



Credit Constraints on Small Businesses: Theory Versus Evidence

Robert Cressy

Cass Business School

Abstract. We describe the prevailing theories of credit constraints, empirical tests of the theories and their policy implications. Theories and empirical tests are then evaluated in the context of small businesses and the likelihood of encountering credit constraints in practice is assessed. Private and public sector responses to perceived credit constraints are then evaluated on both theoretical and empirical grounds and broad conclusions reached about the relevance and effectiveness of policy initiatives. The paper is a written-up version of a lecture that I deliver as part of a course taught to second year students on the BSc degrees at Cass Business School, London. The only prerequisite for the course is that the students should have taken Elementary (Micro and Macro) economics.

Keywords: credit constraints, small business, policy response.

1. Introduction

Billions of dollars are spent annually by governments of a variety of political persuasions around the world to alleviate perceived credit constraints on small, young businesses. The rationale for this activity is certainly questionable, but as Keynes once remarked, it is surprising how much policy-making is influenced by ‘the ideas of some [defunct] economist’. So this paper, a written-up version of a lecture that I deliver as part of a course taught to second year students on the BSc degrees at Cass Business School, London, attempts to outline the work of (largely non-defunct) economists in this area and to draw out whatever policy conclusions seem to be both logical and supported by the facts. The ‘lecture’ itself is based on the concept of ‘active learning’: I intersperse my delivery with class exercises, usually done in pairs (“Talk to your neighbour”), followed by audience response and finally my own answers, which are distributed as part of the Lecture Notes for the course – after the lecture. The only academic prerequisite for the course is that the students should have taken Elementary (Micro and Macro) economics.

2. Learning Objectives

The learning objectives of the topic are as follows:

1. To understand the nature of credit constraints and why small and young firms are more likely to experience them than large and old firms.
2. To be able to understand and evaluate the academic literature on credit constraints.
3. To have the intellectual tools to judge if credit constraints exist in practice.
4. To understand whether policy intervention is required as a result of identifying credit constraints and if so what form this should take.

To be able to achieve these objectives we need to have an understanding of the nature of small businesses and their finances.

3. Definition

There are many definitions of a small firm, but most rely on the numbers of employees of the firm falling below a certain threshold. We shall not attempt to define the concept of a small firm any further in this lecture (an earlier lecture is devoted to the subject) but assume that the reader has an intuitive grasp of what a small firm is and refer him or her to the various government definitions should they wish to understand the focus of government policy in this area.

4. Small and Large Firms

Small firms are profoundly different from large firms. This difference manifests itself in a number of ways, and ultimately impacts on the financing they use, the topic of interest in this lecture.

Firstly, small firms are informationally opaque. This is exemplified (and possibly encouraged) by the fact that the financial reporting requirements on them are lighter. Unincorporated businesses, which are the majority of small businesses, are not obliged to publish accounts information at all. Even limited companies, usually thought of as more sophisticated businesses, are often required to publish only summary accounts – in the UK for example, they do

not have to publish a profit and loss account. Furthermore, small firms' accounts are not eagerly and regularly poured over by teams of stock analysts, attempting to divine their economic destiny, as are the accounts of the PLCs. Even small companies financed by venture capital (a tiny minority of businesses) produce regular financial reports to their financiers only because they are required to do so as a precondition of funding, and their financiers are careful to make sure that such information does not reach the public light of day¹.

Secondly, small firms are often young firms and for this reason have no track record. This extends to their management, sales, costs and borrowing performance. Combined with the absence of information on their accounts, the absence of a track record makes it difficult for banks to assess their prospects when considering a lending proposal.

Thirdly, small, young firms are more risky than big, old firms. The evidence for this proposition has been ground out in a large number of empirical studies and is incontrovertible, though definitions of failure may vary somewhat from one study to another². This excess riskiness often stems from the fact that small firms often rely on one customer, one product, and one supplier. Naturally under these circumstances they are highly vulnerable to changes in the environment. They also have limited sources of finance available in a liquidity crisis, which may occur in the early stages of their lives and is particularly likely in an economic downturn. In the latter situation they will be likely to experience what economists call 'the flight to quality' among the banks, i.e. a switch in lending towards their less risky, larger counterparts as the failure rate amongst small businesses starts to rise and banks try to protect their rates of return.

Case Study: Failure & the Bespoke Tie Manufacturer

An ex-colleague of mine, Andrew, in the training arm of a well-known business school recounted to me the story of the failure of his own business. He had set up in a boom period a firm that manufactured highly artistic, multicoloured, extrovert neckties. These were attractive especially to extrovert wealthy business people. They therefore sold in good numbers at a considerable premium over the mass-produced varieties during the boom period of the late 80s. However, not a high proportion of profits were retained

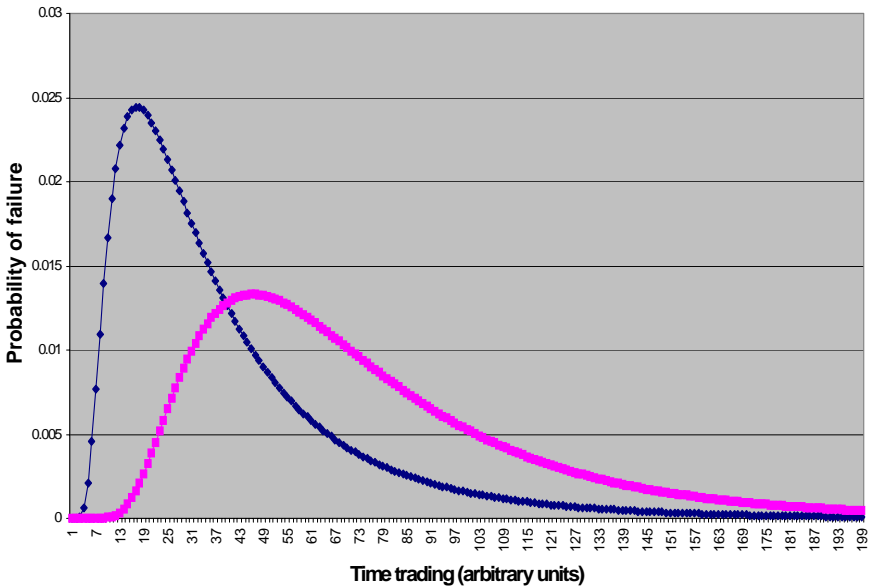
-
1. There is some exception to this: Venture Economics and Venture One, two commercial data agencies make available certain kinds of financial information to their paying subscribers.
 2. In the definition of failure we need to distinguish carefully between solvent business closure and bankruptcy or insolvency. Ceasing to trade is typically not accompanied by bankruptcy for small businesses since the majority of a given cohort of startups does not borrow, even on overdraft.

in the process. When the downturn eventually came (as it inevitably does) in 1990-1 their customers chose either to buy less ties altogether (keeping existing neckties longer) or switched to cheaper, mass-produced varieties e.g. from Marks and Spencer³. Thus demand for Andrew's bespoke ties fell dramatically in a short period of time and the company, which had borrowed to expand in the boom period, was suddenly plunged into bankruptcy. Looking back on the sorry affair Andrew could see that the main reasons for the failure of the business were (a) an excessive dependence on one product; (b) the sensitivity of demand for the product to consumer income levels. On examination Andrew realised that the sensible strategy would have been to have diversified quickly. He might either have developed two products, one mass-produced, the other bespoke, or have developed in addition to the latter an entirely different product, less sensitive to income changes which would provide a minimum turnover in the downturn.

