Financial Contracting With Optimistic Entrepreneurs*

Augustin Landier† David Thesmar‡

August 22, 2006

Abstract

Optimistic beliefs are a source of non pecuniary benefits for entrepreneurs that can explain the "Private Equity Puzzle". This paper looks at the effects of entrepreneurial optimism on financial contracting. When the contract space is restricted to debt, we show the existence of a separating equilibrium where optimists self-select into short-term debt and realists into long-term debt. Long-term debt is optimal for a realist entrepreneur as it smooths payoffs across states. Short-term debt is optimal for optimists for two reasons: (1) "bridging the gap in beliefs" by letting the entrepreneur take a bet on his project's success, and (2) letting the investor impose adaptation decisions in bad states.

We test our theory on a large dataset of French entrepreneurs. First, in agreement with the psychology literature, we find that biases in beliefs may be (partly) explained by individual characteristics and tend to persist over time. Second, as predicted by our model, we find that short-term debt is robustly correlated with "optimistic" expectation errors. Finally, to alleviate endogeneity concerns, we use future expectation errors as an instrument and confirm that optimistic entrepreneurial beliefs lead to a preference for short-term debt.

*This paper owes much to the constructive remarks of the referee. This paper also benefited from insightful discussions with Nicholas Barberis, Marianne Bertrand, Gilles Chemla, Bengt Holmstrom, Steve Kaplan, Ulrike Malmendier, Sendhil Mullainathan, Bernard Salanié, Per Strömberg, Tano Santos, Antoinette Schoar, Enrico Perotti, Raghuram Rajan, Jeremy Stein, Luigi Zingales as well as numerous seminar participants.

†New York University, Stern School of Business, email: alandier@stern.nyu.edu

‡HEC School of Management and CEPR.

1 Introduction

Starting a business is not a profitable activity: Hamilton [2000] documents that median entrepreneurial earnings after ten years of business are 35% less than the predicted alternative wage on a paid-job of the same duration. In addition, because the bulk of their wealth is invested in their own business, entrepreneurs bear a substantial amount of risk that only large private benefits can explain: Moskowitz and Vissing-Jorgensen [2003] estimate that entrepreneurs must enjoy non pecuniary benefits as high as 5 to 20% of their investment every year. These "private benefits of control", as the literature calls them, may correspond to pure hedonic flows: social status, the fun of running a firm or the independence that comes with it. However, in this case, one would be left with the puzzling fact that these benefits amount on average to some 150% of the entrepreneur's annual income.[1]

An alternative interpretation of these findings is that private benefits are pies in the sky: Entrepreneurs do not start new businesses because it is profitable, but because they wrongly believe it is. Many studies show that entrepreneurs typically overestimate the chances that their project will be successful. In their survey, Cooper, Woo and Dunkelberg [1988] find that 68% of entrepreneurs thought their own business would do better than their others' (see also Pinfold [2000]). Experimental evidence suggests that people's optimism about their own ability relative to their competitors' leads to excess entry in a game of entrepreneurship (Camerer and Lovallo [1999]).[2]

This paper examines and documents implications of the fact that entrepreneurial private benefits take the form of optimistic expectations. In a financial contracting framework, we find that differences in opinions between the (optimistic) entrepreneur and the (realistic) investor affect the optimal contract in a fashion similar to differences in objectives (agency conflict): in particular, optimistic entrepreneurs make more use of short term debt. Our results therefore stress the role of differences in opinions as a key determinant of capital structure[3], which has so far mostly been explained through agency consid-

[1] Moskowitz and Vissing-Jorgensen's estimates.

[2] Optimistic expectations about performance can result from the "above average effect", a bias in perception abundantly documented in psychology and particularly strong when uncertainty is high and motivation at stake (Armor and Taylor [2000]). In the case of entrepreneuship, a powerful driver of optimism is also *selection*: Individuals who leave other opportunities to start a new venture tend to be those who, on average, overestimate the prospects of their project. This selection effect creates a natural upward bias in expectations, much like the winner's curse effect set forth in the auction literature (Thaler [1988], Roll [1986]).

[3] We focus here on the maturity of debt because debt is the only means of external finance for most entrepreneurs. Similar insights can, however, be derived within more general contractual environment. When we allow for contingent control transfers for instance, we

erations only. We then go to the data, document the large heterogeneity of entrepreneurial beliefs and find robust, convincing, evidence that short term debt is related to optimism, controlling for its usual determinants.

Our theoretical analysis shows that optimal contracts for optimists are contingent on events that the entrepreneur *does not control* (external risk), but holds overoptimistic expectations about. Two effects are at work: First, optimistic entrepreneurs inefficiently persist in implementing the initially ambitious project even if new information calls for a safer strategy. Hence, optimal contracts for optimists (short-term debt) transfer control to the investor in those states of nature where a realistic decision maker is needed.[4] Secondly, an optimistic entrepreneur is willing to exchange cash flow rights in the low state (that he believes to be unlikely) against claims on the good state (that the investor knows to be unlikely). These differences in valuation across states of nature call for a contract that provides more upsides to the entrepreneur when he/she is optimistic.

Hence, modelling private benefits as optimism allows to reconcile some recent, apparently paradoxical, empirical findings with financial contracting theory. Common agency theory predicts that optimal contracts should *insure* the agent against risks he/she does not control. However, Kaplan and Stromberg [2002] have shown that VC backed entrepreneurs bear much more external risk than should be optimal. Along similar lines, one of the main lessons of CEO compensation literature is the surprising rarity of relative performance evaluation schemes (Murphy [2000], Bertrand and Mullainathan [2001]). These pieces of evidence conflict with common agency theory, but receive a natural interpretation in our framework: entrepreneurs or CEOs overestimate their chances of success. As a result, they have a strong preference for control and cash flow rights contingent on good states of nature.[5]

We then empirically document entrepreneurial optimism and test the major prediction of our model: Entrepreneurial optimism is one of the factors

can prove that difference in opinions give rise to venture capital - like contracts where the entrepreneur loses control when they firm performs poorly.

[4] This effect arises solely from differences in opinion, not from agency problem as in Aghion and Bolton [1992]. These contingent transfers in control are a feature typical of venture-capital contracts (Kaplan and Stromberg [2003]), but they are, in our paper, implemented through debt maturity.

[5] We stress that in equilibrium, investors do not *exploit* optimistic beliefs: They make zero profit on both realistic and optimistic entrepreneurs. The equilibrium is a separating one where, even though entrepreneurial beliefs are not observable, both revelation constraints – for optimists and realists – are non-binding: Optimists prefer short-term debt because it leaves large payments and control to the investor in states that (they believe) are never going to occur. For realists, these financing contracts simply look too risky, and they strictly prefer insurance provided by long term debt.

explaining capital structure, aside from well documented agency considerations. Our dataset comes from two waves of a survey conducted by the French statistical office on a population of entrepreneurs the very year their business was started. This survey contains direct information on (1) entrepreneur *initial expectations* on future business growth, (2) entrepreneurial socio-demographic characteristics and (3) project characteristics. This dataset is then matched with accounting data collected from tax files, which allow us to draw a relationship between the entrepreneur's characteristics, his expectations and the actual venture performance up to seven years after birth.

We draw several conclusions from this empirical analysis. First, we gather evidence that some entrepreneur consistently make positive expectation errors. Expectation errors made by entrepreneurs tend to persist over time and are not well explained by industry shocks. Finally, some observable characteristics are strongly associated with systematic upward expectation biases on the venture's performance. Notably, entrepreneurs with higher education and those who are developing their "own idea" tend to be more optimistic, whereas entrepreneurs who take the business over from someone else tend to be less optimistic. These differences may be understood in the context of a model of "choice-driven" optimism, a la Van den Steen's [2004]. Provided that some agents form beliefs about their entrepreneurial ideas that differ (positively or negatively) from the unbiased expectation, entrepreneurs are optimistic on average about their project, as the "pessimists" don't become entrepreneurs. Interestingly, this simple selection theory of has strong comparative statics implications that we find validated in the data: Those with higher non-entrepreneurial outside options (e.g. higher education) exhibit more optimism, while those receiving more accurate signals on projects have smaller biases (expertise in industry, idea less "novel", firms not actually created but taken over).

Secondly, we find a fairly robust, positive, correlation between optimistic expectation errors and the use of short term debt. The companies we observe are small and use debt as their almost exclusive source of external finance. A natural capital structure variable to look at is therefore the maturity of debt. In a first stage, we simply correlate expectation errors with the use of short-term debt, using two different measures of both. These correlations are strong and robust, and remain so once we control for obvious determinants of expectations that may be correlated with capital structure. We have to acknowledge, however, that these estimates may be biased. We thus propose to instrument expectation errors at the time of creation with expectation errors *three years* after creation. We assume that these future expectation errors are not inherently correlated with capital structure choice at the date of creation. With this methodology, the effect of expectation on capital structure shows up both statistically significant and large.

This paper is part of a growing literature, pioneered by Roll's [1986] analysis of takeovers, that explores the impact of managerial behavioral biases on corporate decision-making. Heaton [2002] shows that managerial optimism offers a unifying view on overinvestment in the presence of free cash-flows (the manager overestimates their NPV) and underinvestment when funds have to be raised by issuing risky securities (the manager believes external finance is too costly). Malmendier and Tate [2002,2003] empirically document, for large, listed firms, the link between CEOs overconfidence and overinvestment using the personal investments of these CEOs in their companies as a measure of overconfidence. Directly related to our topic, De Meza and Southey[1996] show in a model that heterogeneity of beliefs among potential entrepreneurs can explain high failure rates, credit rationing and a preference for debt rather than equity. Coval and Thakor (2003) develop a model where rational agents become financial intermediaries to act as a "beliefs bridge" between the optimists –who become entrepreneurs– and the pessimists –who choose to become investors in the intermediary. Hackbarth [2004] develops a model where optimistic managers exhibit a preference for debt rather than equity, which can increase ex-ante efficiency as it mitigates agency costs.

The paper has four more sections. Section 2 documents in the light of the psychology literature what the most likely sources of differences in beliefs between entrepreneurs and investors are and relates them to observable characteristics. Section 3 outlines a credit-market equilibrium where both realistic and optimistic entrepreneurs coexist. Section 4 is devoted to the empirical analysis: We describe the empirical heterogeneity in beliefs and test for a link between beliefs and debt-contract choice. Section 5, the last one, concludes.

2 Differences in Beliefs and Entrepreneurial Optimism

At the core of our analysis is the assumption that entrepreneurs deviate from rational expectations about the odds of their project succeeding. What are the origins of such deviations? Entrepreneurial projects typically are highy uncertain; because of their novelty, there is very little evidence on which to base future expectations. Under these circumstances, experimental psychologists have shown that agents tend to rely on crude heuristics and that these heuristics may give rise to biased beliefs. At least three psychological mechanisms may be mentionned. The first one is the "above average" effect: the psychology literature documents the fact that, when odds are very difficult to assess, people tend to hold high beliefs on their chances of performing at a given task (Taylor and Brown [1988]). The circumstances under which such self-serving beliefs arise are, however, not well understood, as agents may also display excessively pessimistic beliefs in some settings (Ross and Anderson

[1977]). In the case of entrepreneurship, however, the above average effect may be reinforced by strong motivational factors as positive beliefs help the entrepreneur to commit to a high effort (Armor and Taylor [2000]).

A probably more convincing explanation for entrepreneurial optimism is the *planning fallacy* (Kahneman and Lovallo [1993], Kahneman and Tversky [1979]). A common heuristic used to assess the chances of suceeding is to simulate the environment with chains of events linked together by probabilities. Experiments document the fact that agents have great difficulty in estimating compound probabilities and stick to a simple rule of thumb like taking the average probability of success across nodes, or the probability of success in the first node (Gettys, Kelly and Peterson [1973]). In many experiments, this inference process naturally leads to overoptimism about the probability and the time of completion of a task

In our viewpoint, the strongest source of entrepreneurial optimism is likely to be *selection*: people don't become entrepreneurs by accident but because they perceive that they have a project that dominates their other career choices. If they have noisy assessments of their projects, those who become entrepreneurs hold on average optimistic beliefs. This "choice-driven" theory of over-optimism is developped in Van-den-Steen [2004] and allows to make precise predictions about what observable characteristics we can expect to be correlated with optimism.

