Financial Constraints, Capital Structure and Innovation: An Empirical Investigation

Amaresh K Tiwari*, Pierre Mohnen†, Franz C. Palm‡,
Sybrand Schim van der Loeff§

Abstract

Using a panel data of three waves, within a parametric framework, we investigate the effect of financial constraint on R&D investment and study the determinants of financial constraints. We also investigate the factors that influence or provide incentives for a firm to take up R&D activity. Our findings can be summarized as follows. First, financial constraints adversely affect a firm's R&D intensity as measured by ratio of R&D expenditure to capital asset. Second, firms that are highly leveraged are more likely to be financially constrained, and that highly leveraged firms are less likely to be innovators. Third, the propensity to innovate with respect to leverage is lower when a firm is not financially constrained as compared to a firm that is. Fourth, the propensity to innovate with respect to leverage, conditional on no financial constraint is almost constant, while the propensity to innovate with respect to leverage conditional on being financially constrained, varies over the distribution of firm characteristic such as age, size, and leverage. Fifth, the decision to innovate, the financial constraints faced, and the choice of capital structure are endogenously determined. Sixth, the R&D intensity of firms with different characteristics, conditional on being financially constrained and conditional on not being unconstrained, are different. Seventh, the sensitivity of R&D investment to cash flows is higher for financially constrained firms. The econometric exercise entails using a three step procedure, where expected a posterior (EAP) values of time invariant individual effects obtained from the first stage reduced form are used as substitutes for the time invariant individual effects that are to be controlled for in the structural equations of the second and third stage. The paper provides the theoretical underpinnings for such a procedure.

1 Introduction

Empirically, the study of the effect of financing frictions on investment has broadly followed two approaches. One approach is to ad hoc classify firms into those that are financially constrained and those that are not, see Fazzari, Hubbard, and Petersen (1988) and Kaplan and Zingales (1997), and specify a reduced form accelerator type model for the constrained and unconstrained firms. The extent of financing frictions, controlling for the investment opportunity, is judged by the sensitivity of investment to cash flows, where cash flows realized is taken as measure of internal wealth. Another approach, which is more structural, is to construct an index of financial constraints based on a standard intertemporal investment

*University of Maastricht, A.Tiwari@maastrictuniversity.nl
†University of Maastricht, MERIT and CIRANO, P.Mohnen@MERIT.unu.edu
‡University of Maastricht and CESifo fellow, F.Palm@maastrictuniversity.nl
§University of Maastricht, S.Loeff@maastrictuniversity.nl

model augmented to account for financial frictions, where external financing constraints affect the intertemporal substitution of investment today for investment tomorrow, via the shadow value of scarce external funds, Whited and Wu (2005) . The approach relies on specifying a smooth functional form for adjustment costs of adjusting capital, and specifying the shadow value of external funds as a function of financial state variables. The small number of existing empirical literature that study financing frictions and R&D investment, documented in an excellent review by Hall and Lerner (2010), broadly speaking, follow the two above mentioned approach.

In this chapter we assess the impact of financial constraints on R&D expenditure, and what firm characteristics are associated with a firm being financially constrained, where firms themselves report if they are financially constrained or not, with respect to R&D activity. Using firm's assessment of being financially constrained avoids the need to construct an index of financial constraint. This implies that we do not know whether these firms are constrained due to high cost of external finance or feel constrained because their access to funds is less than what is needed to finance their investment, which a priori classification usually imply: see Moyen (2004) and Hennessy and Whited (2007). However, this also implies that using firm's assessment of being financially constrained allows us to test how important are the classification criteria that distinguish firms as constrained and not constrained.

Secondly, since the test of financing frictions relies on assessing the impact of financial constraints on R&D expenditure, ours is a departure from the reduced form accelerator type models, where the extent of financing frictions is measured by the sensitivity of investment to cash flows for constrained and unconstrained firms, and about which questions have been raised as to whether such a procedure can indeed identify the extent of financing frictions: see for example, Kaplan and Zingales (1997), Gomes (2001), and Hennessy and Whited (2007). In other words, in contrast to reduced form models, ours is a more structural approach, where conditional on firm characteristics, R&D investment is determined by financing frictions and the future expected profitability from R&D investment. The financing frictions as summarized by the reported financial constraints face by the firms could be either due to high cost of external sources of funds or due to the fact that the access to funds is less than what is needed to finance their investment. We find that the impact of financial constraints can be substantial, reducing the R&D intensity, measured by R&D expenditure to capital asset ratio, of an average firm by more than half. Thirdly, using the information on financial constraint as reported by firms, not only allows us to assess the effect of financing frictions on R&D investment, but also, using the framework of switching regression model, allows us to investigate how firms with different characteristics such as maturity and degree of monopoly, behave with respect to innovation and R&D investment under financial constraints and under no financial constraints.

Moreover, papers that a priori classify firms as constrained and unconstrained, take the financial constraints faced by firms as exogenous to investment decisions. However, we know that financial constraints and investment expenditures are determined simultaneously, see Section 6.3 for a detailed discussion. Using information on financial constraints, which are a function of financial state variables and firm characteristics, as reported by firms, allows us to endogenize financial constraints while explaining R&D investment. We do this by specifying a financial constraint equation, whose error term is correlated with the error term of the R&D equation, and by allowing a common time invariant individual effect to affect both the financial constraints faced by the firm and its decision on how much to invest in R&D.

R&D expenditure in our data, however, is observed only for firms that are classified as innovators. Hence, we have to confront the problem of endogenous sample selection. To control for the bias that arises when estimating the R&D equation, due to endogenous sample selection, we specify a selection/innovation equation, the idiosyncratic component of which is correlated to those of the financial constraint and the R&D investment equation. Estimating the innovation equation allows us to test, which type of firms, given their characteristics, are more likely to take up R&D activity. Also, since the decision to innovate and the constraints faced are determined simultaneously, joint estimation of the decision to innovate and the financial constraint faced allows us investigate the implication of financing frictions for a firm's decision to innovate.

Debt, as we know, is an important source of external financing, but given the existence of financing friction, is costly. Given the nature of R&D, discussed in Section 6.2, debt might even be costlier when it comes to financing R&D. Hence, choosing to be an innovator can have an important implication for its choice of capital structure. Consequently leverage, defined as the ratio of long-term debt to capital asset, as variable will have a bearing on both the financial constraint faced by the firm as well as the decision to innovate. However, since there is simultaneity in the financing policy and investment decision of the firm, from a modeling perspective, endogeneity of debt has to be taken into account. Our empirical strategy which is a three step procedure takes into account the endogeneity of the long-term debt while estimating the structural equations relating to the decision to innovate, the financial constraint with respect to R&D that the firm faces, and its R&D expenditure, all of which, again, are endogenously determined.

Costly external finance has important consequences for the dynamics of the firm subject to technological and idiosyncratic shocks, that affect firms with different characteristics differently: see Cooley and Quadrini (2001). Hence, we expect to see financing and investment/innovation policy that vary along the distribution of firms, for example, based on size, age, leverage and degree of financial constraints. Apart from assessing the impact of constraints on R&D expenditure this chapter attempts to explore the implication of financing frictions on the choice of capital structure and the decision to innovate and how such choices or the propensity to innovate with respect to leverage vary over the distribution of firms characteristics such as maturity, size and leverage. Our results suggest that, ceteris paribus, a firm is less inclined to innovate if it is highly leveraged. The marginal propensity to innovate with respect to leverage differs when a firm is financially constrained and when it is not. The propensity to innovate is more affected by leverage, if a firm is not financially constrained than when a firm is. In other words, when a firm is not financially constrained with respect to R&D and if we find that it is highly leveraged, then it is more likely that it is not an innovator. Moreover, if the firm is not financially constrained the marginal propensity to innovate is almost constant, while the propensity to innovate with respect to leverage, when financial constrains bind, varies over the distribution of firm characteristic such as age, size and leverage. While, there have been many papers that have studied financing and innovation policy of firms, see Section 6.2, none to our knowledge, has investigated the behavior of such firms under regimes: (1) when firms are financially constrained and (2) when they are not, and how their financing and innovation policy vary with the distribution of firm characteristics[1].

To summarize, in this chapter we model a firm's behavior, where its choice of capital structure, the decision to innovate, the financial constraint faced, and the amount spent on R&D related activities are determined simultaneously. By doing so, we are able to assess the implication of financing decision on, controlling for other factors, the firms decision to innovate, the financial constraints faced by the firm, and subsequently the effect of financial constraints on R&D investment. We do this by specifying a system of four equations: (a) R&D intensity equation (a switching regression model), (b) financial constraint equation, (c) a selection/innovation equation, and (d) leverage or long-term debt to asset ratio equation. While (a), (b), and (c) are structural equations, we write (d) in reduced form. The simultaneity in all the four above is captured through the correlation of the idiosyncratic terms in

[1]A likely exception is a recent paper by Brown, Fazzari, and Petersen (2009). These authors separate their sample based on age, where they argue that young firms are more likely to face financial distress than large firms. For the subsamples of mature and young firms, they fit an Euler equation augmented with financial state variables. They find that the financial variables indeed matter for young firms, thereby supporting their claim that financial constraints are binding for young firms.

each of the four equations and a common time invariant individual effect that appears in all the equations.

To estimate the above system of four equations, we employ the method of control function, see Blundell and Powell (2003), in which the estimated residuals from the first stage reduced form are used to control for the endogeniety of the debt to asset ratio in the structural equations. However, due to the presence of unobserved time invariant individual effects, of which the residuals are a function of, the residuals are not identified. To this effect, we substitute the expected a posteriori values of the individual effects based on the first stage estimates. The chapter provides the theoretical foundations for such a procedure. Before we end our introductory section, we introduce the remaining sections that follow. Section 6.2 discusses the nature of R&D activity from an economic point of view and its implication for the choice of capital structure. The section also discusses how these innovation and financing policy can vary over the distribution of firm characteristics. In Section 6.3 we briefly review some of the theories of costly external sources of finance, and given the nature of R&D activity, discuss the implications for financial constraint a firm engaging in R&D might face. The rest of the chapter is organized as follows: Section 6.4 discusses the data and the construction of the variables, Section 6.5 discusses the empirical strategy employed, Section 6.6 discusses the results and finally Section 6.7 concludes. The details of the econometric methodology are provided in Appendix A, B, and C.

2 Innovation, Capital Structure, and Firm Dynamics

From the perspective of investment theory, R&D has a number of characteristics that make it different from ordinary investment. Holmstrom (1989) indicates that innovation, and by extension R&D, has five unique characteristics, it is long-term in nature, high risk in terms of the probability of failure, unpredictable in outcome, labor intensive and idiosyncratic. Hall and Lerner (2010) point out that in practice fifty per cent or more of R&D spending is the wages and salaries of highly educated scientists and engineers, and that their efforts create an intangible asset, the firm's knowledge base, from which profits in future years will be generated. To the extent that 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. These characteristics, as we discuss are potential sources of agency issues that can arise between creditors and borrowers.

The high risk involved and unpredictability of outcomes are potential sources of asymmetric information that give rise to agency issues in which the inventor frequently has better information about the likelihood of success and the nature of the contemplated innovation project than the potential investors. Therefore, the marketplace for financing the development of innovative ideas looks like the "lemons" market modeled by Akerlof (1970), and accordingly a higher premium might be charged on the external funds offered. In case of R&D, given its nature, the adverse selection problem can even be more severe, and accordingly, the "lemons" premium for R&D can be higher than that for ordinary investment. This is because investors have more difficulty distinguishing good projects from bad when the projects are long-term R&D investments than when they are more short-term or low-risk projects (Leland and Pyle 1977). 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) and Anton and Yao (2001). 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.

Hall (1993,1994), Opler and Titman (1993, 1994), Blass and Yosha (2001), and Alderson and Betker (1996) are some of the papers that provide empirical evidence that R&D intensive firms are less leveraged than those that are not. More recently, Brown, Fazzari, and Petersen (2009), find similar evidence for a panel of R&D performing US firms. Brown, Fazzari, and Petersen (2009) draw out a financing hierarchy for R&D intensive firms, where equity might be preferred to debt as a means of financing R&D, especially for young firms. Although leverage may be a useful tool for reducing agency costs within a firm, it is of limited value for R&D-intensive firms. 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. Williamson (1988), points out that "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. Titman and Wessels (1988) report that having "unique" assets is associated with lower debt levels. The logic being, first, that consumers will only buy unique products if they are confident that the firm will survive to provide after-sales service. Second, the lack of a secondary market for R&D and the non-collaterability of R&D activity mitigates against debt-financed R&D activity. Also, as Aboody and Lev (2000) argue, the extent of information asymmetry associated with R&D,

however, is larger than that associated with tangible (e.g., property, plant, and equipment) and financial investments because of the relative uniqueness (idiosyncrasy) of R&D. Thus, for example, a failure of a drug under development to pass Phase I clinical tests is a unique event not shared by other pharmaceutical companies. In contrast, a downturn in demand for commercial property for example, will exert a strong common effect on the property values of all real estate companies operating in a given geographical region. Similarly, interest-rate changes will affect systematically the values of bond and stock portfolios of companies. Thus, the relative uniqueness of R&D investments makes it difficult for outsiders to learn about the productivity and value of a given firm's R&D from the performance and products of other firms in the industry, thereby contributing to information asymmetry. Shi (2003) finds evidence that R&D activity, which increases the market value of equity, also increases bond default risk and debt risk premia. Bond holders, ceteris paribus, may be unwilling to hold the risks associated with greater R&D activity.

Given the nature of R&D activity, the capital structure of a firm (debt to asset ratio) amongst other things also contains information on the type of activity a firm is engaged in. After controlling for various factors, such as size, age, degree of monopoly, profitability of the firm and other determinants of capital structure related to corporate governance or reasons related to tax consideration, we would expect that more levered firms, for reasons stated above, are less likely to engage in R&D related activity.

The above discussion relating asymmetric information pertinent to the nature of R&D activity and leverage holds for any given firm. Since, empirically there exists certain regularities in the financial characteristics of firms that are related to their size and age, Cooley and Quadrini (2001) point out that it is important to link patterns of firm growth with their financial decisions. Cooley and Quadrini (2001) introduce financial frictions in a standard model of industry dynamics that generates results which match the empirical findings. Some of results related to growth and financing are that small and young firms (a) experience higher growth and volatility of growth, (b) invest relatively more and exhibit higher risk taking behavior, (c) take on more debt and pay less dividends. Cooley and Quadrini (2001) explain, that in order to understand these phenomena, one has to consider the trade-off that firms face in deciding the optimal amount of debt. On the one hand, more debt allows them to expand the production scale and increase their expected profits; on the other, the expansion of the production scale implies a higher volatility of profits and a higher probability of failure. Given that a large fraction of profits is reinvested, and the firm's future value, by their assumption, is a concave function of equity, the firm's objective is a concave function of profits. This implies that the volatility of profits has a negative impact on the firm's value. Therefore, in deciding whether to expand the scale of production by borrowing more, the firm compares the marginal increase in the expected profits with the marginal increase in its volatility (and therefore, in the volatility of next period equity). Due to diminishing returns, as the firm increases its equity and implements larger production plans, the marginal expected profits from further increasing the production scale decrease. The firm becomes more concerned about the volatility of profits and borrows less in proportion to its equity. Consequently, as the firm grows, the composition of the sources of finance changes in favor of internal sources. As a consequence of higher borrowing, small firms face a higher probability of default. Cooley and Quadrini (2001) model heterogeneity as difference in productivity level. For firms with higher productivity level, the marginal expected profit is higher for each production scale. Consequently, the firm is willing to face higher risk by borrowing more, and expands the scale of production. However, the essential trade off between profits and volatility is still faced.

Our findings that, (i) mature firms are less likely to be innovators, (ii) younger firms are more financially constrained, (iii) marginal propensity with respect to leverage declines with age, (iv) under binding financial constraints the propensity to innovate with respect to leverage varies with the distribution of firm characteristics such as age, size, and leverage, and (v) that large and mature firms are less R&D intensive, suggests that decision to innovate, financing choices, and firm dynamics are not independent. Though we do not study financing choices along the distribution of firm characteristics, some of our findings, for example, small and young firms being more R&D intensive, match the empirical regularities Cooley and Quadrini (2001) study. However, R&D investment is different from scale expansion, in the sense that R&D activity is both risky as well as increases the productivity of the firm. Therefore it might be interesting to incorporate R&D investment in Cooley and Qudarini's (2001) model. For example, suppose a large firm is as R&D intensive as a small firm, and if there are decreasing returns to R&D capital, which is quite plausible, then the large firm will be at a proportionately higher risk than small firms, since not only is it more subject to exogenous shocks, given the scale of its operation, which increases the volatility of profits, but also, since it has tied up more capital in risky R&D ventures, which can make it more default prone. Thus, being an innovator could also affect the amount of debt it can procure for its scale expansion activity.

On the other hand for small and young firms, entry, survival and subsequent growth depends on how innovative they are. Recent research on industry evolution has shown that entry and exit in any given industry are typically high, that new firms start on a very small scale but also that they are not able to remain small forever, and that survival depends on heterogeneous mechanisms which include crucial innovation and growth: see, Audretsch (1995) and Huergo and Jaumandreu (2004). Literature on industrial organization treat entry of firms, as the way in which firms explore the value of new ideas in an uncertain context. Entry, the likelihood of survival, and subsequent growth are determined by barriers to survival. In this framework, entry is innovative and increases with uncertainty, the likelihood of survival is lower the higher the risk is, and the growth subsequent to successful innovation is higher the higher barriers to survival are.

3 R&D Investment and Financial Constraints

Firms are financially constrained when they face a shortage of internal funds needed for their investment and are forced to resort to external sources of financing, which may be costly in the sense that firms are required to pay a high premium on external finance due to capital market imperfections. These capital market imperfections can be a due to a variety of agency and information asymmetry problems, some of which related to R&D investment have been discussed above.

Myers and Majluf (1984), Myers (1984) and Greenwald, Stiglitz, and Weiss (1984), have pointed out that raising equity externally may be costly due to the kind of adverse selection problem identified by Akerlof (1970). Inability to raise new equity would not be a problem if firms could frictionlessly raise unlimited amounts of debt to finance their investment. However, a variety of theories suggest that this is unlikely to be the case. Now, since at any given interest rate managers will be more likely to borrow if their private information suggests that they are relatively prone to default, hence the market for debt could also be subject to the adverse selection problem that afflicts equity market. It is also possible that there can be moral hazard problems, whereby those managers who borrow have an increased incentive to take risks that lead to default. These considerations, as has been shown by Stiglitz and Weiss (1981, 1983), can lead to credit rationing, whereby firms are unable to obtain all the debt financing they would like at the prevailing market interest rate.

Debt overhang as Myers (1977) points out can limit debt finance. Large debt can be a burden on a firm's balance sheet, which can discourage further new investment, particularly if this new investment is financed by issuing claims that are junior to the existing debt. This is because if the existing debt is trading at less than face value, it acts as a tax on the proceeds of the new investment: part of any increase in value generated by the new investment goes to the existing lenders, they being the first claimants, and therefore is unavailable to repay those claimants who put up the new money. Debt overhang models can have two different implications: ex post (once the debt burden is in place) they suggest that highly leveraged firms will be particularly prone to underinvestment. Ex ante, they offer a reason why even more modestly-levered firms, particularly those with attractive future investment opportunities, may be reluctant to raise much debt in the first place, even if this means foregoing some current investment projects. Jensen and Meckling (1976) offer another reason why firms might be unwilling to take on too much debt ex ante: the so-called asset substitution effect, whereby an excessive debt burden can create incentives for managers, acting on behalf of shareholders, to take on risky negative-NPV projects at the expense of lenders.

The above-discussed models of debt and equity finance take the existence of these types of financial claims as given, and then go on to derive implications for investment, capital structure, etc. Another branch of the literature seeks to endogenize the financial contract, typically by positing some specific agency problem (e.g., managers penchant for diverting the firms cashflow to themselves) and asking what sort of claim represents an optimal response to this agency problem. These agency issues as discussed earlier are likely to be even more acute for R&D financing. In much of this work, the optimal contract that emerges resembles a standard debt contract, and there is no outside equity financing. Early examples include Townsend (1979) and Gale and Hellwig (1985), who assume that outside investors can only verify a firm's cash flows by paying some fixed auditing cost. As long as the manager turns over the stipulated debt payments, there is no audit, and the manager gets to keep the rest of the firm's cashflow. However, if the manager/entrepreneur defaults on its obligation, the lender audits, and keeps everything he finds; this can be interpreted as costly bankruptcy. The implications for investment is that, the less wealth the manager/entrepreneur is able to put up, the more he must borrow, and hence the greater is the likelihood of the auditing/bankruptcy cost being incurred. The default cost increases the cost of borrowing. This is because, for a given value of equity, the probability of default increases when the firm borrows more, because then it is more vulnerable to idiosyncratic shocks. This increases the expected default cost and the financial intermediary will, accordingly, demand a higher interest rate.

However, in the above mentioned papers, the borrowing constraints faced by the firms when deriving the optimal contract is specified exogenously. More recent papers seek to endogenize the borrowing constraints faced by the firms. Examples include, Hart and Moore (1994), Almeida and Campello (2002), and Albuquerque and Hopenhayn (2004). In Hart and Moore (1994) the threat of repudiation by the entrepreneur sets an upper bound on the value of debt and debt payments are subject to a cash-flow constraint, in Almeida and Campello

(2002), borrowing constraints are endogenized by conditioning on the firm's ability to raise external finance on its investment spending, while in Albuquerque and Hopenhayn (2004), the borrowing constraints are endogenously derived from limited enforceability problems.

The notion of financial constraint that we employ is one of borrowing constraints, that is, firms are financially constrained when they reach their debt capacity. Depending on the amount of internally available funds, the degree of market imperfection it faces, and the amount of existing debt it services, a firm may or may not have reached its debt capacity. Given that nature of R&D, such as the intangibility and uniqueness of assets and the unpredictability of outcomes, the degree of market imperfections for R&D activity is expected to be even higher. This would imply that external sources of finance could be even more costly when it comes to R&D related activity. Albuquerque and Hopenhayn's (2004) analysis indicates that riskier projects could face tighter constraints. Secondly, due to intangibility and uniqueness of assets, there is a lack of a secondary market for R&D activity and therefore R&D activity and its assets are generally not pledgeable, this naturally reduces the borrowing capacity of an R&D intensive firm. Also, given the uncertainty involved with respect to successfully inventing a new product or idea, the market may not respond favorably at the initial stages of the development of the product. Thus, lower valuation of the firm at initial stages of the development of the project, may also lower the borrowing capacity of the firm.

In general, as some of the above mentioned theories on investment and costly external finance imply, we would expect that given everything else, a more leveraged firm and a firm with lower level of internally available funds to be more financially constrained. Opler and Titman (1994) too find that firms with specialized products are especially vulnerable to financial distress, and that highly leveraged firms that engage in R&D suffer most in economically distressed periods. Here we test, given everything else, whether presence of higher leverage causes a firm to be financially constrained, if internal sources of financing relaxes the borrowing constraint, and how higher cost of external finance, as summarized by the reported financial constraints, affects R&D related expenditures.

4 Data and Construction of Variables

As stated earlier, to assess the impact of financial constraints on R&D expenditure and to assess the impact of the evolution of financial state variables on a firm being financially constrained with respect to R&D activities, we have had to merge two data sets for those years for which information on R&D is available. The data on information related R&D is obtained from the Dutch Community Innovation Surveys which are conducted every two years.

The Innovation Survey data are collected at the enterprise level. A combination of a census and a stratified random sampling is used to collect the data. First, the frame population for the innovation surveys are determined. For enterprises employing 50 persons or more, a census is used. For enterprises employing 10 to 50 persons, a sample is drawn from the general business register. A census of large (250 or more employees) enterprises, and stratified random sample for small and medium sized enterprises from the frame population is used to construct the data set for every survey. The stratum variables are the economic activity and the size of an enterprise, where the economic activity is given by the Dutch standard industrial classification. For our empirical analysis we use three waves of innovation survey data: CIS2.5, CIS3, and CIS3.5 pertaining respectively to the years 1996-98, 1998-2000, and 2000-02.

Information related to financial status of the firms is available at the firm level, which could be constituted of many enterprises consolidated within the firm. The financial data is from the balance sheet of the SFGO (Financial Statistics for Large) and SFKO (Financial Statistics for Small – reported assets less than 23 million Euros – Companies) firms. The data for the SFGO firms are collected by the Central Bureau of Statistics through a questionnaire sent to these firm, consequently the SFGO data is more detailed than the SFKO data, which are compiled from the information sent to the Tax Collectors Office.

However, the SFGO and SFKO data, which are at the company/firm level do not have any information on R&D activity. As mentioned earlier information related to R&D are obtained from the Innovation Surveys. For any given year, our problem here is to infer the size of the relevant R&D variables for a firm when not all enterprises belonging to the firm have been surveyed. To achieve this end, we use the information on the sampling design for the stratified random sampling done by the Central Bureau of Statistics (CBS) of The Netherlands. Below we outline the procedure.

For any given year, let N_T be the total population of the enterprises in the Netherlands and let N be the population of R&D performing enterprise. CBS determines N based on, (a) size class and (b) activity class. Outside N, or the Frame Population as is referred to by the CBS, any enterprise is hardly likely to indulge in any R&D activity. From this frame population a stratified random sampling is done. These strata are again based on size and the activity class. Let S be the total number of strata, and each stratum is indexed by $s = 1, 2, \cdots, S$. Then, $\sum_{s=1}^{S} N_s = N$, where N_s is the population of R&D performing enterprise belonging to stratum s. Let n_s be the sample size of each stratum and let $\Theta_s = \{1, 2, \cdots, i, \cdots, i_s\}$ be the set of enterprises for the s^{th} stratum, that is $|\Theta_s| = n_s$.

Let x be the variable of interest and x_i the value of x for the i^{th} observation in the s^{th} stratum. The mean value of x for an enterprise belonging to the s^{th} stratum is $\bar{x}_s = (\sum_{i \in \Theta_s} x_i)/n_s$. Now consider a firm f. Let N_{fs} be the number of enterprises belonging to the firm f and stratum s and n_{fs} be the number of enterprises belonging to the firm f and stratum s that have been surveyed.

Then the estimated value of x for the firm f, $\hat{x}_f$ is given by

$$\hat{x}_f = \sum_{s=1}^{S} (N_{fs} - n_{fs})\bar{x}_s + \sum_{s=1}^{S} \sum_{k=1}^{n_{fs}} x_{fsk}, \tag{1}$$

where x_{fsk} is the value of x for the k^{th} observation/enterprise belonging to the s^{th} stratum and firm f that has been surveyed and $N_{fs} - n_{fs}$ are the number of enterprise of the f^{th} firm in stratum s that have not been surveyed. It can be shown that $\hat{x}_f$ is an unbiased estimator of the expected value of x for firm f[2]. Table 6.1 below gives, based on size class and 2 digit Dutch Standard Industry Classification (SBI), the number of strata, between which the enterprises surveyed in the CIS surveys were divided.

Table 1: **Number of Enterprise and Number of Strata**

	CIS2.5	CSI3	CIS3.5
Total no. of enterprises	13465	10750	10533
Total no. of strata	240	249	280
These figures are from the original/raw data set.			

The sample of firms used in the estimation are only those for which at least one R&D performing enterprise is present in the innovation surveys. Also, those R&D performing enterprises for whom no firm information was available – firms not present in the SFGO and SFKO data – had to be dropped out. Table 6. 8 in Appendix E tabulates, for the sample of firms used in estimation, $N_f = \sum_{s=1}^{S} N_{fs}$ for each of the three waves, N_f being the total number of potentially R&D performing enterprise belonging to the firm f. The table also shows for each N_f, the number of firms for which, $N_f = n_f$, where n_f is the total number enterprise belonging to the firm f that were surveyed. It also shows the number of firms for which, $(N_f - n_f) > 0$. This tells us the number of firms for which the above procedure was used to get an estimate of the relevant variables for the firm. Here, we would like to point out

[2]Proof:
The proof is based on the assumption that the distribution of x in terms of its expected value is the same for each enterprise in a particular stratum. Let μ_{xf} the population mean of x for the firm f and let μ_{xs} the population mean of x for an enterprise belonging to stratum s. Given our assumption, we know that $\bar{x}_s$ is an unbiased estimator of μ_{xs}, that $\mu_{xf} = \sum_{s=1}^{S} N_{fs}\mu_{xs}$, and that the expected value of $\sum_{s=1}^{S} \sum_{k=1}^{n_{fs}} x_{fsk}$, the second term on the RHS of equation (1), is $\sum_{s=1}^{S} n_{fs}\mu_{xs}$. Taking expectations in (1) and substituting the expected value of $\mathrm{E}(\sum_{s=1}^{S} \sum_{k=1}^{n_{fs}} x_{fsk}) = \sum_{s=1}^{S} n_{fs}\mu_{xs}$ and noting that $\mathrm{E}(\sum_{s=1}^{S} n_{fs}\bar{x}_s) = \sum_{s=1}^{S} n_{fs}\mu_{xs}$, we get $\mathrm{E}(\hat{x}_f) = \mu_{xf} = \sum_{s=1}^{S} N_{fs}\mu_{xs}$.

that the information on N_f was obtained form the Frame Population constructed by the CBS and of course the information on n_f from the CIS surveys. As can be seen from the table, for the sample of firms used for estimation, for 18.06 percent of the total number of firms in CIS2.5, equation (1) was used to get the estimates of the relevant variables for the firms, while these percentages for CIS3 and CIS3.5 were 24.62 and 23.75 respectively. The table also shows that the majority of the firms happen to be single R&D performing enterprises. At least for these firms the above procedure summarized by equation (1), is not applicable. For our sample, the percentage of R&D firms with single R&D performing enterprise are 78.97, 74.01, and 73.87 respectively for CIS2.5, CIS3, and CIS3.5.

The two R&D variables of interest for which the above procedure was used to get an estimate for the firm are R&D expenditure and the share of innovative sales in the total sales of the enterprise. Here we would like to mention that we do not have any information on R&D expenditure of those firms that have been categorized as non-innovator, while share of innovative sales by definition does not exist for non-innovating firms. In order to estimate $\bar{x}_s$ for R&D expenditure, for each stratum s, we have assumed that the R&D expenditure is zero for those enterprises that have been classified as non innovators in the survey [3].