Another reason for the vulnerability of the small firm is that their entrepreneurs often lack human capital: skills, business acumen, vocational qualifications, etc. that are necessary for success. This deficiency conspires against the small firm in financial markets. For example, banks, recognising the excess risk of lending to small firms, will treat their lending proposal differently from one from a large firm, especially in a recessionary environment, often requiring seemingly excessive collateral as security, and charging a much higher margin over Base or Prime⁴. Figure 1 opposite shows a simulation of the failure curve of a small business with the two underlying distributions of profits growth rates⁵. The curves differ only by the initial amount of capital in the firm. It is clear that a better capitalised business (the light curve) has a more favourable profile than a cash-starved startup (the dark curve): the honeymoon period (zero failure rate) is longer, and the peak failure rate is lower⁶.

-
3. Recall the concept of income elasticity of demand from first year Micro. Would you say the demand here is elastic or inelastic? Recollect what happens to revenue if there is a fall in demand and the income elasticity is greater than one. Would you expect it to rise or fall or remain constant?
 4. We should not expect the rates charged by the banks to small and large firms to be the same even if there were no differences in default risk. This is because generally large firms borrow more, and the bank is able to offset the fixed costs of lending to them against a larger loan size, reducing unit costs. This in turn (in a competitive market) will reduce margins relative to those of small firms.
 5. The curve is the failure density as a function of time trading and the area under the curve typically (though for obscure mathematical reasons, not necessarily) sums to one. The height of the curve gives the chance that a firm taken at random will fail at this time during its life. The firm must generally fail at some time between now and eternity.
 6. A theoretical model that produces this curve as a result of a value-maximisation process by the manager can be found in Cressy (2003).

Figure 1: Initial Capitalisation and Failure



Source: Cressy 1999

Fourthly, small firms have fewer collateralisable assets than large firms. This is partly a function of their stage of development: as the firm grows, it may use retained profits to purchase fixed assets, especially if it is manufacturing products. It may also grow by the purchase of other enterprises, acquiring their fixed assets in the process. Therefore larger firms typically have a longer history of asset acquisition and in consequence possess more collateralisable assets as the basis for loans or overdrafts.

Finally, another potential reason for small firms having a smaller proportion of fixed assets is the possibility of capital constraints: Because of the need to raise relatively large amounts of capital to start a manufacturing business startups typically gravitate to the service industries where the proportion of fixed to total assets is lower – for technical reasons.

Case Study: Constraints and the Choice of Business Sector

Peter, a one-time colleague of mine at prestigious British business school had been made redundant from a large firm and decided to start a business in the Tool Hire industry. This was in the early 1980s, when the industry was in its infancy and just about to embark on a growth trajectory. He chose Tool Hire partly because his redundancy money allowed him to adequately capitalise a

business in this sector. (The sector's high returns and growth prospects had also not escaped Peter's notice). If Peter had wanted to go into manufacturing his redundancy lump sum he decided would not have been enough to properly fund his startup.

Question: Does this constitute an example of a capital-constrained entrepreneur or a capital-constrained business or neither?

Answer: Peter started a business, so *he* (the entrepreneur) was not completely constrained by finance. But even if he was not totally constrained we can still ask on grounds of economic efficiency whether he started a business with the *highest* expected return given the limit he perceived on funds. On the one hand, given his limited funds, Peter dismissed the possibility of a manufacturing startup from consideration! On the other, it was clear from Peter's background, predominantly in Service industries rather than Manufacturing, that his skills were more geared to success in Services. So from the point of view of human capital, perhaps Peter did make the right choice after all, and might (optimally) have made the same decision even had he had sufficient financial capital to go into manufacturing.

5. Bank Lending Rules

Most Western banks have simple lending rules for small, young businesses: They typically lend to a firm *in proportion to the value of its assets* e.g. one third of their estimated market value⁷. These assets are required as security or *collateral* for the loan. In 'perfecting' this security the bank will typically take a *fixed* or *floating charge* over the firm's assets. This means that it has a claim on a specific asset (fixed charge) or can choose the asset it wishes to use to pay off the loan when the default occurs (floating charge). Under default it can sell these assets and hope to recover the value of the loan and interest.

What kind of assets will the bank use? Banks are suspicious of intangibles, things that can't be touched or seen and of items with short or passing lives. One bank manager is famous for saying that the problem with service firms is that the assets are their people, and so they tend to go home at 5 o'clock! So banks are likely to ask as the basis for security tangible (rather than intangible) and fixed (rather than current) assets. Classic examples of tangible fixed assets would be land, plant and buildings. Contrast this with intangible assets such as patents and copyrights. A computer program would typically be protected by a copyright (since it consists of lines of *written* code in which the author has obtained a *right to prevent unauthorised copying*.) The code might be worth

7. This is by no means always the case, but in Toivanen and Cressy (2001) we showed that 60% of small business term loans were collateralised. However, loans that were not collateralised were almost certainly to customers with some track record with the bank. See below under Relationship Banking.

literally millions of dollars but its value is often difficult to certify and therefore would not typically form the basis of loan collateral. Likewise, the patent on a new drug would not typically be suitable for this purpose. This again might be ultimately highly valuable, but the bank would tend to regard it as problematic from the point of view of immediate saleability.

So land, machinery and buildings are tangible, fixed assets. They have relatively long lives and larger used values, particularly in the case of land and houses located in 'good' residential areas. Current assets by contrast are items such as stocks of goods and work in progress, debtors and current interest earned. These might in the event of default have a low or zero value. Debtors might be valueless if the company to whom the products or services were sold, and who currently owed money for them in return, were to go bust. Half-finished goods are of little worth on the open market and even finished goods may need to be sold in distressed conditions, reducing their value substantially. Hence, little weight is placed on them by the banker in deciding the value of collateral.

On the principle of matching assets and liabilities⁸ banks may however be prepared to offer short term loans or overdraft facilities on the basis of current assets like debtors. The receipts from these sales are legally valid documents and in the last resort payment can be extracted from the purchaser by court action. On the bank's asset side, overdrafts are in principle repayable on demand: in other words, within 24 hours notice from the bank. Therefore, the risk associated with debtors used as collateral for lines of credit is not great.

