds of income are equal, the inequalities will hold. Only in the case where dividends are not taxed at the corporate level (which was formerly the case in the UK) will the ranking given above not hold.

4. Empirical evidence

Almost all of the empirical evidence for the presence of a wedge between internal and external finance, and the consequent underinvestment in innovation that is discussed in this section uses R&D investment as a proxy for long term investment or investment in innovative activities. There are good reasons for this – R&D is frequently the only measure of innovation that has been observed over long time periods at the firm level and it is highly related to innovative activities, at least in the manufacturing sector. The recent wave of innovation surveys in Europe and elsewhere is beginning to provide us with alternative measures, although thus far it appears that R&D still has greater predictive power for firm performance than the newer measures, at least where it is observed (Mairesse et al. 2006). Later in this section some results obtained using the innovation surveys will be presented.

In the discussion below, the empirical studies in this area are divided into two groups: those based on variations of an investment equation derived from economic theory (Sections 4.1 and 4.2), and other studies – mostly financial-market studies based on stock-price announcement effects on the impact of various corporate-governance settings on the financing of innovation (Section 4.3). The section concludes with a discussion of the evidence on the resulting capital structure of innovating firms.

4.1 Investment-equation studies

For a firm already operating, the presence of “liquidity” constraints on their investment (ordinary as well as innovation and R&D investment) implies that they are unable to obtain funds from external sources to finance all the investment that they would undertake if they had sufficient funds available within the firm. Thus the usual way to examine the empirical relevance of the arguments in the previous section is to estimate R&D investment equations and test whether liquidity constraints or excess sensitivity to cash-flow shocks are present and perhaps more pronounced than for ordinary investment. This approach builds on the extensive literature developed for testing ordinary investment equations for liquidity constraints (Fazzari et al. 1988; Arellano and Bond 1991). It suffers from many of the same difficulties as the estimates in the investment literature (lack of good instruments to isolate supply shocks, measurement error, unobserved differences across firms, and so forth), plus one additional problem that arises from the tendency of firms to smooth R&D spending over time.

The ideal experiment for identifying the effects of liquidity constraints on investment is to give firms additional cash exogenously, and observe whether they pass it on to shareholders or use it for investment and/or R&D. If they choose the first alternative, either the cost of capital to the firm has not fallen, or it has fallen but they still have no good investment opportunities. If they choose the second, then the firm must have had some unexploited investment opportunities that were not profitable using more costly external finance. A finding that investment is sensitive to cash flow shocks that are not signals of future demand increases would allow rejecting the hypothesis that the cost of external funds is the same as the cost of internal funds. However, lack of true experiments of this kind forces researchers to use econometric techniques such as instrumental variables to attempt to control for demand shocks when estimating the investment demand equation, with varying degrees of success.

The methodology for the identification of R&D investment equations is based on a simple supply and demand heuristic, as shown in Figure 1. The curve sloping downward to the right represents the demand for R&D investment funds and the curves sloping upward the supply of funds. Internal funds are available at a constant cost of capital until they are exhausted, at which point it becomes necessary to issue debt or equity in order to finance more investment. When the demand curve cuts the supply curve in the horizontal portion, a shock that increases cash flow (and shifts supply outward) has no effect on the level of investment.
However, if the demand curve cuts the supply curve where it is upward sloping, it is possible for a shock to cash flow to shift the supply curve out in such a way as to induce a substantial increase in R&D investment. Figure 2 illustrates such a case, where the firm shifts from point A to point B in response to a cash flow shock that does not shift the demand curve.

During the past several years, various versions of these methodologies have been applied to data on the R&D investment of US, U.K., French, German, Irish, and Japanese firms and possibly others. The firms examined are typically the largest and most important manufacturing firms in their economy.
For example, Hall (1992) finds a large positive elasticity between R&D and cash flow, using an accelerator-type model and a very large sample of US manufacturing firms. The estimation methodology here controls for both firm effects and simultaneity. Similarly and using some of the same data, Himmelberg and Petersen (1994) look at a panel of 179 US small firms in high-tech industries and find an economically large and statistically significant relationship between R&D investment and internal finance.