While R&D expenditure and the share of innovative sales are continuous variables, we also had to infer about two binary variables for the firm, given the information about these for the enterprises constituting the firm. The two categorical variables happen to (1) a dummy variable on the enterprise being financially constrained with respect to R&D activities and (2) a dummy variable on the decision for an enterprise to be an innovator. Characterizing a firm to be financially constrained was straight forward – a firm is financially constrained if any one of the enterprise is financially constrained with respect to the R&D activities.

Before we discuss the criterion that characterizes a firm into being an innovator or a non-innovator using information on the constituent enterprises being a innovator or a non-innovator, we briefly mention the survey criteria that classifies an enterprise as an innovating or non-innovating enterprise. An enterprise is innovating if it satisfies one of the following criteria: (a) If the enterprise has introduced new product to the market, (b) If the enterprise has introduced new process to the market, (c) If the enterprise has some unfinished R&D project and (d) If the enterprise, began on a R&D project, and abandoned it during the time period that the survey covers. To characterize a firm as an innovator or a non-innovator the following criterion was used. If all the enterprises of a firm have been surveyed and if all the enterprises report themselves to be a non-innovator then the firm is classified as being a

[3] This assumption could possibly lead to a bias in the estimates of R&D expenditure for those firms, for which not all enterprises have been surveyed.

non-innovator. If the number of enterprises present in a firm are more than the number of enterprises that have been surveyed, that is, if $N_f > n_f$, and if one the enterprises surveyed, reports that it has engaged in R&D activities, then we classify this firm to be an innovator. However, if $N_f > n_f$, and if none of the enterprises sampled report to be an innovator, then if any one of the remaining enterprises that have not been surveyed is to be found in a stratum that, based on the CIS survey, has been classified as an innovating stratum, then the firm has been classified as an innovator.

The total number of employees as a measure of the size of the firm was also constructed using information from the CIS data and the General Business Register. As far as the number of employees in a firm is concerned, if all the enterprises belonging to a firm are surveyed, that is if $N_f = n_f$, then we simply add up the number employees of each of the constituent enterprise. However, when $N_f > n_f$, for those enterprise/s that have not been surveyed we take the mid point of the size class of those enterprises that have not been surveyed. The size class to which an enterprise belongs to is available from the General Business Register for every year.

Table 6.2 below shows the number of innovators and number of non-innovators for each of the three waves, and the number firms that are financially constrained with respect to R&D activities[4]. We would also like to point out that most of the information contained

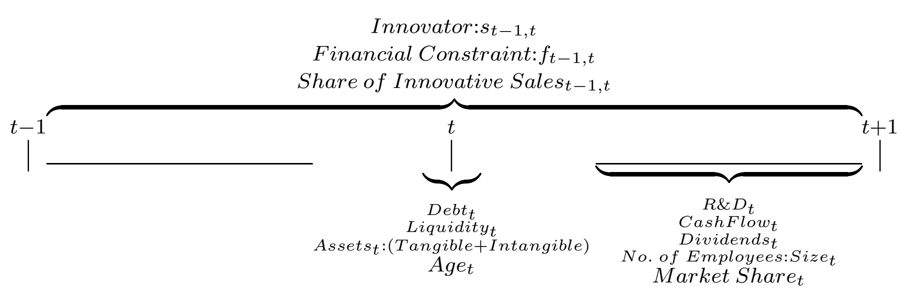
Table 2: **Total number Innovating Firms, Non-Innovating Firms, and Financially Constrained Firms**

	CIS2.5	CIS3	CIS3.5
No. of Non-Innovators	2,405	1,495	2,221
No. of Innovators	2,987	1,933	2,011
No. of Financially Constrained firms	536	416	188
These figures are for the Data Set used in Estimation			

in the Innovation Surveys, cover three time periods. For example, if any enterprise present in CIS3.5, classifies itself to being an innovator, then this implies that over the course of the three year 2000-02, the enterprise carried out any of the R&D related activities that are necessary for it to be classified as an innovator. The same is true of financial constraint, that is, an enterprise if it reports that it is financially constrained, then it is true that in either of the three years, covered by the survey, the enterprise found itself being financially constrained in the proper implementation of any of its R&D projects. This implies that there is an overlapping year between every CIS survey. Hence, if the reporting has to be consistent, then an enterprise reporting itself to be financially constrained, say in CIS3, covering years

[4]For CIS2.5 information on financial constraint is available only for the innovators.

1998-2000, should also report itself as being financially constrained in CIS3.5, covering the years 2000-2002, the year 2000 being the overlapping year. However, we do not find this to be the case. Therefore we assume that the CIS surveys cover two time periods, that is, say for CIS3.5, the time period covered are from the end of 2000 till the end of 2002, covering the years 2001 and 2002. Thus, if a firm reports itself as being innovative or being financially constrained, then we assume that it has been innovating or has been financially constrained for two time periods. The share of innovative sales in the total sales of the enterprise also covers two time periods. However, the figures on R&D expenditure are only for the last year of the two years covered by any of the innovation surveys. Thus, given the nature of the data, we make the assumption that the last year of the two years covered by any of the innovation surveys, is period t and the year preceding it, period $t-1$ and that, without knowing in which of the two periods it was an innovator and which of the two periods it was financially constrained, we make the assumption that the enterprise had been innovating in both periods and that it was financially constrained in both periods. Also, that the share of innovative sales in the total sales of the enterprise had been the same for the two time periods. The figure below elucidates the above discussion and the time line along which the stock and flow variables are realized[5]. The vertical lines, at $t-1$, t, and $t+1$, indicate the beginning of the respective time periods. Having drawn the time line and given the fact that, R&D expenditure is reported only for the last year of the two periods that any CIS covers, then according to the time line defined, this implies that, if the last two periods in our data is indexed T and $T-1$, the R&D expenditure is observed for time periods, T, but not period $T-1$, for time period $T-2$ but not for $T-3$ and so on.



R&D expenditure, the constructed dummy variable on a firm being financial constraint, using information on financial constraints at the enterprise level, and the constructed dummy variable indicating whether the firm is an innovator or not, are three among the four endogenous variable. Our fourth endogenous variable is long-term debt, about which we learn from the SFGO and SFKO data set, which as mentioned earlier are available at the firm level. We

[5]While the number of employees of a firm is a stock variable, the enterprises reported the figures on it for the accounting period, hence the figures on employment are the year end values on the number of employees.

discuss the set of our assumed exogenous variables in the text following this.

5 Empirical Strategy

5.1 R&D Investment, Endogenous Financial Constraint and Endogenous Selection with Endogenous Long Term Debt: A Three step Procedure

To study the choice of capital structure of R&D intensive firms, the effect of financial constraints on R&D expenditure and to account for the features of the data, where R&D expenditure is known only for firms that opt to innovate and we consider a four equation system

$$\begin{aligned} r_{it} &= \mathbf{z}_{it}^{r\prime}\beta + SINS_{it-1,t}\beta_s + \mu\alpha_i \\ &\quad + f_{it-1,t}\beta_f + (f_{it-1,t} \times CF_{it})\beta_{c1} + ((1 - f_{it-1,t}) \times CF_{it})\beta_{c0} \\ &\quad + (f_{it-1,t} \times DIV_{it})\beta_{D1} + ((1 - f_{it-1,t}) \times DIV_{it})\beta_{D0} + \eta_{it}, \end{aligned} \tag{2}$$

$$f_{it-1,t}^{*} = \mathbf{z}_{it}^{f\prime}\varphi + d_{it}\varphi_d + \lambda\alpha_i + \zeta_{it}, \tag{3}$$

$$s_{it-1,t}^{*} = \mathbf{z}_{it}^{s\prime}\gamma + d_{it}\gamma_d + \theta\alpha_i + \upsilon_{it}, \tag{4}$$

$$d_{it} = \mathbf{z}_{it}^{\prime}\delta + \alpha_i + \epsilon_{it}, \tag{5}$$

where r_{it} in equation (2) is the ratio of total R&D expenditure to total capital, (tangible + intangible), asset of the firm and $\mathbf{z}_{it}^{r}$, is vector of strictly exogenous variables. We term equation (2) as the R&D equation. $f_{it-1,t}$, a binary variable that indicates, with value 1, if the firm is financially constrained with respect to innovation or R&D activities in period t and $t-1$. CF_{it} is the cash flows of the firm in period t and DIV_{it} is a dummy variable for positive dividends. We interact, CF_{it} and DIV_{it} with the dummy variables $f_{it-1,t}$ and $1 - f_{it-1,t}$. $SINS_{it-1,t}$ is the share of innovative sales in the total sales of the firm i, for the period $t-1$ and period t. A detailed discussion of the specification for the R&D equation is carried out later when we discuss the third stage estimation.

$f_{it-1,t}^{*}$ is the latent variable underlying $f_{it-1,t}$, that is $f_{it-1,t}$ takes value 1 if $f_{it-1,t}^{*}$ crosses a certain threshold. We term, equation (3) as the financial constraint equation. $\mathbf{z}_{it}^{f}$ is a vector of strictly exogenous variables in equation (3). The specification for financial constraint includes the ratio of outstanding long term debt to book value of capital (tangible + intangible) asset, d_{it}, which is endogenous since the choice of leverage and the financial constraint facing the firm, if we interpret the underlying latent variable $f_{it-1,t}^{*}$ as the shadow price of debt, are determined simultaneously. As we have mentioned above, R&D investment is observed only for the innovators. To rule out possible sample selection bias that could arise because some component (observed or unobserved) of the decision to innovate also determines the outcome – here R&D expenditure – we specify a selection equation, equation (4), where the idiosyncratic error term appearing in the selection equation is correlated with the idiosyncratic error terms appearing in the R&D equation and the financial constraint equations. $s^*_{it-1,t}$ is the latent variable underlying the decision to innovate, $s_{it-1,t}$, which takes value 1 if the firm decides to innovate and 0 otherwise. $s_{it-1,t}$ takes value 1 if $s^*_{it-1,t}$ crosses a certain threshold, which could be 0. $\mathbf{z}^s_i$, is a vector of strictly exogenous variables. The specification for the selection equation includes the endogenous variable, d_{it}. Equation (5) specifies the reduced form for debt to capital ratio, where $\mathbf{z}_{it}$ is a vector of exogenous variables appearing in (5).

Our discussion on R&D investment, financial constraints and debt limits, and the discussion on the decision to innovate and choice of capital structure in the preceding two sections imply that R&D investment, financial constraints, the decision to innovate and leverage are determined simultaneously. Here we also mention that the specification for the financial constraint equation and the selection equation include the stock of liquid assets, LQ_{it}, dummy for dividend payout, DIV_{it}, and share of innovative sales. While these variables could possibly be endogenous or predetermined to the system of equations,(2), (3), (4), and (5), here we assume them to be exogenously given. Hence, subsequently we will assume $\mathbf{z}^r_{it}$, $\mathbf{z}^f_{it}$, $\mathbf{z}^s_{it}$, and $\mathbf{z}_{it}$ to include LQ_{it}, DIV_{it}, and $SINS_{it-1,t}$ [6].

Let Z_{it} be the union of the exogenous variables appearing in $\mathbf{z}^r_{it}$, $\mathbf{z}^f_{it}$, $\mathbf{z}^s_{it}$, and $\mathbf{z}_{it}$ and $Z_i = (Z'_{i1} \dots Z'_{iT_i})'$. α_i is the unobserved individual effect, which we model as a random effect that is normally distributed with mean 0 and variance σ_α, is assumed to be independent of Z_i and is also assumed to be mean independent of η_{it}, ζ_{it}, υ_{it} and ϵ_{it}. Now, it is possible that the unobserved individual specific effect affecting leverage, that is, the choice of capital structure also affects the decision to innovate, the firm's realization that it is financially constrained with respect to R&D activity, and how much R&D expenditure to incur. The factor loadings, μ, λ and θ allow for such a possibility.

The endogeneity of the financial constraint, long-term debt and selection are captured

[6] As an extension of this work we intend to endogenize dividends, share of innovative sales and liquidity by specifying a reduced form for all these variables along with long term debt and condition the idiosyncratic error term of the structural equations (2), (3), and (4) on the vector of reduced form idiosyncratic error term which are correlated to those of the structural form.

through the following error structure, where conditional on Z_i,

$$\begin{pmatrix} \eta_{it} \\ \zeta_{it} \\ \upsilon_{it} \\ \epsilon_{it} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} \sigma_\eta^2 & \rho_{\eta\zeta}\sigma_\eta\sigma_\zeta & \rho_{\eta\upsilon}\sigma_\eta\sigma_\upsilon & \rho_{\eta\epsilon}\sigma_\eta\sigma_\epsilon \\ & \sigma_\zeta^2 & \rho_{\zeta\upsilon}\sigma_\zeta\sigma_\upsilon & \rho_{\zeta\epsilon}\sigma_\zeta\sigma_\epsilon \\ & & \sigma_\upsilon^2 & \rho_{\upsilon\epsilon}\sigma_\upsilon\sigma_\epsilon \\ & & & \sigma_\epsilon^2 \end{pmatrix} \right].$$

We also assume that, each of the error terms, η_{it}, ζ_{it}, υ_{it}, and ϵ_{it} are independently and identically distributed.

We extend the model for R&D expenditure, equation (2), to that of endogenous switching regression model. In this model a switching equation, here the financial constraint equation, sorts the firms over two different regimes, depending on whether $I_{f^*_{it-1,t}>0}$, with only one regime observed. $I_{f^*_{it-1,t}>0}$ is an indicator variable, that takes value 1 if $f^*_{it-1,t} > 0$. Thus, we have

$$r_{1it} = \mathbf{z}^{r\prime}_{it}\beta + \beta_f + CF_{it}\beta_{c1} + DIV_{it}\beta_{D1} + SINS_{it-1,t}\beta_s + \mu\alpha_i + \eta_{1it}, \tag{2a}$$

if $f^*_{it-1,t} > 0$, and

$$r_{0it} = \mathbf{z}^{r\prime}_{it}\beta + CF_{it}\beta_{c0} + DIV_{it}\beta_{D0} + SINS_{it-1,t}\beta_s + \mu\alpha_i + \eta_{0it}, \tag{2b}$$

if $f^*_{it-1,t} \leq 0$, where η_{1it} denotes the distribution of the idiosyncratic term in regime 1, when the firm is financially constrained and η_{0it} denotes the distribution when the firm is not, regime 0. Thus, we have $r_{it} = f_{it-1,t}r_{1it} + (1 - f_{it-1,t})r_{0it}$

$$\begin{aligned} r_{it} &= f_{it-1,t}r_{1it} + (1 - f_{it-1,t})r_{0it} \\ &= \mathbf{z}^{r\prime}_{it}\beta + f_{it-1,t}\beta_f + f_{it-1,t}\beta_1(.) + (1 - f_{it-1,t})\beta_0(.) + \mu\alpha_i + f_{it-1,t}\eta_{1it} + (1 - f_{it-1,t})\eta_{0it}, \end{aligned} \tag{6}$$

where $\beta_1(.) = (CF_{it}\beta_{c1} + DIV_{it}\beta_{D1})$ and $\beta_0(.) = (CF_{it}\beta_{c0} + DIV_{it}\beta_{D0})$ and $SINS_{it-1,t}$ has been subsumed into $\mathbf{z}^{r\prime}_{it}$.

The distribution of the error terms of the system of equations (2a), (2b), (3), (4), and (5) is now given by:

$$\begin{pmatrix} \eta_{1it} \\ \eta_{0it} \\ \zeta_{it} \\ \upsilon_{it} \\ \epsilon_{it} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} \sigma_{\eta_1}^2 & \rho_{\eta_1\eta_0}\sigma_{\eta_1}\sigma_{\eta_0} & \rho_{\eta_1\zeta}\sigma_{\eta_1}\sigma_\zeta & \rho_{\eta_1\upsilon}\sigma_{\eta_1}\sigma_\upsilon & \rho_{\eta_1\epsilon}\sigma_{\eta_1}\sigma_\epsilon \\ & \sigma_{\eta_0}^2 & \rho_{\eta_0\zeta}\sigma_{\eta_0}\sigma_\zeta & \rho_{\eta_0\upsilon}\sigma_{\eta_0}\sigma_\upsilon & \rho_{\eta_0\epsilon}\sigma_{\eta_0}\sigma_\epsilon \\ & & \sigma_\zeta^2 & \rho_{\zeta\upsilon}\sigma_\zeta\sigma_\upsilon & \rho_{\zeta\epsilon}\sigma_\zeta\sigma_\epsilon \\ & & & \sigma_\upsilon^2 & \rho_{\upsilon\epsilon}\sigma_\upsilon\sigma_\epsilon \\ & & & & \sigma_\epsilon^2 \end{pmatrix} \right].$$

We note here that $\rho_{\eta_1\eta_0}$ is not identified, since we do not observe the pair (r_{1it}, r_{0it}) together, but only either one of them. We also restrict σ_ζ^2 and σ_υ^2 to be 1.

To estimate the above model, given by equation (2a), (2b), (3), (4), and (5), we use a three step estimation procedure as an extension of Heckman's classical two step estimation to multivariate selection problems. Heckman (1979) corrects the bias caused by the sample selection using the control function approach, namely, by adding the inverse Mills ratio to main regression equation, obtained from the first stage selection equation. Here we are dealing with two selection problems. One is the endogenous switching and the other one is the sample selection, not to mention, also, the endogeneity of long term debt. To consistently estimate the parameters of equations (2a) and (2b), in Appendix B, we derive the the three correction terms. The first term corrects for the bias due to the endogeneity of long-term debt, while the other two correct for the bias due to endogenous switching and the bias due to endogenous sample selection. These correction terms are obtained for each firm-year observations. Adding these correction terms for each observation, we obtain consistent estimates for the structural equations, (2a) and (2b), using corresponding subsamples.

To define these correction terms, we first consider the conditional distribution of r_{it}, $f^*_{it-1,t}$ and $s^*_{it-1,t}$ given ϵ_{it}. Since η_{1it}, η_{0it}, ζ_{it}, υ_{it} and ϵ_{it} follow a joint normal distribution, the linear projections of η_{1it}, η_{0it}, ζ_{it} and υ_{it} in error form, given ϵ_{it} are given by

$$\eta_{1it} = \rho_{\eta_1\epsilon}\frac{\sigma_{\eta_1}}{\sigma_\epsilon}\epsilon_{it} + \bar{\eta}_{1it}, \quad \eta_{0it} = \rho_{\eta_0\epsilon}\frac{\sigma_{\eta_0}}{\sigma_\epsilon}\epsilon_{it} + \bar{\eta}_{0it}, \quad \zeta_{it} = \rho_{\zeta\epsilon}\frac{\epsilon_{it}}{\sigma_\epsilon} + \bar{\zeta}_{it}, \quad \upsilon_{it} = \rho_{\upsilon\epsilon}\frac{\epsilon_{it}}{\sigma_\epsilon} + \bar{\upsilon}_{it}, \quad (7)$$

where

$$\begin{pmatrix}\bar{\eta}_{1it}\\ \bar{\eta}_{0it}\\ \bar{\zeta}_{it}\\ \bar{\upsilon}_{it}\end{pmatrix} \sim N\left[\begin{pmatrix}0\\0\\0\\0\end{pmatrix}\begin{pmatrix}\sigma^2_{\bar{\eta}_1} & \rho_{\bar{\eta}_1\bar{\eta}_0}\sigma_{\bar{\eta}_1}\sigma_{\bar{\eta}_0} & \rho_{\bar{\eta}_1\bar{\zeta}}\sigma_{\bar{\eta}_1}\sigma_{\bar{\zeta}} & \rho_{\bar{\eta}_1\bar{\upsilon}}\sigma_{\bar{\eta}_1}\sigma_{\bar{\upsilon}}\\ & \sigma^2_{\bar{\eta}_0} & \rho_{\bar{\eta}_0\bar{\zeta}}\sigma_{\bar{\eta}_0}\sigma_{\bar{\zeta}} & \rho_{\bar{\eta}_0\bar{\upsilon}}\sigma_{\bar{\eta}_0}\sigma_{\bar{\upsilon}}\\ & & \sigma^2_{\bar{\zeta}} & \rho_{\bar{\zeta}\bar{\upsilon}}\sigma_{\bar{\zeta}}\sigma_{\bar{\upsilon}}\\ & & & \sigma^2_{\bar{\upsilon}}\end{pmatrix}\right],$$

and are independent of Z_i d_i and α_i[7].

The above then implies that the distribution of r_{1it}, r_{0it}, $f^*_{it-1,t}$ and $s^*_{it-1,t}$ given ϵ_{it} is then given by

$$r_{1it} = \mathbf{z}^{r\prime}_{it}\beta + \beta_f + \beta_1(.) + \mu\alpha_i + \rho_{\eta_1\epsilon}\frac{\sigma_{\eta_1}}{\sigma_\epsilon}\epsilon_{it} + \bar{\eta}_{1it}, \quad (8)$$

if $f^*_{it-1,t} > 0$,

$$r_{0it} = \mathbf{z}^{r\prime}_{it}\beta + \beta_0(.) + \mu\alpha_i + \rho_{\eta_0\epsilon}\frac{\sigma_{\eta_0}}{\sigma_\epsilon}\epsilon_{it} + \bar{\eta}_{0it} \quad (9)$$

if $f^*_{it-1,t} \leq 0$, and

$$f^*_{it-1,t} = \mathbf{z}^{f\prime}_{it}\varphi + d_{it}\varphi_d + \lambda\alpha_i + \rho_{\zeta\epsilon}\frac{\epsilon_{it}}{\sigma_\epsilon} + \bar{\zeta}_{it} = \varphi(\mathbf{z}^{f\prime}_{it}, d_{it}, \alpha_i) + \rho_{\zeta\epsilon}\frac{\epsilon_{it}}{\sigma_\epsilon} + \bar{\zeta}_{it}, \quad (10)$$

[7] Given the joint distribution of η_{1it}, η_{0it}, ζ_{it}, υ_{it} and ϵ_{it}, the conditional variance covariance matrix of η_{1it}, η_{0it}, ζ_{it}, and υ_{it}, given ϵ_{it} can be computed. However, as we will see later the elements of this conditional variance covariance matrix are not estimated.

and

$$s^*_{it-1,t} = \mathbf{z}^{s\prime}_{it}\gamma + d_{it}\gamma_d + \theta\alpha_i + \rho_{\upsilon\epsilon}\frac{\epsilon_{it}}{\sigma_\epsilon} + \bar{\upsilon}_{it} = \gamma(\mathbf{z}^{s\prime}_{it}, d_{it}, \alpha_i) + \rho_{\upsilon\epsilon}\frac{\epsilon_{it}}{\sigma_\epsilon} + \bar{\upsilon}_{it}. \tag{11}$$

Conditioning the endogenous variables of interest on ϵ_{it} allows us to control for the endogeneity of d_{it}, the long-term debt to asset ratio. To estimate the above system of equation, the standard technique is to replace ϵ_{it} by the residuals from the first stage reduced form regression, here equation (5). However, since $\mathrm{E}(\epsilon_{it}|d_i, Z_i, \alpha_i) = d_{it} - \mathbf{z}'_{it}\delta - \alpha_i$, the residuals from the first stage regression are not identified, this is because the α_i's remain unobserved, even though δ and σ_ϵ can be consistently estimated. To estimate the above system of equations, we show in Appendix A that

$$\begin{aligned} E(r_{1it}|d_i, Z_i) &= \int \mathrm{E}(r_{1it}|d_i, Z_i, \alpha_i) f(\alpha_i|d_i, Z_i) d\alpha_i = \mathbf{z}^{r\prime}_{it}\beta + \beta_f + \beta_1(.) + \mu\hat{\alpha}_i + \rho_{\eta_1\epsilon}\frac{\sigma_{\eta_1}}{\sigma_\epsilon}\hat{\epsilon}_{it} \\ &= \mathrm{E}(r_{1it}|d_i, Z_i, \hat{\alpha}_i, \hat{\epsilon}_{it}) \end{aligned} \tag{8a}$$

if $f^*_{it-1,t} > 0$, and

$$\begin{aligned} E(r_{0it}|d_i, Z_i) &= \int \mathrm{E}(r_{0it}|d_i, Z_i, \alpha_i) f(\alpha_i|d_i, Z_i) d\alpha_i = \mathbf{z}^{r\prime}_{it}\beta + \beta_0(.) + \mu\hat{\alpha}_i + \rho_{\eta_0\epsilon}\frac{\sigma_{\eta_0}}{\sigma_\epsilon}\hat{\epsilon}_{it} \\ &= \mathrm{E}(r_{0it}|d_i, Z_i, \hat{\alpha}_i, \hat{\epsilon}_{it}) \end{aligned} \tag{9a}$$

if $f^*_{it-1,t} \leq 0$.

$$\begin{aligned} E(f^*_{it-1,t}|d_i, Z_i) &= \int \mathrm{E}(f^*_{it-1,t}|d_i, Z_i, \alpha_i) f(\alpha_i|d_i, Z_i) d\alpha_i = \varphi(\mathbf{z}^{f\prime}_{it}, d_{it}, \hat{\alpha}_i) + \rho_{\zeta\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon} \\ &= \mathrm{E}(f^*_{it-1,t}|d_i, Z_i, \hat{\alpha}_i, \hat{\epsilon}_{it}), \end{aligned} \tag{10a}$$

and

$$\begin{aligned} E(s^*_{it-1,t}|d_i, Z_i) &= \int \mathrm{E}(s^*_{it-1,t}|d_i, Z_i, \alpha_i) f(\alpha_i|d_i, Z_i) d\alpha_i = \gamma(\mathbf{z}^{s\prime}_{it}, d_{it}, \hat{\alpha}_i) + \rho_{\upsilon\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon} \\ &= \mathrm{E}(s^*_{it-1,t}|d_i, Z_i, \hat{\alpha}_i, \hat{\epsilon}_{it}), \end{aligned} \tag{11a}$$

where $\hat{\epsilon}_{it} = (d_{it} - \mathbf{z}'_{it}\delta - \hat{\alpha}_i)$ and $\hat{\alpha}_i \equiv \hat{\alpha}_i(d_i, Z_i, \delta, \sigma_\epsilon, \sigma_\alpha)$ is the expected a posteriori value of α_i, that are based on the results of the first stage estimation, which we discuss in a later section. To estimate the expected values of r_{1it}, r_{0it}, $f^*_{it-1,t}$ and $s^*_{it-1,t}$ given d_i and Z_i we write the linear projection of r_{1it}, r_{0it}, $f^*_{it-1,t}$ and $s^*_{it-1,t}$ in error form given d_i and Z_i respectively as:

$$r_{it} = f_{it-1,t}r_{1it} + (1 - f_{it-1,t})r_{0it}, \tag{12}$$

where

$$r_{1it} = \mathbf{z}^{r\prime}_{it}\beta + \beta_f + \beta_1(.) + \mu\hat{\alpha}_i + \rho_{\eta_1\epsilon}\frac{\sigma_{\eta_1}}{\sigma_\epsilon}\hat{\epsilon}_{it} + \underline{\eta}_{1it}, \tag{12a}$$

if $f^*_{it-1,t} > 0$,

$$r_{0it} = \mathbf{z}^{r\prime}_{it}\beta + \beta_0(.) + \mu\hat{\alpha}_i + \rho_{\eta_0\epsilon}\frac{\sigma_{\eta_0}}{\sigma_\epsilon}\hat{\epsilon}_{it} + \underline{\eta}_{0it}, \tag{12b}$$

if $f^*_{it-1,t} \leq 0$.

$$f^*_{it-1,t} = \varphi(\mathbf{z}^{f\prime}_{it}, d_{it}, \hat{\alpha}_i) + \rho_{\zeta\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon} + \underline{\zeta}_{it}, \tag{13}$$

and

$$s^*_{it-1,t} = \gamma(\mathbf{z}^{s\prime}_{it}, d_{it}, \hat{\alpha}_i) + \rho_{\upsilon\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon} + \underline{\upsilon}_{it}, \tag{14}$$

where $\underline{\eta}_{1it}$, $\underline{\eta}_{0it}$ $\underline{\zeta}_{it}$ and $\underline{\upsilon}_{it}$ have been defined in Appendix A. The variance covariance matrix of the error component in (12) to (14) is given by,

$$\begin{pmatrix} \underline{\eta}_{1it} \\ \underline{\eta}_{0it} \\ \underline{\zeta}_{it} \\ \underline{\upsilon}_{it} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}\begin{pmatrix} \sigma^2_{\underline{\eta}_1} & \rho_{\underline{\eta}_1\underline{\eta}_0}\sigma_{\underline{\eta}_1}\sigma_{\underline{\eta}_0} & \rho_{\underline{\eta}_1\underline{\zeta}}\sigma_{\underline{\eta}_1}\sigma_{\underline{\zeta}} & \rho_{\underline{\eta}_1\underline{\upsilon}}\sigma_{\underline{\eta}_1}\sigma_{\underline{\upsilon}} \\ & \sigma^2_{\underline{\eta}_0} & \rho_{\underline{\eta_0\zeta}}\sigma_{\underline{\eta}_0}\sigma_{\underline{\zeta}} & \rho_{\underline{\eta_0\upsilon}}\sigma_{\underline{\eta}_0}\sigma_{\underline{\upsilon}} \\ & & \sigma^2_{\underline{\zeta}} & \rho_{\underline{\zeta\upsilon}}\sigma_{\underline{\zeta}}\sigma_{\underline{\upsilon}} \\ & & & \sigma^2_{\underline{\upsilon}} \end{pmatrix}\right],$$

where the respective elements of the matrix have been stated in Appendix A. Again as stated earlier, $\rho_{\underline{\eta}_1\underline{\eta}_0}$, cannot be identified since a firm cannot be at the same time in either of the two regimes. We also note here that, having specified the conditional distribution of r_{1it}, r_{0it}, $f^*_{it-1,t}$ and $s^*_{it-1,t}$ given d_i and Z_i, we can no longer estimate the error structure for $\bar{\eta}_{1it}$, $\bar{\eta}_{0it}$, $\bar{\zeta}_{it}$ and $\bar{\upsilon}_{it}$, though the parameters of interest appearing in equations (12)-(14) can be identified.