To see how, consider a population of potential entrepreneurs. Each agent i has an idea, whose value can be either high (V_H) or low ($V_L = 0$). The *objective* probability that the project is good is α_i, but agents have a prior belief $\widetilde{\alpha}_i$ drawn from a distribution G_i. Let's assume agents are right on average, i.e. $\int \widetilde{\alpha} dG_i = \alpha_i$. Agents become entrepreneurs if their subjective assesment of the project's value, $\widetilde{\alpha}_i V_H$ exceeds the value V_i they get by staying in paid employment. Conditional on becoming an entrepreneur, an agent has a belief which is on average higher than the objective one (α_i) by a factor:

$$\int_{V_i/V_H} \frac{\widetilde{\alpha}}{\alpha_i} dG_i(\widetilde{\alpha}) > 1$$

If all agents were entering entrepreneurship, their average expectations would still be unbiased. But since only those who feel their idea has a value exceeding V_i actually choose to be entrepreneurs, occupational choice leads to an *average overoptimism* of entrepreneurs (the most pessimistic agents remain employed).

This simple model of entrepreneurial optimism generates two comparative statics that will guide us later in our empirical strategy. First, entrepreneurs who have larger outside options in employment (V_i) are on average more overoptimistic about their project's chances of success (because the selection effect described above is stronger). We thus expect that more educated and more experienced agents who select into entrepreneurship should be more optimistic, because they could claim a higher wage on the labor market. Second, agents with less precise information (e.g. in the sense of a mean-preserving spread in G) have a larger over-optimism bias. We thus expect agents with more expertise in the industry to be less optimistic. On the contrary, agents whose motivation is to implement a "novel idea" have a noisier signal and are expected to be more optimistic, provided they choose to become entrepreneurs.

3 Model

We now take this heterogeneity of beliefs as given among entrepreneurs and ask how it affects the credit-market equilibrium, in a model where both realistic and optimistic entrepreneurs coexist, are not distinguishable, and can raise funds for their projects.

3.1 Set-Up

There are three dates, $t = 0, 1, 2$. A cohort of wealthless entrepreneurs, protected by limited liability, raise I at $t = 0$ to finance a project. The returns of the project at time 2 depend on a strategy decision at time $t = 1$ (say, *growth* or *safe*) and on the project's fitness to the market - its type. Projects can be of two types: good or bad. When the entrepreneur chooses the *growth* strategy at time 1, a good project yields R , and a bad one yields zero. If the strategy chosen is *safe*, both types of projects yield L. When the project is a good one, the *growth* strategy is better than the *safe* strategy: $R > L$. When it is a bad one, the *safe* strategy is the better one: $L > 0$.

At time 1, the entrepreneur receives a noncontractible signal about the project's fitness and bases his choice of a strategy on this information. This signal takes the form of an intermediate cash flow generated by the firm at $t = 1$. This cash flow is R with probability 1 if the project is good. If the project is bad, this cash-flow is R with probability p and 0 otherwise. Hence, a zero cash flow is a sure sign that the project is bad, and that the optimal strategy is the safe one (which yields L instead of 0).

The sequence of events is summarized in figure 1. First, investment I is sunk. At date $t = 1$, the interim cash flow is observed. The strategy is chosen by whoever (entrepreneur or investor) holds control of the firm. Last, in $t = 2$, the project generates the final cash flows, depending on its type and the strategy chosen.

A priori, there are as many good as bad projects to pick up. Hence, a given project is good with probability 1/2 and bad with probability 1/2. All entrepreneurs are risk averse with concave VNM utility $u(.)$.

In order to pinpoint the effects of differences in beliefs on financial contracting, we choose here to simply posit that some entrepreneurs are more optimistic than others. In order to make things even clearer, we will make an extreme assumption about differences in beliefs. First, *realists* have correct priors about the project's type. Hence, they ex ante believe that the project is Good with probability 1/2. Once he observes interim cash flows, the realistic entrepreneur incorporates the additional information following Bayes' Rule. His new beliefs at date $t = 1$ are thus given by:

$$\begin{aligned} \text{P(type} &= \text{good}|\text{interim CF } = R) = 1/(1+p) \\ \text{P(type} &= \text{good}|\text{interim CF } = 0) = 0. \end{aligned}$$

Optimists don't have realistic a priori beliefs on the project's type. Ex ante, they believe the project is good with probability 1. Even though the optimistic entrepreneur also uses Bayes' law to update his beliefs at date $t = 1$, he interprets the interim cash flow information differently. Indeed, for an optimist:

$$\begin{aligned} \text{P(type} &= \text{good}|\text{interim CF } = R) = 1 \\ \text{P(type} &= \text{good}|\text{interim CF } = 0) = 1. \end{aligned}$$

In our extreme case, where optimists are *sure* that the project is a good one, they discard all interim information they get about it. Hence, optimists do not update when they see no interim cash flow: this is a limit case, but perfectly consistent with bayesian updating.[6] More precisely, optimists make two kinds of mistakes ex ante: first, they overestimate the probability of a good signal. They think good signals occur with probability 1 (good projects never fail), while realists think good signals occur with probability $(1 + p)/2 < 1$ (bad projects may fail). The second mistake optimists make is that they overestimate the probability of success of the growth strategy (1 versus 1/2). The business plan, as seen by an optimistic entrepreneur, is given in figure 2.

To focus on the important effects, we make the following additional assumptions:

1. Financial markets are competitive and investors hold realistic beliefs.[7]

[6] We consider for its simplicity this limit case of optimism. Proofs available from the authors show how these results can be generalized to moderate optimism, as long as (1)optimism is sufficiently strong and (2) the signal is sufficiently informative about the project's choice.

[7] This assumption is consistent with Coval and Thakor (2003). They develop a theory

2. Conditional on the signal being good, growth is the efficient strategy:

$$\frac{1}{1+p}R > L;$$

of course, this assumption ensures that $R > L$.

3. If the entrepreneur could commit to always choose the safe strategy (whether the signal is good or bad), the NPV of the project would be positive:

$$L > I.$$

4. The project cannot be fully financed by its payoff in the bad state, i.e.:

$$I > \frac{1-p}{2}L.$$

5. The signal is observable but not contractible.

In our data, an overwhelming majority of new ventures are financed by simple debt contracts of either short or long maturity. Venture capital contracts, which specify both contingent repayments *and* control transfers are used by only a very small fraction of new companies. This is not a surprise as the French private equity market is less developed and more late-stage oriented than the US one. For this reason, we analyze the credit-market equilibrium with debt contracting. The debt contract can take two forms: first, a short-term debt contract, that specifies a repayment at date 1. If cash-flow is 0, the entrepreneur has to default and the investor gets control and ownership of the firm. The other type of contract is long term debt, specifying a repayment at $t = 2$. Recall that the signal is observable, so renegotiation may occur in date 1 in order for the investor to induce the entrepreneur to choose the *safe* strategy if he is tempted to play *growth*.

3.2 Results

Our main result is that there is a unique competitive separating equilibrium. In this equilibrium, the optimists choose short-term debt contracts whereas the realists choose long-term debt.

---

of financial intermediation, where rational agents become financial intermediaries to act as a “beliefs bridge” between the optimists –who become entrepreneurs– and the pessimists –who choose to become investors in the intermediary. What matters for our model is that the marginal investor holds beliefs that are "more realistic" than an optimistic entrepreneur.

Proposition 1 *When only debt contracts are available, the equilibrium is separating with optimists choosing a short-term debt contract and realists choosing a long-term debt contract.*

- *The short-term debt contract has a repayment level:*

$$D = \frac{2I - (1-p)L}{1+p}.$$

- *The long-term debt contract has a repayment level:*

$$D = I.$$

- *Investors make zero profit with either type of entrepreneur.*

To prove this result, we proceed in two steps: we first assume that the entrepreneur's beliefs are observable and solve for the optimal contracts. We then show that this pair of contracts is self-selecting.

To understand the logic behind the optimality of short-term debt, it is useful to ask oneself what the optimal contract would be with an optimistic entrepreneur in a frictionless world. Assume that the investor could observe that an entrepreneur is optimistic and that the signal were contractible. First, contrary to the investor, the entrepreneur believes the signal S will be positive for sure. The optimal contract will therefore give him a positive payoff only if $S > 0$ and zero otherwise. Second, the investor knows that in case of a bad signal, value can be created by taking the safe strategy rather than the growth one. A second feature of the optimal contract is therefore to allocate control to the investor in case of a bad signal. Ex ante, the optimistic entrepreneur believes this will not happen, therefore, such a provision in the contract has no cost from his perspective. The benefit is that it increases the project NPV from the investor perspective and therefore the payoff that can be left to the entrepreneur in the good state.

It turns out that short-term debt can implement this first-best contract: indeed, with short-term debt, the investor gets full control and ownership in the bad state. What is the promised repayment D the investor asks for? It is simply given by the zero-profit condition,

$$I = \frac{1+p}{2}D + \frac{1-p}{2}L$$

Note that the short-term debt contract that a realist could get would be exactly the same, as beliefs do not distort strategy choice for this type of contract. But is it the contract a realist would prefer? Consider an entrepreneur who is able to make the case that he is a realist. He is therefore able to commit to choose the safe strategy if $S = 0$ (he knows that not doing so would yield a zero cash-flow). Given that $L < I$, this makes long-term debt risk-free with a realist. The investor can therefore offer a long-term debt contract with repayment $D = I$. Our realist entrepreneur strictly prefers this contract to the short-term debt contract as it smooths cash-flows across states of nature:

$$\frac{1}{2}[u(2R-D)+pu(R-D)+(1-p)u(0)] < \frac{1}{2}[u(2R-I)+pu(R-I)+(1-p)u(L-I)].$$

To finally establish that these contracts are self-selecting, it remains to be shown that an optimist does not want to pretend to be a realist and get a long-term contract.

The revelation contraint for optimists is:

$$u\left(2R - \frac{2I-(1-p)L}{1+p}\right) > u\left(2R - I\right).$$

To see why it always holds, let us write the difference in expected payoffs:

$$\begin{aligned}\Delta &= \left(2R - \frac{2I-(1-p)L}{1+p}\right) - (2R - I) \\ &= -\frac{1-p}{1+p}(L-I) < 0.\end{aligned}$$

From an optimistic's viewpoint, investors lose money with the short-term contract. Short-term debt looks cheaper to them as they get more of the upside of the project.

4 Tests

This section is devoted to testing one premise and one prediction of our model. The premise is that beliefs are heterogenous across entrepreneurs. We use a large dataset on French entrepreneurs that provides us with their expectations. We document that, for some entrepreneurs, expectations tend to be systematically above or below realizations.

Then, we test the main equilibrium prediction of the model: other things equal, optimistic entrepreneurs take on a larger fraction of short term debt. Using our dataset we document a robust correlation between expectation errors and the use of short term debt by the firm. We discuss potential endogeneity biases, and propose various strategies to tackle them.

4.1 A Short Description of the Data

Our dataset consists of the merging of two sources available from the French statistical office (INSEE). The first dataset is a survey on entrepreneurs conducted in 1994 and 1998 by the statistical office. The second source is the Tax Files, which provide us with detailed accounting data at the firm level between 1994 and 2003.

4.1.1 Entrepreneur Dataset

Our first source is the SINE survey on *French entrepreneurs*. In 1994, the French statistical office (INSEE) sent questionnaires to approximately 20% of the entrepreneurs who started or took over a business in France that year. The response rate is high (85%) because answering the survey is mandatory. Thus, we have data for 30,778 firms created/taken over in 1994. In 1997, these firms were re-sent similar questionnaires, but only 18,132 responded, yielding an attrition rate of 41% in three years. Part of this attrition is natural, and part of it is due to firms moving and not being found again by survey managers. The process was then repeated for firms started/taken over in 1998. The 1998 survey wave had 30,068 entrepreneurs surveyed in 1998, and 27,136 still present in 2001.