More recently, J. Brown et al. (2009) have shown that both cash flow and the issuance of public equity are very important for younger US firms during the period 1990-2004, while they have little impact on mature firms’ R&D investment. They focus on the high-technology sector (drugs, office and computing equipment, communications equipment, electronic components, scientific instruments, medical instruments, and software), which accounts for almost all of the increase in R&D during this period, and use Euler-equation methods with fixed firm effects and industry-level year dummies to remove most of the variation due to unobserved differences in firm characteristics and demand shocks across industry. A novel finding in this paper and a companion paper by J. Brown and Petersen (2009) is the increased importance of public-equity issuance in financing R&D in the United States, which doubtless reflects a shift in expectations on the part of investors during this period that lowered the cost of this kind of capital to the firm.

Harhoff (1998) find weak but significant cash-flow effects on R&D for both small and large German firms, although Euler-equation estimates for R&D investment were uninformative due to the smoothness of R&D and the small sample size. Combining limited survey evidence with his regression results, he concludes that R&D investment in small German firms may be constrained by the availability of finance. Bond et al. (1999) find significant differences between the cash flow impacts on R&D and investment for large manufacturing firms in the United Kingdom and Germany. German firms in their sample are insensitive to cash flow shocks, whereas the investment of non-R&D-doing UK firms does respond. Cash flow helps to predict whether a UK firm does R&D, but not the level of that R&D. They interpret their findings to mean that financial constraints are important for British firms, but that those which do R&D are a self-selected group that face fewer constraints. This is consistent with the view that the desire of firms to smooth R&D over time combines with the relatively high cost of financing it to reduce R&D well below the level that would obtain in a frictionless world. That is, some firms do not find it worthwhile to begin an R&D program that they expect will have to be curtailed in the future due to financial constraints.

Mulkay et al. (2001) perform a similar exercise using large French and US manufacturing firms, finding that cash-flow impacts are much larger in the US than in France, both for R&D and for ordinary investment. Except for the well-known fact that R&D exhibits higher serial correlation than investment (presumably because of higher adjustment costs), differences in behaviour are between countries, not between investment types, suggesting that they are due to differences in the structure of financial markets rather than the type of investment, tangible or intangible. This result is consistent with evidence reported in Hall et al. (1999) for the US, France, and Japan during an earlier time period, which basically finds that R&D and investment on the one hand, and sales and cash flow on the other, are simultaneously determined in the United States (neither one "Granger-causes" the other). In the other countries, however, both types of investment are causal for sales and cash flow, with little feedback from sales.

---

7 The Euler equation for investment is an equilibrium condition derived from the firm’s value maximization problem. It expresses the tradeoff between investing today and having one more period of production from the capital versus investing tomorrow and forgoing today’s production.

8 Granger causality is a definition of causality in a time-series context that is due to the late Clive Granger. One variable ‘Granger-causes’ another if its prior values but not its current values help to predict the variable in the presence of the variable’s own history.
and cash flows to the two types of investments. Using a non-structural R&D investment equation together with data for the US, UK, Canada, Europe, and Japan, Bhagat and Welch (1995) find similar results for the 1985-1990 period, with stock returns predicting changes in R&D more strongly for the US and UK firms. Thus in all the countries studied, R&D investment is a predictor of increased performance contemporaneously (measured as sales, cash flow, or stock returns), whereas in the US and the UK, there is immediate feedback from the performance measures to R&D investment, which can be due either to liquidity effects or the positive demand signal from increased sales or cash flow.

Bougheas et al. (2001) examine the effects of liquidity constraints on R&D investment using firm-level data for manufacturing firms in Ireland and also find evidence that R&D investment in these firms is financially constrained, in line with the previous studies of US and UK firms.