To define the correction terms, consider the following conditional mean, $\mathrm{E}(r_{it}|f^*_{it-1,t}, s^*_{it-1,t} > 0, Z_i, d_i, \hat{\alpha}_i, \hat{\epsilon}_{it})$:

$$\begin{aligned} &\mathrm{E}(r_{it}|f^*_{it-1,t}, s^*_{it-1,t} > 0, Z_i, d_i, \hat{\alpha}_i, \hat{\epsilon}_{it}) = \mathrm{E}(f_{it-1,t}r_{1it} + (1-f_{it-1,t})r_{0it}|f^*_{it-1,t}, s^*_{it-1,t} > 0, Z_i, d_i, \hat{\alpha}_i, \hat{\epsilon}_{it}) \\ &= \mathbf{z}^{r\prime}_{it}\beta + f_{it-1,t}\beta_f + f_{it-1,t}\left(\beta_1(.) + \rho_{\eta_1\epsilon}\frac{\sigma_{\eta_1}}{\sigma_\epsilon}\hat{\epsilon}_{it}\right) + (1-f_{it-1,t})\left(\beta_0(.) + \rho_{\eta_0\epsilon}\frac{\sigma_{\eta_0}}{\sigma_\epsilon}\hat{\epsilon}_{it}\right) + \mu\hat{\alpha}_i \\ &+ f_{it-1,t}\mathrm{E}(\underline{\eta}_{1it}|f^*_{it-1,t} > 0, s^*_{it-1,t} > 0, Z_i, d_i, \hat{\alpha}_i, \hat{\epsilon}_{it}) + (1-f_{it-1,t})\mathrm{E}(\underline{\eta}_{0it}|f^*_{it-1,t} \leq 0, s^*_{it-1,t} > 0, Z_i, d_i, \hat{\alpha}_i, \hat{\epsilon}_{it}) \end{aligned} \tag{15}$$

We known that

$$E(\underline{\eta}_{1it}|f^*_{it-1,t} > 0, s^*_{it-1,t} > 0, Z_i, d_i, \hat{\alpha}_i, \hat{\epsilon}_{it}) = \mathrm{E}[\underline{\eta}_{1it}|\underline{\zeta}_{it} > -\varphi(\mathbf{z}^{f\prime}_{it}, d_{it}, \hat{\alpha}_i) - \rho_{\zeta\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon}, \underline{\upsilon}_{it} > -\gamma(\mathbf{z}^{s\prime}_{it}, d_{it}, \hat{\alpha}_i) - \rho_{\upsilon\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon}],$$

and that

$$E(\underline{\eta}_{0it}|f^*_{it-1,t} \leq 0, s^*_{it-1,t} > 0, Z_i, d_i, \hat{\alpha}_i, \hat{\epsilon}_{it}) = \mathrm{E}[\underline{\eta}_{0it}|\underline{\zeta}_{it} \leq -\varphi(\mathbf{z}^{f\prime}_{it}, d_{it}, \hat{\alpha}_i) - \rho_{\zeta\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon}, \underline{\upsilon}_{it} > -\gamma(\mathbf{z}^{s\prime}_{it}, d_{it}, \hat{\alpha}_i) - \rho_{\upsilon\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon}],$$

Define $\hat{\mu}_\zeta = \rho_{\zeta\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon}$ and $\hat{\mu}_\upsilon = \rho_{\upsilon\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon}$. It has been shown in Appendix B that

$$\mathrm{E}[\underline{\eta}_{1it}|\underline{\zeta}_{it} > -\varphi(\mathbf{z}_{it}^{f\prime}, d_{it}, \hat{\alpha}_i) - \hat{\mu}_\zeta, \underline{\upsilon}_{it} > -\gamma(\mathbf{z}_{it}^{s\prime}, d_{it}, \hat{\alpha}_i) - \hat{\mu}_\upsilon] =$$
$$\sigma_{\underline{\eta}_1}\rho_{\underline{\eta}_1\underline{\zeta}}\phi\Big(\frac{\varphi(.)+\hat{\mu}_\zeta}{\sigma_{\underline{\zeta}}}\Big)\frac{\Phi\Big((\frac{\gamma(.)+\hat{\mu}_\upsilon}{\sigma_{\underline{\upsilon}}} - \rho_{\underline{\zeta}\underline{\upsilon}}\frac{\varphi(.)+\hat{\mu}_\zeta}{\sigma_{\underline{\zeta}}})/\sqrt{1-\rho^2_{\underline{\zeta}\underline{\upsilon}}}\Big)}{\Phi_2\Big(\frac{\varphi(.)+\hat{\mu}_\zeta}{\sigma_{\underline{\zeta}}}, \frac{\gamma(.)+\hat{\mu}_\upsilon}{\sigma_{\underline{\upsilon}}}, \rho_{\underline{\zeta}\underline{\upsilon}}\Big)}$$
$$+\sigma_{\underline{\eta}_1}\rho_{\underline{\eta}_1\underline{\upsilon}}\phi\Big(\frac{\gamma(.)+\hat{\mu}_\upsilon}{\sigma_{\underline{\upsilon}}}\Big)\frac{\Phi\Big((\frac{\varphi(.)+\hat{\mu}_\zeta}{\sigma_{\underline{\zeta}}} - \rho_{\underline{\zeta}\underline{\upsilon}}\frac{\gamma(.)+\hat{\mu}_\upsilon}{\sigma_{\underline{\upsilon}}})/\sqrt{1-\rho^2_{\underline{\zeta}\underline{\upsilon}}}\Big)}{\Phi_2\Big(\frac{\varphi(.)+\hat{\mu}_\zeta}{\sigma_{\underline{\zeta}}}, \frac{\gamma(.)+\hat{\mu}_\upsilon}{\sigma_{\underline{\upsilon}}}, \rho_{\underline{\zeta}\underline{\upsilon}}\Big)} \tag{16}$$

and

$$\mathrm{E}[\underline{\eta}_{0it}|\underline{\zeta}_{it} \leq -\varphi(\mathbf{z}_{it}^{f\prime}, d_{it}, \hat{\alpha}_i) - \hat{\mu}_\zeta, \underline{\upsilon}_{it} > -\gamma(\mathbf{z}_{it}^{s\prime}, d_{it}, \hat{\alpha}_i) - \hat{\mu}_\upsilon] =$$
$$-\sigma_{\underline{\eta}_0}\rho_{\underline{\eta}_0\underline{\zeta}}\phi\Big(\frac{\varphi(.)+\hat{\mu}_\zeta}{\sigma_{\underline{\zeta}}}\Big)\frac{\Phi\Big((\frac{\gamma(.)+\hat{\mu}_\upsilon}{\sigma_{\underline{\upsilon}}} - \rho_{\underline{\zeta}\underline{\upsilon}}\frac{\varphi(.)+\hat{\mu}_\zeta}{\sigma_{\underline{\zeta}}})/\sqrt{1-\rho^2_{\underline{\zeta}\underline{\upsilon}}}\Big)}{\Phi_2\Big(-\frac{\varphi(.)+\hat{\mu}_\zeta}{\sigma_{\underline{\zeta}}}, \frac{\gamma(.)+\hat{\mu}_\upsilon}{\sigma_{\underline{\upsilon}}}, -\rho_{\underline{\zeta}\underline{\upsilon}}\Big)}$$
$$+\sigma_{\underline{\eta}_0}\rho_{\underline{\eta}_0\underline{\upsilon}}\phi\Big(\frac{\gamma(.)+\hat{\mu}_\upsilon}{\sigma_{\underline{\upsilon}}}\Big)\frac{\Phi\Big((-\frac{\varphi(.)+\hat{\mu}_\zeta}{\sigma_{\underline{\zeta}}} + \rho_{\underline{\zeta}\underline{\upsilon}}\frac{\gamma(.)+\hat{\mu}_\upsilon}{\sigma_{\underline{\upsilon}}})/\sqrt{1-\rho^2_{\underline{\zeta}\underline{\upsilon}}}\Big)}{\Phi_2\Big(-\frac{\varphi(.)+\hat{\mu}_\zeta}{\sigma_{\underline{\zeta}}}, \frac{\gamma(.)+\hat{\mu}_\upsilon}{\sigma_{\underline{\upsilon}}}, -\rho_{\underline{\zeta}\underline{\upsilon}}\Big)}, \tag{17}$$

where ϕ, Φ, and Φ_2, respectively denote the density function of a standard normal distribution, the cumulative distribution function of a standard normal, and the cumulative distribution function of a standardized bivariate normal.

Define the correction terms $C_{11}(.)_{it}$, $C_{12}(.)_{it}$, $C_{13}(.)_{it}$, $C_{01}(.)_{it}$, $C_{02}(.)_{it}$, and $C_{03}(.)_{it}$ respectively as

$$C_{11}(\Xi, Z_i, d_i, \hat{\alpha}_i, \hat{\epsilon}_{it})_{it} \equiv f_{it-1,t}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon} = f_{it-1,t}\frac{1}{\sigma_\epsilon}(d_{it} - \mathbf{z}_{it}'\delta - \hat{\alpha}_i), \tag{18}$$

$$C_{12}(\Xi, Z_i, d_i, \hat{\alpha}_i, \hat{\epsilon}_{it})_{it} \equiv f_{it-1,t}\phi\Big(\frac{\varphi(.)+\hat{\mu}_\zeta}{\sigma_{\underline{\zeta}}}\Big)\frac{\Phi\Big((\frac{\gamma(.)+\hat{\mu}_\upsilon}{\sigma_{\underline{\upsilon}}} - \rho_{\underline{\zeta}\underline{\upsilon}}\frac{\varphi(.)+\hat{\mu}_\zeta}{\sigma_{\underline{\zeta}}})/\sqrt{1-\rho^2_{\underline{\zeta}\underline{\upsilon}}}\Big)}{\Phi_2\Big(\frac{\varphi(.)+\hat{\mu}_\zeta}{\sigma_{\underline{\zeta}}}, \frac{\gamma(.)+\hat{\mu}_\upsilon}{\sigma_{\underline{\upsilon}}}, \rho_{\underline{\zeta}\underline{\upsilon}}\Big)}, \tag{19}$$

$$C_{13}(\Xi, Z_i, d_i, \hat{\alpha}_i, \hat{\epsilon}_{it})_{it} \equiv f_{it-1,t}\phi\Big(\frac{\gamma(.)+\hat{\mu}_\upsilon}{\sigma_{\underline{\upsilon}}}\Big)\frac{\Phi\Big((\frac{\varphi(.)+\hat{\mu}_\zeta}{\sigma_{\underline{\zeta}}} - \rho_{\underline{\zeta}\underline{\upsilon}}\frac{\gamma(.)+\hat{\mu}_\upsilon}{\sigma_{\underline{\upsilon}}})/\sqrt{1-\rho^2_{\underline{\zeta}\underline{\upsilon}}}\Big)}{\Phi_2\Big(\frac{\varphi(.)+\hat{\mu}_\zeta}{\sigma_{\underline{\zeta}}}, \frac{\gamma(.)+\hat{\mu}_\upsilon}{\sigma_{\underline{\upsilon}}}, \rho_{\underline{\zeta}\underline{\upsilon}}\Big)}, \tag{20}$$

$$C_{01}(\Xi, Z_i, d_i, \hat{\alpha}_i, \hat{\epsilon}_{it})_{it} \equiv (1 - f_{it-1,t})\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon} = (1 - f_{it-1,t})\frac{1}{\sigma_\epsilon}(d_{it} - \mathbf{z}_{it}'\delta - \hat{\alpha}_i), \tag{21}$$

$$C_{02}(\Xi, Z_i, d_i, \hat{\alpha}_i, \hat{\epsilon}_{it})_{it} \equiv (1-f_{it-1,t})\phi\bigg(\frac{\varphi(.)+\hat{\mu}_\zeta}{\sigma_{\underline{\zeta}}}\bigg)\frac{\Phi\bigg((\frac{\gamma(.)+\hat{\mu}_\upsilon}{\sigma_{\underline{\upsilon}}}-\rho_{\underline{\zeta\upsilon}}\frac{\varphi(.)+\hat{\mu}_\zeta}{\sigma_{\underline{\zeta}}})/\sqrt{1-\rho^2_{\underline{\zeta\upsilon}}}\bigg)}{\Phi_2\bigg(-\frac{\varphi(.)+\hat{\mu}_\zeta}{\sigma_{\underline{\zeta}}}, \frac{\gamma(.)+\hat{\mu}_\upsilon}{\sigma_{\underline{\upsilon}}}, -\rho_{\underline{\zeta\upsilon}}\bigg)}, \quad (22)$$

and

$$C_{03}(\Xi, Z_i, d_i, \hat{\alpha}_i, \hat{\epsilon}_{it})_{it} \equiv (1-f_{it-1,t})\phi\bigg(\frac{\gamma(.)+\hat{\mu}_\upsilon}{\sigma_{\underline{\upsilon}}}\bigg)\frac{\Phi\bigg((-\frac{\varphi(.)+\rho_{\zeta\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon}}{\sigma_{\underline{\zeta}}}+\rho_{\underline{\zeta\upsilon}}\frac{\gamma(.)+\hat{\mu}_\upsilon}{\sigma_{\underline{\upsilon}}})/\sqrt{1-\rho^2_{\underline{\zeta\upsilon}}}\bigg)}{\Phi_2\bigg(-\frac{\varphi(.)+\hat{\mu}_\zeta}{\sigma_{\underline{\zeta}}}, \frac{\gamma(.)+\hat{\mu}_\upsilon}{\sigma_{\underline{\upsilon}}}, -\rho_{\underline{\zeta\upsilon}}\bigg)}. \quad (23)$$

With the correction terms defined, we can now write the R&D switching equation (12), conditional on $f^*_{it-1,t}$, $s^*_{it-1,t} > 0$, Z_i, d_i, $\hat{\alpha}_i$, $\hat{\epsilon}_{it}$ as

$$\begin{aligned} r_{it} &= \mathbf{z}^{r\prime}_{it}\beta + f_{it-1,t}\beta_f + f_{it-1,t}\beta_1(.) + (1-f_{it-1,t})\beta_0(.) + \mu\hat{\alpha}_i \\ &\quad + f_{it-1,t}\bigg(\rho_{\eta_1\epsilon}\sigma_{\eta_1}C_{11}(.)_{it} + \rho_{\underline{\eta}_1\underline{\zeta}}\sigma_{\underline{\eta}_1}C_{12}(.)_{it} + \rho_{\underline{\eta}_1\underline{\upsilon}}\sigma_{\underline{\eta}_1}C_{13}(.)_{it}\bigg) \\ &\quad + (1-f_{it-1,t})\bigg(\rho_{\eta_0\epsilon}\sigma_{\eta_0}C_{01}(.)_{it} + \rho_{\underline{\eta}_0\underline{\zeta}}\sigma_{\underline{\eta}_0}C_{02}(.)_{it} + \rho_{\underline{\eta}_0\underline{\upsilon}}\sigma_{\underline{\eta}_0}C_{03}(.)_{it}\bigg) + \tilde{\eta}_{it}, \end{aligned} \quad (24)$$

where $\tilde{\eta}_{it}$ conditional on $f^*_{it-1,t}$, $s^*_{it-1,t}$, Z_i and d_i is distributed with mean zero. The first and second stage of empirical strategy, gives us estimates of Ξ, for constructing the correction terms , where $\Xi = (\varphi\frac{1}{\sigma_{\underline{\zeta}}}, \frac{\varphi_d}{\sigma_{\underline{\zeta}}}, \frac{\lambda}{\sigma_{\underline{\zeta}}}; \gamma\frac{1}{\sigma_{\underline{\upsilon}}}, \frac{\gamma_d}{\sigma_{\underline{\upsilon}}}, \frac{\theta}{\sigma_{\underline{\upsilon}}}; \delta; \rho_{\underline{\zeta\upsilon}}, \frac{\rho_{\zeta\epsilon}}{\sigma_{\underline{\zeta}}}, \frac{\rho_{\upsilon\epsilon}}{\sigma_{\underline{\upsilon}}}, \sigma_\epsilon)$. Plugging these correction terms and $\hat{\alpha}_i$ in the R&D equation, (24), gives consistent estimates of the parameters of the switching model for R&D investment. These parameters can be obtained by running a simple pooled OLS for the sample of selected or innovating firms. $C_{11}(\Xi, .)_{it}$ and $C_{01}(\Xi, .)_{it}$ corrects for the potential bias that can arise due to endogeneity of long-term debt when estimating equation (24). While $C_{11}(\Xi, .)_{it}$ corrects for the bias in the subsample belonging to the regime that is financially constrained, $C_{01}(\Xi.,)_{it}$ corrects for the bias in the subsample firms that are not constrained. $C_{12}(\Xi, .)_{it}$ and $C_{13}(\Xi, .)_{it}$ corrects for the bias that could arise due endogenous switching and endogenous selection while estimating the R&D equation for the firms that are financially constrained, while $C_{02}(\Xi, .)_{it}$ and $C_{03}(\Xi, .)_{it}$ corrects for the bias that could arise due endogenous switching and endogenous selection for the subsample that are not financially constrained.

Apart from the parameters appearing in the R&D switching equations, (12a) and (12b), the other parameters that are estimated while estimating equation (24), are the coefficients of the correction terms. These coefficients are the estimates of the correlation between the unobserved factors affecting R&D expenditure in the two regimes and the unobserved factors that affect the chances of a firm to be financially constrained ($\sigma_{\underline{\eta}_1}\rho_{\underline{\eta}_1\underline{\zeta}}$, $\sigma_{\underline{\eta}_0}\rho_{\underline{\eta}_0\underline{\zeta}}$), the correlation between the unobserved factors affecting R&D expenditure and the unobserved factors affecting the choice of the firm to be an innovator ($\sigma_{\underline{\eta}_1}\rho_{\underline{\eta}_1\underline{\upsilon}}$, $\sigma_{\underline{\eta}_0}\rho_{\underline{\eta}_0\underline{\upsilon}}$), and of course, the correlation between the unobserved factors affecting R&D expenditure and the unobserved factors affecting the choice of capital structure, that is, the extent of leverage measured by long-term debt to capital asset ratio ($\rho_{\eta_0\epsilon}\sigma_{\eta_0}$, $\rho_{\eta_1\epsilon}\sigma_{\eta_1}$).

We estimate the parameters in Ξ in the first and second stage, which we discuss in detail below. In the first stage we estimate equation (5), the reduced form long-term debt equation. Using the estimates from the first stage regression, we jointly estimate equations (13) and (14), which are our conditional financial constraint equation and conditional selection equation. To estimate equations (13) and (14), we use the method of Maximum Likelihood. While such a procedure is quite common for cross sectional data analysis, in panel data we have to account for unobserved firm specific individual effect, α_i, which is correlated with endogenous regressors and appear in each of the equations, (2a), (2b), (3), (4) and (5). With $\hat{\alpha}_i$ in place, the unobserved time individual effect is controlled for and the factor loadings, μ, λ, and θ estimated. The details for estimating the expected a posteriori values of α_i are outlined in Appendix A.

5.1.1 The First Stage: Reduced Form for Long-term Debt

In the first stage of our econometric methodology we estimate equation (5), the reduced form equation for long-term debt.

$$d_{it} = \mathbf{z}'_{it}\delta + \alpha_i + \epsilon_{it} \tag{5}$$

The vector of exogenous variables included in $\mathbf{z}'_{it}$ are: (1) ratio of cash flows of the firm in period t to the capital assets (tangible +intangible) of the firm at the beginning of period t (CF_t), (2) ,market share of the firm measured ratio of the total sales of the firm to the industry total sales in period t, ($MKSH_t$), (3) logarithm of the number of people employed ,($SIZE_t$), (4) the age of the firm (AGE), (5) ratio of cash holdings of the firm to total capital assets (LQ_t), (6) dummy for positive dividends, (DIV_t), (7) dummy if a firm is a multi-enterprise firm, ($DMULTI_t$), (8) dummy for negative realization of cash flows, ($DNFC_t$), (9) share of innovative sales in the total sales of the firms,($SINS_{t-1,t}$), (10) industry or sectoral dummies and finally (11) year dummies.

We can get the estimates of the parameters of the reduced form model, equation (5), we can either use the simple random effects model or maximize the marginal likelihood function. The marginal likelihood function for individual i for the first stage estimation of the reduced form for the debt equation is given by

$$\boldsymbol{L}_i(\delta, \sigma_\epsilon, \sigma_\alpha) = \int_{-\infty}^{\infty} \prod_{t=1}^{T_i} \frac{1}{\sigma_\epsilon \sqrt{2\pi}} \exp\{\frac{(d_{it} - \mathbf{z}_{it}'\delta - \sigma_\alpha \tilde{\alpha}_i)^2}{2\sigma_\epsilon^2}\} \phi(\tilde{\alpha}_i) d\tilde{\alpha}_i \tag{25}$$

where $\tilde{\alpha}_i = \frac{\alpha_i}{\sigma_\alpha}$ and ϕ is the standard normal density function. The estimates of the first stage estimation, thus gives us the estimate of δ, σ_α and σ_ϵ, which we can use to get the estimates of the expected a posteriori values of α_i, $\hat{\alpha}_i$, given d_i and z_i, that would be subsequently used for the second and the third stage estimation.

5.1.2 The Second Stage: Financial Constraint and Selection with Endogenous Long-term Debt

In the second stage we estimate jointly the parameters of financial constraint equation (13) and the selection equation (14). The two equations, assuming liquidity, dividends, and share of innovative sales are exogenous, are stated below. While equation (5), the debt equation, has been written as a reduced form, the financial constraint equation and the selection equation are both structural equations.

$$f_{it-1,t}^* = \varphi(\mathbf{z}_{it}^{f\prime}, d_{it}, \hat{\alpha}_i) + \rho_{\zeta\epsilon} \frac{\hat{\epsilon}_{it}}{\sigma_\epsilon} + \underline{\zeta}_{it}, \tag{13}$$

$$s_{it-1,t}^* = \gamma(\mathbf{z}_{it}^{s\prime}, d_{it}, \hat{\alpha}_i) + \rho_{\upsilon\epsilon} \frac{\hat{\epsilon}_{it}}{\sigma_\epsilon} + \underline{\upsilon}_{it}, \tag{14}$$

where $\hat{\epsilon}_{it} = (d_{it} - \mathbf{z}_{it}'\delta - \hat{\alpha}_i)$ and $\hat{\alpha}_i \equiv \hat{\alpha}_i(d_i, Z_i, \delta, \sigma_\epsilon, \sigma_\alpha)$ is the expected a posteriori value of α_i defined at the population parameters. Given the joint distribution of $\underline{\eta}_{1it}$, $\underline{\eta}_{0it}$, $\underline{\zeta}_{it}$, $\underline{\upsilon}_{it}$, conditional on d_i and Z_i, the marginal distribution of $\underline{\zeta}_{it}$ and $\underline{\upsilon}_{it}$, follows a bivariate normal given by:

$$\begin{pmatrix} \underline{\zeta}_{it} \\ \underline{\upsilon}_{it} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} \sigma_{\underline{\zeta}}^2 & \sigma_{\underline{\zeta}} \sigma_{\underline{\upsilon}} \rho_{\underline{\zeta}\underline{\upsilon}} \\ & \sigma_{\underline{\upsilon}}^2 \end{pmatrix} \right].$$

Thus,

$$\begin{aligned} \Pr(f_{it-1,t} = 1, s_{it-1,t} = 1 | d_i, Z_i, \hat{\epsilon}_{it}, \hat{\alpha}_i) &= \Pr(f_{it-1,t}^* > 0, s_{it-1,t}^* > 0 | d_i, Z_i) \\ = \Pr(\underline{\zeta}_{it} > -\varphi(\mathbf{z}_{it}^{f\prime}, d_{it}, \hat{\alpha}_i) - \rho_{\upsilon\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon}, &\underline{\upsilon}_{it} > -\gamma(\mathbf{z}_{it}^{s\prime}, d_{it}, \hat{\alpha}_i) - \rho_{\upsilon\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon} | d_i, Z_i) \\ &= \Phi_2\Bigg(\frac{\varphi(.) + \rho_{\zeta\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon}}{\sigma_{\underline{\zeta}}}, \frac{\gamma(.) + \rho_{\upsilon\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon}}{\sigma_{\underline{\upsilon}}}, \rho_{\underline{\zeta}\underline{\upsilon}} \Bigg), \end{aligned} \tag{26}$$

where Φ_2 is the cumulative distribution function of a standard bivariate normal. Also,

$$\Pr(f_{it-1,t} = 0, s_{it-1,t} = 1 | d_i, Z_i, \hat{\epsilon}_{it}, \hat{\alpha}_i) = \Phi_2\Bigg(-\frac{\varphi(.) + \rho_{\zeta\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon}}{\sigma_{\underline{\zeta}}}, \frac{\gamma(.) + \rho_{\upsilon\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon}}{\sigma_{\underline{\upsilon}}}, -\rho_{\underline{\zeta}\underline{\upsilon}} \Bigg), \tag{27}$$

$$\Pr(f_{it-1,t}=1, s_{it-1,t}=0|d_i, Z_i, \hat{\epsilon}_{it}, \hat{\alpha}_i) = \Phi_2\bigg(\frac{\varphi(.)+\rho_{\zeta\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon}}{\sigma_{\underline{\zeta}}}, -\frac{\gamma(.)+\rho_{\upsilon\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon}}{\sigma_{\underline{\upsilon}}}, -\rho_{\underline{\zeta\upsilon}}\bigg), \tag{28}$$

and

$$\Pr(f_{it-1,t}=0, s_{it-1,t}=0|d_i, Z_i, \hat{\epsilon}_{it}, \hat{\alpha}_i) = \Phi_2\bigg(-\frac{\varphi(.)+\rho_{\zeta\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon}}{\sigma_{\underline{\zeta}}}, -\frac{\gamma(.)+\rho_{\upsilon\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon}}{\sigma_{\underline{\upsilon}}}, \rho_{\underline{\zeta\upsilon}}\bigg). \tag{29}$$

However for CIS2.5 we do not observe whether $f_{it-1,t}$ is 1 or 0 when $s_{it-1,t}=0$, by integrating out $\underline{\zeta}_{it}$, we get

$$\Pr(s_{it-1,t}=0|d_i, Z_i, \hat{\epsilon}_{it}, \hat{\alpha}_i) = \Phi\bigg(\frac{-(\mathbf{z}_{it}^{s\prime}\gamma + d_{it}\gamma_d + \theta\hat{\alpha}_i + \rho_{\upsilon\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon})}{\sigma_{\underline{\upsilon}}}\bigg). \tag{30}$$

The Conditional Log Likelihood function for individual i in period t given d_i, Z_i, if the time period t corresponds corresponds to CIS3 and CIS3.5, is given by

$$\begin{aligned}&\boldsymbol{L}_{it}(\varphi, \varphi_d, \rho_{\zeta\epsilon}, \lambda, \sigma_{\underline{\zeta}}; \gamma, \gamma_d, \rho_{\upsilon\epsilon}, \theta, \sigma_{\underline{\upsilon}}; \rho_{\underline{\zeta\upsilon}}|d_i, Z_i, \hat{\epsilon}_{it}, \hat{\alpha}_i) = fs\Pr(f=1, s=1)\\ &+(1-f)s\Pr(f=0,s=1) + f(1-s)\Pr(f=1,s=0) + (1-f)(1-s)\Pr(f=0,s=0)\end{aligned}$$

and

$$\begin{aligned}\boldsymbol{L}_{it}(\varphi, \varphi_d, \rho_{\zeta\epsilon}, \lambda, \sigma_{\underline{\zeta}}; \gamma, \gamma_d, \rho_{\upsilon\epsilon}, \theta, \sigma_{\underline{\upsilon}}; \rho_{\underline{\zeta\upsilon}}|d_i, Z_i, \hat{\epsilon}_{it}, \hat{\alpha}_i) &= fs\Pr(f=1,s=1) + (1-f)s\Pr(f=0,s=1)\\ &+ (1-s)\Pr(s=0),\end{aligned}$$

if the time period t corresponds to CIS2.5. Given that error components in (13) and (14) are i.i.d., the Log Likelihood for an individual i, is thus given by

$$\boldsymbol{L}_{i}(\varphi, \varphi_d, \rho_{\zeta\epsilon}, \lambda, \sigma_{\underline{\zeta}}; \gamma, \gamma_d, \rho_{\upsilon\epsilon}, \theta, \sigma_{\underline{\upsilon}}; \rho_{\underline{\zeta\upsilon}}|d_i, Z_i, \hat{\epsilon}_{it}, \hat{\alpha}_i) = \sum_{t=1}^{T_i}\boldsymbol{L}_{it}(.|d_i Z_i, \hat{\epsilon}_{it}, \hat{\alpha}_i). \tag{31}$$

Now, we know that for probit models, the variance of the idiosyncratic components, $\sigma_{\underline{\zeta}}$ and $\sigma_{\underline{\upsilon}}$, are not identified and that the coefficients or the parameters of the model are estimated only up to a scale. Therefore, we can write the Log Likelihood function for an individual i as $\boldsymbol{L}_i(\varphi\frac{1}{\sigma_{\underline{\zeta}}}, \frac{\varphi_d}{\sigma_{\underline{\zeta}}}, \frac{\lambda}{\sigma_{\underline{\zeta}}}, \frac{\rho_{\zeta\epsilon}}{\sigma_{\underline{\zeta}}}; \gamma\frac{1}{\sigma_{\underline{\upsilon}}}, \frac{\gamma_d}{\sigma_{\underline{\upsilon}}}, \frac{\theta}{\sigma_{\underline{\upsilon}}}, \frac{\rho_{\upsilon\epsilon}}{\sigma_{\underline{\upsilon}}}; \rho_{\underline{\zeta\upsilon}}|d_i, Z_i, \hat{\epsilon}_{it}, \hat{\alpha}_i)$.