We thus have a representative *panel* of new firms, half of them started in 1994 - a recession year - and half of them started in 1998 - a year of expansion. This survey of new businesses has information on the entrepreneur's main socio demographic characteristics (age, education, social background), and on his growth expectations as he starts/takes over the business. Other qualitative questions relate to (1) the reasons for which the firm was started (2) the conditions under which it was started (financing, initial research, customer prospectives) and (3) the management of the first three years of operation (change in product line, agressive commercial policy conducted). The first two types of questions correspond to variables collected in the same year the business is started, while the last type of variables corresponds to answers collected three years after.

Using the answers to the questionnaire, we construct the following variables:

1. **AGE:** is the entrepreneur's age, in years. In most regressions, however, we use instead a dummy equal to 1 when the entrepreneur's age is above the median (37).

2. **EDUCATION:** education is broken down into four possible categories: high school dropout (reference), high school graduate (HSG), College

graduate (CG), and Post graduate studies or "*Grande Ecole*" graduate (GE).[8]

3. **SERIAL ENTREPRENEURS**: a dummy equal to 1 when the entrepreneur has started at least one business before this one.

4. **EXPERTISE:** a dummy equal to 1 when the entrepreneur has previous experience within the industry. The exact phrasing of the question is: "In your previous job experiences, did you acquire skills: (1) in the industry you are setting this business in? (2) in a similar activity? (3) in a very different activty? and (4) you have very diverse skills. The EXPERT dummy is equal to 1 when the entrepreneur answers (1).

5. **MOTIVATION: A NEW IDEA:** The question about the entrepreneur's motivation is "was the main motivation that drove you into starting a firm: (1) a new idea (2) a taste for entrepreneurship, for independance (3) an opportunity (4) other entrepreneurs among family or friends (5) until then unemployed. The answers are nonexclusive, but our IDEA dummy equals 1 when the entrepreneur selects (1).

6. **MOTIVATION: AUTONOMY :**Our AUTONOMY dummy equals 1 when the entrepreneur selects (2) in the above question.

7. **"REAL" START-UP:** Some firms are truly created. Others are purchased from another owner, or inherited. The STARTUP dummy is equal to 1 when the firm is truly created.

8. **"DEVELOPMENT" EXPECTATIONS**: The entrepreneur is asked about his expectations for the next 6 or 12 months, roughly one year after the firm is started/taken over (which can be 1994 or 1998 depending on the survey wave). The question is phrased "What is your view of the future?", and the possible answers are: (1) the firm will develop, (2) the firm will keep its current balance, (3) I will have to struggle (4) I will have to shut down the firm (5) I will sell it (6) I do not know. Our EXPGR0 dummy equals to 1 when the entrepreneur answers (1) and 0 when he answers (2), (3) or (4). Entrepreneurs responding (5) or (6) were removed from estimation.

9. **"HIRING" EXPECTATION:** The second expectation variable is related to employement. Again, the entrepreneur is asked about his expectations for the next 6 or 12 months, roughly one year after it is started

[8] Unfortunately, the questionnaire does not allow us to break down this last category into grandes écoles and post graduate studies, which are relatively frequent in France. This is, however, possible using the Labor Force Survey (equivalent to the CPS in the US). Looking at entrepreneurs from the 1991-1993 waves of this survey, we find that more than 80% of the postgraduate-Grande Ecole entrepreneurs are actually graduates from *Grandes Ecoles.*

(which can be 1994 or 1998 depending on the survey wave). The question is phrased "Do you plan to hire in the next 12 months?", and the possible answers are: (1) yes, (2) no or (3) I do not know. Our EXPEMP0 dummy equals to 1 when the entrepreneur answers (1) and 0 when he answers (2). Entrepreneurs responding (3) were removed from estimation.

Sample means on expectation variables (EXPGR0 and EXPEMP0) are reported in Table 1. Both in 1994 and 1998, most entrepreneurs expect their firm to either develop or remain on a steady course, while very few expect "difficulties". This is why we include troubled firms in the reference category when we define EXPGR0. Our results below are not sensitive to that convention. It can also be noticed from Table 1 that while more than 50% of surveyed entrepreneurs expect the firm to "develop", less than a third of them expect to hire a new employee. The firms in our sample, which include many proprietorships and small corporations can be expected to grow without hiring. This suggests that our employment based expectation measure EXPEMP0 is somewhat noisier than the straight "development" expectation. This observations will be confirmed in the following analysis.

Descriptive statistics on the other entrepreneur/project characteristics are reported in Table 2. We compute the mean of each variable - all of them dummies apart from age - using two different ways to split the sample. Columns 1 and 2 split the sample by organizational form. Columns 3 and 4 split the sample into firms with at most one employee and firms with more than one employee. As expected, more educated and more experienced entrepreneurs start/take over larger firms. New ideas are more likely to be implemented in corporations, while autonomy motivated entrepreneurs tend to run smaller business and proprietorships.

Last, the questions on future development and employment expectations were also part of the second questionnaire sent to firms three years after the first one. We therefore constructed "development" (EXPGR1) and "hiring" (EXPEMP1) expectations *three years* after the business was started/taken over. The construction process is identical to initial expectations variables (EXPGR0 and EXPEMP0) as the questions asked were identical. As a result, the panel nature of our dataset - two questionnaires per firm in each wave - allows us to observe two expectations per (surviving) entrepreneur.

4.1.2 Accounting Data

To measure "optimism", we need to compare expectations with realizations. Realizations on growth and employment are retrieved from tax reports (Bénéfices Industriels et Commerciaux), available for all firms making more than

110,000 euros in annual sales. Tax files provide us with balance sheet information, operating income, and employment. They can easily be matched with the SINE surveys as both sources share the same firm identifying number (SIREN). The accounting data are – theoretically – available for every year since the firm first shows up in the tax files, so they allow us to follow the firms from their start. Balance sheet information - hence capital structure - is more detailed for larger firms (essentially, those whose annual sales exceeded 230,000 euros).[9] As a result, the number of observations will drop severely when we look at capital structure.

We match the two datasets, and first remove those firms whose accounts are not reported within their first two years of existence by the tax reports (1994 or 1995 for the first wave, 1998 or 1999 for the second one). We end up with 39,540 firms started either in 1994 or in 1998, present in the SINE surveys, and whose accounts are reported within their first two years of existence. We thus lose almost 20,000 firms in the merging process, but these are overwhelmingly small firms, whose sales are below 110,000 euros.

A little more than half of our sample (23,000) corresponds to newly created firms. The rest are existing firms taken over by new entrepreneurs. The upper panel of Table 3 displays some accounting variables of firms the year they were started or taken over (that is, either 1994 or 1998). In their first year of existence, newly created ventures are small: they typically employ 1.5 workers, and use 35,000 euros of fixed assets, to make up no more than 200,000 euros of total sales. Breaking down the sample into corporations and noncorporations highlights the considerable skewness of firm-size distribution. In contrast, firms that merely changed hands are on average twice as large as newly created firms, consistent with a simple age effect.

Our theory has predictions on the share of short-term loans in outside finance. For a subset of our firms[10], the accounting data allow to break down total debt into (1) short-term bank debt (all loans with maturity of less than two years), (2) long-term bank debt and (3) "other financial debt". For our small firms, this last item mostly consists of loans made to the firm by the owners and their relatives. Given that these loans are likely to be junior to any bank loan, we treat them as equity. In addition, the data provide us with the share of short-term bank debt that takes the form of bank overdrafts.

Unfortunately, the share of bank debt with less than 2 years of maturity

[9] The reason is that small firms in France can choose between two ways of reporting their income to tax authorities: the "simplified" and the "regular" tax regime. The regular tax regime becomes compulsory as soon as annual firm sales exceed 230,000 euros, and requires detailed information about the debt structure. Firms that can and do opt for the "simplified" regime are not required to provide as much detail and just need to report the overall amount of financial debt.

[10] Basically, all firms with turnover above 250,000 euros (see appendix).

includes longer-term loans that will end in less than two years. It is thus a noisy measure of short-term debt, especially for firms taken over: being older, these firms are likely to have accumulated long-term debt in the past. Hence, to measure the level of short-term indebtness, we will use the ratio of credit lines to total bank loans *in the year when the firm is created/taken over*.[11] We divide bank overdrafts by total bank loans because, as mentionned above, they are the almost exclusive source of outside finance.[12] The lower panels of Table 3 provide descriptive statistics on short term debt, depending on whether the firm is a startup and/or a corporation. As it appears, credit lines constitute the bulk of short-maturity debt.

4.2 Measuring Biases in Expectation

Since our ultimate goal is to relate optimism with capital structure choice, we need to find an entrepreneur-level measure of bias in expectations. To do this, we compare the entrepreneur's "development"/"hiring" expectation from the SINE survey with the venture's actual growth that we observe from accounting data. Such an expectation error has *a priori* two components: the bias (which is deterministic) and the true error (which is ex ante random, of zero mean). This section first explains how we compute these expectation errors, and then documents the existence of expectation biases in the data.

4.2.1 Expectation Errors and Bias

To fix ideas, assume an entrepreneur is asked to form expectations on a random variable $\widetilde{Y}$ (for instance sales growth, or future hires), using an information set I. Let us note his subjective expectation as $E_s\left(\widetilde{Y}|I\right)$. We then call *expectation error* the difference between the subjective (reported) expectation and realization $\widetilde{Y}$. Without loss of generality, this error can be written as:

$$\begin{aligned}\Delta &= E_s\left(\widetilde{Y}|I\right) - \widetilde{Y} \\ &= \underbrace{E\left(\widetilde{Y}|I\right) - \widetilde{Y}}_{\text{rational error } \varepsilon} + \underbrace{E_s\left(\widetilde{Y}|I\right) - E\left(\widetilde{Y}|I\right)}_{\text{bias } b} \qquad (1)\end{aligned}$$

where $E\left(\widetilde{Y}|I\right)$ is the "true" expectation of $\widetilde{Y}$, conditional on information set I. The above equation shows that the difference between the reported expectation

[11] We ran - but do not report - separate regressions using bank loans with less than two years of maturity; they delivered results similar to – albeit sometimes weaker than – the ones with credit lines we provide in the following analysis.

[12] Only a negligible fraction of our firms are financed through venture capital.

and realization is the sum of two components. The first component is a random variable $\widetilde{\varepsilon}$ of mean zero. It is the error that a rational agent would make, and is by assumption unpredictable with information I. The second component b is deterministic, and is equal to zero if and only if the entrepreneur is rational. If $b > 0$, the entrepreneur is optimistic in the sense that he systematically overestimate the future mean of Y.

Both expectations and realizations are available for each entrepreneur in our data. Our empirical strategy will therefore be to compute $\Delta = b + \widetilde{\varepsilon}$ as a proxy for b at the entrepreneur level. Our measures of expectations are retrieved from the SINE surveys and described above. Unfortunately, these measures are discrete: EXPGR0 is equal to 1 when the entrepreneur expects development, and zero else. EXPEMP0 is equal to 1 if the entrepreneur expects to hire, and zero else. These expectations are measured in the year of creation/take over (hence in 1994 or 1998). We then compare"development" expectations EXPGR0 with actual sales growth and hiring expectations EXPEMP0 with actual employment changes. Since expectation variables are discrete, we also discretize realizations of sales growth and employment such that initial expectation errors are given by:

$$\Delta_S = EXPGR0 - 1_{(\Delta \ln(\text{SALES})_0 > 3\% \text{ and firm survives})} = b_S + \widetilde{\varepsilon}_S \quad (2)$$
$$\Delta_E = EXPEM0 - 1_{(\Delta(\text{Employment 1994 - 1996})_0 \geqslant 1 \text{ and firm survives})}. = b_E + \widetilde{\varepsilon}_E \quad (3)$$

Clearly, these measures are noisy proxies of the entrepreneur's bias in expectation. First of all because a potentially large part of this error consists in a "rational" error ε that has no reason to be equal to zero. Second, because expectations are discrete, the 3% threshold to define actual sales growth is arbitrary – what do entrepreneurs mean by growing instead of stagnating ? Our choice matches the average consumer price increase of the French economy over the period, and is therefore very conservative: an entrepreneur with a positive expectation error did not overestimate growth if his business's growth was more than zero in real terms. We therefore underestimate the magnitude of optimism in the sample, if "reasonable growth" is understood as being above zero.