Exercise: Constraints & Bank Lending Rules (I)

Consider the following: A firm has no assets at all to place as security for a loan it requests. If the bank lends half the value of assets what will be the value of the loan? Is the firm credit-constrained?

Answer: The answer to the first part is of course "zero". For the second part the questions to ask yourself are: Does the firm have a viable project? Has it approached all potential sources of finance? Assuming that both these are true, then the firm *will* be credit-constrained. The problem, of course, is: *Who* makes the assessment of the project's viability? *When* is the assessment made?

8. This is a well known technique for reducing financial risk. See any standard text (e.g. Brealey and Myers, (2003) for a discussion of the details.

Exercise: Constraints & Bank Lending Rules (II)

Two entrepreneurs, one with high productivity (High) and one with low productivity (Low), have searched efficiently amongst a set of banks for funding for their *viable* projects. They find that only one will lend to them. In the judgement of the bank the each firm has £10,000 worth of fixed assets. The bank's judgement is based on its estimate of the distressed value of these assets when sold on the market. The entrepreneurs on the other hand claim to have discounted profits to the value of £20,000 and £40,000 respectively. The bank has a lending rule that says it will lend to an entrepreneur on the basis of 1/2 the value of assets. It thus offers a loan of £5,000 to both. However, the High entrepreneur wants a loan of £8k whilst the Low entrepreneur wants one of merely £3k.

Is either entrepreneur credit-constrained?

Answer: It all depends. Consider the two entrepreneurs.

1. High Productivity Entrepreneur

To maximise profits this entrepreneur wants to borrow *more* than the bank will allow given her assets, i.e. £8,000 > £5,000. So she *is* credit-constrained: raising her asset base would allow more borrowing and *increase* the NPV of the firm. This means that the economic arrangement cannot have been optimal i.e. value-maximising.

2. Low Productivity Entrepreneur

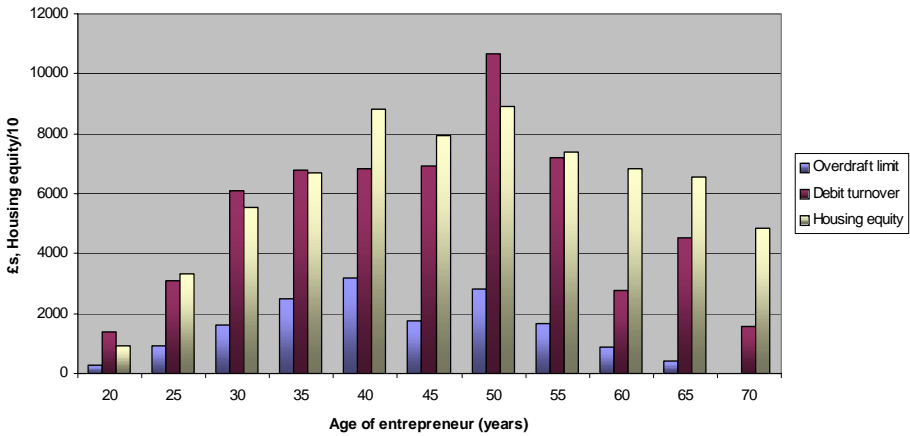
To maximise profits this entrepreneur wants to borrow *less* than the bank will allow given her assets, i.e. £3,000 < £5,000. So she is *not* credit-constrained: her profits are as high as they could be *regardless of the bank's lending rule*. In the language of linear programming, the constraint doesn't 'bite': Offering her more money would not increase her profits. The money would presumably be used for some non-profit making activity e.g. buying a BMW for personal use.

6. Theory and Tests of Credit Constraints

Evans and Jovanovic (1989) developed an empirical test for the existence of credit constraints. This was based on the idea of a bank lending rule discussed above.

They argued that firms were credit constrained if and only if a relaxation of the lending rule (or equivalently an unanticipated increase in fixed assets) increased business survival rates⁹. What might be the reasoning behind the Evans and Jovanovic (EJ) argument?

Chart 1: Relation Between Overdraft, House Equity, Turnover and Entrepreneurial Age, 1988



If on average an increase/decrease in assets increases/decreases business survival then bank borrowing based on asset values keeps profits artificially low. Hence for the ‘marginal’ business (the one just slipping into bankruptcy) an increase in assets allows an increase in borrowing. This increases profits and prevents failure. Hence, in summary: A relaxation of the lending rule reduces business failure rates or equivalently increases business survival rates.

7. Spurious Correlation? Alternative Explanations

The Evans and Jovanovic theory argues that on average bank lending rules *limit borrowing below economically desirable levels*. But what if observed correlation is caused by a third factor, age of the entrepreneur? This would explain the observed relationship of:

- *Borrowing to age* (Chart 1): Older borrowers (until about the age of 50) borrow more because they have more housing assets to place as collateral. Younger borrowers and those reaching retirement have either not accumulated assets or have started to consume their assets. In either case they have less assets to place as collateral and so (being constrained) borrow less.

9. More accurately they argued that credit rationing exists if and only if more assets increased the probability of a person moving into self-employment from some other labour market state. This is the mirror image of the proposition discussed in the text which is couched in terms of survival (probability of staying in business) rather than moving into business.

- *Age to survival* (Chart 2): Older entrepreneurs with more maturity, realism and access to networks are more likely to survive in business (partly perhaps because their wealth enables them to borrow more).

On this theory, experience is the prerogative of age, and energy of youth. The productivity effects of these two forces (one positive and one negative) will cancel out in the mid-50s age range. Thus we should expect an inverse-U shaped relationship of survival to age of the entrepreneur (Chart 2). Because more experienced entrepreneurs will sell more, and require more working capital in consequence, borrowing levels will be positively related to sales; and hence increase with age up to a certain point, declining thereafter as declining energy begins to outweigh greater experience and skills (Chart 1). Assets likewise will behave according to the lifecycle consumption theory implying an inverse U-shaped relation of assets to age (Chart 1): Older entrepreneurs have accumulated more assets; these begin to be consumed after a certain age is reached. Thus we have explained why survival, sales and borrowing are positively correlated with each other via the age factor.