As mentioned earlier, the main objective of this chapter is to assess the impact of financial constraints on R&D expenditure and to determine what causes financial constraint. We interpret financial constraints as high cost or price of external finance that a firm would be required to pay for a variety of reasons, discussed in Section 6.2 and 6.3, relating to lack of internal funds or information asymmetry. To this effect, $f^*_{it-1,t}$, the latent variable underlying $f_{it-1,t}$, can reflect the shadow price of external funds, d_{it}, the ratio of long-term debt to capital assets. Controlling for other factors, a firm is likely to be more financially constrained or more likely to to have reached its debt capacity the higher the debt asset ratio is, and consequently face a binding borrowing constraint, in optimally deciding its investment and financial policy. However, it quite possible that a firm may be financially constrained not only because of high cost of external funds, but also due to high need for funds. Quantity constraints, due to credit rationing, can be as important as cost of external finance. Kaplan and Zingales (1997) use a classification scheme that a priori distinguish firms that are not financially constrained, firms that are likely to be financially constrained due to high need of funds, and firms that are indeed financially constrained. The firms that are not constrained are those that have liquidity reserves more than what would be needed to finance their investment or have initiated paying dividends. Firms that are likely to be financially constrained because of need for high funds mention having difficulties in obtaining financing. For example, they include firm-year observation in which firm's postpone an equity or convertible debt offering due to adverse market conditions, or claim that they need equity capital but are waiting for improved market conditions. Generally these firms are prevented from paying dividends and have little cash available. Firms that cut dividends are more likely to fall in this category. Finally, in their classification scheme firms that are indeed financially constrained are firms that are in violation of debt covenants, have been cut out of their usual sources of credit, are renegotiating debt payments and have declare that they are forced to reduce investment because of liquidity problems. Here, we would like to mention that the information that is available with us are not as detailed as are for the firms that Kaplan and Zingales analyze. Consequently, we do not have a rich specification for the financial constraint equation that can also account for financial constraint that arises from high need of funds, and indeed, test if such variables do indicate whether firms are financially constrained on account of high need of funds. However, our specification does include variables like cash reserve and dividend payout which, to an extent, can control for the fact that firms are financially constrained due to high need for funds.

The list of exogenous variables included in $\mathbf{z}_{it}^{f\prime}$ are: (1) share of innovative sales in the total sales of the firm in period $t-1, t(SINS)$, (2) ratio of cash flows of the firm to the book value of capital assets (tangible +intangible) of the firm in period t (CF), (3) the ratio of liquidity holdings to the book value of capital assets of the firm in period t (LQ), (4) dummy for positive dividends in period t (DIV), (5) market share of the firm in period $t(MKSH)$, (6) logarithm of the number of people employed in period $t(SIZE)$, (7) dummy for negative cash flows in period t, $(DNCF)$, (8) ratio of intangible asset to the total capital asset of the firm in period t ($RAINT$) [8], (9) age of the firm (AGE), (10) industry dummies and (11) year dummies.

In the specification for the financial constraint equation, we need to control for future expected profitability of the R&D related activity. This is because, if a firm finds that the R&D project for the product or the process that it is willing to undertake will reap higher profits in future, then its need for R&D expenditure today would be high. However, if its internal funds are not sufficiently high and if capital market imperfections facing the firm are acute, then with a higher need for R&D expenditure, the firm would find it self more likely to be facing financial constraint. We include share of innovative sales in the total sales of the firm, ($SINS$), as a proxy for future expected profitability of the the firm, profitability, which is related to R&D activities. Share of innovative sales, ($SINS$), measures of how important R&D activity is to a firm. Assuming, that there is time lag between starting of an R&D project, successful realization of the product, and incurring some fixed cost to implement or produce the new process or product, share of innovative sales in the total sales, ($SINS$), is a result of past innovative activity, and therefore, at most be predetermined, if not strictly exogenous. This implies that variation in ($SINS$), will be a result of past R&D endeavors as well as due to high or low demand for the newly introduced product to the market. Therefore, ($SINS$) can potentially signal demand for R&D related activity and investment. In other words higher percentage of innovative sales can trigger a high demand for R&D investment, which in turn implies, given every thing else, a firm is more likely to find itself reporting financial constraints.

The specification for the financial constraint equation includes cash flows, (CF), and liquidity, (LQ), holdings of the firm. Healthy realization of cash flows can be beneficial for firms that are financially constrained on the margin. The realized cash flows for firms that are financially constrained could be used as a source of internal equity, thus easing the constraint faced by the firm. On the other hand, since cash flows are repositories of demand signals, realized cash flows can also trigger demand for R&D capital. Thus, increase in cash flows can also indicate a high demand for R&D expenditures, which ceteris paribus, can lead a firm to be reporting it self as financially constrained.

Hall and Lerner (2010) and Brown, Fazzari, and Petersen (2009) point out that most of the R&D spending is in form of payments to highly skilled workers, who often require a great deal of firm-specific knowledge and training. The effort of the skilled workers create the knowledge

[8] This variable too is likely to be predetermined if not endogenous. However, as mentioned earlier we intend to enogenize the potentially endogenous variables by specifying a system of reduced form equations for them and estimate a system of seemingly unrelated regression.

base of the firm, and is therefore embedded in the human capital of the firms. This knowledge base is lost once they are fired. The implication of this is that firms choose to smooth their R&D spending, if only to avoid laying off their knowledge workers. This in turn implies that R&D intensive firms behave as if they have large adjustment cost. Thus, when faced with high adjustment costs, and when it is not sure about a steady supply of positive cash flows, an R&D intensive firm is likely to practice precautionary savings to reduce its chances of being financially constrained at any given time. This argument implies that among the R&D intensive firms, those that maintain liquidity reserve, (LQ), are less likely to be financially constrained. We include dummy for dividends paid out, (DIV), since it is likely that cash distributions to shareholders are evidence of the availability of internally generated funds. Therefore, it is more likely that a firm that pays out dividends is unlikely to be affected by costly external finance and that at least some firms that do not distribute are affected. Fama and French (2000) found the the ratio of dividends to total assets to be a good predictor of profitability of the firm. Also, Fama and French (2002) point out that since it is expensive to finance investment with new risky securities, dividends are less attractive for firms with less profitable assets in place, large current and expected investments, and high leverage. Thus, controlling for other effects, more profitable firms pay out more of their earnings as dividends, indicating that such firms are less likely to be financially constrained.

We include size of the firm, $(SIZE)$, since it might be that large firms have more internal funds at their disposal to carry out R&D related activities, and that small firms may be typically young and less well known, hence more vulnerable to capital market imperfections. Hennessey and Whited (2007) find large differences between the cost of external funds for small and large firms. Large firms behave as if they face small indirect costs of external finance, and small firms behave as if they face large indirect costs of external finance. Market share, $(MKSH)$, of the firm is included in the specification, since, it may be possible that firms enjoying a higher degree of monopoly have a better access to capital markets. While size and may be correlated with market share of the firm it may that, for certain products, relatively smaller firms enjoy a higher degree of monopoly. Now, since the secondary market for intangible asset is fraught with more frictions and generally does not exists, hence firms with higher percentage of intangible asset have less amount of pledgeable support to borrow. It could also be possible that such firms face higher bankruptcy costs. Thus it seems likely that firms with high ratio of intangible asset to total capital asset, $(RAINT)$, are more likely to be financially constrained. The age of the firm, (AGE), has been an important determinant of financial constraint as suggested by many studies: for an example, Brown, Fazzari, and

Petersen (2009). Old firms, given that they have survived, have probably build reputation over the years that allow them to have better access to capital markets and are thus less likely to be financially constrained, given every thing else.

The specification for the selection equation includes d_{it}, long term debt to capital asset ratio. The idea being, as discussed earlier, to test if highly leveraged firms are innovative or not. The list of variables that we include in $\mathbf{z}_{it}^{s\prime}$, for the selection equation are: (1) the ratio of liquidity holdings to the capital assets of the firm (LQ), (2) dummy for positive dividends (DIV), (3) market share of the firm $(MKSH)$, (4) logarithm of the number of people employed $(SIZE)$, (5) dummy that takes value 1 if the number of enterprises consolidated within a firm is more than one, $(DMULTI)$, (6) age of the firm (AGE), (7) the ratio of intangible assets to total (tangible +intangible) assets $(RAINT)$, (8) industry dummies, and (9) year dummies.

The liquidity holdings of the firm, (LQ), are included in the specification for the selection equation. The reason being, R&D intensive firms, for reasons related the nature of R&D activity and it implications for capital market imperfections faced by them, have to bear high cost of external finance. Also, as discussed earlier, to avoid high adjustment cost related to hiring and firing its knowledge workers, R&D intensive firms tend to smooth their investment spending over time. This necessitates them to maintain large amounts of cash reserves. Including (LQ) in the specification allows us to test, if this is indeed true or not. Again, for reasons related to capital market imperfections facing R&D activity, it might be less attractive for for firms indulging in R&D activity to distribute cash as dividends, (DIV). We include $(RAINT)$ in the selection/innovation equation, since for any two firms that have the same capital base, the one with more intangible asset in its capital base is more likely to be an innovator. This is because a large part of the capital of an R&D intensive firm resides in the knowledge base of the firm which is intangible. Though, it might seem that the percentage of intangible asset in the total asset base of a firm is more of an outcome of a firm's decision to innovate rather than its having any bearing on the innovation decision of the firm. However, we know that there is persistence in innovation activity of a firm, see Wladimir et al. (2009), or in other words, innovation decision exhibit certain degree of path dependency. $(RAINT)$ being the outcome of past innovation activity, captures the persistence in the innovation decision of the the firm.

In the Schumpeterian tradition, it makes sense to include size, $(SIZE)$, as an explanatory variable in the selection equation. It can also be argued that if there are fixed costs of investing, then as Cohen and Klepper (1996) argue, large firms have a higher incentive to engage in innovative activities because they can amortize these costs by selling more units of output. Nilsen and Schiantarelli (2003) find strong statistical evidence of this relationship, including much greater incidences of zero investments in small versus large plants. They attribute this relevance of plant size, to the presence of fixed costs and to potential indivisibilities in investment. As stated earlier market share,($MKSH$), of the firm is a proxy for concentration or the degree of monopoly. We include this variable in the selection equation, since in the Schumpeterian tradition it has been argued that firms that enjoy monopoly position are more incited to innovate, if only to prevent entry of potential rivals. We also include age, (AGE), of the firm in the selection equation, since it is likely that young firms, in order to survive, have a larger incentive to innovate new products. Also see Section 6.2 for a detailed discussion on maturity and innovation practices of firms.

5.1.3 The Third Stage: R&D Investment

In the third and the final stage we estimate the R&D equation, equation (24), which is a switching model with the added correction terms, $C_{11}(\Xi,.)_{it}$, $C_{12}(\Xi,.)_{it}$, $C_{13}(\Xi,.)_{it}$, $C_{01}(\Xi,.)_{it}$, $C_{02}(\Xi,.)_{it}$, and $C_{03}(\Xi,.)_{it}$, which control for the bias that can arise due to endogeneity of debt, financial constraint and selection.

$$
\begin{aligned}
r_{it} = &\ \mathbf{z}_{it}^{r\prime}\beta + f_{it-1,t}\beta_f + f_{it-1,t}\beta_1(.) + (1 - f_{it-1,t})\beta_0(.) + \mu\hat{\alpha}_i \\
&+ f_{it-1,t}\left(\rho_{\eta_1\epsilon}\sigma_{\eta_1}C_{11}(.)_{it} + \rho_{\underline{\eta}_1\underline{\zeta}}\sigma_{\underline{\eta}_1}C_{12}(.)_{it} + \rho_{\underline{\eta}_1\underline{\upsilon}}\sigma_{\underline{\eta}_1}C_{13}(.)_{it}\right) \\
&+ (1 - f_{it-1,t})\left(\rho_{\eta_0\epsilon}\sigma_{\eta_0}C_{01}(.)_{it} + \rho_{\underline{\eta}_0\underline{\zeta}}\sigma_{\underline{\eta}_0}C_{02}(.)_{it} + \rho_{\underline{\eta}_0\underline{\upsilon}}\sigma_{\underline{\eta}_0}C_{03}(.)_{it}\right) + \tilde{\eta}_{it},
\end{aligned}
\tag{24}
$$

To control for the individual effects, α_i, we substitute the expected a posteriori values, $\hat{\alpha}_i(.)$, computed from the first stage reduced form estimation. With $\hat{\alpha}_i(.)$ in place to control for time invariant individual specific effect, a simple pooled OLS, of equation (24), for the subsample of innovating firms leads to consistent estimates of the parameters.

As mentioned earlier, one of the goals of this chapter is to assess the impact of financial constraint on R&D investment. Interpreting $f^*_{it-1,t}$ as the shadow price of external funds, implies that a firm in question finds it self being financially constrained, $f_{it-1,t} = 1$, if the price that it is required is high or if it exceeds a certain threshold. Interpreting the latent variable $f^*_{it-1,t}$ as the price of external funds, underlying $f_{it-1,t}$, allows us test whether the plethora of reasons, discussed earlier, does indeed make make external financing costly and if such high cost does indeed impact R&D activity adversely.

Following Fazzari, Hubbard, and Petersen (1988), many studies have concluded that firms, a priori, classified as financially more constrained have higher investment-cash flow sensitivities. Empirically, the existence of financial frictions for innovative firms has been investigated, either by examining the sensitivity of R&D investment to financial factors, including measures of internal finance such as cash flows, that uses reduced form of accelerator type models of investment as in Fazzari, Hubbard, and Petersen (1988), or by using the structural framework of Euler equations as in Bond and Meghir (1994), see Hall and Lerner (2010) for a survey. However, Kaplan and Zingales (1997) , Kaplan and Zingales (2000) and Cleary (1999) have provided evidence that cash flow sensitivity need not identify liquidity constrained firms, that is, sensitivity is not monotonic in the degree of constraints. Cash flows provide information about future investment opportunities, hence, investment cash flow sensitivity may equally occur because firms are sensitive to demand signals. On the theoretical side, Gomes (2001) and Alti (2003) simulate dynamic investment models, demonstrating that significant cash flow coefficients are not necessarily generated by financing frictions. Conversely, Gomes (2001) shows that financing frictions are not sufficient to generate significant coefficients on cash flow.

The distinguishing feature of our model is the switching regression model for R&D expenditure that allows us endogenize financial constraint with respect to R&D related activity, which as mentioned earlier, is an information that the firms themselves report. Second, though cash flows are included in the specification and has different coefficients in the two different regimes, a test for the existence of financial frictions does not rest on the sensitivity of R&D investment to cash flows for the financially constrained and unconstrained firms, but rather through the test of the effect of reported financial constraints on R&D investment, and as mentioned in the last section, its implications for the decision to innovate.

Here, we would like to emphasize that in order to see the effect of firm's reportage of financial constraints on R&D investment, we would like to fix the firm's investment opportunity. Since, we do not have any information on the market valuation of the firms, we can not construct average "q" for our firms or for that matter, any such measure related to the firm's R&D investment. To this end, in our specification we include cash flows and share of innovative sales in the total sales of the firm, ($SINS$), which can be indicative of demand signals. Moyen (2004), simulates series from dynamic models of constrained and unconstrained firms, and pools them together to represent a theoretical sample. She does this to investigate whether some of the empirical results relating to investment cash flow sensitivity, can be replicated in her theoretical sample. Among other things, she finds that Tobin's "q" is a poor proxy for investment opportunities, cash flow is an excellent proxy, and that cash flow is an increasing function of the income shock. Since, it is quite likely that the cash flows of the firm is correlated to, or constitutes of the cash flows that emanates from R&D related activity, accordingly, in our model the role of cash flows, (CF), is to control for the R&D investment opportunity of firms.

Now, cash flows in the specification for the R&D equation, pertain to the firm as a whole. A measure to control for the investment opportunity for R&D related activity should be based on a measure such as Tobin's "q" for R&D related activity or cash flows that result from R&D output. However, in the absence of any such measure, we use share of innovative sales in the total sales of the firm, ($SINS$), which as explained in the discussion of the second stage specification, can potentially signal demand for R&D related activity.

In our results, we do obtain different effect of changes in cash flows on R&D investment for firms that are financially constrained and firms that are not. This could be because, cash flows are correlated with investment opportunity set and observables such as ($SINS$), which we include in the specification, to control for future expected profitability, cannot perfectly control for the firm's investment opportunity related to future expected profitability. Other reasons, why the sensitivity of cash flows across constrained and unconstrained firms might differ, could be due to the nature of costs related to issuing new equity and cost of borrowing and how well proxies that classify firms as constrained or otherwise are reflective of these costs. However, within the framework of our model, we cannot test for the exact mechanism that drives the results on R&D investment-cash flow sensitivity across constrained and unconstrained firms. For a detailed discussion on investment-cashflow sensitivity, we refer to Moyen (2004), Hennessey and Whited (2007), Whited (2008), Alti (2003), and Gomes (2001) and the references therein[9].

The specification for the R&D equation does not include any financial stock variables such as long-term debt to asset ratio or cash reserves to asset ratio. This is so because in the structural model for R&D investment, R&D expenditures are determined only by the degree of constraints a firm faces and the expected profitability from R&D investment. Therefore, it seems unlikely that leverage and cash holdings will have an independent effect, other than through financial constraints affecting the firm. However, we do include dummy for dividend payout, (DIV), during the period, in which R&D expenditure was incurred, to investigate the behavior of firms with respect to R&D spending for firm's paying out dividends. The other variables that are included in the specification are: size of the firm, ($SIZE$), market share,($MKSH$), and the age of the firm (AGE). The rationale, for including these variables

[9]Though, these papers do not specifically model R&D investment, the results in these papers can have relevance in explaining R&D investment-cash flow coefficients, or the models in these papers could be augmented to allow for features of R&D investment.

is essentially the same as for the selection equation. However, among the set of innovating firms, it can be possible that smaller firms invest relatively more in R&D than larger firms. This is because, it has been argued that smaller firms have a higher marginal "q" than large firms, which could even be true of R&D capital. Smaller firms also experience higher growth rates than large firms and more likely to take risk than larger firms, see Cooley and Quadrini (2001), Gomes (2001) and the discussion in Section 6.2. The specification for the third stage also includes a dummy that takes value 1 if the number of enterprises consolidated within a firm is more than one, ($DMULTI$).

5.2 Identification and Estimation of Average Partial Effects

Estimation of average partial effects (APE) in presence of endogeneity and unobserved heterogeneity has recently been of concern the in econometric literature, see Wooldridge (2004, 2005). Our empirical strategy allows us to take into account, the enodogeneity of the explanatory variables as well as unobserved heterogeneity. We do this by conditioning the endogenous variable of interest on the idiosyncratic component of the reduced form equation of the endogenous explanatory variable, thereby controlling for the endogeneity of our endogenous regressor. Estimation of the structural parameters, as we know requires substitution of the idiosyncratic component of the reduced form equation in the structural equations by their conditional expected value, the residuals from the reduced form estimation. However, the residuals, $\mathrm{E}(\epsilon_{it}|d_i, Z_i, \alpha_i)$, from the first stage reduced form equation are not identified due to presence of time invariant individual effect, which are not observed. In Appendix A, however, we have shown that even though $\mathrm{E}(\epsilon_{it}|d_i, Z_i, \alpha_i)$ is not identified, $\mathrm{E}(\epsilon_{it}|d_i, Z_i)$ is. This allows us to replace α_i with their expected a posterior values $\hat{\alpha}_i(d_i, Z_i, \hat{\delta}, \hat{\sigma}_\alpha, \hat{\sigma}_\epsilon)$, which helps us in identifying the parameters of interest, the details are provided in Appendix A. In this section we only state the functional form of the APE of a typical variable, continuous or categorical, for the second and the third stage estimation.

5.2.1 Average Partial Effects for the Second Stage

Suppose, we are interested in changes in the probability of a firm being financially constrained for small changes in d_{it} or z_{kit} where z_k is any of the exogenous regressors that appear in the financial constraint equation. From the result on identification of parameters, discussed in Appendix A, we know that

$$\Pr(f_{it-1,t} = 1|d_i, Z_i, \hat{\epsilon}_{it}, \hat{\alpha}_i) = \Phi\left(\mathbf{z}_{it}^{f\prime}\tilde{\varphi} + d_{it}\tilde{\varphi}_d + \tilde{\lambda}\hat{\alpha}_i + \tilde{\rho}_{\zeta\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon}\right)$$

where $\tilde{\varphi} = \varphi\frac{1}{\sigma_{\underline{\varsigma}}}$, $\tilde{\varphi}_d = \frac{\varphi_d}{\sigma_{\underline{\varsigma}}}$, $\tilde{\lambda} = \frac{\lambda}{\sigma_{\underline{\varsigma}}}$, $\tilde{\rho}_{\zeta\epsilon} = \frac{\rho_{\zeta\epsilon}}{\sigma_{\underline{\varsigma}}}$. The APE of the change in z_{kit} from $\bar{z}_k$ to $\bar{z}_k + \kappa$ on the probability of being financially constrained given $d_{it} = \bar{d}$ and $\mathbf{z}^f_{it_{-k}} = \bar{\mathbf{z}}^f_{-k}$ as shown in Appendix A is given by

$$\frac{\Delta \Pr(f_{it-1,t} = 1|\bar{z}_k, \bar{Z}_{-k}, \bar{d})}{\Delta z_k} = \frac{\int \Phi\Big(\bar{z}_k + \kappa, \bar{\mathbf{z}}^f_{-k}, \bar{d}, \hat{\alpha}, \hat{\epsilon}\Big) dG(\hat{\alpha}, \hat{\epsilon}) - \int \Phi\Big(\bar{z}_k, \bar{\mathbf{z}}^f_{-k}, \bar{d}, \hat{\alpha}, \hat{\epsilon}\Big) dG(\hat{\alpha}, \hat{\epsilon})}{\kappa}. \tag{32}$$

In the limit when $\Delta z_k = \kappa$ goes to zero, since the integrand is a smooth function of its arguments we can change the order of differentiation and integration, to get

$$\frac{\partial \Pr(f_{it-1,t} = 1|\bar{z}_k, \bar{Z}_{-k}, \bar{d})}{\partial z_k} = \int \phi\Big(\bar{z}_k, \bar{\mathbf{z}}^f_{-k}, \bar{d}, \hat{\alpha}, \hat{\epsilon}\Big) \varphi_k dG(\hat{\alpha}, \hat{\epsilon}). \tag{33}$$

The sample analog of the RHS of the above equation can be computed as follows:

$$\frac{1}{\sum_{i=1}^{N} T_i} \sum_{i=1}^{N} \sum_{t=1}^{T_i} \phi\Big(\bar{z}_k, \bar{\mathbf{z}}^f_{-k}, \bar{d}, \hat{\alpha}_i, \hat{\epsilon}_{it}\Big) \varphi_k. \tag{34}$$

Suppose, z_k is an exogenous dummy variable D taking values 0 and 1, then the APE of change of D_{it} from 0 to 1 on the probability of a firm being financially constrained, given other covariates, is given by

$$\int \Phi\Big(D = 1, \bar{\mathbf{z}}^f_{-k}, \bar{d}, \hat{\alpha}, \hat{\epsilon}\Big) dG(\hat{\alpha}, \hat{\epsilon}) - \int \Phi\Big(D = 0, \bar{\mathbf{z}}^f_{-k}, \bar{d}, \hat{\alpha}, \hat{\epsilon}\Big) dG(\hat{\alpha}, \hat{\epsilon}), \tag{35}$$

whose sample analog, given (34), can be computed in the same way.

We may also be interested in the APE of a variable on the conditional probability of an event or compare APE of a variable on the probability of an event conditional on two mutually exclusive events. For example, we may be interested in the marginal effect of long-term debt to asset ratio on the probability of firm being an innovator, conditional on being financially constrained as compared to the marginal effect of long-term debt on the probability of being innovator, conditional on the firm not being financially constrained. For the above mentioned example, the APE of changing long-term debt from $d_{it} = \bar{d}$ to $d_{it} = \bar{d} + \kappa$ on the probability of a firm being financially constrained, conditional on being a innovator and given $Z_i = \bar{Z}$, is given by

$$\begin{aligned}\frac{\Delta \Pr(s_{it-1,t} = 1|f_{it-1,t} = 1, \bar{d}, \bar{Z})}{\Delta d} =& \Bigg[\int \left(\frac{\Phi_2(\bar{\varphi}(\bar{d} + \kappa, .), \bar{\gamma}(\bar{d} + \kappa, .), \rho_{\underline{\varsigma}\underline{\upsilon}})}{\Phi(\bar{\varphi}(\bar{d} + \kappa, .))}\right) dG(\hat{\alpha}, \hat{\epsilon}) \\ &- \int \left(\frac{\Phi_2(\bar{\varphi}(\bar{d}, .), \bar{\gamma}(\bar{d}, .), \rho_{\underline{\varsigma}\underline{\upsilon}})}{\Phi(\bar{\varphi}(\bar{d}, .))}\right) dG(\hat{\alpha}, \hat{\epsilon})\Bigg] / \kappa,\end{aligned} \tag{36}$$

where $\bar{\varphi}(.) = \frac{\varphi(.) + \rho_{\zeta\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon}}{\sigma_{\underline{\varsigma}}}$ and $\bar{\gamma}(.) = \frac{\gamma(.) + \rho_{\upsilon\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon}}{\sigma_{\underline{\upsilon}}}$. The sample analog of which can be constructed by averaging the integrand over $\hat{\alpha}_i$ and $\hat{\epsilon}_{it}$. However, in the limit as κ tends to zero, since the integrand is a smooth function of its argument, we can change the order of differentiation and integration to obtain the APE as

$$\frac{\partial \Pr(s_{it-1,t}=1|f_{it-1,t}=1,\bar{d},\bar{Z})}{\partial d}=\int\frac{\partial}{\partial d}\left(\frac{\Phi_2(\bar{\varphi}(.),\bar{\gamma}(.),\rho_{\underline{\zeta\upsilon}})}{\Phi(\bar{\varphi}(.))}\right)dG(\hat{\alpha},\hat{\epsilon}), \tag{37}$$

The right hand side of the above involves taking derivative of cumulative distribution function of a standard bivariate normal with respect to $\bar{\varphi}(.)$ and $\bar{\gamma}(.)$. If we want to compare the APE of increase in leverage on the probability of being financially constrained, conditional on being a non-innovator, then we have

$$\frac{\partial \Pr(s_{it-1,t}=1|f_{it-1,t}=0,\bar{d},\bar{Z})}{\partial d}=\int\frac{\partial}{\partial d}\left(\frac{\Phi_2(\bar{\varphi}(.),-\bar{\gamma}(.),-\rho_{\underline{\zeta\upsilon}})}{1-\Phi(\bar{\varphi}(.))}\right)dG(\hat{\alpha},\hat{\epsilon}). \tag{38}$$

The derivative of the $\Phi_2(\bar{\varphi}(.),\bar{\gamma}(.),\rho_{\underline{\zeta\upsilon}})$ and $\Phi_2(\bar{\varphi}(.),-\bar{\gamma}(.),-\rho_{\underline{\zeta\upsilon}})$ with respect to $\bar{\varphi}(.)$ and $\bar{\gamma}(.)$ are stated in Greene (2003).

5.2.2 Average Partial Effects for the Third Stage

The average effect of financial constraint on R&D intensity, for any individual, i, in time period, t, is computed as the difference in the expected R&D expenditure between the two regimes, financially constrained and non-financially constrained, averaged over $\hat{\alpha}$ and $\hat{\epsilon}$:

$$\begin{aligned}\Delta\mathrm{E}(r_{it}|Z_i,d_i)=&\int \mathrm{E}(r_{1it}|Z_i,d_i,\hat{\alpha},\hat{\epsilon})dG(\hat{\alpha},\hat{\epsilon})\\&-\int \mathrm{E}(r_{0it}|Z_i,d_i,\hat{\alpha},\hat{\epsilon})dG(\hat{\alpha},\hat{\epsilon}).\end{aligned} \tag{39}$$

From equations (8a) and (9a) we know that

$$\mathrm{E}(r_{1it}|d_i,Z_i,\hat{\alpha}_i,\hat{\epsilon}_{it})=\mathbf{z}_{it}^{r\prime}\beta+\beta_f+\beta_1(.)+\mu\hat{\alpha}_i+\rho_{\eta_1\epsilon}\frac{\sigma_{\eta_1}}{\sigma_\epsilon}\hat{\epsilon}_{it}$$

if $f^*_{it-1,t}>0$, and

$$\mathrm{E}(r_{0it}|d_i,Z_i,\hat{\alpha}_i,\hat{\epsilon}_{it})=\mathbf{z}_{it}^{r\prime}\beta+\beta_0(.)+\mu\hat{\alpha}_i+\rho_{\eta_0\epsilon}\frac{\sigma_{\eta_0}}{\sigma_\epsilon}\hat{\epsilon}_{it}$$

if $f^*_{it-1,t}\leq 0$. Given the above, the average effect of financial constraint on R&D intensity, for an individual, i, in time period t, is given by:

$$\Delta\mathrm{E}(r_{it}|Z_i,d_i)=\beta_f+\beta_1(.)-\beta_0(.)+\int\left(\rho_{\eta_1\epsilon}\frac{\sigma_{\eta_1}}{\sigma_\epsilon}\hat{\epsilon}_{it}-\rho_{\eta_0\epsilon}\frac{\sigma_{\eta_0}}{\sigma_\epsilon}\hat{\epsilon}_{it}\right)dG(\hat{\alpha},\hat{\epsilon}). \tag{40}$$

Since, we assume that the time invariant individual effects, has the same effect in both the regimes, the $\hat{\alpha}$ term drops out. The sample analog of the above can be obtained by taking the average over $\hat{\epsilon}_{it}$ for given $Z_i=\bar{Z}$ and $d_i=\bar{d}$.

The APE of the other variables on R&D intensity will be different for the two regimes, defined as financially constrained and not financially constrained. The APE's, are composite of the direct effect on R&D intensity, if the variable is included in the specification for R&D intensity, and its effect through the correction terms, that are different for the two regimes. Including the indirect effect, that comes through the correction terms, implies that APE's obtained are conditional APE's, that is, conditional on being selected. Without the indirect effect that comes through the correction terms, that correct for the bias due to sample selection, the unconditional APE's are simply the coefficient estimates.