Table 4 presents the sample distributions of Δ_S and Δ_E using different thresholds, computed in the year of creation. The top panel computes the distribution of Δ_S and Δ_E for firms started in 1994. The bottom panel focuses on firms started in 1998. Reassuringly, the distributions of errors are similar across years. Columns 1-3 report the distribution of "employment" based expectation errors Δ_E using 0, 1 or 2 new employees as thresholds of net increase in employment. Columns 4-6 look at sales based "development" expectation errors Δ_S using 0%, 10% and 20% sales growth as arbitrary thresholds. As is apparent from columns 4-6, the choice of threshold does not affect the distribution of Δ_S too much.

Whatever threshold we chose, the fraction of entrepreneurs whose expectations exceeded realizations is always above 12%. The fraction of entepreneur whose expectations were below realization is never below 24%. This suggests that there exists a stable group of entrepreneurs for which $b + \widetilde{\varepsilon} > 0$ as well as a stable group of entrepreneurs for which $b + \widetilde{\varepsilon} < 0$. The picture is different for the distribution of Δ_E: the group of entrepreneurs for which ex post realizations of employment increase ended up being smaller than ex ante expectations is not very different as long as the threshold for employment change is 1 or 2 employees (13 versus 18% in 1994, 16 versus 21% in 1998). It drops, however, significantly as soon as we consider that overall employment stability (zero net increase) as the norm. In this case, very few entrepreneurs end up with $b + \widetilde{\varepsilon} > 0$ (5% in 1994, 4% in 1998). This is not too surprising: as mentionned above, the question asked is "do you plan to *hire* in the next twelve months ?". With this very low threshold, expectation can only beat realization when the entrepreneur plans to hire, while total employment ended up decreasing by at least one employee. This event is rare in great part because many (nearly 50%) firms have zero employees in the data in their year of creation. Setting this particular problem aside, it seems that our distribution of expectation errors is reasonably stable across thresholds. We have, however, checked the sensitivity of all our regression results to the threshold chosen. They turn out to be robust, (we report some of these tests in the paper, others are available from the authors upon request).

Obviously, another lesson to be drawn from Table 4 is that many entrepreneurs fail to predict the future correctly. Depending on the variable chosen, and on the threshold, between 30 and 45% of the entrepreneurs end up having realization above or below expectations. This is not too surprising, as even rational individuals make forecast errors in the presence of uncertainty. But this suggests that "rational" expectation errors ε are going to add noise to our measures of bias.

One clear risk at this stage is that the noise ε is too big compared to the actual bias. To show that expectation errors are not "pure noise", we present in Table 5 the correlation between Δ_E and Δ_S. Column 1 regresses Δ_E on Δ_S controlling by industry dummies interacted with year dummies (column 4 does the reverse). Obviously, the coefficient on Δ_S is highly significant, suggesting that both errors capture a common factor. This common factor may, however, be a common bias (b_E and b_S are correlated), or a common rational error (ε_E and ε_S are correlated). In other words, Δ_E and Δ_S could be correlated for two very different reasons. First, entrepreneurs that are optimistic on sales may also tend to be optimistic on hires. Second, entrepreneurs who are surprised by an unexpected boom in their industry also end up hiring more workers than expected.

To tackle this critique, we rely on the panel structure of the SINE survey to look at the persistence of optimism. Using the second questionnaire of each wave, we computed *second period* expectation errors, using expectations formed three years after creation:

$$\begin{aligned}\Delta'_S &= EXPGR1 - 1_{(\Delta \ln(\text{SALES})_1 > 3\% \text{ and firm survives})} = b'_S + \widetilde{\varepsilon}'_S \\ \Delta'_E &= EXPEM1 - 1_{(\Delta(\text{Employment})_1 \geqslant 1 \text{ and firm survives})}. = b'_E + \widetilde{\varepsilon}'_E\end{aligned}$$

Let us start with second period "development" expectation errors Δ'_S. For firms started/taken over in 1994, $EXPGR1$ is the "future development" expectation as formed in 1997. This variable is constructed exactly like $EXPGR0$, because the 1997 question about future development is identical to the 1994 one. $EXPGR1$ is then compared to actual sales growth over 1997-1999. Sales realization is discretized using the 3% threshold. For firms started in 1998, we compare "future development" expectations formed in 2001 to actual sales growth in 2001-2003. Second period "hiring" expectation errors Δ'_E are computed in the exact same way.

The good news is that, by definition, the "rational" errors $\widetilde{\varepsilon}'$ have mean zero *conditionnal* on all information available to the entrepreneur when the second-period expectation is formed. As a result, $\widetilde{\varepsilon}'$ is orthogonal to $\widetilde{\varepsilon}$, i.e. $E(\widetilde{\varepsilon}'|\widetilde{\varepsilon}) = 0$. This suggests that if second period expectation errors Δ' are correlated with first period ones Δ, the correlation has to come through the biases b and b', not through the rational errors $\widetilde{\varepsilon}$ and $\widetilde{\varepsilon}'$.

We report these correlations in Table 5, columns 2-3 and 5-6. In column 2, we regress initial employment expectation errors Δ_E on second period employment expectation errors Δ'_E. We also control for interactions of year and industry dummies. The correlation is statistically very significant. In column 3, we include as regressors both second period expectation errors Δ'_E and Δ'_S. As it turns out, the future development expectation errors has some additional explanatory power on its own, although the statistical significance is less strong. In columns 5 and 6, we repeat this analysis using initial "development" expectation errors Δ_S as the dependent variable. Results reported in these columns confirm that errors are strongly persistent. As "rational" expectation errors ε and ε' should be uncorrelated, this indicates that the bias shows some persistence over time. In other words, results from Table 5 are consistent with entrepreneurs failing to converge to fully rational expectation three years after firm creation. Some of their initial bias seems to remains.

4.2.2 The Determinants of Optimism

The above analysis suggests that some entrepreneurs make consistent errors on their expectations. This section seeks to provide additional evidence of such biases in expectation by looking at entrepreneur characteristics. Our strategy is now to regress expectation errors Δ_S and Δ_E on initial entrepreneur and firm characteristics, that are observable when the expectation is formed. Let X be such a set of characteristics (say, age, education etc). Since X is part of the information set I on which expectations are formed, it must be that:

$$
\begin{aligned}
E(\Delta|X) &= E(b|X) + E(\varepsilon|X) \\
&= E(b|X)
\end{aligned}
$$

Since the "rational" error cannot be predicted by observables, the mean expectation error conditionnal on observables has to be equal to the mean bias. For instance, if the expectation errors of highly educated entrepreneurs are on average positive, we take this as evidence that educated entrepreneurs are more optimistic than less educated ones.

Which observables should we expect to (1) be available from the data and (2) be correlated with a bias in expectation ? First come the entrepreneur's characteristics. Educated entrepreneurs enjoy a larger outside option on the labor market: hence, those who choose to start a firm must have received a better private signal, other things equal; Hence, as outlined in our model in section 2, they are more likely to be optimistic. On the other hand, more educated entrepreneurs may simply be more "rational". In general, psychology theory is ambiguous about possible biases arising from education. First, general education gives entrepreneurs a view on the "big picture" which according to Kahneman and Lovallo [1993] leads to more unbiased expectations. More specific to France and interesting to us, is the highly selective "grande école" system. Provided these successful students suffer from *base rate neglect*, they might overattribute their academic success to their own ability, and end up for this reason overestimating their odds of success as entrepreneurs. We also include gender: using a dataset on positions and trading records for some 35,000 investors, Barber and Odean [2001] show that the turnover rate of common stocks for men is one and half times larger than that of women. They rely on evidence from psychological literature to interpret this difference as evidence that men are more overconfident (i.e. overestimate the precision of their information) than women. Combined with selection (into entrepreneurship), the overweighting of private information translates into greater optimism. Hence, if they are overconfident, male entrepreneurs should also be more optimistic.

We also include various measures of the entrepreneur's experience. We use entrepreneur's age. Experience is likely to increase entrepreneurs' outside options on the labor market. Thus, like education, age may have a positive impact on optimism (see again our model section 2). But it could also be argued that experienced entrepreneurs are likely to observe more precise signals. In this case, optimism should be less prevalent among older entrepreneurs. We also measure the fact that the entrepreneur has already started a firm before. Serial entrepreneurs may, or may not, have been successful in the past. Psychology documents the fact that agents tend to attribute success to their own ability and failures to bad luck (Zuckerman [1979]). The pool of repeat entrepreneurs is therefore likely to exhibit higher optimism than new entrants. Moreover only the most optimistic among entrepreneurs are likely to "try again", a selection effect that reinforces the previous one. Finally, we include in our analysis the entrepreneur's expertise in the industry. In the management literature, Russo and Shoemaker [1992] provide statistical evidence that expertise allows one to "know what one does not know", i.e. to exhibit less optimism in the field of expertise. Many psychologists do, however, argue otherwise. Self declared areas of expertise are those areas where the agent is personally commited the most, and personal commitment is likely to foster optimism (Weinstein [1980]). Slovic, Fischhoff and Lichtenstein [1980] argue that experts tend to be *overconfident*, i.e. they always overestimate the precision of their knownledge, which leads them to underweight outside information (Kahneman and Tversky [1979] recall that experts are also subject to the *planning fallacy*).

Another set of observables consists of project characteristics. First we include a dummy equal to one when the project is a "new idea". As discussed in section 2, when faced with a high level of uncertainty, entrepreneurs are more likely to use heuristics that are biased toward optimism. A large body of literature in cognitive psychology documents that uncertainty fosters in general optimistic expectation. A sizeable management literature confirms it in the case of entrepreneurs (see for instance Busenitz and Barney [1997]). As a second measure of project "novelty", we ask if the firm is a "true" start-up, or whether the entrepreneur bought it/inherited it. Indeed, a little less than half of the sample consists of entrepreneurs taking over an already existing firm. These entrepreneurs are likely to face less uncertainty, because the firm – its customers or at least its assets – already exists. Moreover their selection into entrepreneurship might be more exogenous (e.g. inheriting the business). For these two measures of novelty, evidence from experimental psychology gives a concordant and unambiguous insight: we expect them to be correlated with optimism.

Our last variable is the entrepreneur's motivation to achieve independance. A priori, this can affect optimism in both directions. A desire for independance is likely to magnify the "inside view effect" (the underweighting of external information) and therefore to be correlated with higher optimism. However entrepreneurs who value independance might have a lower subjective outside option in paid employment, which could mitigate the optimism of this category.

4.2.3 Potential Concerns

Our strategy of regressing expectation errors on observables raises some methodological concerns. First, we do not measure actual expectation errors, as Δ_S and Δ_E are discrete. As a result, the average expectation error for each category of entrepreneur can be either positive, or negative, depending on the threshold that we choose on sales growth/employment change. This is why we will not interpret a positive coefficient on education as evidence of skilled entrepreneurs being optimistic, but rather a evidence that skilled entrepreneurs are *more optimistic* than unskilled ones. We do not know which category is biased, but at least one has to be.

A second concern is potentially more serious: since we observe initial expectation errors only for two years (1994 and 1998), we cannot rule out the possibility that aggregate shocks may have stronger effects on some classes of entrepreneurs. As a result, a correlation between an observable and expectation errors may be the result of the particular realization of a shock on one particular group of entrepreneurs. Assume for instance that skilled entrepreneurs are purely rational – $b_i = 0$ – but that they tend to cluster in the software industry. Assume further that the software industry is hit by a negative shock, which was not expected to occur with probability 1. In this case, all skilled entrepreneurs are going to have a large, positive, expectation error and the naive procedure is going to attribute it to skilled entrepreneurs' biases.