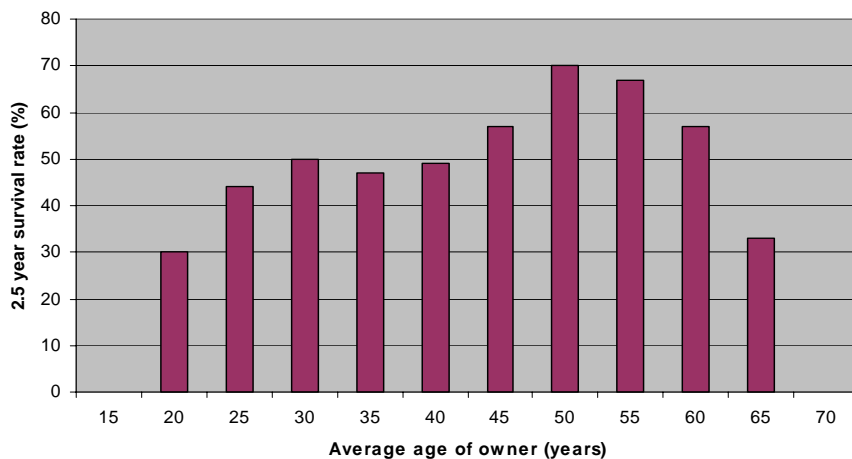
Exercise: Effects of Age on Business Characteristics

Compare the position of two firms A and B which differ simply in terms of the average age of the owners, 20 years and 50 years respectively using chart 1. How do business survival, sales, housing equity and overdraft borrowing figures differ between the firms?

But what does this imply for the existence of capital constraints. Returning now to the discussion of the Evans and Jovanovic theory, we may ask: Under the alternative hypothesis of human capital driving the correlation of assets and survival, what would we expect to find in the data? We would expect of course to find the concave relationships with respect to age for all three variables we have observed above. However, using more sophisticated statistical techniques, we should expect also to find *one important difference*: once we hold human capital *constant* the effect of assets on survival should drop out. Why? Because it occurs *only because of the correlation of assets and survival with age*. If this happens then we should conclude that capital constraints do *not* exist: the correlation of assets and survival is spurious.¹⁰ The correlation of human capital and survival is the primary *causal* relationship.

10. This is a technical term meaning that the simple correlation between X and Y is positive but the partial correlation (controlling for relevant variables) is zero.

Chart 2: Business Survival by Age of Proprietor



So what are the empirical facts? In Cressy (1996) I performed a regression of survival on Assets alone; Human capital (HC) alone; and finally Assets and HC together. As expected both assets and HC individually were positively correlated with survival, but once the effect of HC was controlled for, assets *dropped out of the equation*. I concluded that startups were not financial capital-constrained, but might indeed be human capital-deficient. Other studies have concluded much the same thing, e.g. Aston (1991) (at most 6% constrained); Cambridge (1998) (no quantitative estimate but little evidence of constraints); Cressy (2003b) (at most 3% constrained).¹¹

8. Decreasing Absolute Risk Aversion (Dara)

You don't have to take my word for it! There is another explanation for the EJ finding also which does not imply credit constraints either. It depends merely on some plausible assumptions and limited evidence about human tolerance of risk. It is commonly believed (and there is evidence to show) that people in general *dislike* risk. Studies of the stock market show that people need to be offered higher returns to invest in more risky securities. This is consistent with risk aversion. Likewise, most people take out some kind of insurance policy against fire, theft etc. which involves the payment of a premium. This also

11. Whilst there is empirical evidence for the positive effects of inheritance on business startups controlling for measures of human capital (e.g. Burke, Fitzroy and Nolan, 2000) this is not necessarily evidence for the existence of capital and specifically credit constraints. The phenomenon of control aversion may well explain these effects (see later).

suggests dislike of risk since by the mechanism of insurance the risk is transferred to another party.¹²

Imagine, then, that when I increase your assets you become less risk averse, i.e. you become more willing to take risks. This is quite a plausible phenomenon if you think about it: What it means is that if I offer you simultaneously an increase in your wealth W by £1 and a bet which yields +£1 with probability $\frac{1}{2}$ and -£1 with probability $\frac{1}{2}$, *with your additional assets* you are now *more* likely to take the bet than before. In the language of economics this means your utility of income function displays decreasing absolute risk aversion (or DARA)! Bearing in mind that entrepreneurship tends to be more income-risky than employment, this means that the marginal entrepreneur (one for whom the expected costs *just outweigh* the expected benefits) would switch into self-employment should she receive a windfall gain. There is furthermore some empirical evidence to support the assumption that entrepreneurs have decreasing absolute risk aversion (see Guiso, and Paiella, 1999).

9. Asymmetric Information and Its Effects

An understanding of much of the modern literature on credit constraints presupposes a grasp the concept of asymmetric information and its implications for credit markets. The following provides some of the basics for the lay reader.

Perfect or symmetric information in a market is a situation where the market participants know all the relevant facts regarding the motivations of the participants, the objects traded, the prices at which these objects are traded and the participants costs of trading. In a credit (rather than product or factor) market this means that firms wishing to borrow money know the price of credit (interest rates, charges etc), the collateral requirements (if any), the repayment schedule, etc, of all the banks (and other potential sources of finance¹³) in the market. They also know their own chances of repaying the loan and their own cost function (including of course, input prices etc). In addition to that they know (if relevant) how many other applicants there are in the market for funds, and so on. *Symmetry* of information arises because the banks and other competing financial institutions know *exactly the same things*. By contrast, *asymmetric* information is a regime in which one side to the (potential) transaction *knows less than the other*. Thus for example, if the firm knows more about its own projects than the bank (e.g. because the bank finds it too

12. There are of course counterexamples. The most glaring is the fact that huge numbers (millions) of people, often the poorest, engage in regularly in an unfair bet, namely the national lottery. This is inconsistent with risk aversion.

13. E.g. hire purchase rates.

costly to monitor the firm's activities closely) then this might mean that the bank does not know the firm's probability of default but the firm does.¹⁴ This is in fact a common assumption of models of the credit market including the celebrated Stiglitz-Weiss (1981) model of credit rationing.

So what are the implications of this possible asymmetry of information? There are two main phenomena that this may give rise to. The first is called *adverse selection*, a phenomenon that arises, for example, where 'bad' borrowers (those with high probability of default) dissemble, pretending to be 'good' borrowers (those with a low probability of default). At first the bank cannot tell the difference and offers them good and bad alike the same contract terms (e.g. a low rate of interest). Since in reality the default probability of the bad borrowers is higher than the bank imagined, this policy results in the bank making an overall loss on its lending. To solve this problem the bank will redesign the good borrower's contract to make it less attractive to the bad borrower. This is typically done with collateral requirements for the loan. The good borrower then ends up as the scapegoat, placing collateral on an otherwise uncollateralised loan, since collateral requirements dissuade the bad borrower from dissembling. Thus by the use of collateral the bank is able to separate the sheep (good borrowers) from the goats (bad borrowers).