Given the R&D equation (24), with added correction terms, the APE, conditional on being selected, of a change in exogenous variable, z_{kit}, from $\bar{z}_k$ to $\bar{z}_k + \kappa$, given $Z_{i_{-k}} = \bar{Z}_{-k}$ and $d_i = \bar{d}$, on R&D intensity for a firm, i, that is financially constrained in time period t is given by:

$$\frac{\Delta}{\Delta z_k} \mathrm{E}(r_1 | f^* > 0, s^* > 0, \bar{z}_k, \bar{Z}_{-k}, \bar{d}) = \beta_{1k} \\ + \frac{\int \Big[\mathrm{E}(\underline{\eta}_1 | f^* > 0, s^* > 0, \bar{z}_k + \kappa, \bar{Z}_{-k}, \bar{d}, \hat{\alpha}, \hat{\epsilon}) - \mathrm{E}(\underline{\eta}_1 | f^* > 0, s^* > 0, \bar{z}_k, \bar{Z}_{-k}, \bar{d}, \hat{\alpha}, \hat{\epsilon}) \Big] dG(\hat{\alpha}, \hat{\epsilon})}{\kappa}. \tag{41}$$

That is,

$$\begin{aligned} &\frac{\Delta}{\Delta z_k} \mathrm{E}(r_1 | f^* > 0, s^* > 0, \bar{z}_k, \bar{Z}_{-k}, \bar{d}) = \beta_{1k} \\ &+ \Big[\int \Big(C_{12}(., \bar{z}_k + \kappa, \hat{\alpha}, \hat{\epsilon}) \sigma_{\underline{\eta}_1} \rho_{\underline{\eta}_1 \underline{\varsigma}} + C_{13}(., \bar{z}_k + \kappa, \hat{\alpha}, \hat{\epsilon}) \sigma_{\underline{\eta}_1} \rho_{\underline{\eta}_1 \underline{\upsilon}} \Big) dG(\hat{\alpha}, \hat{\epsilon}) \\ &- \int \Big(C_{12}(., \bar{z}_k, \hat{\alpha}, \hat{\epsilon}) \sigma_{\underline{\eta}_1} \rho_{\underline{\eta}_1 \underline{\varsigma}} + C_{13}(., \bar{z}_k, \hat{\alpha}, \hat{\epsilon}) \sigma_{\underline{\eta}_1} \rho_{\underline{\eta}_1 \underline{\upsilon}} \Big) dG(\hat{\alpha}, \hat{\epsilon}) \Big] / \kappa, \end{aligned} \tag{42}$$

where β_{1k} in the above two equations is the direct effect of z_k on the R&D intensity, and the remaining term the indirect effect. The sample analog of the RHS of (42) is given by:

$$\begin{aligned} &\frac{\Delta}{\Delta z_k} \mathrm{E}(r_1 | f^* > 0, s^* > 0, \bar{z}_k, \bar{Z}_{-k}, \bar{d}) = \beta_{1k} \\ &+ \Big[\frac{1}{\sum_{i=1}^N T_i} \sum_{i=1}^N \sum_{t=1}^{T_i} s_{it-1,t} \Big(C_{12}(., \bar{z}_k + \kappa, \hat{\alpha}, \hat{\epsilon}) \sigma_{\underline{\eta}_1} \rho_{\underline{\eta}_1 \underline{\varsigma}} + C_{13}(., \bar{z}_k + \kappa, \hat{\alpha}, \hat{\epsilon}) \sigma_{\underline{\eta}_1} \rho_{\underline{\eta}_1 \underline{\upsilon}} \Big) \\ &- \frac{1}{\sum_{i=1}^N T_i} \sum_{i=1}^N \sum_{t=1}^{T_i} s_{it-1,t} \Big(C_{12}(., \bar{z}_k, \hat{\alpha}, \hat{\epsilon}) \sigma_{\underline{\eta}_1} \rho_{\underline{\eta}_1 \underline{\varsigma}} + C_{13}(., \bar{z}_k, \hat{\alpha}, \hat{\epsilon}) \sigma_{\underline{\eta}_1} \rho_{\underline{\eta}_1 \underline{\upsilon}} \Big) \Big] / \kappa, \end{aligned}$$

Replacing the 1's in the subscripts with 0's in the RHS of the equation gives APE of z_{kit} on R&D intensity for a firm that belongs to the regime that is not financially constrained. Thus, we have

$$\begin{aligned} &\frac{\Delta}{\Delta z_k} \mathrm{E}(r | f^*, s^* > 0, \bar{z}_k, \bar{Z}_{-k}, \bar{d}) = \\ &f \frac{\Delta}{\Delta z_k} \mathrm{E}(r_1 | f^* > 0, s^* > 0, \bar{z}_k, \bar{Z}_{-k}, \bar{d}) + (1 - f) \frac{\Delta}{\Delta z_k} \mathrm{E}(r_0 | f^* \leq 0, s^* > 0, \bar{z}_k, \bar{Z}_{-k}, \bar{d}). \end{aligned} \tag{43}$$

If z_k happens to be a binary variable, D, taking values 0 and 1, then we simply replace $z_k + \kappa$ with $D = 1$ and z_k with $D = 0$ in equations (41-42). For a continuous variable, z_k, in the limit when κ goes to 0, given functional form of the correction terms, which are smooth functions of their arguments, we can again change the order of integration and differentiation in (41-42) to obtain the APE.

6 Results

We discuss the results of our structural model under two subsections. In the first subsection we discuss the results of the Second Stage, in which we jointly estimate the financial constraint and sample selection equation. In the second subsection we discuss the results of the R&D switching regression model.

6.1 Second Stage

Table 6.3 illustrates the results of the joint estimation of the financial constraint equation and the selection equation. We present the results of three specification of the joint estimation, in which the specification for the selection or the innovation equation remains the same, but the specification for the financial constraint equation is varied. For reasons which we discuss later, our most preferred specification is the parsimonious specification presented in columns **(C0)** and **(C1)**. Table 6.4 presents the APE's of the respective variables for the specification illustrated in columns **(C0)** and **(C1)**. To compute the APE's, the values of all the covariates were fixed at the mean value, the mean taken over all the firms and the years. We also note here that many of the covariates are dummy variables taking values 0's and 1's, therefore, the mean of these variables will be fractions lying between 0 and 1.

Our results suggest that controlling for other factors, highly leveraged firms do not take up R&D activity. This validates the claim about the risky nature of R&D related activities, which in general involves assets and products that are both highly intangible in nature, and bank-based institutional lenders being more averse to provide finance to new and high risk investments. This only implies that R&D intensive firms have a preference for equity based financing, which can be outside equity or retained earnings. Allen and Gale (2000) argue that in economies that better manage asymmetric information problems R&D intensive firms tend to be more valuable. Also, market-based financial systems are more likely to encourage R&D activity and consequently will value R&D more highly than bank-based economies. Market-based financial systems have well developed information distribution mechanisms and are able to allocate capital to activities even when investors hold diverse views on the value of the investment. This corroborates the findings of Brown, Fazzari, and Petersen (2009), who also find a preference for equity based financing for high-tech firms and in particular firms that are young among them.

We find a significant negative coefficient for dummy for dividends, (DIV). This suggests that firms that pay out dividends are less likely to be R&D intensive. Now, given the nature of R&D activity, which makes borrowing costly, internal funds may be more preferable. Therefore, innovative firms are less likely to distribute cash as dividends. Again, for reasons related to high adjustment costs, that compels R&D intensive firms to smooth out investment spending, and of course, the need to avert costly borrowing, we wanted to test if innovative firms maintain high cash reserves to finance a foreseeable R&D investment project in the future during which it may or may not be financially constrained. Our results suggests that innovative firms are very likely to hold cash reserves (LQ).

We find that younger firms are more innovative. This corroborates the findings of other studies that find that young firms in their bid to survive take up more innovative activity. Entry, typically in the literature on Industrial Organization, see Audretsch (1995) and Huergo and Jaumandreu (2004), is envisaged as the way in which firms explore the value of new ideas in an uncertain context. Entry, the likelihood of survival and subsequent conditional growth are determined by barriers to survival, which differ by industries according to technological opportunities. In this framework, entry is innovative and increases with uncertainty, the likelihood of survival is lower the higher the risk is, and the growth subsequent to successful innovation is higher the higher barriers to survival are.

Our findings also suggests that large firms are more probable to innovate. This finding, is consistent with the idea that large firms have a higher incentive to engage in innovative activities because they can amortize the large fixed costs of investing costs by selling more units of output. Moreover, larger firms can be expected to have a larger stock of knowledge base, and therefore are expected to produce more innovation than small firms. We find that firms with large market share are more innovative. This result confirms the fact that a firm is more incited to innovate if it enjoys a monopoly position to prevent entry of potential rivals. We also find that firms that have many enterprises consolidated within them, ($DMULTI$), are more likely to be innovative. Now, these enterprises are more or less independent entities, working under a parent firm or a company. It has been found that firms that have many enterprise consolidated within them, are more likely to be globally engaged and have better network of customers and suppliers. Such firms are also multinational in nature. These firms are also large both in terms of employment and sales. These considerations make multi enterprise firms more likely to have better access to knowledge and also hold a larger stock of knowledge, hence such firms are more likely to be innovative. A significant positive coefficient on ($RAINT$), stress the fact that, given every thing else, a firm with higher ratio of intangible asset in the total asset base of the firm is more likely to be innovative. Now, to the extent that this ratio is a result of past innovation decision, it captures the fact that there is persistence in the innovation decision of the firms.

We find that the scaled correlation between υ_{it}, the unobserved idiosyncratic component of the selection equation and ϵ_{it} the idiosyncratic component of the reduced form debt equation, $\frac{\rho_{\upsilon\epsilon}}{\sigma_{\underline{\upsilon}}}$, to be high and significant, suggesting strong simultaneity in the decision to innovate and choice of leverage. In addition we also find that, θ, the factor loading in the selection equation to be significant, suggesting that the unobserved individual specific term that influence the choice of capital structure also affect the decision to innovate. We find the χ^2 test for $\rho_{\underline{\zeta\upsilon}}$ to be significant, suggesting a strong simultaneity in the decision to innovate and the financial constraint with respect to R&D activity that the firm faces. This could be because the firm's ability to access outside financing may be conditional on its innovative activity. For example, suppose there is a firm that has not been innovative in the past, but now has an R&D project that with certain degree of success, has a positive net worth. However, since the capital employed by the firm to embark on the R&D project can be highly intangible, its value, in the event of default, in the second hand market might be significantly low. Since lenders can not observe or verify the progress or the importance of R&D project, are only willing to lend up to the value of the new investment at liquidation. Thus, the decision to innovate can have a bearing on the firm's ability to raise external means of financing if its wealth is not sufficiently high. While the above might be an extreme example, innovative firms too with new R&D projects can face similar situations.

Figure 6.1 illustrates the plots of APE of long-term debt to asset ratio on the unconditional probability of innovation or selection against age, size, and leverage. We find that the APE of leverage on the probability of innovation to be negative and decreasing in the age of the firm. This suggests that once we condition on size and other covariates, older firms are more averse to engage in R&D activity with higher debt in their capital structure as compared to young firms. This finding in a way confirms some of the findings, both theoretical and empirical, that younger firms borrow more, face higher probability of default, grow faster, and in general display risky behavior. Figure 6.1(b) plots the APE of long-term debt to asset ratio against size. We find that the effect is decreasing in absolute terms with size. This is due to the fact that large firms any way have a higher propensity to innovate, hence a marginal increase in leverage is not as effective in affecting innovation of large firms as compared to small or intermediate sized firms. We also find that, as Figure 6.1(c) attests, the partial effect of leverage on the propensity to innovate is negative and decreasing in the size of leverage.

As stated earlier, Table 6.3 presents three specifications for the financial constraint equation. The specification in column **(A0)** has cash flows (CF) and ($RAINT$), ratio of intangible assets to total (tangible + intangible). The specification in column **(B0)** does not include ($RAINT$), while the specification in column **(C0)**, has neither of them. We included ($RAINT$) in the specification for financial constraint, since secondary markets for intangible asset is fraught with more frictions and generally does not exist, hence firms with higher percentage of intangible asset have less amount of pledgeable support to borrow, and thus can be expected to be more financially constrained. However, we do not find it to be significantly explaining financial constraints that a firm faces with respect to R&D related activity. Moving from the specification in column **(A0)** to **(A1)**, we find no change in the sign or the significance of the rest of the variables.

We had included (CF) as measures of change in the internal networth of a firm. An increase in the cash flows of the firm was expected to ease the financial constraints faced by a firm. Instead, what we find is that cash flows, (CF), has a significant positive sign. This can be explained by the fact that the cash flows of the firms are correlated with, or are derived from the product or the process the firm endeavors to improve or find a better substitute for. Since the cash flows realized contains signals about the demand for the firms's product, cash flows are correlated with the R&D investment opportunity set. This in turn implies that an increase in the cash flows signals a higher demand for R&D investment, thus ceteris paribus, an increase in cash flows can lead to a firm reporting itself being financially constrained, especially since we control for negative realization of cash flows.

Though cash flows can signal demand for investment needs of a firm, a measure of how important R&D activity to a firm, as explained when discussing the specification of the second stage estimation, is the share of innovative sales in the total sales of the firm, ($SINS$), which can signal demand for R&D related activity and investment. In other words higher percentage of innovative sales can be indicative of a high demand for R&D investment, which in turn implies, given every thing else, a firm is more likely to find itself reporting financial constraints. Also, since ($SINS$) is positive only for innovators, its inclusion in the financial constraint equation controls for the fact that it is the innovators, who are the ones who mostly report financial constraints with respect to R&D[10]. Since the dummy for negative cash flows,

[10]Not including ($SINS$) in specification for the financial constraint equation resulted in some of the variable turning insignificant and some becoming marginally significant.

($DNCF$), is derived from the realized cash flows, removing cash flows from the specification results in a significant coefficient for ($DNCF$). That is, we find that negative realization of cash flows, leads to a significant increase in the probability of a firm being financially constrained. We also tried a specification, where we removed ($DNCF$). This resulted in cash flows being marginally significant with a positive sign.

Now, we began with the question whether "shifts" in the supply of internal equity finance, that is, whether changes in the cash flows realized, can affect the financing constraints faced by the firms. Instead, what we find is that realized cash flows proxies shifts in demand for investment, which given every thing else, leads a firm to report it self as financially constrained. However, given the fact that negative realization of cash flows leads a firm to report it self as more financially constrained, it seems that dummy for negative cash flows does a better job in accounting for shifts in supply of internal finance than cash flows. This is because, while on the margin a small positive income shock can turn negative realization of cash flows into positive realization, on the average a very large income shock, which is an extreme event, will be need to turn negative realization of cash flows into positive realization. The point, therefore is that dummy variable, ($DNCF$), is by and large purged of the demand signals, and hence performs as a better variable in accounting for shifts in the supply of internal equity finance. To the extent that share of innovative sales, ($SINS$), can control for the demand signals and ($DNCF$) can account for the shifts in supply of internal funds, our preferred specification for the second stage is the one in columns **(C0)** and **(C1)**.

Though it is true that innovative firms are averse to debt financing, it may be that it does not have have the requisite amount of internal finance to finance their R&D activity. In such a situation a firm has to resort to external sources of financing. Borrowing as we know, is a prevalent means of financing investment, but is costly due to capital market imperfections. Given the nature of R&D, such as the intangibility and uniqueness of assets and the unpredictability of outcomes, the degree of market imperfection related R&D is expected to be even higher, implying that external sources of finance could be even more costly when it comes to R&D related activity. The notion of financial constraint that we employ is one of borrowing constraints, that is, firms are financially constrained when they reach their debt capacity. Depending on the amount of internally available funds, the degree of market imperfection it faces, and the amount of existing debt it services, a firm may or may not have reached its debt capacity. We had stated earlier that we assume the latent variable $f^*_{it-1,t}$, underlying $f_{it-1,t}$, as the shadow price of debt. This implies that if borrowing constraints constraints for R&D firms are indeed binding, then controlling for other factors,

$f^*_{it-1,t}$ will be high. Now, controlling for other factors, a firm with higher leverage is more likely to have reached its debt limit, which makes the borrowing constraints more binding. Our results suggests that this is indeed the case, that is, we find a highly leveraged firm to be more likely to be financially constrained. We find that the scaled correlation between ζ_{it}, the unobserved idiosyncratic component of the financial constraint equation, and ϵ_{it}, the idiosyncratic component of the reduced form debt equation, $\frac{\rho_{\zeta\epsilon}}{\sigma_{\underline{\zeta}}}$, to be significant, suggesting a strong simultaneity in the choice of leverage/capital structure and the financial constraint with respect to R&D investment. This shouldn't be surprising, since after all, we are talking about shadow price of external finance related to endogenous borrowing constraints in the firm's optimal financing and real decisions.

We also find that firms that maintain a higher amount of liquidity reserve are less financially constrained. In anticipation of financing needs and the need to avoid costly external means of finance, firms do practice precautionary savings. Given that nature of R&D, as discussed earlier, the need to finance R&D investment, through retained earnings might even be higher. However, it is also possible that high cash holdings, simply reflect the fact that these firms are profitable firms with holdings in excess of their financing needs for real investment. We also find that firms that pay dividends are significantly less constrained, indicating that such firms are profitable firms with internally generated funds more than the financing needs of the firm.

In our specification for financial constraint equation we also included market share, size and age of the firm as a proxy for degree of informational asymmetry, which we would like to control for. Large, mature and firms enjoying higher degree of monopoly are generally well placed to mitigate informational asymmetry problems such as adverse selection. Hennessey and Whited (2007) find that large firms face small cost of external finance as compared to small firms. However, what we find is that size has a significant positive sign, indicating that large firms are more financially constrained. Here we would like to point out that, it is mostly the innovators who answer to the questions relating to financial constraints: 100% for CIS2.5, 82% for CIS3, and 83% for CIS3.5. Therefore, somehow we need to control for the fact that the information on being financial constraint reported is the constraint with respect to R&D, which only the innovators are more likely to answer. The strong positive sign for size in the financial constraint equation is due to the fact that size of the firm, to a large extent explains the innovation decision of the firm, and is therefore able to control for the fact that it is the innovators who mostly report on being financially constrained. To confirm this, we drop size, not reported here. Since size and age of firms are positively correlated, dropping size, results in age being significantly positive.

Our results suggests that firms with large market share are less financially constrained with respect to R&D. Higher degree of monopoly implies that demand for its product will be more inelastic, which in turn can enable the firm to be more profitable by exercising discretionary pricing. Also, firms enjoying a higher degree of monopoly could have a better access to capital markets. We also find that matured firms are less financially constrained. This is for the simple reason that old firms having survived through time have built reputation over the years, and are therefore less likely to face adverse information asymmetry problems, as compared to young firms.

Finally, we discuss a result that pertains to Figure 6.2. Figure 6.2 plots the average partial effect of leverage on the propensity to innovate conditional on being financially *constrained* and conditional being financially *unconstrained*. We plot the APE of leverage against size, age and leverage. We find that conditional on *not* being financially constrained, the APE of leverage on innovation to be negative and almost constant over the the distribution of size, age and leverage. In contrast, the APE of leverage on innovation conditional on being financially constrained varies widely over the distribution of age, size and leverage, and is less in absolute terms when compared to the APE of leverage on innovation conditional on not being financially constrained. This indicates that under no financial constraints innovative firms, regardless of size, maturity, and existing level of debt, would *uniformly*, given the nature of R&D related activity, be less inclined to finance itself with debt. In other words, when borrowing constraints do not bind and debt is accessible on easier terms, and if for some reason the firm has to finance itself with debt, then it is very unlikely that it will do it for innovation purposes or it is unlikely that it is an innovating firm. The following scenario might explain: suppose there is a profitable firm, that has a substantial amount of cash holdings, that it can distribute to its share holders. Being profitable, it is likely that it has a rather large debt capacity and suppose its existing debt levels are such that it has not reached its debt capacity. In such a situation, the firm can distribute cash and borrow more to finance its investment. However, if it decides to innovate or spend more on R&D related activity, then as our results suggests, it would be less inclined to distribute cash as dividends, be more inclined to maintain a high cash reserves and not borrow more, in other words, finance itself with retained earnings.

When financial constraints set it, innovating firms, though still averse to debt financing, do borrow as is reflected in the relatively higher marginal propensity to innovate with respect to marginal increase in leverage as compared to the marginal propensity to innovate of the unconstrained firms. Now, under financial constraints, as Lambrecht and Myers (2008) explain, there can be two possibilities: (a) postpone investment or (b) borrow more to invest. Given the fact that most of the firms that report being financially constrained are innovators, it is true that these firms have not entirely abandoned innovative activity. Therefore, given our result that under constraint, the propensity to innovate with respect to leverage is relatively higher than under no financial constraints, it suggests that some projects might have been valuable enough to pursue by borrowing, even if that implied a higher cost. However, what we find is that the marginal propensity to innovate with respect to leverage varies with the distribution of size, age and leverage conditional on being constrained. This is because under financial constraints, the relative cost of external financing perceived by the firm depends on its age, size and the existing levels of debt. Take for example, the age of the firm, now even though we find that younger firms are more financially constrained, it is the young firms that are more innovative. This is because for younger firms survival and subsequent growth depends on their innovation. Hence, under financial constraints young firms are more willing to borrow than matured firms. Consequently, we find the marginal propensity to innovate with respect to leverage, though negative, is higher compared to a matured firm. Another line of reasoning supporting this finding is that of Cooley and Quadrini (2001), who argue that young firms typically operate on a smaller scale compared to mature firms, and therefore younger firm's profit are less susceptible to exogenous shocks as compared to mature firms. Given that a firm's objective function is concave in profits, higher volatility in profits imply a lower valuation of the firm, which makes further borrowing by matured firms operating on larger scale more problematic.

Conditional on being financially constrained, the marginal or the partial propensity to innovate with respect to marginal increase in leverage, asymptotes to zero as size of the firm increases. This is because, conditional on other covariates, as the size of the firm increases, for reasons discussed earlier, a firm becomes more certain to innovate and consequently financial constraints and marginal increase in leverage have little effect on the marginal propensity to innovate. In other words as the size of the firm increases, the effect of size dominates. However, under constraint small and medium sized firms, are still are quite unlikely to engage in R&D activity with higher debt in their capital structure.

6.2 Third Stage: R&D Switching Regression

In this section we discuss the structural estimates of the R&D equation switching model. The obtained coefficients are shown in Table 6.5 and the APE's of the variables in Table

6. As explained earlier, the switching regression model is augmented with correction terms, obtained after first and second stage estimation. The correction terms correct for the bias arising due to endogeneity of the choice of leverage, financial constraints faced and the decision to innovate. The correction terms constructed are on the basis of first stage reduced form regression, discussed earlier, and second stage specification, column **(C0)** and **(C1)** reported in Table 6.3.

We find that financial constraints do adversely and significantly affect R&D expenditure of firms. Given our interpretation of $f^*_{it-1,t}$ as shadow price of external finance, faced with borrowing constraints, the result suggests that costly external sources of finance do have a bearing on R&D activity. Table 6.5 also shows that for financially constrained firms R&D investment is more sensitive to cash flows as compared to the firms that are not financially constrained. The sensitivity of R&D investment to cash flows is significant for the constrained firms but not for the unconstrained firms. However, as stressed in an earlier discussion, the test for the impact of financing frictions on investment expenditure in our model does not rely on the comparison of sensitivity of investment to cash flows for the constrained and unconstrained firms. Such differences in the sensitivity of investment to cash flows can arise, since, as Gomes (2001) argues constrained firms have a higher marginal productivity of capital and invest more than those unconstrained. They also have a higher value of Tobin's "q" on average. Also, as Hennessy and Whited (2007) point out, inability of proxies, in our model share of innovative sales in the total sales of the firms, $SINS$, to perfectly control for investment opportunity can give predictive power to cash flows in investment regression equations, when cash flows correlated with the investment opportunity set of the firms. Therefore, instead of interpreting the higher and significant coefficient of cash flows for financially constrained firms, it should be interpreted as controlling for investment opportunity for constrained firms.

We find that firms whose current share of innovative sales are high, are more likely to be R&D intensive. Not including ($SINS$) in the specification, not reported here, resulted in a larger coefficient for cashflows, insignificant positive sign for $f_{it-1,t}$, and insignificant coefficients for some of the correction terms. This suggests that share of innovative sales in the total sales of the firm, ($SINS$), among others, is a also a function of demand signals or income shocks[11].

We also find that firms that payout dividends invest less in R&D. This seems to be more true for firms that are financially constrained with respect to R&D related activity. This corroborates the result of negative relation between dividend payout and innovation obtained

[11] We also estimated the system of equations with ($SINS_{it,t+1}$), reducing the number of observations to about half, however, the results remained qualitatively the same.

in the second stage. The prime reason for the the negative correlation being that innovation is a risky venture, and therefore lenders demand a premium on loans offered. Consequently, R&d intensive firms prefer internal source of financing, which makes them much less likely to distribute cash. The argument, it seems is strengthened by the fact that firms that are financially constrained with respect to R&D related activity, and if for some reasons are distributing cash or have to, are even less likely to take up R&D activity.

We also find that firms with large market share invest more in R&D, if only to maintain their monopoly position in the market. However, as suggested by the APE's in column **(b)** of Table 6.6, under financial constraints firms enjoying monopoly power could also be less R&D intensive. We find that large firms are less R&D intensive, even though as the second stage results suggest, are more likely to be innovators. This should not be surprising, since smaller firms are the ones that are more profitable and experience higher growth. In their bid to grow, they exhibit risky behavior both in terms of investment and borrowing, and consequently also experience higher volatility of growth. Also, for larger firms, investing as-much-as or proportionately more would imply subjecting themselves to higher risk. Since large firms operate on a larger scale and are therefore more subject to exogenous shocks, investing as-much-as or proportionately more than small firms would imply, tying up more capital in some risky venture, when they already are more subject to exigencies of nature. Our findings also suggests that young firms are more R&D intensive, corroborating the result of the second stage, where we found them to more innovative. The arguments presented earlier in the discussion of the second stage results, for young firms being more innovative, also apply here. Firms that have a number of enterprises consolidated within them, $DMULTI$, are more likely to R&D intensive. Now, such firms can be large, but as stated earlier, the various enterprises constituting the firm are more or less legal entities, with certain degree of autonomy. These enterprises can have their innovation policy. Thus, such firms are expected to be engaged in diverse activity and products. Hence, such firms, even being large, are found to be more R&D intensive. However, as the the APE's in column **(b)** of Table 6.6 suggests, the $DMULTI$ firms, under financial constraints, can behave otherwise.

In our analysis we also find that the various correlation term as $\rho_{\eta_1\epsilon}\sigma_{\eta_1}$, $\sigma_{\underline{\eta}_1}\rho_{\underline{\eta}_1\underline{\zeta}}$, $\sigma_{\underline{\eta}_1}\rho_{\underline{\eta}_1\underline{\upsilon}}$, $\rho_{\eta_0\epsilon}\sigma_{\eta_0}$, $\sigma_{\underline{\eta}_0}\rho_{\underline{\eta}_0\underline{\zeta}}$, and $\sigma_{\underline{\eta}_0}\rho_{\underline{\eta}_0\underline{\upsilon}}$ to be significant, suggesting strong simultaneity in the decision to innovate, the choice of capital structure and the amount of R&D expenditure. The significance of μ, the factor loading in the switching model also suggest that time invariant individual effects that affect the decision to innovate, the choice of capital structure also affect decision on the amount of R&D expenditure.

7 Conclusion

In this chapter we presented an empirical strategy for modelling a firm that chooses to be an innovator, decides on the extent leverage, and how much it wants to spend on R&D, where all these decisions are determined endogenously. In its decision to innovate and in deciding how much to spend on R&D activity, the firm faces endogenous financial constraints. The strategy entails estimating a system of four simultaneous equations, estimated in three steps. The first step being, estimation of the reduced form leverage equation. The second step involves, conditional on first step estimates, joint estimation of the structural equations relating to the decision to innovate and the financial constraints faced. Finally, in the third stage we estimate a switching regression model for R&D investment, conditional on being an innovator and the first and second stage estimates. While such a multi-step procedure is not uncommon for cross sectional data, in panel data one has to account for common time invariant individual effect that affects all of the firm's decision and the financial constraints it faces. Our approach, in which we replace the time invariant individual effect in the structural equations by its expected a posteriori value, obtained from the first stage estimates, allows us to consistently estimate the parameters of the subsequent stages.

Econometric methodology aside, the aim of the chapter has been to study what firm characteristics are associated with a firm's decision to innovate and the financial constraints it faces with respect to R&D activity and how financial constraints affect R&D investment. Among other firm characteristics, the chapter investigates, how firm's leverage is related to the firms's decision to innovate and the financial constraints it faces with respect to R&D activity. We find that a highly leveraged firm is less likely to be an innovator and more likely to be financially constrained. We find that financial constraints adversely affects a firm's R&D expenditure, reducing the R&D expenditure by almost two thirds, for an average firm.

We find that large, young and firms enjoying a higher degree of monopoly are more likely to be innovators. Also, firms that have many enterprises consolidated within them are more likely to be innovators. We also find that young firms are more likely to be facing financial constraints, while firms that enjoy a monopoly position are less constrained. As far a R&D investment is concerned, we find that it is the small and young firms that are more R&D intensive than large and mature firms and that firms with large market share and multi-enterprise firms are more R&D intensive.

Also, the behavior of firms, conditional on being financially constrained and conditional on being unconstrained, with regard to innovation and R&D intensity can be characteristically different. Under no financial constraints, the propensity to innovate with respect to leverage is negative and lower as compared to a situation in which firms find them selves financially constrained. Also, the propensity to innovate under no financial constraints, almost does not vary with firm characteristics such as maturity, size and leverage, while under financial constraints, the propensity to innovate with respect to leverage varies with the distribution of firm characteristics. As far as R&D intensity is concerned, while unconditionally multi-enterprise and firms enjoying monopoly power are likely to be highly R&D intensive, under constraints the R&D intensity of such firms decreases. The results help us in understanding how incentives to innovate and capital market imperfections, with respect to financing R&D activity, facing the firms interact, and how these are distributed across firm characteristics.

Finally, given our results that financial constraints adversely affect R&D investment, and the fact that we have utilized almost a third of the CIS data, it seems to us that the effect of financial constraints for aggregate R&D investment for the economy can be considerable, which in turn can impediment aggregate growth.

For future research, within the econometric technique we have developed, we seek to endogenize some of the variables that could be potentially endogenous and compute the standard errors of the APE.

References

Aboody, D., and B. Lev (2000): "Information Asymmetry, R&D, and Insider Gains," *Journal of Finance*, 55, 2747 – 2766.

Akerlof, G. (1970): "The Market for "Lemons": Quality, Uncertainty, and the Market Mechanism," *The Quarterly Journal of Economics*, 84, 488–500.