A first way of adressing this concern consists of looking at the explanatory power of aggregate and industry shocks. In non reported regressions, we have investigated the explanatory power of industry and year dummies on the expectation errors Δ_E and Δ_S. As it turns out, year dummies alone only explain 0.1% of the variance of both expectation errors. This result is not surprising in light of the strong stability of the distribution of expectation errors across years shown in Table 4. When included, industry dummies raise the R^2 to some 4% for "development" expectation errors, and only to 1% for "hiring" expectation errors. Thus, industry *trends* capture a small part of the variance in expectation errors. One possibility is that entrepreneurs in different industries have different ideas about what "development" means. Finally, we added to the regressions interactions between industry and year dummies. These interactions turn out to add only 0.1% and 0.3% to the R^2. Industry shocks, like aggregate shocks, explain only a tiny fraction of the dispersion in expectation errors. In the following regressions, we do however control for industry $\times$ year dummies.

One other possibility is to look at the stability of the coefficients on observables across years. This exercise is informative because these years are located at two very different points of the business cycle: in 1994, the French economy barely emerges from its worst recession year since the 1970s (GDP growth 2.2%). In constrast, 1998 is a year of strong recovery (economic growth is 3.5%) Assume for instance that the effect of education on expectation errors turns out to be similar in 1994 and 1998. In such a case, this correlation would be spurious if and only if industries with lots of skilled entrepreneurs received a positive shock *both* in 1994 and 1998. In the following, we will therefore report, for all our regressions, results using the pooled 1994/1998 sample, but also separate results for each wave of the survey. We will also test the equality of coefficients across waves. As a further robustness check, we will also compare the effect of observables on initial (1994,1998) expectation errors to their effect on second period (1997,2001) expectation errors. These effects, we argue below, should be similar.

One last concern with our approach is related to our measure of "hiring" expectation error. This measure Δ_E is likely to be noisier than Δ_S because it is not directly related to the *outcome* of the venture, but to the use of inputs. For instance, entrepreneurs may not necessarily need more employees to "grow". As a result, while they may be making optimistic expectations on the development of their business, they may be perfectly realistic in terms of employment. In other words, when employment is a fixed cost, future development is not always tied to new hires. This concern is particularly stringent in a country like France where labor regulation is tight and with newly created firms which often have no employee. The main consequence is that Δ_E is likely to be a noisier estimate of biases than Δ_S, which may make our results weaker. In the main text, we will therefore focus on Δ_S, and refer to results pertaining to "employment" expectation errors Δ_E as robustness checks. They are in general marginally less strong but still significant most of the time.

4.2.4 Regression Results

We are now ready to regress the initial "development" expectation errors on the various observables described above. Estimates are reported in Table 6. Column 1 pools all the observations available in our dataset. Column 2 restricts the sample to firms started/taken over in 1994, while column 3 focuses on the 1998 wave. We use a linear probability model to make results easier to read, though a logit model does not deliver different results. Given our above discussion, we control for industry shocks by interacting 90 two-digit industry dummies with year dummies. The estimation of standard errors allows for broad form of correlation of error terms across firms of the same industry $\times$ year. Column 4 presents tests of the null hypothesis that coefficients are identicial across creation years. These tests are obtained by regressing Δ_S on observables interacted with year-of-creation dummies, and testing the null that these interactions are equal to zero. For each explanatory variable, we report the p-value of the test in column 4.

Before we turn to the effect of each explanatory variable, two general remarks are in order. First, our observables have a low explanatory power ($R^2 = 0.07$ in the pooled regression). Notice however that there is no reason for us to expect a high R^2, because expectation errors include both a potential bias and a "rational" expectation error ($\Delta = b + \varepsilon$). When the degree of ex ante idiosyncratic uncertainty faced by agents is large, the variance of ε can be a large fraction of the overall variance of Δ. As a result, even if we could perfectly predict b with our model and observables, the R^2 of the regressions of $b + \varepsilon$ on b could be mechanistically low.

Second, entrepreneur characteristics come out jointly and separately significant. We interpret this as evidence of different biases across classes of entrepreneurs. This interpretation rests, however, on the assumption that aggregate shocks do affect all types of entrepreneurs to a similar extent. Fortunately, estimation results are also very consistent *across* years of creation.[13] A student test does not reject equality of coefficients for 9 out of 10 explanatory variables. 6 out of 10 explanatory variables come out with a strong statistical significance (1%) for both years. We interpret such evidence as showing that these variables are not highly correlated with industry-wide shocks, and that they really describe the heterogeneity of beliefs. We ran, but do not report, identical regressions using expectation errors on "employment" as the dependent variable: again, 9 out of ten explanatory variable have the same coefficient across years. These estimates are, however, slightly less strong statistically: only 2 variables ("real start-up" and "new idea") out of 10 come out with strong statistical significance (1%) in both 1994 and 1998. In the pooled regression, 7 out of 10 variables turn out to be strongly significant, with the same signs as for "development" expectation errors.[14]

By and large, across definitions of errors and across years of creation, the education variables and the novelty of the project (real start-up, implementing a new idea) have the strongest effects. In all specifications, education seems to be positively correlated with high expectations when compared to realizations. Economically, the effect is not very large, but still worth considering: the *grande école* coefficient is approximately one tenth of the sample standard deviation of the expectation error (0.7). This effect is consistent with education giving self-confidence and endowing potential entrepreneurs with better outside options, thus compelling them to choose the project when their subjective evaluation is higher. The other very robust results concerns the novelty

[13] We also ran, in unreported results, separate regressions for small and large firms, corporations and sole proprietorships, and start-ups and nonstart-ups. Results were very consistent with the ones we present in table 6, so we chose not to present them to save space.

[14] This table is available from the authors upon request.

of the project. Novelty props up expectations. Entrepreneurs implementing their own new idea tend to systematically overestimate their growth prospects. Entrepreneurs with some experience in the industry tend to be more realistic. Both coefficients hover around one tenth of standard deviation of the expectation error. The coefficient on "real start-up" also is positive, strongly significant, but twice as large. All in all, entrepreneurial optimism arises in situations of high uncertainty, as cognitive psychology predicts. All these results also arise when we use "hiring" expectation errors as the dependent variable.

In constrast, serial entrepreneurs are consistently more optimistic (the result is less strong with "hiring" expectations). The size of the effect is again roughly one tenth of the standard deviation of expectation error. It is consistent with serial entrepreneurs not updating rationally: they discard their failures as "bad luck", and attribute their successes to themselves. Those who are most able to think this way self select into serial entrepreneurship. Last result, entrepreneurs motivated by autonomy tend, in general, to be optimistic (effect located in 1994 only, but significant in 1998 with "hiring" expectation errors).

As a further robustness check, we use second period expectation errors Δ'_E and Δ'_S. Theoretically, second period expectation errors should, again, be the sum of two components: a second period bias b'_S, and a second period "rational" error ε'_S. By definition, ε'_S is uncorrelated with entrepreneur and project observables at year of creation - or even three years after creation, when second period expectations are formed. The second period bias b'_S is likely to be positively correlated with the initial bias b_S. Indeed, even if the entrepreneur updates his beliefs using Bayes' rule, an initially more optimistic entrepreneur is more likely to be optimistic three years after. Thus, rational (bayesian) updating is likely to reduce the initial bias, but will not destroy it.

As a result, regressing second period expectation errors on year-of-creation observables should generate estimates that are noisier than, but comparable to those from Table 6. We present such estimates in Table 7 for second period "development" expectation errors. The estimation strategy and regressors are the same as in Table 6. To make reading easier, column 1 reports the regression using *first* period expectation error (as in Table 6, column 1), while column 2 uses *second* period expectation error on "future development". Column 3 reports the p value of t-tests of equality for each coefficient (this statistic is obtained by estimating the two equations of columns 1 and 2 as simultaneous equations).

For the second period error, all estimates are indeed noisier than with initial expectation errors, but go in the same direction and are often statistically significant. Variables measuring the "novelty" of the project (implementing a new idea, the firm being a real start-up) still come out significant, but somewhat smaller than in column 1. This also true for variables related to education: high school graduates still seem to be systematically upward biased, even three year after creation. The effect of college graduates becomes less significant, but the coefficient is as large as for high school graduates. For both initial and second period expectations, expertise of the industry significantly reduces - and to the same extent for both periods - the chance that expectations exceed realizations. All in all, observables that predict a bias in expectation at year of creation still predict a bias three years after creation, although the bias seems smaller. This is consistent with certain classes of optimistic entrepreneurs updating their beliefs, but remaining optimistic.[15]

One last concern is related to the fact that our expectation variables were qualitative. Therefore, we had to make assumptions about what entrepreneurs meant when they said they expected their venture to "develop". In all our tables, we assume 3% growth in sales is the threshold below which a firm cannot be meant to "develop", and that entrepreneurs expecting development without reaching that threshold could safely be assumed to make a positive expectation error. Given the arbitrariness of this threshold, however, we need to check the sensitivity of our results. Notice, however, that we saw in Table 4 that the distribution of expectation errors is fairly stable across thresholds.

In Table 8, we report OLS estimates using three different thresholds for "development". They appear extremely stable. The observables, sample and estimation method are identical to Table 6. Column 1 uses 0% as the value of sales growth to be compared to expectation of "development". Column 2 uses 10%, and column 3 uses 20% as the threshold. The coefficient estimates who were statistically significant in Table 6 remain so with all three definitions. Their values are very stable across regressions, and do not differ statistically. The only exception is the coefficient on "Real start-up", which tends to decrease as we move the threshold up. It is, however, always strongly statistically significant, and still large, even with the 20% cutoff.[16]

[15] This overall diagnosis is confirmed when we use second period "hiring" expectation errors as the dependent variable (results not reported to save space). Estimates are, however, much less significant. One possible reason for this is that "hiring" errors are less well correlated intertemporally (correlation coefficient of 0.06) than "development" errors (correlation coefficient of 0.14). As we mentioned earlier, entrepreneurs might well expect development without new hires. This discrepancy is likely to be larger after three years of existence.

[16] We perform a similar robustness check on "employment" expectation errors. Estimates still appear robust and significant, although they tend to vary slightly more. One possible explanation is that the question asked is slightly more precise, since the entrepreneur is asked if he "plans to hire". One other possibility, confirmed by our other results on this measure, is that it is simply less reliable as a measure of optimism.

4.3 Optimism and Short Term Debt

We are now set to test the relation between optimism and the use of short term debt that is the main prediction of our model. We start with a naive assessment of the correlation between the bias in expectation and the use of short term debt.

4.3.1 OLS Evidence

As can be seen from equations (2) and (3), expectation errors Δ_S and Δ_E are noisy measures of the biases b_S and b_E; in this setting, the difference between unbiased expectations and realizations can be considered as measurement error. A simple approach is therefore to regress our measures of short term debt on Δ_S and Δ_E:

$$STD_i = \alpha + \beta.\Delta_i + Z_i.\delta + \upsilon_i \tag{4}$$

where Z_i include standard determinants of the use of short term debt.

Plain OLS estimates of (4) may be biased toward underestimating or overestimating the coefficient β. First, as we just said, expectation errors Δ_i are noisy measures of the expectation biases. Measurement error is likely to bias our results toward zero provided the "rational" error ε_i is uncorrelated with equation (4)'s residual υ_i.

But there are reasons to think that ε_i and υ_i may be positively correlated. Assume for instance that bankers only lend short term to some classes of entrepreneurs. These firms have high υ_i. Assume further that a negative shock hits all firms. One possibility is that firms with short term debt on their balance sheet will suffer the most. Their owners' expectations errors ε_i will mechanistically turn out to be large. This generate a positive correlation between ε_i and υ_i. But this story works even if short term debt does not directly affect performance. Assume for instance that bankers lend short term to risky firms (high υ_i), and that the shock turns out to be negative. Then high υ_i firms are the most exposed and are going underperform (high ε_i).