The second effect of asymmetric information, is a phenomenon, familiar to students of insurance for many years, called *moral hazard*. This arises when borrower effort is unobservable to the bank. Effort is important to the bank as it influences the chances that its loan will be repaid. Because effort is unobservable, however, the lending contract cannot specify that a certain level of effort is necessary from the borrower to get the interest rate offered. The result is once more that the side with more information can exploit this. In consequence the entrepreneur will be able to get a loan on the basis of how hard she will work, but afterwards buys a BMW, puts her feet on the desk and allows the business to slide into bankruptcy. She can do this with some confidence, knowing full well that the bank will not detect her behaviour and that limited liability and the absence of collateral will protect her from the bailiffs.

The bank is, however, once more playing a repeated 'game', and will soon wise up to the moral hazard problem as it once more experiences losses on its loan portfolio. The manager will find after some thought that collateral is again a way to get the entrepreneur's (suspected) feet off the desk. This is a particularly powerful instrument of persuasion if you require that the loan is secured on the entrepreneur's house!¹⁵ So once again moral hazard issues can be resolved by collateralising the loan. Of course, like the use of collateral to

14. More generally, the bank may have a less precise estimate of that probability than the firm.

15. Typically a personal rather than business asset ('outside' collateral), thus overcoming limited liability issues.

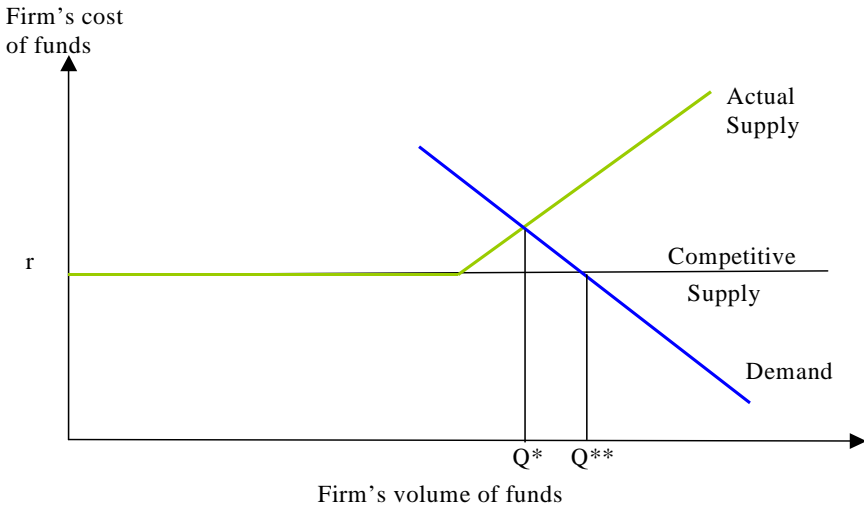
solve the adverse selection problem, its use in solving the moral hazard dilemma is not costless: as a result of the good borrower needing to place collateral, she will have lower expected utility or profits. Thus, there is a loss of utility to the 'innocent' party and to society as a whole, by comparison with the symmetric information situation.

In the next section we will see how the costs of asymmetric information are built into an analysis of credit constraints.

10. The Investment-Cashflow Relationship and Rationing

Fazzari, Hubbard and Petersen (1987) (henceforth FHP) found a positive correlation of small business' cash flows and their investment decisions. This, it was argued, was consistent with the hypothesis that internal funds were cheaper than external ones and firms would tend to substitute the latter for the former as the constraint on internal funds was relaxed (as would happen when the value of the firm increased with the extra cash inflow.). Moreover, their results showed that the marginal effect of cash flow on investment varied systematically with the size of firm, so that smaller firms seemed to face tighter constraints. This would seem to be consistent with the idea that smaller firms are constrained by imperfect capital markets to rely on internal finance to fund their operations. Figure 2 opposite shows the model underlying this hypothesis. We note that the competitive supply curve of funds is horizontal at the bank's cost of capital, r . The actual supply curve by contrast starts to turn upwards once internal funds have been exhausted and as moral hazard and adverse selection premia are priced into the cost of capital (at Q^*). The optimal amount of lending (Q^{**}) is therefore greater than the actual amount.

Figure 2 : Effects of Markets Solving Informational Problems on the Pricing and Availability of Credit (after Fazzari et al, 1987)



Does the existence of an upward sloping supply-curve for funds imply the need for government intervention? Not necessarily, according to the authors, since the costs of asymmetric information that are priced into this curve are *real costs*. They cannot be wished away. Government intervention might seem justified if the cost to society of the intervention was less than the welfare loss from doing nothing. This, however, is not necessarily the case. Therefore, in some sense we may be observing an *optimal* allocation of resources given the real costs of dealing with asymmetric information (see Cressy, 2002 and the symposium papers contained therein for a discussion of these issues).

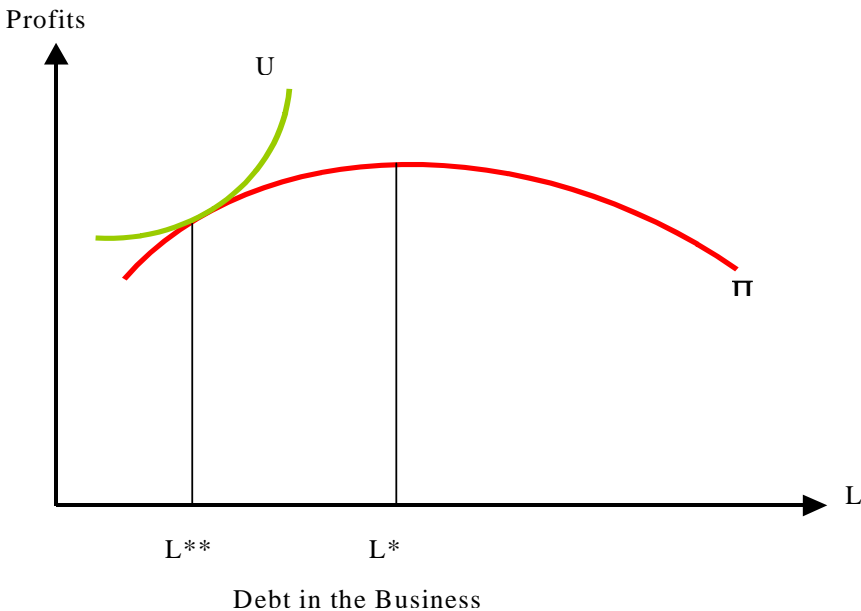
11. Alternative Explanations: Control Aversion

It is possible however, to interpret the FHP results in a different way. Entrepreneurs of smaller firms are well-known to be *control averse*: they do not like any kind of interference in their operations, in particular by Big Brother in the form of the local bank manager. (Cressy, 1996). For this (and other reasons) they tend to borrow little¹⁶. In the language of economics this

16. In Cressy, 1993) I showed that only 1/3 of businesses borrowed even on overdraft at startup. This grew to 1/2 within 3 years, but was still a minority of (surviving) businesses. Indeed, the attrition rate in the sample was considerable (many businesses closed within 3 years) but the propensity to borrow amongst survivors, and the average amount borrowed, increased over time.