Albuquerque, R., and H. A. Hopenhayn (2004): "Optimal Lending Contracts and Firm Dynamics," *Review of Economic Studies*, 71, 285–315.

Alderson, M., and B. L. Betker (1996): "Liquidation Costs and Accounting Data," *Financial Management*, 25, 25–36.

Allen, F., and D. Gale (2000): *Comparing Financial Systems*. MIT Press.

Almeida, H., and M. Campello (2002): "Financial Constraints and Investment-Cash Flow Sensitivities: New Research Directions," Working Paper, New York University, Stern Business School.

Alti, A. (2003): "How Sensitive is Investment to Cash Flow when Financing is Frictionless?," *Journal of Finance*, 58, 707–722.

ANTON, J., AND D. YAO (2001): "The Sale of Intellectual Property: Strategic Disclosure, Property Rights, and Incomplete Contracts," Working paper, The Wharton School, University of Pennsylvania.

AUDRETSCH, D. (1995): *Innovation and Industry Evolution*. MIT Press.

BHATTACHARYA, S., AND J. R. RITTER (1983): "Innovation and Communication: Signaling with Partial Disclosure," *Review of Economic Studies*, 50, 331–346.

BLASS, A., AND O. YOSHA (2001): "Financing R&D in Mature Companies: An Empirical Analysis," Working Paper, Bank of Israel, Tel Aviv University, and CEPR.

BLUNDELL, R., AND J. POWELL (2003): "Endogeneity in Nonparametric and Semiparametric Regression Models," in *Advances in Economics and Econonometrics: Theory and Applications, Eighth World Congress, Vol. II*, ed. by L. H. M. Dewatripont, and S. Turnovsky. Cambridge University Press, Cambridge.

BOND, S., AND C. MEGHIR (1994): "Dynamic Investment Models and the Firm's Financial Policy," *Review of Economic Studies*, 61, 197–222.

BROWN, J. R., S. M. FAZZARI, AND B. C. PETERSEN (2009): "Financing Innovation and Growth: Cash Flow, External Equity, and the 1990's R&D Boom," *Journal of Finance*, 64, 151–185.

CLEARY, S. (1999): "The Relationship between Firm Investment and Financial Status," *Journal of Finance*, 54, 673–692.

COHEN, W. M., AND S. KLEPPER (1996): "A Reprise of Size and R&D," *The Economic Journal*, 106, 925–951.

COOLEY, T. F., AND V. QUADRINI (2001): "Financial Markets and Firm Dynamics," *American Economic Review*, 91, 1286–1310.

FAMA, E. F., AND K. R. FRENCH (2000): "Forecasting Profitability and Earnings," *Journal of Business*, 73, 161–175.

——— (2002): "Testing Trade-off and Pecking Order Predictions about Dividends and Debt," *Review of Financial Studies*, 15, 1–37.

FAZZARI, S. M., R. G. HUBBARD, AND B. PETERSEN (1988): "Financing Constraints and Corporate Investment," *Brookings Papers on Economic Activity*, 1, 144–195.

Gale, D., and M. Hellwig (1985): "Incentive Compatible Debt Contracts: The One Period Problem," *Review of Economic Studies*, 52, 647–663.

Gibbons, R. D., D. Hedeker, S. C. Charles, and P. Frisch (1994): "A Random-Effects Probit Model for Predicting Medical Malpractice Claims," *Journal of the American Statistical Association*, 89, 760–767.

Gomes, J. F. (2001): "Financing Investment," *American Economic Review*, 91, 1263–1285.

Greene, W. H. (2002): *Econometric Analysis*. Prentice Hall, 4^{th} edn.

Greenwald, B., J. Stiglitz, and A. Weiss (1984): "Informational Imperfections in the Capital Market and Macroeconomic Fluctuation," *American Economic Review*, 74, 194–199.

Hall, B. (1994a): "Corporate Capital Structure and Investment Horizons in the United States,1976-1987," *Business History Review*, 68, 110–143.

——— (1994b): "R&D Tax Policy During the Eighties: Success or Failure?," *Tax Policy and the Economy*, 7, 1–36.

Hall, B., and J. Lerner (2010): "The Financing of R&D and Innovation," in *Handbook of the Economics of Innovation*, ed. by B. H. Hall, and N. Rosenberg. Cambridge University Press, Cambridge, forthcoming.

Hart, O., and J. Moore (1994): "A theory of Debt based on the Inalienability of Human Capital," *Quarterly Journal of Economics*, 109, 841–879.

Heckman, J. (1979): "Sample selection bias as a specification error," *Econometrica*, 47, 153–161.

Hennessy, C., and T. M. Whited (2007): "How Costly is External Financing? Evidence from a Structural Estimation," *Journal of Finance*, 62, 1705–1745.

Holmstrom, B. (1989): "Agency Costs and Innovation," *Journal of Economic Behavior and Organization*, 12, 305–327.

Huergo, E., and J. Jaumandreu (2004): "How does Probability of Innovation change with Firm Age?," *Small Business Economics*, 22, 193–207.

Jensen, M., and W. Meckling (1976): "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," *Journal of Financial Economics*, 3, 305–360.

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

——— (2000): "Investment-Cash Flow Sensitivity are not Valid Measures of Financing Constraints," *Quarterly Journal of Economics*, 115, 707–712.

Lambrecht, B., and S. C. Myers (2008): "Debt and Managerial Rents in a Real-Options Model of the Firm," *Journal of Financial Economics*, 89, 209–231.

Leland, H. E., and D. H. Pyle (1977): "Informational Asymmetries, Financial Structure, and Financial Intermediation," *Journal of Finance*, 32, 371–387.

Moyen, N. (2004): "Investment-Cash Flow Sensitivities: Constrained versus Unconstrained Firms," *Journal of Finance*, 69, 2061–2092.

Murphy, K. M., and R. H. Topel (1985): "Estimation and Inference in Two-Step Econometric Models," *Journal of Business & Economic Statistics*, 3, 370–379.

Myers, S. (1977): "Determinants of corporate borrowing," *Journal of Financial Economics*, 5, 147–175.

——— (1984): "The Capital Structure Puzzle," *Journal of Finance*, 39, 575–592.

Myers, S., and N. Majluf (1984): "Corporate Financing and Investment Decisions when Firms have Information that Investors do not have," *Journal of Financial Economics*, 13, 187–222.

Newey, W. K. (1984): "A Method of Moment Interpretation of Sequential Estimators," *Economics Letters*, 14, 201–206.

Newey, W. K., and D. McFadden (1986): "Large Sample Estimation and Hypothesis Testing," in *Handbook of Econometrics*, ed. by R. F. Engle, and D. McFadden, pp. 2111–2245. Elsevier.

Nilsen, O. A., and F. Schiantarelli (2003): "Zeros And Lumps In Investment: Empirical Evidence On Irreversibilities and Nonconvexities," *The Review of Economics and Statistics*, 85, 1021–1037.

Opler, T., and S. Titman (1993): "The Determinants of Leveraged Buyout Activity: Free Cash Flow vs. Financial Distress Costs," *Journal of Finance*, 48, 1985–1999.

——— (1994): "Financial Distress and Corporate Performance," *Journal of Finance*, 49, 1015–1040.

Raymond, W., P. Mohnen, F. Palm, and S. S. van der Loeff (2007): "Persistence of innovation in Dutch manufacturing: Is it spurious?," *Review of Economics and Statistics*, forthcoming.

Shi, C. (2003): "On the Trade-off between the Future Benefits and Riskiness of R&D: A Bondholders' Perspective," *Journal of Accounting and Economics*, 35, 227–254.

Stein, J. C. (2003): "Agency, Information and Corporate Investment," in *Handbook of the Economics of Finance*, ed. by M. H. George Constantinides, and R. Stulz, pp. 111–165. Elsevier.

Stiglitz, J., and A. Weiss (1981): "Credit Rationing and Markets with Imperfect Information," *American Economic Review*, 71, 393–411.

——— (1983): "Incentive Effects of Terminations: Applications to the Credit and Labor Markets," *American Economic Review*, 73, 912–927.

Titman, S., and R. Wessels (1988): "The Determinants of Capital Structure Choice," *Journal of Finance*, 43, 1–19.

Townsend, R. (1979): "Optimal Contracts and Competitive Markets with Costly State Verification," *Journal of Economic Theory*, 21, 265–293.

Whited, T. M. (2008): "What Can Cash Shortfalls and Windfalls Tell Us About Finance Constraints?," in *The Economics of Imperfect Markets*, ed. by G. Calcagnini, and E. Saltari. Springer, New York, forthcoming.

Whited, T. M., and G. Wu (2005): "Financial Constraints Risk," *Review of Financial Studies*, 19, 531–559.

Williamson, O. E. (1988): "Corporate Finance and Corporate Governance," *Journal of Finance*, 43, 567–591.

Wooldridge, J. M. (2004): "Estimating Average Partial Effects Under Conditional Moment Independence Assumptions," Working Paper, Michigan State University, Department of Economics.

——— (2005): "Unobserved Heterogeneity and Estimation of Average Partial Effects," in *Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, ed. by D. Andrews, and J. Stock, pp. 27–55. Cambridge University Press, Cambridge.

Appendix A: Identification of Structural Parameters with Expected a Posteriori Values of Individual Effects

In this section we discuss the identification of the parameters for the second stage . Though the second stage estimation involves a joint estimation of the probability of a firm being financially constrained and the probability of the firm being an innovator, here we discuss identification only for the parameters that are pertinent to the financial constraint equation. The arguments carry over to the joint estimation of the financial constraint and the selection equations and the third stage R&D switching regression model.

Suppose, we are interested only in the financial constraint affecting R&D, where the firm's leverage explains the financial constraint the firm faces. To obtain the structural estimates for financial constraint, consider the financial constraint equation, equation (3), and the reduced form Debt equation, equation (5),

$$f^*_{it-1,t} = \mathbf{z}^{f\prime}_{it}\varphi + d_{it}\varphi_d + \lambda\alpha_i + \zeta_{it} \tag{3}$$

$$d_{it} = \mathbf{z}'_{it}\delta + \alpha_i + \epsilon_{it}. \tag{5}$$

Since leverage is endogenous to the financial constraint that the firm faces, the idiosyncratic components ζ_{it} and ϵ_{it} are correlated. To obtain the structural estimates of the financial constraint equation, a two step procedure can be employed, in which consistent estimates of the reduced form parameters obtained in the first stage can be used to obtain parameters of the structural equation in the second stage by conditional likelihood or any other method of moment.

Now, given that ζ_{it} and ϵ_{it} follow a normal distribution with correlation $\rho_{\zeta\epsilon}$, we can write the linear projection of ζ_{it} on ϵ_{it} in error form as $\zeta_{it} = \rho_{\zeta\epsilon}\frac{\epsilon_{it}}{\sigma_\epsilon} + \bar{\zeta}_{it}$. Then conditional on d_i, Z_i, ϵ_{it} and α_i the projection of $f^*_{it-1,t}$, can be written as

$$f^*_{it-1,t} = \mathbf{z}^{f\prime}_{it}\varphi + d_{it}\varphi_d + \lambda\alpha_i + \rho_{\zeta\epsilon}\frac{\epsilon_{it}}{\sigma_\epsilon} + \bar{\zeta}_{it}, \tag{A-1}$$

where $\bar{\zeta}_{it}$, is independent of d_i, Z_i, ϵ_{it} and α_i, and is normally distributed with mean zero and variance $(1-\rho^2_{\zeta\epsilon})$[12]. The probability of a firm being financially constrained, in period t conditional on d_i, Z_i, α_i and ϵ_{it} is then given by

$$\begin{aligned} \Pr(f_{it-1,t} = 1 | d_i, Z_i, \alpha_i, \epsilon_{it}) &= \Phi\left(\frac{\mathbf{z}^{f\prime}_{it}\varphi + d_{it}\varphi_d + \lambda\alpha_i + \rho_{\zeta\epsilon}\frac{\epsilon_{it}}{\sigma_\epsilon}}{\sqrt{(1-\rho^2_{\zeta\epsilon})}}\right) \\ &= \Phi\left(\frac{\mathbf{z}^{f\prime}_{it}\varphi + d_{it}\varphi_d + \lambda\alpha_i + \rho_{\zeta\epsilon}(\frac{d_{it}-\mathbf{z}'_{it}\delta-\alpha_i}{\sigma_\epsilon})}{\sqrt{(1-\rho^2_{\zeta\epsilon})}}\right), \end{aligned} \tag{A-2}$$

[12]This essentially implies that conditional on ϵ_{it}, ζ_{it} is independent of d_{it}.

where Φ is cumulative density function of a standard normal. The log likelihood function of a firm being financially constrained in period t is given by

$$\begin{aligned}\boldsymbol{L}_{itf}(\varphi, \varphi_d, \rho_{\zeta\epsilon}, \lambda; \delta, \sigma_\epsilon | d_i, Z_i, \alpha_i, \epsilon_{it}) &= f_{it-1,t} \ln \Pr(f_{it-1,t} = 1 | d_i, Z_i, \alpha_i, \epsilon_{it}) \\ &+ (1 - f_{it-1,t}) \ln \Pr(f_{it-1,t} = 0 | d_i, Z_i, \alpha_i, \epsilon_{it}).\end{aligned}$$

The parameters, φ, φ_d, λ, $\rho_{\zeta\epsilon}$ in (A-2), could have been identified if we knew α_i, δ, σ_ϵ and σ_α, where σ_α the standard deviation of the marginal distribution of α_i . Such a two step procedure is quite standard in models for cross sectional data, where estimates of the first stage errors are employed to control for the endogeneity of the endogenous regressors. The problem, however, with models for panel data is that the estimates of $\mathrm{E}(\epsilon_{it}|d_i, Z_i, \alpha_i) = d_{it} - \mathbf{z}'_{it}\delta - \alpha_i$ remain unidentified since we do not observe α_i, even though δ, σ_ϵ and σ_α can be consistently estimated from the first stage reduced form estimation of equation (5). Here we show that using a two step procedure, where δ, σ_ϵ and σ_α are estimated in the first stage, $\mathrm{E}(\epsilon_{it}|d_i, Z_i)$ can still be identified, which can then be used for the identification of the structural parameters.

Before we discuss identification for our non-linear model, we first consider identification for a linear model. Let $f(\alpha|d, Z)$ be the conditional distribution of time invariant individual effect α_i conditional on d and Z. Assume for the moment that $f^*_{it-1,t}$ is some observed continuous variable. Now, since d_i, Z_i, and α_i are in the conditioning set, of which ϵ_{it} is a function of, we can suppress ϵ_{it}, while writing the conditional expectation of $f^*_{it-1,t}$. Thus we have $E(f^*_{it-1,t}|d_i, Z_i, \epsilon_{it}, \alpha_i) = E(f^*_{it-1,t}|d_i, Z_i, \alpha_i)$, which is given is given by

$$\begin{aligned}E(f^*_{it-1,t}|d_i, Z_i, \alpha_i) &= \mathbf{z}^{f\prime}_{it}\varphi + d_{it}\varphi_d + \lambda\alpha_i + \rho_{\zeta\epsilon}\frac{\epsilon_{it}}{\sigma_\epsilon} + E(\bar{\zeta}_{it}|d_i, Z_i, \alpha_i) \\ &= \mathbf{z}^{f\prime}_{it}\varphi + d_{it}\varphi_d + \lambda\alpha_i + \rho_{\zeta\epsilon}\frac{(d_{it} - \mathbf{z}'_{it}\delta - \alpha_i)}{\sigma_\epsilon} \\ &= \mathbf{z}^{f\prime}_{it}\varphi + d_{it}\varphi_d + \rho_{\zeta\epsilon}\frac{(d_{it} - \mathbf{z}'_{it}\delta)}{\sigma_\epsilon} + (\lambda - \frac{\rho_{\zeta\epsilon}}{\sigma_\epsilon})\alpha_i \qquad \text{(A-3)}\end{aligned}$$

For any individual, i, taking expectation of the above with respect to the conditional distribution of α, $f(\alpha|d, Z)$ we obtain

$$\begin{aligned}E(f^*_{t-1,t}|d, Z) = \int E(f^*_{t-1,t}|d, Z, \alpha)f(\alpha|d, Z)d(\alpha) = &\quad \mathbf{z}^{f\prime}_{t}\varphi + d_t\varphi_d + \rho_{\zeta\epsilon}\frac{(d_t - \mathbf{z}'_t\delta)}{\sigma_\epsilon} \\ + &\quad (\lambda - \frac{\rho_{\zeta\epsilon}}{\sigma_\epsilon}) \int \alpha f(\alpha|d, Z)d(\alpha) \qquad \text{(A-4)}\end{aligned}$$

Using Bayes rule we can write $f(\alpha|d, Z)$ as

$$f(\alpha|d, Z) = \frac{f(d, Z|\alpha)g(\alpha)}{h(d, Z)}, \qquad \text{(A-5)}$$

where g, and h are density functions. The above can be written as

$$f(\alpha|d, Z) = \frac{f(d|Z, \alpha)p(Z|\alpha)g(\alpha)}{h(d|Z)p(Z)},$$

Since by our assumption the time invariant individual effects, α_i, are independent of the exogenous variables Z_i, hence $p(Z|\alpha) = p(Z)$, that is,

$$f(\alpha|d, Z) = \frac{f(d|Z\alpha)g(\alpha)}{h(d|Z)} = \frac{f(d|Z,\alpha)g(\alpha)}{\int f(d|Z,\alpha)g(\alpha)d\alpha}, \tag{A-6}$$

Hence,

$$\begin{aligned}
\int \alpha f(\alpha|d, Z)d(\alpha) &= \int \frac{\alpha f(d|Z,\alpha)g(\alpha)d\alpha}{\int f(d|Z,\alpha)g(\alpha)d\alpha} \\
&= \frac{\int \alpha \prod_{t=1}^{T} f(d_t|Z,\alpha)g(\alpha)d\alpha}{\int \prod_{t=1}^{T} f(d_t|Z,\alpha)g(\alpha)d\alpha} \\
&= \frac{\int \alpha \prod_{t=1}^{T} \frac{1}{\sigma_\epsilon}\phi(\frac{d_t - \mathbf{z}_t'\delta - \alpha}{\sigma_\epsilon})g(\alpha)d\alpha}{\int \prod_{t=1}^{T} \frac{1}{\sigma_\epsilon}\phi(\frac{d_t - \mathbf{z}_t'\delta - \alpha}{\sigma_\epsilon})g(\alpha)d\alpha} \\
&= \frac{\int \sigma_\alpha\tilde{\alpha} \prod_{t=1}^{T} \frac{1}{\sigma_\epsilon}\phi(\frac{d_t - \mathbf{z}_t'\delta - \sigma_\alpha\tilde{\alpha}}{\sigma_\epsilon})\phi(\tilde{\alpha})d\tilde{\alpha}}{\int \prod_{t=1}^{T} \frac{1}{\sigma_\epsilon}\phi(\frac{d_t - \mathbf{z}_t'\delta - \sigma_\alpha\tilde{\alpha}}{\sigma_\epsilon})\phi(\tilde{\alpha})d\tilde{\alpha}},
\end{aligned} \tag{A-7}$$

where the second and the third equality follows from the fact that conditional on Z and α, each of the d_t, $d_t \in \{d_1, \ldots, d_T\}$ are independently normally distributed with mean $\mathbf{z}_t'\delta + \alpha$ and standard deviation σ_ϵ. $g(\alpha)$ by our assumption is normally distributed with mean zero and variance σ_α^2 and $\tilde{\alpha} = \frac{\alpha}{\sigma_\alpha}$ follows a standard normal distribution. Given that ϕ is a smooth exponential function of its argument and denominator of the right hand side of equation (A-8) is the estimated marginal maximum likelihood, which converges to its true value on the LHS, it can be shown that

$$\frac{\int \sigma_\alpha\tilde{\alpha} \prod_{t=1}^{T} \frac{1}{\sigma_\epsilon}\phi(\frac{d_t - \mathbf{z}_t'\delta - \sigma_\alpha\tilde{\alpha}}{\sigma_\epsilon})\phi(\tilde{\alpha})d\tilde{\alpha}}{\int \prod_{t=1}^{T} \frac{1}{\sigma_\epsilon}\phi(\frac{d_t - \mathbf{z}_t'\delta - \sigma_\alpha\tilde{\alpha}}{\sigma_\epsilon})\phi(\tilde{\alpha})d\tilde{\alpha}} \overset{asy}{\equiv} \frac{\int \hat{\sigma}_\alpha\tilde{\alpha} \prod_{t=1}^{T} \frac{1}{\hat{\sigma}_\epsilon}\phi(\frac{d_t - \mathbf{z}_t'\hat{\delta} - \hat{\sigma}_\alpha\tilde{\alpha}}{\hat{\sigma}_\epsilon})\phi(\tilde{\alpha})d\tilde{\alpha}}{\int \prod_{t=1}^{T} \frac{1}{\hat{\sigma}_\epsilon}\phi(\frac{d_t - \mathbf{z}_t'\hat{\delta} - \hat{\sigma}_\alpha\tilde{\alpha}}{\hat{\sigma}_\epsilon})\phi(\tilde{\alpha})d\tilde{\alpha}}. \tag{A-8}$$

The right hand side of (A-8) is the estimated expected a posteriori value of α, $\hat{\alpha}(d, Z, \hat{\delta}, \hat{\sigma}_\epsilon, \hat{\sigma}_\alpha)$. To estimate the RHS of equation (A-8), numerical integration with respect to $\tilde{\alpha}_i$ can be carried out using Gauss-Hermite quadratures. Therefore under asymptotic equivalence we can write (A-4) as

$$\int E(f_{t-1,t}^*|d, Z, \alpha)f(\alpha|d, Z)d(\alpha) \overset{asy}{\equiv} \mathbf{z}_t^{f\prime}\varphi + d_t\varphi_d + \rho_{\zeta\epsilon}\frac{(d_t - \mathbf{z}_t'\delta)}{\sigma_\epsilon} + (\lambda - \frac{\rho_{\zeta\epsilon}}{\sigma_\epsilon})\hat{\alpha}(d, Z, \hat{\delta}, \hat{\sigma}_\epsilon, \hat{\sigma}_\alpha), \tag{A-9}$$

and since

$$\frac{d_{it} - \mathbf{z}_{it}'\delta - \hat{\alpha}_i}{\sigma_\epsilon} \overset{asy}{\equiv} \frac{d_{it} - \mathbf{z}_{it}'\hat{\delta} - \hat{\alpha}_i}{\hat{\sigma}_\epsilon},$$

hence

$$\int E(f_{t-1,t}^*|d, Z, \alpha)f(\alpha|d, Z)d(\alpha) \overset{asy}{\equiv} \mathbf{z}_t^{f\prime}\varphi + d_t\varphi_d + \lambda\hat{\alpha} + \rho_{\zeta\epsilon}(\frac{d_t - \mathbf{z}_t'\hat{\delta} - \hat{\alpha}}{\hat{\sigma}_\epsilon}). \tag{A-10}$$

Therefore, if population parameters, δ, σ_ϵ and σ_α were known, the above implies that we could write the linear predictor of $f^*_{t-1,t}$, given d and Z in error form as

$$\begin{aligned} f^*_{t-1,t} &= \mathbf{z}^{f\prime}_t\varphi + d_t\varphi_d + \rho_{\zeta\epsilon}\frac{(d_t - \mathbf{z}'_t\delta - \hat{\alpha})}{\sigma_\epsilon} + \lambda\hat{\alpha} + \tilde{\zeta}_t \\ &= \mathbf{z}^{f\prime}_t\varphi + d_t\varphi_d + \rho_{\zeta\epsilon}\frac{\hat{\epsilon}_t}{\sigma_\epsilon} + \lambda\hat{\alpha} + \tilde{\zeta}_t, \end{aligned} \tag{A-11}$$

where $\hat{\epsilon}_t = d_t - \mathbf{z}'_t\delta - \hat{\alpha} = \mathrm{E}(\epsilon_t|d, Z)$, which is identified, as claimed earlier. Conditional of d and Z, $\tilde{\zeta}_t$ is i.i.d. and normally distributed with mean 0 and variance $\sigma^2_{\tilde{\zeta}}$ and $\hat{\alpha}$ in the above equation is defined at population parameters. For liner models, say if $f^*_{t-1,t}$ was observed and continuous, the parameters of interest, φ, φ_d and $\rho_{\zeta\epsilon}$, can be consistently estimated by running a pooled OLS regression, with estimated a posteriori values, $\hat{\alpha}$ substituted for α and consistent first step estimates $\hat{\sigma}_\epsilon$ and $\hat{\delta}$ for σ_ϵ and δ respectively. The test of exogeneity of d_t can be carried carried out with by testing the significance of the estimates of $\rho_{\zeta\epsilon}$ and λ. The average partial effect of any exogenous variable z_k and endogenous d are respectively given by

$$\frac{\partial E(f^*_{it-1,t}|d_i, Z_i)}{\partial z_{kit}} = \varphi_k \qquad \& \qquad \frac{\partial E(f^*_{it-1,t}|d_i, Z_i)}{\partial d_{it}} = \varphi_d. \tag{A-12}$$

Now let us consider our non-linear model, where $f^*_{t-1,t}$, is a latent variable underlying $f_{t-1,t}$ which takes value 0 or 1. It follows from (A-11) and the fact that in probit models the the parameters are identified only up to a scale, probability of being financially constrained, given d_i an Z_i, is given by

$$\begin{aligned} \Pr(f_{it-1,t} = 1|d_i, Z_i, \hat{\epsilon}_{it}, \hat{\alpha}_i) &= \Pr(f^*_{it-1,t} > 0|d_i, Z_i, \hat{\epsilon}_t, \hat{\alpha}_i) \\ &= \Phi\left(\mathbf{z}^{f\prime}_{it}\varphi\frac{1}{\sigma_{\tilde{\zeta}}} + d_{it}\frac{\varphi_d}{\sigma_{\tilde{\zeta}}} + \frac{\lambda}{\sigma_{\tilde{\zeta}}}\hat{\alpha}_i + \frac{\rho_{\zeta\epsilon}}{\sigma_{\tilde{\zeta}}}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon}\right). \end{aligned} \tag{A-13}$$

However, $\Pr(f_{it-1,t} = 1|d_i, Z_i, \hat{\epsilon}_{it}, \hat{\alpha}_i) \neq \Pr(f_{it-1,t} = 1|d_i, Z_i, \epsilon_{it}, \alpha_i)$, and for any individual i, the mean effect or the average partial effect (APE) of changing a variable say z_k in time period t from z_{kt} to $z_{kt} + \kappa$ is given by

$$\begin{aligned} \frac{\Delta \Pr(f_{t-1,t} = 1)}{\Delta z_{kt}} =& \Big[\int \Pr(f_{t-1,t} = 1|d, Z_{-k}, z_{k_{-t}}, (z_{kt} + \kappa), \alpha, \epsilon_t)dg(\alpha, \epsilon_t) \\ &- \int \Pr(f_{t-1,t} = 1|d, Z_{-k}, z_{k_{-t}}, z_{kt}, \alpha, \epsilon_t)dg(\alpha, \epsilon_t)\Big]/\kappa, \end{aligned} \tag{A-14}$$

where $g(\alpha, \epsilon_t)$ is the distribution function of α and ϵ_t. Hence, to recover the above measure in (A-14), like Chamberlain (1984), we assume a distribution for α_i conditional on d_i and Z_i:

$$\alpha_i = \Psi(d_i, Z_i) + \bar{\alpha}_i,$$

where $\bar{\alpha}_i$ is normally distributed with mean 0, variance $\sigma^2_{\bar{\alpha}}$ and is independent of every thing else. However, for any individual i, we have shown above that,

$$\mathrm{E}(\alpha|d, Z) = \int \alpha f(\alpha|d, Z) d(\alpha) = \Psi(d, Z) = \hat{\alpha}(d, Z, \delta, \sigma_\epsilon, \sigma_\alpha).$$

The above implies that conditional on d and Z, ϵ_t is distributed as

$$\epsilon_t = d_t - \mathbf{z}'_t\delta - \hat{\alpha} - \bar{\alpha}_i = \hat{\epsilon}_t - \bar{\alpha}_i.$$

Hence, under the assumption about the conditional distribution of α_i, we can write the linear projection of $f^*_{it-1,t}$ as

$$f^*_{it-1,t} = \mathbf{z}^{f\prime}_{it}\varphi + d_{it}\varphi_d + \lambda\hat{\alpha}_i + \rho_{\zeta\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon} + (\lambda - \frac{\rho_{\zeta\epsilon}}{\sigma_\epsilon})\bar{\alpha}_i + \bar{\zeta}_{it}, \tag{A-15}$$

with the composite error term $(\lambda - \frac{\rho_{\zeta\epsilon}}{\sigma_\epsilon})\bar{\alpha}_i + \bar{\zeta}_{it} = \underline{\zeta}_{it}$, which is uncorrelated with any of the covariates and is normally distributed with mean zero and variance $(\lambda - \frac{\rho_{\zeta\epsilon}}{\sigma_\epsilon})^2\sigma^2_{\bar{\alpha}} + \sigma^2_{\bar{\zeta}} = \sigma^2_{\underline{\zeta}}$. Therefore, the above assumption about the conditional distribution of α_i implies that,

$$\Pr(f_{it-1,t} = 1|d_i, Z_i) = \Phi\left(\mathbf{z}^{f\prime}_{it}\varphi\frac{1}{\sigma_{\underline{\zeta}}} + d_{it}\frac{\varphi_d}{\sigma_{\underline{\zeta}}} + \frac{\lambda}{\sigma_{\underline{\zeta}}}\hat{\alpha}_i + \frac{\rho_{\zeta\epsilon}}{\sigma_{\underline{\zeta}}}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon}\right). \tag{A-16}$$

Since $\sigma_{\tilde{\zeta}}$ in (A-13) and $\sigma_{\underline{\zeta}}$ in (A-16) both remain unidentified, assuming both of them to be 1, we obtain the same estimate of the parameters.