These plausible mechanisms generate a positive correlation between capital structure and our measure of optimism, for reasons that have nothing to do with our theory. If firm exposure to (ex post) negative shocks was observable, however, we could include it in the Z and control for it. In a first step, we will content ourselves with doing just this. In the next section, we will propose an identifying strategy to deal as efficiently as possible with this problem.

Which controls Z_i do we choose? Analyzing the determinants of debt maturity among listed US corporations, Barclay and Smith [1995] argue that firms with higher growth prospects and less collateral should make more use of short-term debt. To control for collateral, we use the year-of-creation share of fixed assets in total assets. To control for growth opportunities, we include a dummy equal to one when the firm is a real start-up, as well as 2-digit industry fixed effects. We also interact industry fixed effects with year-of-creation dummies to account for potential changes in the yield curve across years, and their possibly different effects across sectors. Last, Barclay and Smith [1995] also include firm size as a control for firm quality. We therefore add the log of initial total assets to our list of controls.

Table 9 gathers all results using the year-of-creation share of lines of credit in bank debt as a dependent variable, and "development" expectation error as our measure of bias. In all regressions, we allow error terms to be correlated across observations within each industry $\times$ year group. Estimates of these regressions using "hiring" expectation errors are not reported, but delivered very similar results to Table 9. We also checked the robustness of the results we present to the threshold on sales growth chosen to compute expectation errors. They are pretty much insensitive to the threshold chosen, as long as it is between 0 and 20%.

Regression in Table 9, column 1, is the baseline specfication. As can be noted, compared to regressions presented in Table 6, the number of observations drops dramatically because initial capital structure is available only for a subset of firms (those large enough at the end of the creation year to fill in detailed tax forms). In line with results from Barclay and Smith [1995], larger firms use less short-term debt, as do projects with more tangible assets. Real start-up use more short term debt. More importantly to us, the expectation error about future performance is strongly positively correlated with use of short term debt. The coefficient is significant and stable across specifications, but hovers around 0.03, which is economically small. Given that the standard deviation is 0.6 on expectation errors and 0.4 on shares of short term debt, approximately one twentieth of the variation in short-term debt is explained by optimism. This, however, could be due to measurement error of the bias. The identification strategy we develop below will allow us to deal with this question.

What can we do about the possible endogeneity biases that have discussed above ? Column 2 and 3 provide separate estimates for 1994 and 1998, assuming that shocks are likely to differ in these two years. The coefficient changes somewhat, being equal to 0.04 in 1994 and 0.03 in 1998. They do not differ statistically, but the estimate is slightly weaker in 1998.

In column 4, we try to control directly for firm specific risk by including a dummy equal to 1 when the firm disappears before its third birthday.[17] Our

[17] We also ran regressions using firm profitability after two years (return on assets) as a measure of subsequent performance. Results were not very different. We chose not to report them because banks are likely to focus more on risk rather than overall profitability when

working hypothesis is that this "death" dummy is likely to include information about firm risk that is known to the banker, but not to the econometrician at the date of creation. This seems to be the case, since the coefficient on future death is highly significant and positive on short term debt. This suggests that either bankers have more information about the future of the firm than econometricians, or that short term debt exposes firms to liquidity crises and causes bankruptcy. Either way, this variable thus controls for at least part of firm risk. Yet, its inclusion does not affect the correlation between short term debt and expectation error. The coefficient remains unchanged at 0.03. This is comforting, but still insufficient since firm risk may not be *totally* controlled for with this approach.

One last possibility is to include as many observables Z as possible, to control for firm specific risk and more generally the firm's propensity to be lent short term. In column 5, we added as additional controls in equation (4) all entrepreneur and project specific variables used in Table 6 to explain optimism. Regressions results are displayed in Table 9, column 3. As it turns out, the coefficient on development expectation error remains statistically significant and is not affected by the inclusion of our controls. What remains unclear, though, is whether these added controls stand for their own direct effects on debt maturity or for their effects through optimism (we saw for instance that education or novelty could trigger overoptimism).

All in all, these first results are encouraging and suggest that our upward bias may not be too large. This, however, remains to be confirmed by a proper instrumenting strategy, to which we now turn our attention.

4.3.2 Optimism and Short Term Debt: An Instrumental Variable Approach

We propose in this section a strategy to instrument b, i.e. the *bias* in expectations that entrepreneurs have. Such an instrument would have (1) to be correlated with b, and (2) not to be correlated with unobservables affecting short term debt v. We propose here to use *second period* expectation errors Δ' as instruments for b. But before looking at the results, let us discuss whether this variable satisfies the conditions under which it is a valid instrument.

First, there are strong reasons to think that second period expectation errors are good predictors of initial expectation errors. We already know from Table 5 that second period expectation errors Δ'_E and Δ'_S and initial expectation errors are correlated with a high level of statistical significance. For the "employment" expectation, the t-statistic is 4.7 (Table 5, column 2); for "development" expectations, it is as high as 13.6 (from Table 5, column 5).

making their lending decisions.

Such evidence is confirmed by the results from Table 7, where it is apparent that observables correlated with initial expectation errors tend also to be correlated with second period expectation errors. What we need to check at this stage is whether the correlation between initial and second expectation error is still strong enough for the subsample of firms for which capital structure is available, and whether it remains so once we control for the determinants of short term debt. We provide such a formal test below.

The second condition is an *identifying* assumption that cannot be tested: $\Delta' = b' + \varepsilon'$ should not be correlated with the residual υ. υ represents the information, available to the entrepreneur and his banker, but not to the econometrician, used to determine the level of short term debt that the firm takes on, *in the first year* of its existence. The second period expectation error Δ' is the difference between expectation of future growth formed *three years after creation* and its realization. We make here the assumption that this future error does not reflect any information that the two parties use when they sign the debt contract. What can be taken for granted is that the future "rational" error ε' has to be independent from all information available three years after creation, in particular year-of-creation capital structure and therefore υ. The more debatable assumption is to posit that the future bias b' does not contain any information related to initial capital structure, apart from its predictive power on the initial bias b. Given the three year lag between future expectations and initial capital structure, this is a leap of faith we are ready to make.

We focus here on "development" expectation errors, which so far has provided us with the strongest and most stable estimates.[18] Three regressions are reported in Table 10. Column 1 presents the reduced form estimate: the direct regression of initial capital structure on second period expectation errors, controlling for observable determinants of debt maturity. These controls are the same as in Table 9, column 1. Given that we need to compute second period expectation errors, we need to restrict ourselves to firms still alive five years after creation, which reduces the sample somewhat. As column 1 shows, initial capital structure and future expectation errors are positively and significantly correlated. The coefficient is 0.03, identical to the coefficient of initial expectation error reported in Table , column 1. It is statistically significant at 1.8%. The correlation between capital structure and expectation error therefore seems fairly stable.

We then move to the instrumenting strategy per se. Table 10, column 2

[18] For the subsample of firms for which capital structure is available, second period "employment" expectations errors turn out to be weak instruments of initial expectation errors on "hiring". This is not surprising: the sample correlation between the two variables was already weaker (0.05) than for "development" expectation errors (0.14). So a proper instrumenting strategy is not valid for this measure of expectation errors.

looks at the first stage: we regress initial expectation errors on future expectation errors. There are two differences with the results reported in Table 5. First, the sample is restricted to firms for which capital structure information is available. Second, we include as controls the determinants of short term debt that will be used in the second stage regression (tangible assets, industry $\times$ year dummies, size, real start-up). The coefficient on future error turns out to be 0.10, very similar to the coefficient obtained in Table 3, with a much larger sample and no controls. Economically, the size of the coefficient is not huge: a one s.d. increase in future error triggers a 0.1 s.d. increase in initial error. Yet, the correlation is statistically very significant, with a t-statistic of 4.9, and a F statistic of 23.9. This means we can safely assume that the our instrument is not weak, i.e. that the bias under TSLS is small compared to the bias under OLS (see for instance Bound, Jaeger and Baker, 1995).

Table 10, column 3 presents the second stage estimation. There, initial expectation errors are instrumented using second period expectation errors. The estimated coefficient is now equal to 0.28 and statistically significant at 5%. It is ten times as large as the OLS estimate. This suggests that our OLS estimates were downward biased because of measurement error, not upward biased because of systematic shocks being correlated with capital structure and expectation errors. Remember that measurement error is likely to be fairly large in our context given that, in addition to his bias, entrepreneurs may be making "rational" (i.e. unbiased) errors which prevent us from accurately measuring biases from expectation errors. This, and the fact that our instrument does not seem to suffer from statistical weakness, suggest that we can take the new 0.28 coefficient seriously. The magnitude of the effect of bias on capital structure becomes large. A one-standard-deviation increase in expectation error (+0.7) leads to an approximate increase of credit lines by 20 percentage points of total bank debt, some 50% of its sample standard deviation. Quantitatively, optimistic beliefs seem to have a large effect on the willingness to take on short term debt.

5 Conclusion

This paper argues that differences in beliefs exist, have real effects, and therefore do matter in the design of financial contracts. We test a simple model of contracting with optimistic entrepreneurs with data on entrepreneurial expectations and outcomes. We show that there is substantial heterogeneity in beliefs in the data and that this heterogeneity can be partly explained by sociodemographic and psychological characteristics of the entrepreneur. We then establish a positive, robust correlation between optimism and the use of short term debt, using an instrumental variables identification strategy. This correlation is consistent with the main prediction of our financial contracting model.

6 References

Aghion, Philippe and Bolton, Patrick [1992], "An Incomplete Contracts Approach to Financial Contracting", *Review of Economic Studies*, vol 59, N°3, pp 473-494

Armor, D. and Taylor, S.E. [2000], "When Prediction Fail: The Dilemna of Unrealistic Optimism", in Gilovich, Griffin and Kahneman Eds. *Heurstics and Biases: The Psychology of Intuitive Judgement.*

Barberis, N. and R. Thaler [2002], "A survey of Behavioral Finance", NBER working paper N°9222

Barber, Brad, and Odean, T. [2001], "Boys Will be Boys: Gender, Overconfidence and Common Stock Investment", *Quarterly Journal of Economics,* vol , N°1, pp 261-293

Barclay, Michael and Smith, C. [1995], "The Maturity Structure of Corporate Debt", *The Journal of Finance*, Vol 50, N°2, pp 609-631

Bertrand, Marianne and Mullainathan, Sendhil. [2001] "Are CEOs Rewarded for Luck? The Ones Without Principals Are." *Quarterly Journal of Economics*, vol 116, n°3, pp. 901-32

Bound, John, Jaeger, D. and Baker, R. [1995], "Problems With Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak", *Journal of the American Statistical Association*, Vol 90, n°430, pp 443-450

Busenitz, Lowell, and Barney, J. [1997], "Differences Between Entrepreneurs and Managers in Large Organization: Biases and Heuristics in Strategic Decision Making", *Journal of Business Venturing*, vol 12, pp 9-30

Camerer, Colin and Lovallo, D. [1999], "Optimism and Excess Entry: An Experimental Approach", *American Economic Review*, Vol 89, N°1, pp 306-318

Coval, Joshua D., and Anjan Thakor,. "Financial Intermediation as a Beliefs-Bridge Between Optimists and Pessimists", mimeo, HBS.

De Meza, David, and Southey, C. [1996], "The Borrower's Curse: Optimism, Finance and Entrepreneurship", *The Economic Journal*, Vol. 106, N° 435, pp. 375-386

Gettys, Charles, Kelly III, C, Peterson, C. [1973], "The Best Guess Hypothesis on Multistage Inference", *Organizational Behavior and Human Per-*

formance, Vol 10, pp 364-373.

Hackbarth, D. [2004], "Managerial Traits and Capital Structure Decision", mimeo Washington University.

Hamilton, Barton H. [2000], "Does Entrepreneurship Pay? An empirical Analysis of the Returns to Self-Employment", *Journal of Political Economy*, vol. 108, pp. 604-630

Heaton, J.B., "Managerial Optimism and Corporate Finance", *Financial Management*, Summer 2002, pp33-45.