4.2 Studies using innovation survey data

Recently a number of authors, mostly in Europe, have combined innovation survey data and firm financial data in order to look more closely at the relationship between innovative activities and the presence of financial constraints. This endeavour is welcome given the fact that the prior work was confined only to R&D investment, but it has suffered to some extent from the fact that in order for a firm to experience financial constraints related to innovation, they must at least want to be an innovator, leading to substantial simultaneity between the two. Various approaches are taken to mitigate this problem in the studies described below.

Savignac (2008) and Hajivassiliou and Savignac (2008) examine a simultaneous model of the two discrete states (facing financial constraints and innovating) directly, using a large sample of French firms. They do indeed find simultaneity between the two: binding financing constraints discourage innovation and at the same time innovative firms are more likely to face binding financing constraints. Using data on UK firms, Canepa and Stoneman (2008) find that the cost and availability of finance matters for innovation, especially for high-technology firms and for the smaller firms in their sample.

W. Brown (1997) argues that existing tests of the impact of capital market imperfections on innovative firms cannot distinguish between two possibilities: (i) capital markets are perfect and different factors drive the firm’s different types of expenditure or (ii) capital markets are imperfect and different types of expenditure react differently to a common factor (shocks to the supply of internal finance). He then compares the sensitivity of investment to cash flow for innovative and non-innovative firms in the UK. The results support the hypothesis that capital markets are imperfect, finding that the investment of innovative firms is more sensitive to cash flow.

Magri (2009) also compares the cash flow sensitivity of innovative and non-innovative firms, this time for Italian firms. She finds that investment in small innovative firms is more sensitive to cash flow than in small non-innovative firms, but that there is no difference for large firms. Note that this analysis is for ordinary investment rather than for innovation investment itself, but it is still suggestive of financial constraints for smaller innovating firms.

Jensen and Webster (2009) look at the cyclical response to macro-economic conditions of success in commercialization of new products and services by Australian firms. The most important macro-economic determinant of success is the cost of bank overdrafts (lower is better, obviously), followed
by the interest rate, and the cost of R&D. In general, commercial success is pro-cyclical in their sample, although others have found some evidence that research itself (as opposed to development) is counter-cyclical.

4.3 Corporate-governance structure and the financing of R&D

The challenge for direct tests of the relationship between management and ownership structures and firm investment strategies is that in any given firm, the two have co-evolved to be adapted to one another, and therefore observed correlations do not admit much of a causal story. In particular, any attempt to look at variations in performance that result from variations in strategy and governance will struggle with the fact that selection for “fitness” tends to eliminate some of the experiments one might like to see.

The usual way around this problem is to look for announcement effects, that is, for market reaction to information surprises or news that firms have not yet adapted to. The event study methodology is described in detail in Campbell et al. (1997). Briefly, it involves computing the returns to an investor from holding a share of stock in a firm during the 3 to 5 day period around the time that some news that affects the firm is revealed publicly. The returns computed are “abnormal,” that is, they are adjusted for the overall market returns during the same period.

Both Alam and Walton (1995) and Zantout (1997) find higher abnormal returns to firm shares following new debt issues when the firm is more R&D-intensive. The argument is that the acquisition of new sources of financing is good news when the firm has an asymmetric-information problem because of its R&D strategy. Similarly, Szewczyk et al. (1996) find that investment opportunities (as proxied by Tobin’s q) explain R&D-associated abnormal returns, and that these returns are higher when the firm is highly leveraged, implying a higher required rate of return for debt finance in equilibrium. Of course, in both of these cases, the evidence is non-experimental and the links between the theory and the empirical tests somewhat tenuous, but the results are suggestive.

Evidence on the importance of agency costs as they relate to R&D takes several forms. Several researchers have studied the impact of anti-takeover amendments (which arguably increase managerial security and willingness to take on risk while reducing managerial discipline) on R&D investment and firm value. Johnston and Rao (1997) find that such amendments are not followed by cuts in R&D, while Pugh et al. (1999) find that adoption of an Employee Stock Ownership Plan (ESOP), which is a form of anti-takeover protection, is followed by R&D increases. Cho (1992) finds that R&D intensity increases with the share that managerial shareholdings represent of the manager’s wealth and interprets this as incentive pay mitigating agency costs and inducing long-term investment.