Now, having assumed the conditional distribution of α_i, for any individual i, we now have

$$\Pr(f_{t-1,t} = 1|d, Z, \alpha, \epsilon_t) = \Pr(f_{t-1,t} = 1|d, Z, \hat{\alpha}(.), \hat{\epsilon}_t, \bar{\alpha})$$

and

$$\int \Pr(f_{t-1,t} = 1|d, Z, \hat{\alpha}(.), \hat{\epsilon}_t, \bar{\alpha}) dF(\hat{\alpha}(.), \hat{\epsilon}_t, \bar{\alpha}) = \int\int \Pr(f_{t-1,t} = 1|d, Z, \hat{\alpha}(.), \hat{\epsilon}_t, \bar{\alpha}) dG(\hat{\alpha}(.), \hat{\epsilon}_t) h(\bar{\alpha}) d\bar{\alpha},$$

where $F(\hat{\alpha}(.), \hat{\epsilon}_t, \bar{\alpha})$ is the joint distribution function of the arguments and $G(\hat{\alpha}(.), \hat{\epsilon}_t)$ is the joint distribution of $\hat{\alpha}(.)$ and $\hat{\epsilon}_t$. The equality above follows, since $\hat{\alpha}(.)$ and $\hat{\epsilon}_t$ by our assumption are independent of $\bar{\alpha}$. Also, since $\bar{\alpha}$ is independent of d and Z, we can write the above as

$$\begin{aligned}
&\int \Pr(f_{t-1,t} = 1|d, Z, \hat{\alpha}(.), \hat{\epsilon}_t, \bar{\alpha}) dF(\hat{\alpha}(.), \hat{\epsilon}_t, \bar{\alpha}) \\
&= \int\int \Pr(f_{t-1,t} = 1|d, Z, \hat{\alpha}(.), \hat{\epsilon}_t, \bar{\alpha}) h(\bar{\alpha}|d, Z, \hat{\alpha}(.), \hat{\epsilon}_t) d\bar{\alpha} dG(\hat{\alpha}(.), \hat{\epsilon}_t) \\
&= \int \Pr(f_{t-1,t} = 1|d, Z, \hat{\alpha}(.), \hat{\epsilon}_t) dG(\hat{\alpha}(.), \hat{\epsilon}_t)
\end{aligned} \tag{A-17}$$

Thus we have shown that

$$\int \Pr(f_{t-1,t} = 1|d, Z, \alpha, \epsilon_t) dg(\alpha, \epsilon_t) = \int \Pr(f_{t-1,t} = 1|d, Z, \hat{\alpha}(.), \hat{\epsilon}_t) dG(\hat{\alpha}(.), \hat{\epsilon}_t) \tag{A-18}$$

To obtain the the sample analog of RHS of (A-18), for fixed $d = \bar{d}$ and $Z = \bar{Z}$, we can compute

$$\frac{1}{\sum_{i=1}^{N} T_i} \sum_{i=1}^{N} \sum_{t=1}^{T_i} \Pr(f_{it-1,t} = 1 | \bar{d}, \bar{Z}, \hat{\alpha}_i(.), \hat{\epsilon}_{it}). \tag{A-19}$$

With (A-19) we can now compute (A-14), the mean effect of changing a variable say z_k from $\bar{z}_k$ to $\bar{z}_k + \kappa$ given $d = \bar{d}$ and $Z_{-k} = \bar{Z}_{-k}$.

Now consider the selection equation. The projection of $s^*_{it-1,t}$, the latent variable underlying selection, in error form conditional on ϵ_{it} is given by

$$s^*_{it-1,t} = \mathbf{z}^{s\prime}_{it}\gamma + d_{it}\gamma_d + \theta\alpha_i + \rho_{\upsilon\epsilon}\frac{\epsilon_{it}}{\sigma_\epsilon} + \bar{\upsilon}_{it}.$$

By the same logic that we used to arrive at equation (A-15) for the financial constraint equation, we can write the selection equation as

$$s^*_{it-1,t} = \mathbf{z}^{s\prime}_{it}\gamma + d_{it}\gamma_d + \theta\hat{\alpha}_i + \rho_{\upsilon\epsilon}\frac{\hat{\epsilon}_{it}}{\sigma_\epsilon} + \underline{\upsilon}_{it}, \tag{A-20}$$

where $\underline{\upsilon}_{it} = (\theta - \frac{\rho_{\upsilon\epsilon}}{\sigma_\epsilon})\bar{\alpha}_i + \bar{\upsilon}_{it}$ which is uncorrelated with any of the covariates and is normally distributed with mean zero and variance $(\theta - \frac{\rho_{\upsilon\epsilon}}{\sigma_\epsilon})^2\sigma^2_{\bar{\alpha}} + \sigma^2_{\bar{\upsilon}} = \sigma^2_{\underline{\upsilon}}$. However, $\underline{\zeta}_t$ and $\underline{\upsilon}_t$ are correlated, with the correlation coefficient being

$$\rho_{\underline{\zeta}\underline{\upsilon}} = \frac{(\lambda - \frac{\rho_{\zeta\epsilon}}{\sigma_\epsilon})(\theta - \frac{\rho_{\upsilon\epsilon}}{\sigma_\epsilon})\sigma^2_{\bar{\alpha}} + \rho_{\bar{\zeta}\bar{\upsilon}}}{\sigma_{\underline{\zeta}}\sigma_{\underline{\upsilon}}}.$$

Having assumed a conditional distribution for α_i, the R&D equations for the R&D switching regression model can now be written as

$$r_{1it} = \mathbf{z}^{r\prime}_{it}\beta + \beta_f + \beta_1(.) + \mu\hat{\alpha}_i + \rho_{\eta_1\epsilon}\frac{\sigma_{\eta_1}}{\sigma_\epsilon}\hat{\epsilon}_{it} + \underline{\eta}_{1it}, \tag{A-21a}$$

if $f^*_{it-1,t} > 0$,

$$r_{0it} = \mathbf{z}^{r\prime}_{it}\beta + \beta_0(.) + \mu\hat{\alpha}_i + \rho_{\eta_0\epsilon}\frac{\sigma_{\eta_0}}{\sigma_\epsilon}\hat{\epsilon}_{it} + \underline{\eta}_{0it}, \tag{A-21b}$$

if $f^*_{it-1,t} \leq 0$. The error terms $\underline{\eta}_{1it}$ and $\underline{\eta}_{0it}$ are respectively given by $\underline{\eta}_{1it} = (\mu - \frac{\rho_{\eta_1\epsilon}}{\sigma_\epsilon})\bar{\alpha}_i + \bar{\eta}_{1it}$ and $\underline{\eta}_{0it} = (\mu - \frac{\rho_{\eta_0\epsilon}}{\sigma_\epsilon})\bar{\alpha}_i + \bar{\eta}_{0it}$, which by assumption, are normally distributed with mean zero and variances $\sigma^2_{\underline{\eta}_1} = (\mu - \frac{\rho_{\eta_1\epsilon}}{\sigma_\epsilon})^2\sigma^2_{\bar{\alpha}} + \sigma^2_{\bar{\eta}_1}$ and $\sigma^2_{\underline{\eta}_0} = (\mu - \frac{\rho_{\eta_0\epsilon}}{\sigma_\epsilon})^2\sigma^2_{\bar{\alpha}} + \sigma^2_{\bar{\eta}_0}$ respectively. We finally note that $\underline{\eta}_{1it}$ and $\underline{\eta}_{0it}$ are correlated with $\underline{\zeta}_{it}$ and $\underline{\upsilon}_{it}$. The correlation between $\underline{\eta}_{1it}$ and $\underline{\zeta}_{it}$ and between $\underline{\eta}_{1it}$ and $\underline{\upsilon}_{it}$ are respectively given by

$$\rho_{\underline{\eta}_1\underline{\zeta}} = \frac{(\mu - \frac{\rho_{\eta_1\epsilon}}{\sigma_\epsilon})(\lambda - \frac{\rho_{\zeta\epsilon}}{\sigma_\epsilon})\sigma^2_{\bar{\alpha}} + \rho_{\bar{\eta}_1\bar{\zeta}}}{\sigma_{\underline{\eta}_1}\sigma_{\underline{\zeta}}} \quad \& \quad \rho_{\underline{\eta}_1\underline{\upsilon}} = \frac{(\mu - \frac{\rho_{\eta_1\epsilon}}{\sigma_\epsilon})(\theta - \frac{\rho_{\upsilon\epsilon}}{\sigma_\epsilon})\sigma^2_{\bar{\alpha}} + \rho_{\bar{\eta}_1\bar{\upsilon}}}{\sigma_{\underline{\eta}_1}\sigma_{\underline{\upsilon}}}.$$

Replacing the 1's above with 0's we get the correlation between $\underline{\eta}_{0it}$ and $\underline{\zeta}_{it}$ and between $\underline{\eta}_{0it}$ and $\underline{\upsilon}_{it}$ respectively.

Appendix B: Derivation of the Correction Terms

To avoid complicating the notations, we denote the error components, $\underline{\eta}_1$, $\underline{\eta}_0$, $\underline{\zeta}$ and $\underline{\upsilon}$ of equations (12a), (12b), (13) and (14) respectively as η_1, η_0, ζ and υ. We know that the conditional expectation of η, where η is either η_1 or η_0, given ζ and υ, $\mathrm{E}[\eta|\zeta,\upsilon]$, is given by

$$\mathrm{E}[\eta|\zeta,\upsilon] = \mu_\eta + \frac{\sigma_\eta(\rho_{\eta\zeta} - \rho_{\eta\upsilon}\rho_{\zeta\upsilon})(\zeta - \mu_\zeta)}{\sigma_\zeta(1-\rho_{\zeta\upsilon}^2)} + \frac{\sigma_\eta(\rho_{\eta\upsilon} - \rho_{\eta\zeta}\rho_{\zeta\upsilon})(\upsilon - \mu_\upsilon)}{\sigma_\upsilon(1-\rho_{\zeta\upsilon}^2)}.$$

Since, $\mu_\eta = \mu_\zeta = \mu_\upsilon = 0$ we have,

$$\mathrm{E}[\eta|\zeta,\upsilon] = \frac{\sigma_\eta(\rho_{\eta\zeta} - \rho_{\eta\upsilon}\rho_{\zeta\upsilon})(\zeta)}{\sigma_\zeta(1-\rho_{\zeta\upsilon}^2)} + \frac{\sigma_\eta(\rho_{\eta\upsilon} - \rho_{\eta\zeta}\rho_{\zeta\upsilon})(\upsilon)}{\sigma_\upsilon(1-\rho_{\zeta\upsilon}^2)}.$$

Define, $\bar{\zeta} = \frac{\zeta}{\sigma_\zeta}$ and $\bar{\upsilon} = \frac{\upsilon}{\sigma_\upsilon}$, then

$$\mathrm{E}[\eta|\zeta,\upsilon] = \frac{\sigma_\eta(\rho_{\eta\zeta} - \rho_{\eta\upsilon}\rho_{\zeta\upsilon})\bar{\zeta}}{(1-\rho_{\zeta\upsilon}^2)} + \frac{\sigma_\eta(\rho_{\eta\upsilon} - \rho_{\eta\zeta}\rho_{\zeta\upsilon})\bar{\upsilon}}{(1-\rho_{\zeta\upsilon}^2)},$$

which can be written as

$$\mathrm{E}[\eta|\zeta,\upsilon] = \frac{\sigma_\eta\rho_{\eta\zeta}}{(1-\rho_{\zeta\upsilon}^2)}(\bar{\zeta} - \rho_{\zeta\upsilon}\bar{\upsilon}) + \frac{\sigma_\eta\rho_{\eta\upsilon}}{(1-\rho_{\zeta\upsilon}^2)}(\bar{\upsilon} - \rho_{\zeta\upsilon}\bar{\zeta}). \tag{B-1}$$

Hence,

$$\mathrm{E}[\eta|\zeta > -a, \upsilon > -b] = \mathrm{E}[\eta|\bar{\zeta} > \frac{-a}{\sigma_\zeta}, \bar{\upsilon} > \frac{-b}{\sigma_\upsilon}] = \frac{\int_{\frac{-b}{\sigma_\upsilon}}^{\infty}\int_{\frac{-a}{\sigma_\zeta}}^{\infty} \mathrm{E}[\eta|\bar{\zeta},\bar{\upsilon}]\phi_2(\bar{\zeta},\bar{\upsilon},\rho_{\zeta\upsilon})d\bar{\zeta}d\bar{\upsilon}}{\Phi_2\left(\frac{a}{\sigma_\zeta}, \frac{b}{\sigma_\upsilon}, \rho_{\zeta\upsilon}\right)}$$

$$= \frac{1}{\Phi_2\left(\frac{a}{\sigma_\zeta}, \frac{b}{\sigma_\upsilon}, \rho_{\zeta\upsilon}\right)} \frac{\sigma_\eta\rho_{\eta\zeta}}{(1-\rho_{\zeta\upsilon}^2)} \int_{\frac{-b}{\sigma_\upsilon}}^{\infty}\int_{\frac{-a}{\sigma_\zeta}}^{\infty} (\bar{\zeta} - \rho_{\zeta\upsilon}\bar{\upsilon})\phi_2(\bar{\zeta},\bar{\upsilon},\rho_{\zeta\upsilon})d\bar{\zeta}d\bar{\upsilon}$$

$$+ \frac{1}{\Phi_2\left(\frac{a}{\sigma_\zeta}, \frac{b}{\sigma_\upsilon}, \rho_{\zeta\upsilon}\right)} \frac{\sigma_\eta\rho_{\eta\upsilon}}{(1-\rho_{\zeta\upsilon}^2)} \int_{\frac{-b}{\sigma_\upsilon}}^{\infty}\int_{\frac{-a}{\sigma_\zeta}}^{\infty} (\bar{\upsilon} - \rho_{\zeta\upsilon}\bar{\zeta})\phi_2(\bar{\zeta},\bar{\upsilon},\rho_{\zeta\upsilon})d\bar{\zeta}d\bar{\upsilon}, \tag{B-2}$$

where, ϕ_2 and Φ_2 denote respectively the density and cumulative density function function of a standard bivariate normal. Now, consider the expression $\int_{\frac{-b}{\sigma_\upsilon}}^{\infty}\int_{\frac{-a}{\sigma_\zeta}}^{\infty}(\bar{\zeta} - \rho_{\zeta\upsilon}\bar{\upsilon})\phi_2(\bar{\zeta},\bar{\upsilon},\rho_{\zeta\upsilon})d\bar{\zeta}d\bar{\upsilon}$, of the RHS in (B-2). Given that $\phi_2(\bar{\zeta},\bar{\upsilon},\rho_{\zeta\upsilon}) = \phi(\bar{\zeta})\frac{1}{\sqrt{(1-\rho_{\zeta\upsilon}^2)}}\phi\left(\frac{\bar{\upsilon} - \rho_{\zeta\upsilon}\bar{\zeta}}{\sqrt{(1-\rho_{\zeta\upsilon}^2)}}\right)$, the concerned expression can be written as

$$\int_{\frac{-b}{\sigma_\upsilon}}^{\infty}\int_{\frac{-a}{\sigma_\zeta}}^{\infty}(\bar{\zeta} - \rho_{\zeta\upsilon}\bar{\upsilon})\phi(\bar{\zeta})\frac{1}{\sqrt{(1-\rho_{\zeta\upsilon}^2)}}\phi\left(\frac{\bar{\upsilon} - \rho_{\zeta\upsilon}\bar{\zeta}}{\sqrt{(1-\rho_{\zeta\upsilon}^2)}}\right)d\bar{\zeta}d\bar{\upsilon} =$$

$$\int_{\frac{-a}{\sigma_\zeta}}^{\infty}\bar{\zeta}\phi(\bar{\zeta})\left(1 - \Phi\left(\frac{\frac{-b}{\sigma_\upsilon} - \rho_{\zeta\upsilon}\bar{\zeta}}{\sqrt{(1-\rho_{\zeta\upsilon}^2)}}\right)\right)d\bar{\zeta} - \rho_{\zeta\upsilon}\int_{\frac{-b}{\sigma_\upsilon}}^{\infty}\int_{\frac{-a}{\sigma_\zeta}}^{\infty}\bar{\upsilon}\phi(\bar{\zeta})\frac{1}{\sqrt{(1-\rho_{\zeta\upsilon}^2)}}\phi\left(\frac{\bar{\upsilon} - \rho_{\zeta\upsilon}\bar{\zeta}}{\sqrt{(1-\rho_{\zeta\upsilon}^2)}}\right)d\bar{\zeta}d\bar{\upsilon}. \tag{B-3}$$

Now, let $y = \frac{\bar{\upsilon} - \rho_{\zeta\upsilon}\bar{\zeta}}{\sqrt{(1-\rho_{\zeta\upsilon}^2)}}$, then $dy = \frac{d\bar{\upsilon}}{\sqrt{(1-\rho_{\zeta\upsilon}^2)}}$. Having defined y, the right hand side of (B-3) can now be written as

$$\int_{\frac{-a}{\sigma_\zeta}}^{\infty}\bar{\zeta}\phi(\bar{\zeta})\left(1 - \Phi\left(\frac{\frac{-b}{\sigma_\upsilon} - \rho_{\zeta\upsilon}\bar{\zeta}}{\sqrt{(1-\rho_{\zeta\upsilon}^2)}}\right)\right)d\bar{\zeta} - \rho_{\zeta\upsilon}\int_{\frac{\frac{-b}{\sigma_\upsilon} - \rho_{\zeta\upsilon}\bar{\zeta}}{\sqrt{(1-\rho_{\zeta\upsilon}^2)}}}^{\infty}\int_{\frac{-a}{\sigma_\zeta}}^{\infty}(y\sqrt{(1-\rho_{\zeta\upsilon}^2)} + \rho_{\zeta\upsilon}\bar{\zeta})\phi(\bar{\zeta})\phi(y)d\bar{\zeta}dy$$

$$= \int_{\frac{-a}{\sigma_\zeta}}^{\infty}\bar{\zeta}\phi(\bar{\zeta})\left(1 - \Phi\left(\frac{\frac{-b}{\sigma_\upsilon} - \rho_{\zeta\upsilon}\bar{\zeta}}{\sqrt{(1-\rho_{\zeta\upsilon}^2)}}\right)\right)d\bar{\zeta}$$

$$-\rho_{\zeta\upsilon}\int_{\frac{\frac{-b}{\sigma_\upsilon}-\rho_{\zeta\upsilon}\bar{\zeta}}{\sqrt{(1-\rho^2_{\zeta\upsilon})}}}^{\infty}\int_{\frac{-a}{\sigma_\zeta}}^{\infty} y\sqrt{(1-\rho^2_{\zeta\upsilon})}\phi(\bar{\zeta})\phi(y)d\bar{\zeta}dy-\rho^2_{\zeta\upsilon}\int_{\frac{\frac{-b}{\sigma_\upsilon}-\rho_{\zeta\upsilon}\bar{\zeta}}{\sqrt{(1-\rho^2_{\zeta\upsilon})}}}^{\infty}\int_{\frac{-a}{\sigma_\zeta}}^{\infty}\bar{\zeta}\phi(\bar{\zeta})\phi(y)d\bar{\zeta}dy \tag{B-4}$$

$$=(1-\rho^2_{\zeta\upsilon})\int_{\frac{-a}{\sigma_\zeta}}^{\infty}\bar{\zeta}\phi(\bar{\zeta})\left(1-\Phi\left(\frac{\frac{-b}{\sigma_\upsilon}-\rho_{\zeta\upsilon}\bar{\zeta}}{\sqrt{(1-\rho^2_{\zeta\upsilon})}}\right)\right)d\bar{\zeta}-\rho_{\zeta\upsilon}\sqrt{(1-\rho^2_{\zeta\upsilon})}\int_{\frac{\frac{-b}{\sigma_\upsilon}-\rho_{\zeta\upsilon}\bar{\zeta}}{\sqrt{(1-\rho^2_{\zeta\upsilon})}}}^{\infty}\int_{\frac{-a}{\sigma_\zeta}}^{\infty}y\phi(\bar{\zeta})\phi(y)d\bar{\zeta}dy$$

$$=(1-\rho^2_{\zeta\upsilon})\int_{\frac{-a}{\sigma_\zeta}}^{\infty}\bar{\zeta}\phi(\bar{\zeta})\Phi\left(\frac{\frac{b}{\sigma_\upsilon}+\rho_{\zeta\upsilon}\bar{\zeta}}{\sqrt{(1-\rho^2_{\zeta\upsilon})}}\right)d\bar{\zeta}-\rho_{\zeta\upsilon}\sqrt{(1-\rho^2_{\zeta\upsilon})}\int_{\frac{\frac{-b}{\sigma_\upsilon}-\rho_{\zeta\upsilon}\bar{\zeta}}{\sqrt{(1-\rho^2_{\zeta\upsilon})}}}^{\infty}\int_{\frac{-a}{\sigma_\zeta}}^{\infty}y\phi(\bar{\zeta})\phi(y)d\bar{\zeta}dy. \tag{B-5}$$

Now, note that $\bar{\zeta}\phi(\bar{\zeta})d\bar{\zeta}=-d\phi(\bar{\zeta})$ and $\phi(\bar{\zeta})=\phi(-\bar{\zeta})$, hence using integration by parts, the first part of the last equation of (B-5) can now be written as

$$\begin{aligned}
&(1-\rho^2_{\zeta\upsilon})\int_{\frac{-a}{\sigma_\zeta}}^{\infty}\bar{\zeta}\phi(\bar{\zeta})\Phi\left(\frac{\frac{b}{\sigma_\upsilon}+\rho_{\zeta\upsilon}\bar{\zeta}}{\sqrt{(1-\rho^2_{\zeta\upsilon})}}\right)d\bar{\zeta}=(1-\rho^2_{\zeta\upsilon})\int_{\frac{-a}{\sigma_\zeta}}^{\infty}-d\phi(\bar{\zeta})\Phi\left(\frac{\frac{b}{\sigma_\upsilon}+\rho_{\zeta\upsilon}\bar{\zeta}}{\sqrt{(1-\rho^2_{\zeta\upsilon})}}\right)\\
&=-(1-\rho^2_{\zeta\upsilon})\phi(\bar{\zeta})\Phi\left(\frac{\frac{b}{\sigma_\upsilon}+\rho_{\zeta\upsilon}\bar{\zeta}}{\sqrt{(1-\rho^2_{\zeta\upsilon})}}\right)\Bigg|_{\frac{-a}{\sigma_\zeta}}^{\infty}+\rho_{\zeta\upsilon}\sqrt{(1-\rho^2_{\zeta\upsilon})}\int_{\frac{-a}{\sigma_\zeta}}^{\infty}\phi(\bar{\zeta})\phi\left(\frac{\frac{b}{\sigma_\upsilon}+\rho_{\zeta\upsilon}\bar{\zeta}}{\sqrt{(1-\rho^2_{\zeta\upsilon})}}\right)d\bar{\zeta}\\
&=(1-\rho^2_{\zeta\upsilon})\phi(\frac{a}{\sigma_\zeta})\Phi\left(\frac{\frac{b}{\sigma_\upsilon}-\rho_{\zeta\upsilon}\frac{a}{\sigma_\zeta}}{\sqrt{(1-\rho^2_{\zeta\upsilon})}}\right)+\rho_{\zeta\upsilon}\sqrt{(1-\rho^2_{\zeta\upsilon})}\int_{\frac{-a}{\sigma_\zeta}}^{\infty}\phi(\bar{\zeta})\phi\left(\frac{\frac{b}{\sigma_\upsilon}+\rho_{\zeta\upsilon}\bar{\zeta}}{\sqrt{(1-\rho^2_{\zeta\upsilon})}}\right)d\bar{\zeta}.
\end{aligned} \tag{B-6}$$

The second expression of the last line in equation (B-5) can be written as

$$\begin{aligned}
&-\rho_{\zeta\upsilon}\sqrt{(1-\rho^2_{\zeta\upsilon})}\int_{\frac{\frac{-b}{\sigma_\upsilon}-\rho_{\zeta\upsilon}\bar{\zeta}}{\sqrt{(1-\rho^2_{\zeta\upsilon})}}}^{\infty}\int_{\frac{-a}{\sigma_\zeta}}^{\infty}y\phi(\bar{\zeta})\phi(y)d\bar{\zeta}dy=\rho_{\zeta\upsilon}\sqrt{(1-\rho^2_{\zeta\upsilon})}\int_{\frac{-a}{\sigma_\zeta}}^{\infty}\int_{\frac{\frac{-b}{\sigma_\upsilon}-\rho_{\zeta\upsilon}\bar{\zeta}}{\sqrt{(1-\rho^2_{\zeta\upsilon})}}}^{\infty}d\phi(y)\phi(\bar{\zeta})d\bar{\zeta}\\
&=\rho_{\zeta\upsilon}\sqrt{(1-\rho^2_{\zeta\upsilon})}\int_{\frac{-a}{\sigma_\zeta}}^{\infty}\phi(y)\Bigg|_{\frac{\frac{-b}{\sigma_\upsilon}-\rho_{\zeta\upsilon}\bar{\zeta}}{\sqrt{(1-\rho^2_{\zeta\upsilon})}}}^{\infty}\phi(\bar{\zeta})d\bar{\zeta}=-\rho_{\zeta\upsilon}\sqrt{(1-\rho^2_{\zeta\upsilon})}\int_{\frac{-a}{\sigma_\zeta}}^{\infty}\phi\left(\frac{\frac{b}{\sigma_\upsilon}+\rho_{\zeta\upsilon}\bar{\zeta}}{\sqrt{(1-\rho^2_{\zeta\upsilon})}}\right)\phi(\bar{\zeta})d\bar{\zeta}.
\end{aligned} \tag{B-7}$$

Plugging the results obtained in (B-6) and (B-7) into (B-4), we obtain

$$\int_{\frac{-b}{\sigma_\upsilon}}^{\infty}\int_{\frac{-a}{\sigma_\zeta}}^{\infty}(\bar{\zeta}-\rho_{\zeta\upsilon}\bar{\upsilon})\phi_2(\bar{\zeta},\bar{\upsilon},\rho_{\zeta\upsilon})d\bar{\zeta}d\bar{\upsilon}=(1-\rho^2_{\zeta\upsilon})\phi(\frac{a}{\sigma_\zeta})\Phi\left(\frac{\frac{b}{\sigma_\upsilon}-\rho_{\zeta\upsilon}\frac{a}{\sigma_\zeta}}{\sqrt{(1-\rho^2_{\zeta\upsilon})}}\right).$$

Similarly, it can be shown that

$$\int_{\frac{-b}{\sigma_\upsilon}}^{\infty}\int_{\frac{-a}{\sigma_\zeta}}^{\infty}(\bar{\upsilon}-\rho_{\zeta\upsilon}\bar{\zeta})\phi_2(\bar{\zeta},\bar{\upsilon},\rho_{\zeta\upsilon})d\bar{\zeta}d\bar{\upsilon}=(1-\rho^2_{\zeta\upsilon})\phi(\frac{b}{\sigma_\upsilon})\Phi\left(\frac{\frac{a}{\sigma_\zeta}-\rho_{\zeta\upsilon}\frac{b}{\sigma_\upsilon}}{\sqrt{(1-\rho^2_{\zeta\upsilon})}}\right).$$

Hence,

$$\mathrm{E}[\eta|\bar{\zeta}>\frac{-a}{\sigma_\zeta},\bar{\upsilon}>\frac{-b}{\sigma_\upsilon}]=\frac{\sigma_\eta\rho_{\eta\zeta}\phi(\frac{a}{\sigma_\zeta})}{\Phi_2\left(\frac{a}{\sigma_\zeta},\frac{b}{\sigma_\upsilon},\rho_{\zeta\upsilon}\right)}\Phi\left(\frac{\frac{b}{\sigma_\upsilon}-\rho_{\zeta\upsilon}\frac{a}{\sigma_\zeta}}{\sqrt{(1-\rho^2_{\zeta\upsilon})}}\right)+\frac{\sigma_\eta\rho_{\eta\upsilon}\phi(\frac{b}{\sigma_\upsilon})}{\Phi_2\left(\frac{a}{\sigma_\zeta},\frac{b}{\sigma_\upsilon},\rho_{\zeta\upsilon}\right)}\Phi\left(\frac{\frac{a}{\sigma_\zeta}-\rho_{\zeta\upsilon}\frac{b}{\sigma_\upsilon}}{\sqrt{(1-\rho^2_{\zeta\upsilon})}}\right). \tag{B-8}$$

Now, consider

$$\mathrm{E}[\eta|\zeta \leq -a, \upsilon > -b] = \mathrm{E}[\eta|\bar{\zeta} \leq \frac{-a}{\sigma_\zeta}, \bar{\upsilon} > \frac{-b}{\sigma_\upsilon}] = \frac{\int_{\frac{-b}{\sigma_\upsilon}}^{\infty} \int_{-\infty}^{\frac{-a}{\sigma_\zeta}} \mathrm{E}[\eta|\bar{\zeta}, \bar{\upsilon}]\phi_2(\bar{\zeta}, \bar{\upsilon}, \rho_{\zeta\upsilon}) d\bar{\zeta} d\bar{\upsilon}}{\Phi_2\left(\frac{-a}{\sigma_\zeta}, \frac{b}{\sigma_\upsilon}, -\rho_{\zeta\upsilon}\right)}$$

$$= \frac{1}{\Phi_2\left(\frac{-a}{\sigma_\zeta}, \frac{b}{\sigma_\upsilon}, -\rho_{\zeta\upsilon}\right)} \frac{\sigma_\eta \rho_{\eta\zeta}}{(1-\rho_{\zeta\upsilon}^2)} \int_{\frac{-b}{\sigma_\upsilon}}^{\infty} \int_{-\infty}^{\frac{-a}{\sigma_\zeta}} (\bar{\zeta} - \rho_{\zeta\upsilon}\bar{\upsilon})\phi_2(\bar{\zeta}, \bar{\upsilon}, \rho_{\zeta\upsilon}) d\bar{\zeta} d\bar{\upsilon}$$

$$+ \frac{1}{\Phi_2\left(\frac{-a}{\sigma_\zeta}, \frac{b}{\sigma_\upsilon}, -\rho_{\zeta\upsilon}\right)} \frac{\sigma_\eta \rho_{\eta\upsilon}}{(1-\rho_{\zeta\upsilon}^2)} \int_{\frac{-b}{\sigma_\upsilon}}^{\infty} \int_{-\infty}^{\frac{-a}{\sigma_\zeta}} (\bar{\upsilon} - \rho_{\zeta\upsilon}\bar{\zeta})\phi_2(\bar{\zeta}, \bar{\upsilon}, \rho_{\zeta\upsilon}) d\bar{\zeta} d\bar{\upsilon}. \tag{B-9}$$

By a method analogous to that used in deriving (B-8), it can be shown that

$$\mathrm{E}[\eta|\bar{\zeta} \leq \frac{-a}{\sigma_\zeta}, \bar{\upsilon} > \frac{-b}{\sigma_\upsilon}] = \frac{-\sigma_\eta \rho_{\eta\zeta} \phi(\frac{a}{\sigma_\zeta})}{\Phi_2\left(\frac{-a}{\sigma_\zeta}, \frac{b}{\sigma_\upsilon}, -\rho_{\zeta\upsilon}\right)} \Phi\left(\frac{\frac{b}{\sigma_\upsilon} - \rho_{\zeta\upsilon}\frac{a}{\sigma_\zeta}}{\sqrt{(1-\rho_{\zeta\upsilon}^2)}}\right) + \frac{\sigma_\eta \rho_{\eta\upsilon} \phi(\frac{b}{\sigma_\upsilon})}{\Phi_2\left(\frac{-a}{\sigma_\zeta}, \frac{b}{\sigma_\upsilon}, -\rho_{\zeta\upsilon}\right)} \Phi\left(\frac{\frac{-a}{\sigma_\zeta} + \rho_{\zeta\upsilon}\frac{b}{\sigma_\upsilon}}{\sqrt{(1-\rho_{\zeta\upsilon}^2)}}\right). \tag{B-10}$$

Appendix C: Asymptotic Variance Covariance Matrix of the Second and Third Stage Estimates

In this section we give the asymptotic variance covariance matrix of the coefficients of the second stage and third stage R&D switching regression model. Newey (1984) has shown that sequential estimators can be interpreted as members of a class of Method of Moments (MM) estimators and that, this interpretation facilitates derivation of asymptotic covariance matrices for multi-step estimators. Let $\Theta = \{\Theta_1, \Theta_2, \Theta_3\}$, where Θ_1, Θ_2, and Θ_3 are respectively the parameters to be estimated in the first, second and third step estimation of the sequential estimator. Following Newey (1984) we write the first, second, and third step estimation as an MM estimation based on the following population moment conditions:

$$E(L_{i1\Theta_1}) = E\frac{\partial \ln L_{i1}(\Theta_1)}{\partial \Theta_1} = 0 \tag{C-1}$$

$$E(L_{i2\Theta_2}) = E\frac{\partial \ln L_{i2}(\Theta_1, \Theta_2, \hat{\alpha}_i(.,\Theta_1), \hat{\epsilon}_{it}(.,\Theta_1))}{\partial \Theta_2} = 0 \tag{C-2}$$

and

$$E[F_{it}(X_{it}, \Theta_1, \Theta_2, \Theta_3)] = E[s_{it-1,t} X_{it}(r_{it} - X_{it}'\Theta_3)] = 0 \tag{C-3}$$

where $L_{i1}(\Theta_1)$ is the likelihood function for individual i, for the first step reduced form long-term debt equation. $L_{i2}(\Theta_1, \Theta_2, \hat{\alpha}_i(.,\Theta_1))$ is the likelihood function for the second step estimation in which the joint probability of a firm being an innovator and the firm being financially constraint constrained is estimated. Equation (C-3) is the first order condition for minimizing the sum of squared error for the pooled OLS regression of X_{it} on r_{it} for those firms, that have been selected, $s_{it-1,t} = 1$. $X_{it} = \{f_{it-1,t}$, $\mathbf{z}^{r\prime}_{it}$, CF_{it}, LQ_{it}, $SINS_{it-1,t}$, DIV_{it}, $\hat{\alpha}_i(.,\Theta_1)$, $C_{11}(.,\Theta_1)_{it}$, $C_{12}(.,\Theta_1,\Theta_2)_{it}$, $C_{13}(.,\Theta_1,\Theta_2)_{it}$, $C_{01}(.,\Theta_1)_{it}$, $C_{02}(.,\Theta_1,\Theta_2)_{it}$, $C_{03}(.,\Theta_1,\Theta_2)_{it}$, Industry Dummies, Time Dummies $\}$ and $\hat{\alpha}_i(.,\Theta_1) \equiv \hat{\alpha}_i(d_i, Z_i, \Theta_1)$ is the expected a posteriori (EAP) values, based on the first stage reduced form estimates.