means that the psychological costs of borrowing outweigh the benefits (at the margin) for the entrepreneur of the smaller firm. As firms get larger things get less personal, management tends to be rewarded by salaries rather than simply profits, and the aversion to perceived bank interference starts to wane. But at the level of the Micro business (one with less than 10 employees) control-aversion restricts borrowing *from the demand side*. The equilibrium tradeoff is illustrated in Figure 3 where the black line indicates profits of the firm as a function of borrowing. This represents the utility function of the financial manager of a larger firm. By contrast the grey line represents an indifference curve for the entrepreneur of a small firm. Whilst profits are a ‘good’ (yield positive marginal utility) borrowing is ‘bad’ (yield negative marginal utility). Thus the indifference curve is upward-sloping – its slope being the ratio of the marginal utility of borrowing to that of profits¹⁷. Utility is therefore increasing as we move to the North-West of the diagram with higher profits and lower borrowing. The highest indifference curve attainable with the black profit constraint is the grey one. The optimum for the larger firm is where profits are maximised, at L^* . The optimum of the control-averse entrepreneur, equates the marginal disutility of borrowing with the marginal utility of profits, yielding the smaller borrowing amount L^{**} .

Figure 3: Effects of Control Aversion on the Amount of Borrowing



17. The standard formula for the slope of an indifference curve is $-\text{MU}_x/\text{MU}_y$ where x and y are the two commodities yielding utility to the consumer.

12. Rationing, Bank Heterogeneity and Borrower Search

One feature of the definition of credit rationing is that we assume that the potential borrower has approached all potential lending sources and discovered that none will lend to her. In the simplest competitive model of course, all banks are identical, potential borrowers are costlessly aware of offers of funds and therefore search becomes unnecessary. However, we can imagine a world (the real world!) in which banks are heterogeneous, information is costly to the entrepreneur to acquire and that as a result a she approaches only a subset of potential lenders.¹⁸ This might arise for example, if manager ability (experience) is in short supply and distributed unevenly across banks. Then some banks will be more informed than others¹⁹. This may in turn lead to one bank rejecting the firm's request and another to accepting it, when approached. But once again, if the supply of credit from any bank is elastic, and search is costless, firms will find the banks that are more informed about their profitable project and borrow from them, so the banks that are less informed will make losses²⁰. It is only if firms do not know the 'best' bank and there are costs to finding this out that they may end up starved of funds for good projects. But this does not have to be a market imperfection: if at the margin the costs of search are equated to the expected benefits, then just the right number of projects will be funded.

13. Market Solutions: (I) Relationship Banking

Relationship banking is the situation in which a bank manager through his/her relationship with the entrepreneur(s) acquires information about the firm that will be material to his/her lending decisions. Empirically, small firms rarely switch banks – the cost of so doing is too high. And the result is that the typical firm has a relationship of some 10 years with its bank. Relationship banking would seem to have the advantage to a high quality firm in that it should expect better credit facilities and perhaps less rationing when the economy is in recession. By contrast a low quality firm might expect progressively less favourable treatment over time. In practice, the evidence is that relationship banking results in a lower requirement for collateral, and more credit availability but little affect on the interest rates charged. (Petersen and Rajan (1996), Berger and Udell (1998)). This empirical finding seems consistent with

18. Indeed empirical research in the UK shows that it is highly likely that banks do *not* all have the same beliefs about a given project.

19. Information is of course a two edged sword: one can learn that a firm's project is good or bad.

20. We can imagine that the anticipation of such losses would lead to banks competing up the price of scarce informed managers by their attempting to attract them away. And so on.

theory. In a simple model (Cressy, 2001) it can be shown that relationship banking will eliminate any initial information asymmetry with the result that firms will *never* switch banks, and all banking is therefore ‘relationship banking’. The argument goes as follows. Suppose there are two types of borrowers, High and Low, with qualities (success probabilities) $p_H > p_L$. Each project has a two-point distribution of returns, as usual, and requires \$1 to get off the ground. Borrowers join a bank now with the prospect of borrowing in 1 year’s time to fund their project. Hs and Ls occur with equal probability in the borrower population. To fix ideas suppose that $p_H=3/4$ and $p_L=1/4$. There are n competitive banks each with opportunity cost of lending r . Wealth of borrowers, and therefore collateral, is zero. Banks know the distribution of borrower quality and can learn about their own borrowers’ quality after lending to them. Under complete information on borrower types the bank prices according to the rule $\alpha = r/p$ ²¹ where r is the bank’s cost of capital and α the interest rate it charges on the risky loan. This yields $\alpha_H = 4r/3$ and $\alpha_L = 4r$. In the present, borrowers are allocated randomly to banks. In one period’s time the firm’s own bank learns with certainty their quality whilst other banks know only the average quality of the borrower pool. It would seem that the Hs would stay with their own banks whilst the Ls would prefer to search, since the latter’s return to switching would be a lower interest rate $\alpha'_L = 2r < 4r$. However, *the fact of switching itself conveys information about borrower type*: only the Ls switch. Hence the welcoming bank instead of charging $2r$ will now charge $4r$ to the new applicant – just as did the firm’s own bank. This means that both the Hs and Ls in fact stay put and *all* lending is done on a relationship basis²². What has happened, however, is that perceptions of firms credit risk have changed and this has enabled the bank to make more informed judgements about its clients loans and to price them accordingly²³.

-
21. This assumes that lending is a ‘fair bet’ i.e. that on average banks only cover their cost of funds, as would be the case under a competitive banking system.
 22. This model relies of course on a number of simplifying assumptions. For example, the length of relationship a borrowing firm has with its bank (namely the time taken to gain perfect knowledge about the borrower) is arbitrarily set here at 1 period and the bank learns with certainty the firm’s quality at the end of the period. We have also assumed that the cost to the bank of learning the information about borrower quality is zero. Most of these assumptions can be relaxed without altering the qualitative conclusions. For example, making the cost to the bank of being informed a fixed cost per \$ lent simply alters the bank’s cost of funds from r to $r+f$ and interest rates to $(r+f)/p$. All else remains the same. Likewise the discrete distribution of borrower quality is inessential: if the distribution of borrower quality is made continuous we get exactly the same result. This means that under SW assumptions the marginal borrower is profitable and so there is initially deficient credit in the market. Once relationships are established the marginal borrower is then offered funds that are priced correctly and no rationing occurs. Under DMW assumptions this same logic leads to elimination of surplus funds.
 23. This would predict that the average credit spread for a bank increases with the length of the relationship. I am not aware of any empirical test of this hypothesis however.