Some have argued that institutional ownership of the managerial firm can be conducive to more R&D by reducing the agency costs due to free-riding by owners that is a feature of the governance of firms with diffuse ownership structure. Others however have held that such ownership pays too much attention to short-term earnings and therefore discourages long-term investment.10

There is some limited evidence that this may indeed be the case. Eng and Shackell (2001) find that firms adopting long-term performance plans for their managers do not increase their R&D spending but that institutional ownership is associated with higher R&D. In addition, R&D-performing firms tend not to be held by banks and insurance companies, at least in the US. These types of firms are presumed

10 Institutions such as mutual and pension funds often control somewhat larger blocks of shares than individuals, making monitoring firm and manager behaviour a more effective and more rewarding activity for these organizations.
to be poor at monitoring R&D performance. Majumdar and Nagarajan (1997) find that high institutional-investor ownership does not lead to short-term behaviour on the part of the firm; in particular, it does not lead to cuts in R&D spending. Francis and Smith (1995) find that diffusely held firms are less innovative, implying that monitoring alleviates agency costs and enables investment in innovation.

In a study of German manufacturing firms, Czarnitzki and Kraft (2009) use patent data as an indicator of innovative activity and find the opposite result. Companies with dispersed ownership are more active in innovation as measured by patent applications, but patenting falls as leverage increases. They interpret this result as suggesting that leverage acts as a disciplinary device that restrains overinvestment by managers. Although they attempt to control for the endogeneity of leverage using instrumental variables, the instruments (lagged cash flow and tangible-asset intensity) are arguably also related to unobserved determinants of innovative activity; in the absence of independent measures of over- or under-investment, it is difficult to know whether leverage is a discipline device or merely an indicator that the firm is not as innovation intensive as some others.

In general, the evidence summarized above is fairly clear and indicates that long-term incentives for managers can encourage R&D and that institutional ownership does not necessarily discourage R&D investment. However, it is fairly silent on the magnitude of these effects, and whether these governance features truly close the agency cost-induced gap between the cost of capital and the return to R&D.

4.4 Capital structure and R&D

Another way to look at the financing-innovation nexus is to examine the capital structure that results from the financing decisions of R&D-intensive firms. Here the evidence that debt is disfavoured is quite clear. Work using US data such as Friend and Lang (1988) and Hall (1992) shows a clear negative correlation between R&D intensity and leverage. The same is true of unpublished results for European firms (Hall et al. 2009).

Although leverage may be a useful tool for reducing agency costs in the firm, it is of limited value for R&D-intensive firms. As discussed above, because the knowledge asset created by R&D investment is intangible, partly embedded in human capital, and ordinarily very specialized to the particular firm in which it resides, the capital structure of R&D-intensive firms customarily exhibits considerably less leverage than that of other firms. Banks and other debt holders prefer to use physical assets to secure loans and are reluctant to lend when the project involves substantial R&D investment rather than investment in plant and equipment. In the words of Williamson (1988), “redeployable” assets (that is, assets whose value in an alternative use is almost as high as in their current use) are more suited to the governance structures associated with debt. Additional empirical support for this idea is provided by Alderson and Betker (1996), who find that liquidation costs and R&D are positively related across firms. The implication is that the sunk costs associated with R&D investment are higher than that for ordinary investment.