The estimates for Θ_1, Θ_2, and Θ_3 are obtained by solving the sample analog of the above population moment conditions. The sample analog of moment condition for the first step estimation is given by

$$\frac{1}{N} L_{1\hat{\Theta}_1} = \frac{1}{N}\sum_{i=1}^{N} L_{i1\hat{\Theta}_1} = \frac{1}{N}\sum_{i=1}^{N} \frac{\partial \ln L_{i1}(\hat{\Theta}_1)}{\partial \Theta_1} \tag{C-4}$$

where $L_{i1}(\Theta_1)$ is given by equation (25), $\Theta_1 = \{\delta, \sigma_\epsilon, \sigma_\alpha\}$ and N is the total number of individuals/firms. Under standard regularity conditions, the first step, reduced form ML estimate, $\hat{\Theta}_1$, is consistent and the asymptotic variance of $\hat{\Theta}_1$ coincides with asymptotic variance obtained in the ML estimation.

The sample analog of population moment condition for the second step estimation is given by

$$\frac{1}{N} L_{2\hat{\Theta}_2} = \frac{1}{N}\sum_{i=1}^{N} L_{i2\hat{\Theta}_2} = \frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T_i} \frac{\partial \ln L_{it2}(\hat{\Theta}_1, \hat{\Theta}_2, \hat{\alpha}_i(.,\hat{\Theta}_1), \hat{\epsilon}_{it}(.,\hat{\Theta}_1))}{\partial \Theta_2} \tag{C-5}$$

where $L_{it2}(\Theta_1, \Theta_2, \hat{\alpha}_i(.,\Theta_1), \hat{\epsilon}_{it}(.,\Theta_1))$ is given by equation (31) and $\Theta_2 = \{\varphi\frac{1}{\sigma_{\underline{\varsigma}}}$, $\frac{\varphi_d}{\sigma_{\underline{\varsigma}}}$, $\frac{\lambda}{\sigma_{\underline{\varsigma}}}$, $\frac{\rho_{\varsigma\epsilon}}{\sigma_{\underline{\varsigma}}}$; $\gamma\frac{1}{\sigma_{\underline{\upsilon}}}$, $\frac{\gamma_d}{\sigma_{\underline{\upsilon}}}$, $\frac{\theta}{\sigma_{\underline{\upsilon}}}$, $\frac{\rho_{\upsilon\epsilon}}{\sigma_{\underline{\upsilon}}}$; $\rho_{\underline{\varsigma\upsilon}}\}$. In Appendix A, we have shown, that with α_i substitute by their EAP values $\hat{\alpha}_i(d_i, z_i, \Theta_1)$ still leads to the identification of Θ_2. Let Θ_1^* and Θ_2^*, respectively be the true values of Θ_1, Θ_2. Under the assumptions we make, maximizing $L_{i2}(\hat{\Theta}_1, \Theta_2, .)$ is asymptotically equivalent to maximizing $L_{i2}(\Theta_1^*, \Theta_2, .)$, where $\hat{\Theta}_1$ are consistent first step estimates. Murphy and Topel (1985) and Newey (1984), have derived the asymptotic distribution of the second step estimates of a two step sequential estimator.

Murphy and Topel have shown that the second step ML estimate of Θ_2 is consistent and asymptotically normally distributed with asymptotic covariance matrix given by

$$V_2^* = \frac{1}{N}[V_2 + V_2[R_2 V_1 R_2' - R_1 V_1 R_2' - R_2 V_1 R_1'] V_2], \tag{C-6}$$

where V_1 is the asymptotic covariance matrix of $\sqrt{N}(\hat{\Theta}_1 - \Theta_1^*)$ based on maximization of $L_1(\Theta_1)$, V_2 is the asymptotic covariance matrix of $\sqrt{N}(\hat{\Theta}_2 - \Theta_2^*)$ based on maximization of $L_2(\Theta_1, \Theta_2, \alpha_i)$ conditional on Θ_1 and α_i,

$$R_1 = E\left[\frac{1}{N}\left(\frac{\partial \ln L_2}{\partial \Theta_2}\right)\left(\frac{\partial \ln L_1}{\partial \Theta_1'}\right)\right]$$

and

$$R_2 = E\left[\frac{1}{N}\left(\frac{\partial \ln L_2}{\partial \Theta_2}\right)\left(\frac{\partial \ln L_2}{\partial \Theta_1'}\right)\right].$$

The sample analog of R_1 and R_2 are given by

$$\hat{R}_1 = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\partial \sum_{t=1}^{T_i} \ln L_{it2}}{\partial \Theta_2}\right)\left(\frac{\partial \ln L_{i1}}{\partial \Theta_1'}\right)$$

and

$$\hat{R}_2 = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\partial \sum_{t=1}^{T_i} \ln L_{it2}}{\partial \Theta_2}\right)\left(\frac{\partial \sum_{t=1}^{T_i} \ln L_{it2}}{\partial \Theta_1'}\right),$$

where the derivatives are evaluated at $\hat{\Theta}_1$ and $\hat{\Theta}_2$. Let Θ_{2f} and Θ_{2s} respectively denote the parameters of the Financial Constraint and the Selection equation of the second step. Since the second step is essentially a combination of Heckman and bivariate probit, $\frac{\partial \ln L_2}{\partial \Theta_2}$ and $\frac{\partial \ln L_2}{\partial \Theta_1}$ both involve taking the derivative of cumulative density function of a standard bivariate normal with respect to Θ_{2f}, Θ_{2s}, $\rho_{\underline{\zeta\upsilon}}$ and Θ_1, and evaluated at $\hat{\Theta}_{2f}$, $\hat{\Theta}_{2s}$, $\hat{\rho}_{\underline{\zeta\upsilon}}$, and $\hat{\Theta}_1$. The score functions of a bivariate probit are stated in Greene (2002). Murphy and Topel (1985) and Newey (1984), derive the asymptotic distribution of the second step estimates of a two step sequential estimator only for a single cross section of the data. However, since we are dealing with panel data, the derivative of EAPs, $\hat{\alpha}_i(d_i, z_i, \delta, \sigma_\epsilon, \sigma_\alpha)$ with respect to $\Theta_1 = \{\delta, \sigma_\epsilon, \sigma_\alpha\}$, that appear in the likelihood $L_{it2}(\Theta_1, \Theta_2, .)$ is also evaluated. That is, the derivative of $\hat{\alpha}_i(d_i, z_i, \delta, \sigma_\epsilon, \sigma_\alpha)$,

$$\frac{\partial \hat{\alpha}_i(d_i, z_i, \delta, \sigma_\epsilon, \sigma_\alpha)}{\partial \Theta_1} = \frac{\partial}{\partial \Theta_1}\left(\frac{\int \sigma_\alpha \tilde{\alpha}_i \prod_{t=1}^{T_i} \frac{1}{\sigma_\epsilon}\phi(\frac{d_{it} - z_{it}\delta - \sigma_\alpha \tilde{\alpha}_i}{\sigma_\epsilon})\phi(\tilde{\alpha}_i) d\tilde{\alpha}_i}{\boldsymbol{L}_{i1}(\delta, \sigma_\epsilon, \sigma_\alpha)}\right)$$

as an argument of $L_{i2}(\Theta_1, \Theta_2, \alpha_i)$ is evaluated at $\hat{\Theta}_1 = \{\hat{\delta}, \hat{\sigma}_\epsilon, \hat{\sigma}_\alpha\}$ for each individual.

To derive, the asymptotic distribution of the third step estimates Θ_3 consider the Taylor's expansion of $L_{1\Theta_1}(\Theta_1)$, $L_{2\Theta_2}(\Theta_1, \Theta_2, .)$ and $F(\Theta_1, \Theta_2, \Theta_3)$, where

$$\frac{1}{N}F(\hat{\Theta}_1, \hat{\Theta}_2, \hat{\Theta}_3) = \frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T_i} F_{it}(X_{it}, \hat{\Theta}_1, \hat{\Theta}_2, \hat{\Theta}_3) = \frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T_i} s_{it}X_{it}(\hat{\Xi})(r_{it} - X_{it}(\hat{\Xi})'\hat{\Theta}_3)$$

is the sample analog of the population moment condition given in (C-3). A series of Taylor's expansion of $L_{1\Theta_1}(\Theta_1)$, $L_{2\Theta_2}(\Theta_1, \Theta_2)$ and $F(\Theta_1, \Theta_2, \Theta_3)$ gives

$$\frac{1}{N}\begin{bmatrix} L_{1\Theta_1\Theta_1} & 0 & 0 \\ L_{2\Theta_1\Theta_2} & L_{2\Theta_2\Theta_2} & 0 \\ F_{\Theta_1} & F_{\Theta_2} & F_{\Theta_3} \end{bmatrix}\begin{bmatrix} \sqrt{N}(\Theta_1 - \Theta_1^*) \\ \sqrt{N}(\Theta_2 - \Theta_2^*) \\ \sqrt{N}(\Theta_3 - \Theta_3^*) \end{bmatrix} = -\frac{1}{\sqrt{N}}\begin{bmatrix} L_{1\Theta_1} \\ L_{2\Theta_2} \\ F \end{bmatrix}$$

In matrix notation the above can be written as

$$B_N\sqrt{N}(\Theta - \Theta^*) = -\frac{1}{\sqrt{N}}\Lambda_N$$

Under the standard regularity conditions for Generalized Method of Moments (GMM), see Hansen (1982) and Newey (1984), B_N converges in probability to the lower block triangular matrix $B_* = \lim EB_N$ and $\frac{1}{\sqrt{N}}\Lambda_N$ convergence in distribution to an asymptotically normal random variable with mean zero and a covariance matrix $A_* = \lim E\frac{1}{N}\Lambda_N\Lambda_N'$, where A_* is given by

$$A_* = \begin{bmatrix} V_{L_1L_1} & V_{L_1L_2} & V_{L_1F} \\ V_{L_2L_1} & V_{L_2L_2} & V_{L_2F} \\ V_{FL_1} & V_{FL_2} & V_{FF} \end{bmatrix},$$

where a typical element of A_*, say $V_{L_1L_2}$ is given by $V_{L_1L_2} = \frac{1}{N}E[L_{1\Theta_1}(\Theta_1)L_{2\Theta_2}(\Theta_1,\Theta_2)']$. Under the regularity conditions $\sqrt{N}(\Theta - \Theta^*)$ is asymptotically normal with zero mean and covariance matrix given by $B_*^{-1}A_*B_*^{-1\prime}$. To derive the asymptotic distribution of $\sqrt{N}(\Theta_3 - \Theta_3^*)$ consider the partitioned matrix of B_*

$$B_* = \frac{1}{N}\begin{bmatrix} \begin{bmatrix} L_{1\Theta_1\Theta_1} & 0 \\ L_{2\Theta_1\Theta_2} & L_{2\Theta_2\Theta_2} \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} F_{\Theta_1} & F_{\Theta_2} \end{bmatrix} & \begin{bmatrix} F_{\Theta_3} \end{bmatrix} \end{bmatrix}$$

where with a slight abuse of notation we denote by $L_{1\Theta_1\Theta_1}$ the expectation of $L_{1\Theta_1\Theta_1}$. Then by an application of the partitioned inverse formula and by using the facts that $(\frac{1}{N}L_{1\Theta_1\Theta_1})^{-1}V_{L_1L_1}(\frac{1}{N}L_{1\Theta_1\Theta_1})^{-1} = V_1$ and $(\frac{1}{N}L_{2\Theta_2\Theta_2})^{-1}V_{L_2L_2}(\frac{1}{N}L_{2\Theta_2\Theta_2})^{-1} = V_2$, where V_1 is the asymptotic covariance matrix of $\sqrt{N}(\hat{\Theta}_1 - \Theta_1^*)$ based on maximization of $L_1(\Theta_1)$, V_2 is the asymptotic covariance matrix of $\sqrt{N}(\hat{\Theta}_2 - \Theta_2^*)$ based on maximization of $L_2(\Theta_1, \Theta_2, \alpha_i)$ conditional on Θ_1 and α_i, we arrive at the asymptotic distribution of $\sqrt{N}(\Theta_3 - \Theta_3^*)$, V_3^*, given by

$$\begin{aligned} V_3^* = F_{\Theta_3}^{-1}\{[F_{\Theta_1} - [F_{\Theta_2}V_2L_{\Theta_2\Theta_1}]V_1([F_{\Theta_1} - [F_{\Theta_2}'V_2L_{\Theta_2\Theta_1}']] + V_{L_1L_2}V_2L_{\Theta_2\Theta_2}' - V_{L_1F}) \\ +F_{\Theta_2}V_2[V_{L_2L_1}V_1[F_{\Theta_2}'V_2L_{\Theta_2\Theta_1}'] + F_{\Theta_2}' - V_{L_2F}] - V_{FL_1}V_1[F_{\Theta_2}'V_2L_{\Theta_2\Theta_1}'] - V_{FL_2}V_2F_{\Theta_2}'\}F_{\Theta_3}^{-1\prime} + F_{\Theta_3}^{-1}\{V_{FF}\}F_{\Theta_3}^{-1\prime} \end{aligned} \quad \text{(C-7)}$$

Appendix D: Second and Third Stage Estimation Results

Table 3: **Second Stage Estimates**

	Finance	**Innovation**	**Finance**	**Innovation**	**Finance**	**Innovation**
	(A0)	**(A1)**	**(B0)**	**(B1)**	**(C0)**	**(C1)**
$SINS_{t-1,t}$	0.010***		0.010***		0.009***	
	(0.002)		(0.002)		(0.002)	
$DEBT_t$	0.471***	-0.250***	0.468***	-0.250***	0.123**	-0.248***
	(0.149)	(0.061)	(0.148)	(0.061)	(0.052)	(0.061)
CF_t	0.009**		0.009**			
	(0.004)		(0.004)			
LQ_t	-0.102***	0.048***	-0.102***	0.048***	-0.030***	0.048***
	(0.031)	(0.013)	(0.031)	(0.012)	(0.012)	(0.012)
DIV_t	-0.031	-0.109**	-0.031	-0.109**	-0.178***	-0.110***
	(0.081)	(0.042)	(0.081)	(0.043)	(0.055)	(0.043)
$MKSH_t$	-0.015**	0.212***	-0.015**	0.212***	-0.002	0.215***
	(0.006)	(0.012)	(0.006)	(0.019)	(0.004)	(0.012)
$SIZE_t$	0.180***	0.101***	0.179***	0.101***	0.189***	0.101***
	(0.012)	(0.022)	(0.012)	(0.022)	(0.012)	(0.021)
AGE_t	-0.001	-0.007***	-0.001	-0.007***	-0.003*	-0.007***
	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
$RAINT_t$	-0.042	0.441***		0.447***		0.447***
	(0.112)	(0.104)		(0.103		(0.103)
$DNCF_t$	0.104		0.104		0.227***	
	(0.075)		(0.075)		(0.056)	
$DMULTI_t$		2.071***		2.071***		2.060***
		(0.055)		(0.055)		(0.055)
$\frac{\lambda}{\sigma_{\underline{\zeta}}}$	-0.447***		-0.444***		-0.099**	
	(0.1486)		(0.148)		(0.051)	
$\frac{\theta}{\sigma_{\underline{\upsilon}}}$		0.270***		0.270***		0.269***
		(0.062)		(0.062)		(0.062)
$\frac{\rho_{\zeta\epsilon}}{\sigma_{\underline{\zeta}}}$	-1.363***		-1.355***		-0.354**	
	(0.432)		(0.432)		(0.153)	
$\frac{\rho_{\upsilon\epsilon}}{\sigma_{\underline{\upsilon}}}$		0.612***		0.612***		0.607***
		(0.173)		(0.173)		(0.173)
$\rho_{\underline{\zeta\upsilon}}$	0.528 ***		0.528 ***		0.530 ***	
Likelihood-Ratio test: χ^2_1	(412.049)		(412.094)		(411.582)	
Total Number of Observations: 13034						

Significance levels : $*$: 10% $**$: 5% $***$: 1%

Table 4: **Average Partial Effects for Second Stage**

The APE's in this table are for the specification in column **(C0)** and **(C1)** of Table 3.

	Finance		**Innovation**	
	Coefficient	**APE**	**Coefficient**	**APE**
$SINS_{t-1,t}$	0.009***	0.001		
	(0.002)			
$DEBT_t$	0.123**	0.016	-0.248***	-0.091
	(0.052)		(0.061)	
LQ_t	-0.030***	-0.004	0.048***	0.017
	(0.012)		(0.012)	
DIV_t	-0.178***	-0.022	-0.110***	-0.013
	(0.055)		(0.043)	
$MKSH_t$	-0.002	-0.0003	0.215***	0.037
	(0.004)		(0.012)	
$SIZE_t$	0.189***	0.025	0.101***	0.079
	(0.012)		(0.021)	
AGE_t	-0.003*	-0.0005	-0.007***	-0.003
	(0.002)		(0.001)	
$RAINT_t$			0.447***	0.163
			(0.103)	
$DNCF_t$	0.227***	0.034		
	(0.056)			
$DMULTI_t$			2.060***	0.545
			(0.055)	

Significance levels : $*$: 10% $**$: 5% $***$: 1%

Table 5: **Third Stage: R&D Switching Equation**

$f_{t-1,t}$	-0.3077**
	(0.1539)
$SINS_{t-1,t}$	0.0122***
	(0.0012)
$f_{t-1,t} * CF_t$	0.0993***
	(0.0162)
$(1 - f_{t-1,t}) * CF_t$	0.0001
	(0.001)
$f_{t-1,t} * DIV_t$	-0.2101***
	(0.0728)
$(1 - f_{t-1,t}) * DIV_t$	-0.0396
	(0.0245)
$MKSH_t$	0.0074***
	(0.0017)
$SIZE_t$	-0.0429***
	(0.0109)
AGE_t	-0.0095***
	(0.0009)
$DMULTI_t$	0.2297***
	(0.4413)
μ	0.0155***
	(0.0053)
$C_{11}(.)_t$	0.3278***
	(0.0803)
$C_{12}(.)_t$	0.2245**
	(0.0895)
$C_{13}(.)_t$	1.1382***
	(0.1613)
$C_{01}(.)_t$	-0.0687***
	(0.0176)
$C_{02}(.)_t$	-0.2492**
	(0.1161)
$C_{03}(.)_t$	0.2943***
	(0.0519)

Total Number of Observations: 13034
Number of Censored Observations: 6767

Table 6: **APE's for the Third Stage**

The table reports the Average Partial Effects of some of the variables on R&D intensity. While the APE of $f_{t-1,t}$, the reported financial constraint faced by the firms is unconditional, the APE's of the rest of the variables in column (a) is conditional on being selected. The APE's of the variables reported in column (b) are conditional on being selected and conditional on being financially constrained. The APE's of the variables reported in column (c) are conditional on being selected and conditional on being not financially constrained.

$f_{t-1,t}$		-0.2635	
	a	**b**	**c**
$SIZE_t$	-0.0482	-0.0632	-0.0456
$MKSH_t$	0.7013	-0.0137	0.7026
AGE_t	-0.0076	-0.0036	-0.0077
DIV_t	-0.0305	-0.0866	-0.0205
$DMULTI_t$	0.0881	-0.0945	0.1192

Appendix E: Descriptive Statistics

Table 7: **Descriptive Statistics of Variables for Censored and Non-Censored Variables**

	CIS2.5	CSI3	CIS3.5
r	0.5061	0.3377	0.1914
$SINS$	8.5349	8.8979	5.668
No. of Censored obs.	(2946)	(1844)	(1981)
$DEBT$	0.7359	0.712	0.9155
CF	0.8601	0.9671	0.3978
$DMULTI$	0.2107	0.261	0.2607
LQ	0.9644	0.9407	0.9420
DIV	0.1825	0.1832	0.1772
$MKSH$	0.5396	0.7318	0.6443
$LOG(SIZE)$	4.5745	4.1161	4.328
AGE	20.71	23.51	22.99
$RAINT$	0.036	0.0386	0.0506
$DNCF$	0.0873	0.0929	0.127
Total no. of obs.	(5362)	(3424)	(4248)

The figures reported are the means of the respective variables

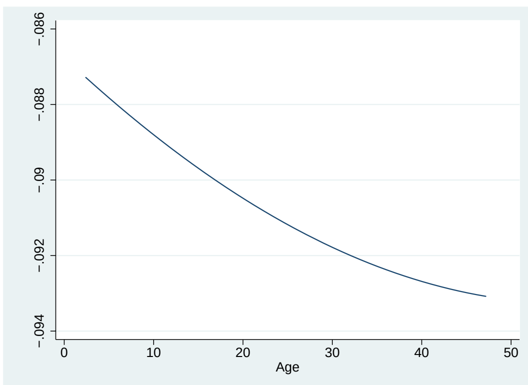
Table 8: **Total number of Enterprise, N_f, and number of Enterprise within a firm Surveyed, n_f**

The table illustrates the number of firms, in each of the three CIS waves, for which the number of number of enterprise surveyed is equal to the number of enterprise present in the firm, $N_f = n_f$, and the number of firms, for which the number of enterprises present in the firm exceeds the number of enterprises surveyed. These figures pertain to the CIS data set prior to merging with the SFGO and the SFKO data set. Since not all the CIS firms are in the SFGO and the SFKO data set, the data used for estimation after cleaning, is a bit less than half the size of the original data set.

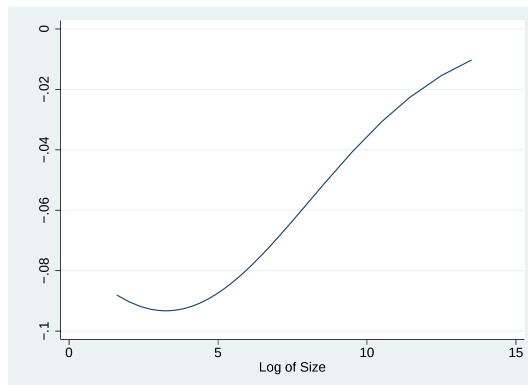
CIS2.5			CSI3			CIS3.5		
	No. of firms	No. of firms		No. of firms	No. of firms		No. of firms	No. of firms
N_f	$N_f = n_f$	$N_f > n_f$	N_f	$N_f = n_f$	$N_f > n_f$	N_f	$N_f = n_f$	$N_f > n_f$
1	9400	0	1	6155	0	1	7096	0
2	151	1255	2	67	823	2	137	978
3	20	608	3	4	424	3	24	553
4	3	316	4	3	237	4	2	290
5	3	247	5	2	108	5		222
6		149	6		115	6		122
7		107	7		48	7		105
8		60	8		77	8		50
9	2	93	9		58	9		77
10		83	10		39	10		82
11		106	11		63	11		50
12		49	12		39	12		58
13		43	13		15	13		49
14		59	14		50	14		46
15		46	15		17	15		25
16		31	16		28	16		51
17		62	17		15	17		15
18		36	18		26	18		55
19		37	19		13	19		8
20		29	20		21	20		28
21		13	21		2	21		43
22		23	22		27	22		36
23		15	24		5	23		18
25		34	25		9	24		25
26		46	26		8	25		11
27		4	27		21	27		17
29		14	28		13	28		19
30		14	29		8	29		11
31		18	30		8	30		15
32		15	31		3	31		7
33		11	32		16	32		16
34		18	34		22	33		25
37		15	40		10	37		21
38		15	45		14	38		13
43		15	48		18	39		20
44		17	50		19	40		9
45		14	57		16	41		10
48		20	60		16	46		15
49		22				50		16
51		28				53		47
56		19				55		16
66		33						
85		41						

Figure 1: APE of long-term Debt to Asset Ratio on Unconditional Probability of Innovation.

(a)



(b)



(c)

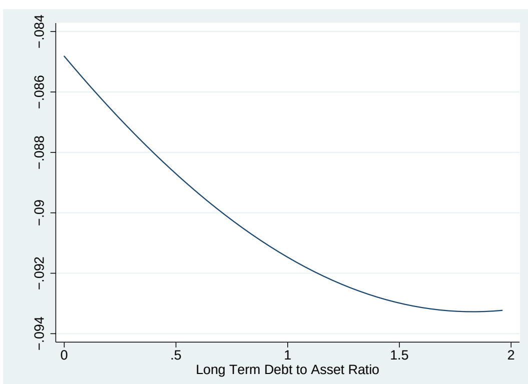
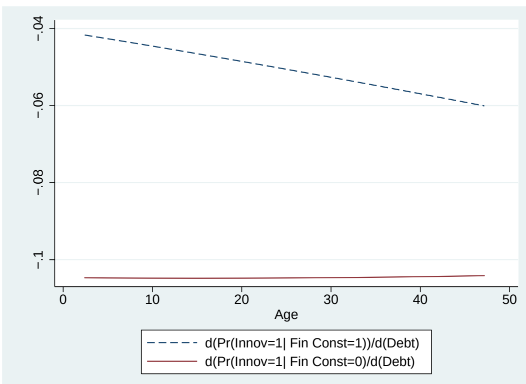
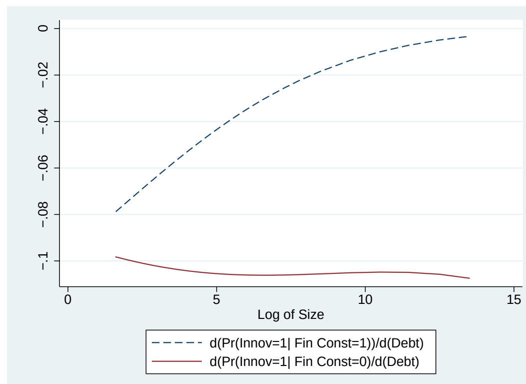


Figure 2: APE of long-term Debt to Asset Ratio on Probability of Innovation conditional on being Financially Constrained, $\frac{\partial \Pr(s_{it-1,t}=1|f_{it-1,t}=1)}{\partial d_{i,t}}$, and APE of long-term Debt to Asset Ratio on Probability of Innovation conditional on *not* being Financially Constrained, $\frac{\partial \Pr(s_{it-1,t}=1|f_{it-1,t}=0)}{\partial d_{i,t}}$.

(a)



(b)



(c)