14. Market Solutions (II): Mutual Guarantee Schemes

A mutual guarantee scheme (MGS) is an arrangement whereby firms in a given locality or industry pay a membership fee to join an association one of the main benefits of which is insurance against loan default. The arrangement is that when any member firm approaches a bank to borrow, the MGS will guarantee the firm's risk to the bank. Obviously the chances of the MGS not being able to pay the bank are less than that of any individual firm, though the precise probability that an MGS itself will default depends on how large its membership is and the upper limit to borrowing by any given member. MGS societies have a strong incentive to monitor potential and actual members to avoid unscrupulous behaviour and scrupulous risk. However, such monitoring is feasible only within a reasonably well-defined and tightly knit group of firms, e.g. in a given locality and industry or craft. Furthermore, these 'collective' approaches to borrowing may be more successful in some countries than others due to cultural reasons. For example, it appears that the Southern countries of the EU together with France (which straddles Northern and Southern regions) are more open to MGS than the Northern ones. The colder countries of Northern Europe, where collective action is less of a tradition, are less willing hosts to this kind of organisation.

15. Market Solutions: Outside Equity?

If debt cannot easily be raised by a small firm due to absence of collateral then one might imagine that equity would be the alternative and indeed more suitable form of finance. Outside equity funding involves the purchase by an outside organisation or individual of shares in the firm. However, despite its seeming attractiveness there are insurmountable problems with this proposal.

Firstly, outside equity is by definition irrelevant to the majority of small businesses who are unincorporated and hence cannot (legally) issue equity.

Secondly, control aversion operates even more strongly in the case of equity rather than debt to discourage most small firms from gaining finance this way.²⁴ If they dislike even the minimal interference associated with a bank loan, what will the owners of a family business do if outsiders propose to muscle in and take seats on the Board?! One dares not think about it.

Thirdly, venture capitalists or Business Angels, the likely source of such finance are not interested in buying equity in the vast majority of small firms (Limited Companies) as they offer no prospects of capital gain of the order

24. Cressy and Olofsson (1997) found that some small Swedish firms would rather sell the business altogether rather than give up a share to an outsider! The aversion to outside equity declined with younger firms in the service industries.

they are used to. VCs are always concerned about exit routes, in other words, ways of harvesting their investments and gaining profits. Traditionally they have (*ex ante*) rates of return of 30% or above on a per annum basis on their investments²⁵. These rates of return are only possible however if the firm grows very fast and in a short time (3-5 years) ends up with a stock market flotation or a trade sale. The VC will typically exit her investment (sell her shares in the company) in 5-7 years. Needless to say, the overwhelming majority of small companies do not shape up to these requirements, run by relatively unambitious management teams and being slow- or no-growth in orientation and performance.

16. Government Solutions: Loan Guarantee Schemes

If relationship banking only works within a time horizon of 5 years or more and MGS schemes require specific local and cultural conditions to be successful in the solution of credit shortage, we might conclude that government intervention is desirable at least for the well-defined subset of young, small firms under consideration. And one of the most popular remedies proposed by governments eager to please (their electorate?!) is the Loan Guarantee Scheme or LGS. But we shall see that these schemes are by no means bereft of theoretical difficulties, whatever their practical usefulness may turn out to be.

Under a loan guarantee scheme the government agrees to indemnify the bank up to a certain proportion of its loan to a borrower without collateral, with the interest rate charged 'on purely commercial grounds', in return for the borrower paying an insurance premium on the loan of 1 or 2 % of its value. The objective is to get the bank to lend to borrowers to whom it would not otherwise lend in view of the borrower's lack of collateral. The theoretical problem with such schemes is that they do not address the issue of adverse selection and moral hazard on which they were predicated. Adverse selection, as we have seen above, under conditions of unobservable borrower quality (talent), is dealt with by the bank by making better quality borrowers place more collateral. Moral hazard, in the context of unobservable borrower effort, is dealt with by the bank by asking for collateral from borrowers. This, as we have seen above, creates an incentive amongst lazy borrowers to put in effort, since by so doing, they are more likely to avoid losing their house!

Unfortunately, neither of these issues is dealt with by the government loan guarantee scheme since by definition the borrower has not been required to place collateral on her loan, and so the bank is unable to charge the good quality applicant differentially from the bad. Likewise, greater effort cannot be

25. *Ex post* may be a different story.

engineered by the LGS for the same reason. Of course it is true that some loans are made without collateral; but *ex hypothesi* this is not the case here.

This might seem to be the end of the matter, the last nail in the coffin of the LGS. However, recent empirical research has impugned the oft-quoted maxim of adverse selection in (small) business banking, and whilst the finding was that moral hazard is still with us, the need to use collateral as a sorting device is perhaps no longer quite so compelling (Toivanen and Cressy, 2002). Thus we may need only to provide the incentive to effort otherwise missing in the LGS scheme. Furthermore, several rather competent studies in the UK suggest that LGS performs better than might have been predicted from theoretical considerations alone. Thus the Department of Trade and Industry (DTI) has shown rather convincingly that there is financial and economic additivity²⁶, at modest levels, in the workings of the LGS: firms are getting money they would not have otherwise got, and producing output that other firms would not have produced, in the absence of the scheme.

17. The Information Regime

Most of the analysis in the theoretical finance literature assumes that (a) asymmetric information regime holds and that (b) this is characterised either by adverse selection and/or by moral hazard. An implication of this theory is that collateral is used only under a regime of *asymmetric* information²⁷. This in turn makes the test for the information regime impossible since in practice collateral usage is pervasive. Thus, in order to test for the regime type we need a theory in which collateral can play a role even under *symmetric* information. In Toivanen and Cressy (2000) the first test of the information regime underlying the credit market was made. We allowed for the possibility that collateral might play a role in either a symmetric or an asymmetric regime. We also allowed for the role of market power in determining contract characteristics alongside information asymmetry, a feature of real-world markets largely rejected since the beginning of the information revolution.

So what were the results? Using a critical test for the information regime as to whether the probability of bankruptcy is exogenous (symmetric info) or endogenous (asymmetric info), and a measure of relative bargaining strength

-
26. Financial additivity occurs if the funds provided by LGS would not have been provided by other private sector financial institutions. Economic additivity occurs if the output from projects financed by LGS does not 'crowd out' private sector output.
 27. Symmetric information combined with deadweight loss from bankruptcy (deadweight is a loss to both parties to the contract) is enough to ensure that the parties to the credit contract bargain collateral down to zero. This is because the firm values the marginal unit of collateral higher than the bank for all positive levels of collateral. By contrast their valuation of interest payments is the same. In consequence the firm will bribe the bank to substitute interest payments for collateral until the latter is reduced to zero.

to measure monopoly power of firm and bank, we found that (a) the regime is indeed asymmetric information rather than symmetric; but (b) the dominant form is moral hazard rather than adverse selection; and finally that (c) monopoly power plays a significant role in the division of the surplus produced by the project with the bank taking the lion's share. The role of collateral is thus primarily shown to be one of creating incentives amongst borrowers to put in effort to their projects, since collateral is lost in the event of default. By contrast its role as a mechanism to avoid dissembling by the poor quality borrowers was found to be negligible. The conclusion from the analysis is simple: government should pay more attention to the problems of monopoly power in lending relationships rather than to informational asymmetry, particularly to the role of adverse selection a la Stiglitz and Weiss. Government attention was indeed directed to the monopoly power of the banks in the Cruickshank report (HM Treasury, 2000).