In addition, servicing debt usually requires a stable source of cash flow, which makes it more difficult to find the funds for an R&D investment program that must be sustained at a certain level in order to be productive. For both these reasons, firms are either unable or reluctant to use debt finance for R&D investment, which may raise the cost of capital, depending on the precise tax treatment of debt versus equity. Confirming empirical evidence for the idea that limiting free cash flow in R&D firms by issuing debt is a less desirable method of reducing agency costs is provided by Chung and Wright (1998), who find that financial slack and R&D spending are positively correlated with the value of growth firms, but not correlated with that of other firms. Czarnitzki and Kraft (2009) find that more leveraged German firms have lower innovation output (measured by patents), especially when ownership of the firm is dispersed.
In the view of some observers, the LBO wave of the 1980s in the United States and the United Kingdom arose partly because high real interest rates meant that there were strong pressures to eliminate free cash flow within firms (Blair and Litan 1990). For firms in industries where R&D is an important form of investment, such pressure should have been reduced by the need for internal funds to undertake such investment and indeed Hall (1993, 1994) and Opler and Titman (1993) find that firms with high R&D intensity were much less likely to do an LBO. Opler and Titman (1994) find that R&D firms that were leveraged suffered more than other firms when facing economic distress, presumably because leverage meant that they were unable to sustain R&D programs in the face of reduced cash flow.

In related work using data on Israeli firms, Blass and Yosha (2003) report that R&D-intensive firms listed on the US stock exchanges use highly equity-based sources of financing, whereas those listed only in Israel rely more on bank financing and government funding. The former are more profitable and faster-growing, which suggests that the choice of where to list the shares and whether to finance with new equity is indeed sensitive to the expected rate of return to the R&D being undertaken. That is, investors supplying arms-length finance require higher returns to compensate them for the risk of a “lemon”.

4.5 Young innovative companies

Before leaving this topic, I briefly review the evidence on the financing of young innovative firms and start-ups. At the macro-economic level and beginning with Rajan and Zingales (1998), a number of studies have related the overall financial development level with measures such as the entry of new firms in a country. Using a set of 16 countries including the US, European, and mid-level developing countries in Eastern Europe and Latin America, Aghion et al. (2007) show that the level of financial development in a country (private credit and the size of market capitalization) strongly influences entry and post-entry growth of small firms but has little impact on large-firm growth. These authors highlight the differences in the level of market capitalization between the US, the U.K., and a few Nordic countries on the one hand and continental Europe on the other. Although to some extent the growth of small firms and stock markets are endogenously related in that they grow together, the methodologies chosen by Rajan-Zingales and Aghion et al. are designed to minimize the identification problem that results, which makes their results robust. This work does not relate directly to innovative firms per se, but it is suggestive, as these are often very prominent among new firms, especially in the manufacturing sectors that are studied by these authors.

An entirely different approach that looks directly at the cash flow impact on entrepreneurial firms was pursued by Holtz-Eakin et al. (1994). Using tax data on US entrepreneurs (sole proprietors) some of whom received inheritances, they were able to show that the reception of funds from inheritance increased entrepreneurial survival and that those that survived grew faster, suggesting that such firms were financially constrained prior to the cash influx. Turning to R&D-doing and innovative small firms in particular, in most of the work reviewed in Section 4.1, when the authors have looked they have found greater financial constraints for these firms than for the larger firms in their samples. For examples, see Himmelberg and Petersen (1994) and J. Brown et al. (2009).

Of course, looking at smaller firms that have successfully entered but face some financial constraints is only part of the story. We would also like to know about entry that never takes place, or failure to innovate due to constraints. Here the innovation surveys can be helpful, especially if they survey non-innovative firms carefully. For example, Canepa and Stoneman (2008) found that finance mattered especially for the smaller innovative firms in their UK sample, as does Magri (2008) for Italian firms. Audretsch and Lehmann (2004) examined the financing of young innovative firms on the German
Neuer Markt, finding a negative correlation between bank debt and VC financing. They also show that the VC-backed firms have much higher growth rates. They argue that these two findings show the importance of financiers willing to risk capital for the success of young innovative firms, especially in a country like Germany where bank finance dominates.