Kahneman, Daniel, and Lovallo, D. [1993], "Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking", *Management Science*, Vol 39, N°1, pp 17-31

Kahneman, Daniel and Tversky, A. [1973], "On The Psychology of Prediction", *Psychological Review*, Vol 80, pp 237-251

Kahneman, Daniel and Tversky, A. [1979], "Intuitive Predictions: Biases and Corrective Procedures", in S. Makrdakis and S.C. Wheelwright Eds, TIMS Studies in Management Science, Vol 12, pp 313-327.

Kaplan, Steve and Stromberg, Per, [2003] "Financial Contracting Theory Meets the Real World: An Empirical Analysis of Venture Capital Contracts", *the Review of Economic Studies,* .

Kaplan, Steve and Stromberg, Per [2002], "Characteristics, Contracts, and Actions: Evidence from Venture Capitalist Analyses," NBER Working Papers N°8764

Malmendier, Ulrike and Tate, Geoffrey, [2002], "CEO optimism and Corporate Investment", mimeo Wharton, Stanford

Malmendier, Ulrike and Tate, Geoffrey, [2003],"Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction",mimeo Wharton, Stanford

Manove, Michael, and Padilla, J. [1999], "Banking (Conservatively) With Optimists", *Rand Journal of Economics,* Vol 30, N°2, pp 324-350

Moskovitz, Tobias and Vissing-Jorgensen, A. [2002], "The Private Equity Puzzle", *Americain Economic Review*

Pinfold, John [2001], "The Expectations of New Business Founders", *Journal of Small Business Management*, Vol 39, N°3, pp 279-285

Rajan, Raghuram [1992], "Insiders and Outsiders: The Choice Between Informed and Arm's Length Debt", *Journal of Finance*, Vol 47, N°4, pp 1367-1400

Roll, R. [1986], "The Hubris Hypothesis of Corporate Takeovers", *The Journal of Business*, Vol. 59, No. 2, Part 1., pp. 197-216.

Ross, Lee and Anderson, Craig [1977], "Shortcomings in the Attribution Process: On The Origins and Maintenance of Erroneous Social Assessments", *Advances in Experimental Psychology*

Russo, Edward, and Shoemaker, P., [1992] "Managing optimism", *Sloan Management Review*, Winter

Scarpetta, Stefano, Hemmings, P., Tressel, T. and Woo, J. [2002], "The Role of Policy and Institutions for Productivity and Firm Dynamics: Evidence From Micro and Industry Data", OECD working paper N°329

Shane, Scott and Venkataraman, S., [2000] "The Promise of Entrepreneurship as a Field of Research", *Academy of Management Review*, vol 25, N°1, pp 217-226

Simon, Mark, Houghton, S. and Aquino, K. [1999], "Cognitive Biases, Risk Perception and Venture Formation: How Individuals Decide to Start Companies", *Journal of Business Venturing*, Vol 15, pp 113-134

Slovic, Paul, Fischhoff, B. and Lichtenstein, S. [1980], "Facts Versus Fears: Understanding Perceived Risk", in R. Schwing and W. Albers Eds *Societal Risk Assessment: How Safe is Safe Enough ?* New York: Plenum Press.

Souleles, Nicholas [2001], "Consumer Sentiment: Its Rationality and Usefulness in Forecasting Expenditure - Evidence from the Michigan Micro Data", NBER WP N°8410.

Thaler, Richard [1988], "Anomalies: The Winner's Curse", *The Journal of Economic Perspectives,* Vol. 2, No. 1., pp. 191-202.

Tversky, Amos, and Kahneman, D. [1974], "Availability: A Heuristic for Judging Frequency and Probability", *Cognitive Psychology*, Vol 4, pp 207-232

Van den Steen, Eric [2004], "Rational Overoptimism (and Other Biases)", American Economic Review, september 2004.

Weinstein, N. [1980], "Unrealistic Optimism About Future Life Events", *Journal of Personality and Social Psyochology*, vol 39, pp 806-820

Zuckerman, M. [1979]: "Attibution of success and failure revisited, or: The motivational bias is alive and well in attribution theory," *Journal of Personality*, 47, 245– 287.

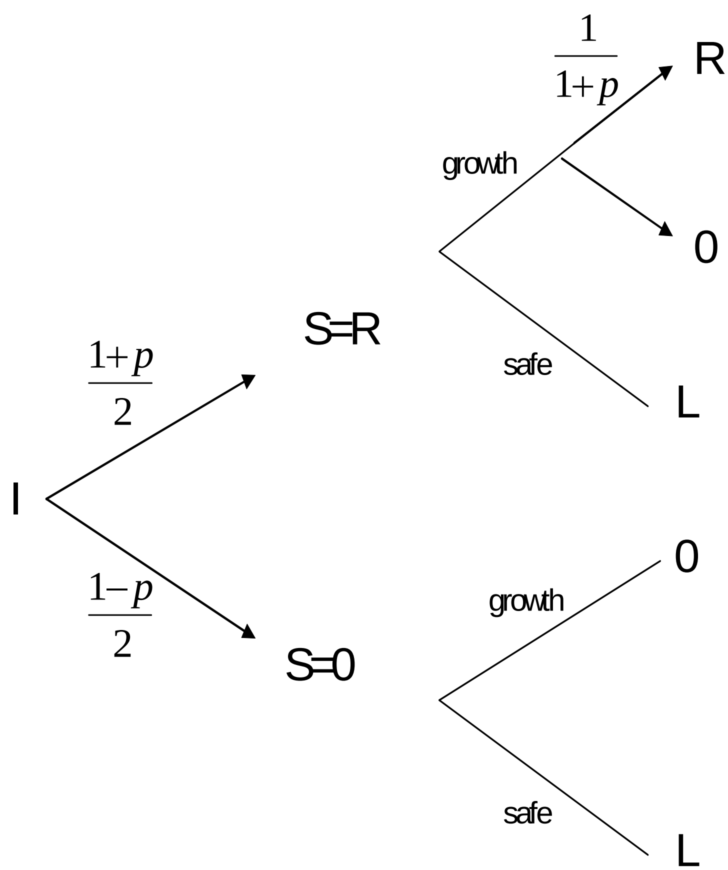

Figure 1: The Business Plan as seen by a Realist

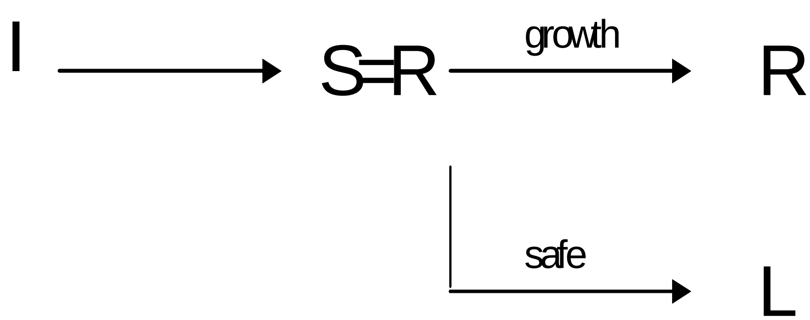

Figure2: The Business Plan as seen by an Optimist

Table 1: Summary Statistics on Expectations

	1994 Survey	1998 Survey
Plans to hire within a year	0.26	0.31
Expects "development"	0.54	0.58
Expects "difficulties"	0.06	0.06
Observations	19,069	11,794

Source: 1994 and 1998 SINE surveys. The first row reports the sample means of a dummy equal to 1 if the entrepreneur replies "Yes" to the question "do you plan to hire within the next 12 months ?". The second row reports the sample means of a dummy equal to 1 if the entrepreneur replies "the firm will develop" to the question "what is your view of the future ?". The third row reports the sample mean of a dummy equal to 1 when the entrepreneur answers "I will have to struggle with a difficult situation", or "I will have to shut down the firm" to the question "what is your view of the future ?". Column 1 focuses on firms created in 1994. Column 2 focuses on firms created in 1998.

Table 2: Summary Statistics on Entrepreneur Characteristics

	Non Corporation	Corporation	Small	Big
Has already started one business	0.02	0.15	0.07	0.13
Experience in the industry	0.59	0.56	0.55	0.65
Motive: Desire to implement own idea	0.09	0.22	0.15	0.16
Motive: Desire for autonomy	0.58	0.48	0.54	0.50
Entrepreneurs in family	0.44	0.45	0.45	0.44
High school graduate	0.14	0.20	0.17	0.18
College graduate	0.08	0.17	0.11	0.14
Post graduate studies or "Grandes Ecoles"	0.03	0.14	0.08	0.10
Age (years)	35	39	36	38
Male Entrepreneur	0.75	0.77	0.75	0.78
Observations	10,929	10,493	16,360	5,063

Source: 1994 and 1998 SINE surveys. Most variables are dummies, so that the reported means stand for percentage in the category. The only exception is age. Columns 1 and 2 split the sample into firms that are sole proprietorships and corporations (whose owners enjoy formal limited liability). Columns 3 and 4 split the sample into firms with at most one employee at the year of creation, and firms with at least two employees.

Table 3: Size and Capital Structure of New Firms in 1994 and 1998

	Firm really created		Firm changing hands	
	Sole Prop.	Corp.	Sole Prop.	Corp.
Employment (employees)	0.4	2.5	1.0	5.4
Fixed assets (000 euros)	17	52	71	85
Total Sales (000 euros)	117	244	142	484
Observations	11,007	12,179	8,181	7,415
Equity / (debt + equity)	0.69	0.63	0.52	0.66
Observations	4,639	12,083	10,828	5,181
Short term loans / Bank loans	0.45	0.46	0.27	0.39
Credit lines / Bank loans	0.34	0.36	0.19	0.31
Observations	250	2,750	536	2,305

Source: 1994 and 1998 SINE surveys and tax files. Size indicators and capital structure are measured at the year of firm creation. We restricted ourselves to firms that were first present in the tax file during the survey year (hence 1994 for the first wave, and 1998 for the second one). There are fewer observations for the detailed capital structure because the tax files do not report detailed financing for small businesses (with sales below 230,000 euros). "Corporations" corresponds to firms whose owner enjoys formal limited liability.

Table 4: The Distribution of Expectation Errors When the Threshold Varies

	Expectation Error on					
	Employment			"Development"		
	(1)	(2)	(3)	(1)	(2)	(3)
For firms created in 1994, share of entrepreneurs for which (%):						
Realization exceeds expectations	62	14	5	24	28	34
Realization matches expectations	33	72	77	56	55	52
Realization are below expectations	5	13	18	20	17	13
For firms created in 1998, share of entrepreneurs for which (%):						
Realization exceeds expectations	61	14	5	22	27	36
Realization matches expectations	34	69	73	54	53	53
Realization are below expectations	4	16	21	23	20	12
Threshold (employ. change, sales growth)	0	1	2	0%	10%	20%

Source: 1994 and 1998 SINE surveys and Tax Files. This table presents the distribution of initial expectation errors, i.e. differences between initial expectations and subsequent realizations. Columns 1-3 focus on "hiring" exp. errors. Expectation is then equal to 1 when the entrepreneur answers "Yes" to the question "do you plan to hire over the next 12 months ?". Realization is equal to one if, over its first two years of existence, firm labor force increases by at least 0 employee (col. 1), at least 1 employee (col. 2) or at least 2 employees (col. 3). Columns 4-6 look at "development" exp. errors. Expectation is then equal to 1 when the entrepreneur answers "the firm will develop" to the question "what is your view of the future ?". Realization is equal to one if, over its first two years of existence, firm sales grow by more than 0% % (col. 4), by more than 10% (col. 5) or by more than 20% (col 6.). The top panel focuses on firms created/taken over in 1994. The bottom panel focuses on the sample of firms created/taken over in 1998.