18. Conclusion

Are there credit constraints on small businesses in the real world? Is credit rationing an important phenomenon? These questions can only be answered, the above discussion suggests, by a detailed *empirical* examination of the characteristics of firms, the sectors of the economy in which they operate, the specific the time period or part of the macro cycle in question and the nature of the information regime in which all this is embedded. It seems a tall order! Theories of rationing are often abstract constructs based on questionable assumptions and theories that are difficult to test empirically. So we cannot always rely on theories to provide the guide to policy. For my own part, the empirical results that seem to me convincing rely less on theory for justification than straightforward questioning of participants together with cross-checks from other sources. By and large, these kinds of studies suggest that credit constraints are not a widespread phenomenon and that effective government intervention in the small minority of cases where it may exist is cheap and effective. Therefore despite the mountain of theoretical literature suggesting the abstract possibility of credit constraints, it does not appear in general to be an important empirical phenomenon. Often the major issue is rather the competitiveness of the banking system and the amount of information on the profitability of small business lending *to banks*. Some personal experience is instructive here. In a recent conference held in Zagreb, former Yugoslavia, which focussed on the role of loan and mutual guarantee schemes in former Eastern Bloc countries, it became clear that in many of the countries present there was a dominant oligopoly in the banking system which preferred to concentrate on large international firms rather than small local ones. The banks therefore had little awareness of the potential market for loans

to their entrepreneurial firms – a market long exploited by the Western banks. Part of the problem was clearly the issue of the availability of collateral, and the solution to which as we have seen might have been the introduction of loan or mutual guarantee schemes. But what really lay at the root of the problem was ignorance of the big banks of small firms and the restrictive practices of banking oligopolies that made them complacent about addressing the SME market²⁸. Banking information and reform was therefore in my view the first priority and without it these governments would be simply ‘rearranging the deckchairs on the Titanic’.

Reading: Financial Structure

- Brealey, Richard, Stuart Myers, (2003), *Principles of Corporate Finance (7th Edition)*, McGraw-Hill.
- Cosh, Andy and Alan Hughes, (1994) “Size, financial structure and profitability”, in Hughes and Storey, eds, *Financing small firms*, Routledge, 1994.
- Cressy, Robert and Christer Olofsson, (1997) “European SME financing: An overview”, *Small Business Economics* 9,87-96, 1997.
- Toivanen, Otto and Robert Cressy, (2000), “Lazy firms or dominant banks? An analysis of the UK market for SME loans” (summary version only), *Venture Capital: Special Issue*.

Reading: Financial Constraints

- Aston Business School, (1990), *Constraints on the Growth of Small Firms*, Department of Trade and Industry, UK.
- Berger A. & Udell G. (1998) The economics of small business finance: The roles of private equity and debt markets in the financial growth cycle; *Journal of Banking and Finance*; 22; pp. 613-673.
- Burke, A. E., FitzRoy, F. R. and Nolan, M.A. (2000), When Less is More: Distinguishing between Entrepreneurial Choice and Performance. *Oxford Bulletin of Economics and Statistics*, 62(5): 565-587.
- Chan, Yuk-shee, and Anjan V. Thakor, (1987) Collateral and Competitive Equilibria with Moral Hazard and Private Information, *Journal of Finance XLII, No. 2, June*.
- Cressy, Robert C., (1993), *The Startup Tracking Exercise: Third Year Report*, prepared for National Westminster Bank of Great Britain, November.
- Cressy, Robert and David Storey, (1994), *New firms and their bank*, National Westminster Bank of Great Britain.
- Cressy, Robert and Otto Toivanen, (2001), “Lazy firms or dominant banks? An analysis of the UK market for SME loans”, *Venture Capital: Special Issue*.
- Cressy, Robert (Ed.), (2002), “Funding Gaps: A symposium”, *The Economic Journal, February*
- Cressy, Robert, (1996), “Are business startups debt-rationed?”, *The Economic Journal* 106 (September), 1253-1270.

28. Even large Western European banks with many years dealing with the small business sector may be unsure of the profitability of small firms to the bank. To find an answer to this question was one of the objectives of the Startup Tracking Exercise, embarked on in 1990. The conclusion was positive!

- Cressy, Robert C., (2003a), "Why do most firms die young?", *Small Business Economics* (forthcoming).
- Cressy, Robert, (2003b), Upper bounds to random credit rationing frequency from the Startup Tracking Exercise dataset, *Note*, Cass Business School, London.
- de Meza, David and David Webb, (1987), Too Much Investment: A Problem of Asymmetric Information, *Quarterly Journal of Economics*, 102, 281-292.
- Evans, David and Boyan Jovanovic. (1989). "An Estimated Model of Entrepreneurial Choice Under Liquidity Constraints." *Journal of Political Economy* 97,4: 808-827.
- Small Business Research Centre, (1998), *The State of British Enterprise: Growth, Innovation and Competitive Advantage in Small and Medium-sized Firms*, Department of Economics, University of Cambridge.
- Stiglitz Joseph E and Andrew Weiss, (1981), Credit Rationing in Markets with Imperfect Information, *American Economic Review* 71,3, 393-410.

Reading: Alternative Explanations

- Cressy, Robert, (1995), Borrowing and Control: A Theory of Business Types, *Small Business Economics* 7: 1-10.
- Cressy, Robert, (1998), Credit rationing or risk aversion? An alternative explanation for the Evans-Jovanovic finding, *Economics Letters*.
- Guiso, Luigi and Monica Paiella, (1999) "Risk aversion, wealth and background risk", *Manuscript*, Birkbeck College, London, 1999.

Reading: Government Intervention

- Cowling, Marc and Peter Mitchell, (1996), *A Review of the Loan Guarantee Scheme*, Warwick Business School, Occasional Paper.
- National Economic Research Associates (NERA), (1989), *An Evaluation of the Loan Guarantee Scheme*, Research Paper No. 74, Department of Employment, 1990.
- Treasury, HM, (2000), Competition in UK Banking: A report to the Chancellor of the Exchequer, by Donald Cruickshank, March (This report can be downloaded from: http://www.hmTreasury.gov.uk/documents/financial_services/banking/bankreview/fin_bank_reviewfinal.cfm).