5. Conclusions and policy implications

A number of conclusions have emerged from this body of theoretical and empirical work, conclusions that are fairly consistent across the various samples and methodological variations. It is perhaps important to emphasize that the market failures catalogued and explored here are confined to those due to the separation of owner and manager, or financier and entrepreneur. However, the principle public policy argument for subsidizing innovative firms may be different, as it is based on the social benefits to be derived from the unpriced knowledge spillovers from such firms. That is, the problem of inappropriability of the returns to investment in knowledge that was pointed out long ago by Arrow (1962) and Nelson (1959). The arguments and evidence here are in addition to these.

First, there is solid evidence that debt is a disfavoured source of finance for R&D investment. Second, the “Anglo-Saxon” economies, with their thick and highly developed stock markets and relatively transparent ownership structures, typically exhibit more sensitivity and responsiveness of R&D to cash flow than continental economies. Third, and much more speculatively, this greater responsiveness may arise because they are financially constrained, in the sense that they view external sources of finance as much more costly than internal, and therefore require a considerably higher rate of return to investments done on the margin when they are tapping these sources. However, it is perhaps equally likely that this responsiveness occurs because firms are more sensitive to demand signals in thick financial markets; a definitive explanation of the “excess sensitivity” result awaits further research.

Finally, there is now considerable evidence that young and/or small firms are more likely to face financial constraints than large established firms, a result that is not surprising, if the source of the problem is a “lemons” premium. Presumably such firms have a weaker track record on which investors can base their evaluations.

There are a number of policy implications from the results discussed in this paper. One implication is for governments to design policies conducive to lowering the cost of financing innovation, which many governments have done. The chief instrument in this area is a simple tax credit or subsidy, in some cases targeted towards small firms. Others include various programmes that are project related or targeted towards pre-commercial research and development (e.g., see David et al. (2000) for an international survey of this literature; Czarnitzki and Hussinger (2004) for a detailed study of R&D subsidies in Germany; and Hall and Maffioli (2008) for a survey of such programs in Latin America and their evaluation).

One of the problems in designing policy in this area is that different countries face different problems. For example, although one can argue that financing constraints are not absent in the United States, it is fairly clear that the rise of the VC industry has mitigated the problem, at least for high technology start-ups (including those in green technologies). In other countries, creating such an industry may be more of a challenge. In Israel, for example, it appears to have been kick-started successfully by the government, but only on the second try (Gilson 2003; Avnimelech and Teubal 2004).

Several European governments have also attempted to behave somewhat as VCs towards their own start-ups and small innovative firms. In Germany, more than 800 federal and state government financing...
programs have been established for new firms in the recent past (OECD 1995). In 1980, the Swedish
government established the first of a series of investment companies (along with instituting a series
of measures such as reduced capital-gains taxes to encourage private investment in start-ups), partly
on the United States model. By 1987, the government share of VC funding was 43 percent (Karaomeriolu
and Jacobsson 1999). Recently, the UK has instituted a series of government programs under the
Enterprise Fund umbrella which allocate funds to small and medium-sized firms in high technology
and certain regions, as well as guaranteeing some loans to small businesses (Bank of England 2001).

In a recent theoretical paper, Takalo and Tanayama (2008) argue that government R&D subsidies based
on screening of firms and projects can serve as important signals to private-equity markets and venture
capitalists of the quality of the recipients, which will reduce to some extent the “lemons” premium
Innovation Research (SBIR) grants and concluded the same: Receiving such a grant enabled firms to
obtain more resources elsewhere, and in addition such firms grew faster afterwards, when compared
to a matched sample of firms that did not receive funds from SBIR. This suggests an important role for
government policy in screening firms for funding.

For established firms, estimates of cash-flow sensitivity point to differences between the US and UK
on the one hand and continental European countries on the other, with the former more sensitive than
the latter, so in principle US and UK firms should be more subject to financial constraints. But there is
little evidence that this leads to lower innovative activity, in fact somewhat the contrary, so other forces
must be at work, too. The largest of these firms in all countries and the ones that perform the greatest
amount of R&D tend to compete with each other in international markets, so that it is not likely that
the behaviour of firms from different countries can diverge too far.
References