Table 5: Correlations between Various Measures of Optimism

	Expectation Error on					
	Employment			"Development"		
	(1)	(2)	(3)	(1)	(2)	(3)
Employment expectation error	-	-	-	0.291***	-	-
				(0.012)		
Development expectation error	0.197***	-	-	-	-	-
	(0.015)					
Employment expectation error	-	0.056***	0.047***	-	-	0.025***
(3 years after creation)		(0.012)	(0.014)			(0.009)
Development expectation error	-	-	0.020**	-	0.106***	0.101***
(3 years after creation)			(0.011)		(0.008)	(0.008)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year of creation FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry x Year of creation FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.06	0.02	0.02	0.10	0.07	0.07
Observations	36,147	13,978	10,891	36,147	10,743	10,743

Source: 1994 and 1998 SINE surveys and Tax Files. This table reports the correlation between the various measures of expectation errors. In columns 1-3, the dependent variable is the initial "employment" expectation error, i.e. the difference between initial hiring expectation and realization. Initial hiring expectation is then equal to 1 when, at the end of the first year of the firm's existence, the entrepreneur answers "Yes" to the question "do you plan to hire over the next 12 months ?". Realization is equal to one if, over its first two years of existence, firm labor force increases by at least 1 employee. Column 1 reports the OLS regression of this variable on the initial "development" expectation error (see below) as well as industry dummies interacted with year-of-creation dummies. Column 2 regresses this same variable on the "second period" employment expectation error, which is computed using the same method, but on the basis of expectations and realizations over the two years following the third year of the firm's existence. Column 3 further adds the "second period" "development" expectation. Columns 4-6 use the "development" exp. errors as the dependent variable. Expectation is then equal to 1 when the entrepreneur answers "the firm will develop" to the question "what is your view of the future ?". Realization is equal to one if, over its first two years of existence, firm sales grow by more than 3%. Error terms are assumed to be correlated for firms created/taken over the same year in the same industry. * means significant at 10%, ** at 5% and *** at 1%.

Table 6: Explanatory Power of Observables on Initial Expectation Errors

	Expectation Error on "Development"			
	All	1994	1998	P-value equality
High school graduate	0.07***	0.06***	0.07***	.62
	(0.01)	(0.01)	(0.01)	
College graduate	0.11***	0.10***	0.12***	.54
	(0.02)	(0.02)	(0.03)	
Grande Ecole graduate	0.12***	0.10***	0.14***	.41
	(0.02)	(0.04)	(0.03)	
Age > 38 years	-0.02	-0.00	-0.03***	.13
	(0.01)	(0.02)	(0.01)	
Entrepreneur is male	-0.02*	-0.04**	-0.01*	.27
	(0.01)	(0.02)	(0.02)	
Serial entrepreneur	0.08***	0.08***	0.07***	.64
	(0.01)	(0.01)	(0.02)	
Experience in industry	-0.05***	-0.06***	-0.04***	.25
	(0.01)	(0.01)	(0.01)	
Motive: new idea	0.08***	0.07***	0.09***	.48
	(0.01)	(0.03)	(0.01)	
Motive: autonomy	0.03***	0.05***	0.01	.02**
	(0.01)	(0.01)	(0.01)	
Real start-up	0.18***	0.19***	0.16***	.29
	(0.01)	(0.02)	(0.02)	
Year of creation x Industry FE	Yes	Yes	Yes	
R^2	0.07	0.06	0.08	
Observations	31,832	14,415	17,417	

Source: 1994 and 1998 SINE surveys and Tax Files. This table investigates the explanatory power of entrepreneur and project observables on the dispersion of intial expectation errors. All of the explanatory variables are dummies. All regressions control for industry dummies interacted with year-of-creation dummies, to control for industry shocks. Initial expectation errors are computed in the year the firm is created/taken over using a 3% threshold in sales growth (see Table 5 for description). Column 1 looks at the whole sample of firms started/taken over in 1994 and 1998. Column 2 restrict the sample to firms started/taken over in 1994; column 3 focuses on 1998. Column 4 tests the equality of coefficients on each observables across the two sub-samples. The approach there is to run an OLS regression on the whole sample, including as regressors the entrepreneur/project observables interacted with year-of-creation dummies. Column 4 reports the t-probabilities of a student test that these coefficient are equal to zero. In all specifications, error terms are assumed to be correlated for firms created/taken over the same year in the same industry. * means significant at 10%, ** at 5% and *** at 1%.

Table 7: Explanatory Power of Observables on Second Period Expectation Errors

	Expectation Error on "Development"		
	At year of creation	3 yrs after creation	p value equality
High school graduate	0.07***	0.05***	.46
	(0.01)	(0.02)	
College graduate	0.11***	0.05*	.02**
	(0.02)	(0.03)	
Grande Ecole graduate	0.12***	0.01	.01***
	(0.02)	(0.03)	
Age > 38 years	-0.02	-0.00	.37
	(0.01)	(0.01)	
Entrepreneur is male	-0.02*	0.02	.02**
	(0.01)	(0.02)	
Serial entrepreneur	0.08***	0.02	.00***
	(0.01)	(0.02)	
Experience in industry	-0.05**	-0.05***	.97
	(0.01)	(0.02)	
Motive: new idea	0.08***	0.04**	.07*
	(0.01)	(0.02)	
Motive: autonomy	0.03***	0.01	.08*
	(0.01)	(0.01)	
Real start-up	0.18***	0.11***	.00***
	(0.02)	(0.02)	
Industry x year of creation FE	Yes	Yes	
Observations	32,263	9,444	

Source: 1994 and 1998 SINE surveys and Tax Files. This table investigates the explanatory power of entrepreneur and project observables on the dispersion of second period expectation errors. All regressions control for industry dummies interacted with year-of-creation dummies, to control for industry shocks. Second period expectation errors are computed three years after the firm is created/taken over (see Table 5 for description). Column 1 reports the regression estimte of initial expectation error on entrepreneur/project observables and is identical to Table 6, col. 1. Column 2 reports the same regression using second period, instead of initial, expectation errors as the dependent variable. Column 3 tests the equality of coefficients on each observables for each dependent variable. The approach there is to run the two regressions of columns 1 and 2 as a system of seemingly unrelated equations - thus allowing for error terms in both equations to be correlated within a given firm. Column 3 reports the tests that the coefficients on each observable are identical in both equations. In all specifications, error terms are assumed to be correlated for firms created/taken over the same year in the same industry. * means significant at 10%, ** at 5% and *** at 1%.

Table 8: Effect of Observables on Initial Expectation Error: Robustness to the Threshold

	Expectation error on "Development"		
	(1)	(2)	(3)
High school graduate	0.06***	0.06***	0.05***
	(0.01)	(0.01)	(0.01)
College graduate	0.11***	0.11***	0.11***
	(0.02)	(0.02)	(0.02)
Grande Ecole graduate	0.13***	0.11***	0.11***
	(0.03)	(0.02)	(0.02)
Age > 38 years	-0.02*	-0.01	0.00
	(0.01)	(0.01)	(0.01)
Entrepreneur is male	0.02*	0.02*	-0.02*
	(0.01)	(0.01)	(0.01)
Serial entrepreneur	0.07***	0.09***	0.08***
	(0.01)	(0.01)	(0.01)
Experience in industry	-0.06***	-0.04***	-0.03***
	(0.01)	(0.01)	(0.01)
Motive: new idea	0.09***	0.08***	0.07***
	(0.01)	(0.01)	(0.01)
Motive: autonomy	0.04***	0.03***	0.03***
	(0.01)	(0.01)	(0.01)
Real start-up	0.20***	0.13***	0.09***
	(0.02)	(0.01)	(0.01)
Threshold: "development" means sales growth greater than	0%	10%	20%
Industry x year of creation FE	Yes	Yes	Yes
R^2	0.08	0.06	0.05
Observations	31,832	31,832	31,832

Source: 1994 and 1998 SINE surveys and Tax Files. This table repeats the results of Table 6, using several alternative definitions of the initial "development" expectation error. Expectation is then equal to 1 when the entrepreneur answers "the firm will develop" to the question "what is your view of the future ?". Realization is equal to one if, over its first two years of existence, firm sales grow by more than 0% % (col. 1), by more than 10% (col. 2) or by more than 20% (col. 3). All regressions control for industry dummies interacted with year-of-creation dummies, to control for industry shocks. In all specifications, error terms are assumed to be correlated for firms created/taken over the same year in the same industry. * means significant at 10%, ** at 5% and *** at 1%.

Table 9: Short Term Bank Debt and Optimism: Main Results Using "Development" Expectation Errors

	Credit Lines / Bank Debt				
	All	1994	1998	All	All
Expectation Error	0.03***	0.04***	0.03**	0.03***	0.03***
based on "development"	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Log ($assets_0$)	-0.03***	-0.03***	-0.03***	-0.03***	-0.04***
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
Tangible $assets_0$/$assets_0$	-0.51***	-0.48***	-0.54***	-0.51***	-0.51***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Real start-up	-0.01	-0.03*	0.01	-0.01	0.02
	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)
Death in 2 years	-	-	-	0.14***	0.14***
				(0.03)	(0.03)
Table 5 regressors incl.	No	No	No	No	Yes
Industry FE × year FE	Yes	Yes	Yes	Yes	Yes
R^2	0.13	0.13	0.17	0.15	0.18
Observations	5,474	2,932	2,542	5,474	4,349

Source: 1994 and 1998 SINE surveys and Tax Files. This Table regresses the ratio of credit lines in total bank credit on the initial expectation error on "development" and some controls. The initial expectation error uses a 3% sales growth threshold and is constructed exactly as in Tables 5 and 6. Controls are measured in the year of firm creation/take over. They include: (1) the share of tangible assets in total assets, (2) the log of total assets, (3) a dummy equal to 1 when the firm is a "real" start-up and (4) a dummy equal to 1 if the firm disappears from the incorporation files after two years of existence. All regressions also control for industry dummies interacted with year-of-creation dummies, to control for industry shocks. Column 1 estimates the model on the whole sample without the death dummies. Column 2 estimates the same model on the sub-sample of firms created/taken over in 1994, while column 3 focuses on 1998 creations. Column 4 goes back to the full sample, but includes the "death dummy". Column 5 also adds the entrepreneur/project observables used in Table 6. In all specifications, error terms are assumed to be correlated for firms created/taken over the same year in the same industry. * means significant at 10%, ** at 5% and *** at 1%.

Table 10: Short Term Bank Debt and Optimism: IV Estimates

Dependant variable:	Reduced Form Line of credit / bank debt	First Stage Year of creation expectation error	Second Stage Line of credit / bank debt
"Development" exp. error,	-	-	0.28**
year of creation (instrumented)			(0.14)
"Development" exp. error	0.03**	0.10***	-
3 yrs after creation	(0.01)	(0.02)	
Log ($assets_0$)	-0.01*	0.02	-0.02**
	(0.01)	(0.01)	(0.01)
Tangible $assets_0/assets_0$	-0.49***	-0.03	-0.48***
	(0.03)	(0.04)	(0.03)
Real start-up	-0.03**	0.21***	-0.09***
	(0.02)	(0.02)	(0.03)
Industry x Year of creation FE	Yes	Yes	Yes
R^2	0.15	0.10	-
F statistic	-	23.9	-
Observations	3,099	3,099	3,099

Source: 1994 and 1998 SINE surveys and Tax Files. This Table redoes the estimates of table 9, column 1 using second period "development" expectation errors as an instrument for initial expectation errors. The initial expectation error uses a 3% sales growth threshold and is constructed exactly as in Tables 5 and 6. Second period expectation error is constructed as in Table 7. Controls are measured in the year of firm creation/take over. They include: (1) the share of tangible assets in total assets, (2) the log of total assets, (3) a dummy equal to 1 when the firm is a "real" start-up. All regressions also control for industry dummies interacted with year-of-creation dummies, to control for industry shocks. Column 1 is the "reduced form": it regresses the share of credit lines directly on second period expectation errors. Column 2 is the first stage: it regresses initial expectation errors on second period ones, including the second stage controls as in Table **??**. We also present the Fisher statistic under the null that the coefficient second period exp. errors is equal to zero. Column 3 is the two stage least squares estimate of the effect of initial exp. errors on the share of credit lines, instrumenting initial with second period expectation errors. In all specifications, error terms are assumed to be correlated for firms created/taken over the same year in the same industry. * means significant at 10%, ** at 5% and *** at 1%.