ETLA
ELINKEINOELÄMÄN TUTKIMUSLAITOS
THE RESEARCH INSTITUTE OF THE FINNISH ECONOMY
Lönnrotinkatu 4 B 00120 Helsinki Finland Tel. 358-9-609 900
Telefax 358-9-601 753 World Wide Web: http://www.etla.fi/

Keskusteluaiheita – Discussion papers

No. 987

Ari Hyytinen – Mika Pajarinen

WHY ARE ALL NEW ENTREPRENEURS BETTER THAN AVERAGE? EVIDENCE FROM SUBJECTIVE FAILURE RATE EXPECTATIONS

We would like to thank Anssi Rantala and Otto Toivanen for helpful discussions. We gratefully acknowledge the financial support provided by the National Technology Agency of Finland (Tekes, project 579/31/03). Corresponding author: Ari Hyytinen, Etlatieto Ltd, Lönnrotinkatu 4 B FIN-00120 Helsinki, Tel: +358-400-777086, Fax: +358-9-601753, E-mail: ari.hyytinen@etla.fi. All opinions expressed are those of the authors.

ISSN 0781-6847 23.06.2005

HYYTINEN, Ari – **PAJARINEN,** Mika, **WHY ARE ALL NEW ENTREPRENEURS BETTER THAN AVERAGE? EVIDENCE FROM SUBJECTIVE FAILURE RATE EXPECTATIONS.** Helsinki: ETLA, Elinkeinoelämän Tutkimuslaitos, The Research Institute of the Finnish Economy, 2005, 34 p. (Keskusteluaiheita, Discussion Papers, ISSN, 0781-6847; No. 987).

ABSTRACT: There is considerable consensus in the literature that entrepreneurs are prone to unrealistic optimism. Our new field evidence from a sample of four-month-old start-ups echoes this finding: as many as 87% of new entrepreneurs expect to survive at least three years in business, whereas the actual statistical survival rate is around 60%. Our field study allows us to rule out certain previously overlooked explanations for the often-documented optimistic bias in new entrepreneurs' judgment: It is not due to a framing effect nor driven by rash and sloppy survey responses by busy entrepreneurs. Nor can we relate it to the risk preferences of entrepreneurs or to a difficulty in understanding algebra of probability. We also study how entrepreneurs update their failure rate expectations: When prompted to rethink, 33% of the entrepreneurs update their risk beliefs. We reject the formal restrictions of a Bayesian learning model, in part because of the presence of optimistic entrepreneurs.

JEL: D81, D84, D21.

KEYWORDS: entrepreneurship, new firms, default, expectations, survival.

HYYTINEN, Ari – **PAJARINEN,** Mika, **MIKSI KAIKKI UUDET YRITTÄJÄT OVAT KESKIMÄÄRÄISTÄ PAREMPIA? TULOKSIA SUBJEKTIIVISISTA ELOONJÄÄMISTODENNÄKÖISYYKSISTÄ**. Helsinki: ETLA, Elinkeinoelämän Tutkimuslaitos, The Research Institute of the Finnish Economy, 2005, 34 s. (Keskusteluaiheita, Discussion Papers, ISSN, 0781-6847; No. 987).

TIIVISTELMÄ: Uudet yrittäjät ovat tutkimusten mukaan usein erittäin optimistisia yrityksensä menestymismahdollisuuden ja erityisesti sen "eloonjäämistodennäköisyyden" suhteen. Noin neljän kuukauden ikäisille yrityksille tehty kyselymme tukee tätä tulosta: keskimäärin peräti 87% uusista yrittäjistä odotti yrityksensä olevan toiminnassa vielä kolmen vuoden kuluttua, kun tilastollisesti vastaava todennäköisyys on vain noin 60%. Pyrimme tutkimuksessamme sulkemaan pois eräitä syitä sille, miksi yrittäjät ovat kyselytutkimuksien mukaan keskimäärin varsin optimistisia. Osoitamme, että ylioptimistisuus ei liity kysymystapaan (so. siihen, kuinka kysymys on muotoiltu) tai kiireisten yrittäjien huolimattomiin vastauksiin. Ylioptimistisuus ei näyttäisi myöskään liittyvän yrittäjien riskipreferensseihin tai vaikeuksiin ymmärtää todennäköisyyslaskentaa. Tutkimme myös, kuinka yrittäjät päivittävät käsitystään yrityksensä eloonjäämistodennäköisyydestä. Kun yrittäjille kerrotaan arvioita uusien yritysten keskimääräisestä elonjäämistodennäköisyydestä ja pyydetään sen jälkeen harkitsemaan uudelleen omaa käsitystään yrityksensä eloonjäämistodennäköisyydestä, 33% yrittäjistä muuttaa vastaustaan. Käyttämällä yrittäjien alkuperäisiä ja päivitettyjä vastauksia mallinnamme yrittäjien tapaa päivittää odotuksiaan. Hylkäämme hypoteesin Bayesilaisen mallin mukaisesta odotusten päivityksestä, osin sen vuoksi, että osa yrittäjistä on ylioptimistisia.

JEL-luokittelu: D81, D84, D21.

AVAINSANAT: yrittäjyys, uudet yritykset, epäonnistuminen, odotukset, eloonjääminen.

Table of Contents

1 Introduction

There seems to be considerable consensus in the literature that entrepreneurs are prone to unrealistic optimism. This consensus says, roughly, that the founders of new businesses tend to overestimate their chances of success and think that their businesses are (almost) failure-proof both in absolute and in relative terms. Survey-based field evidence reported in Cooper, Woo, Dunkelberg (1988) and Pinfold (2001) seems to provide the most direct empirical support for the consensus, for these papers focus on entrepreneurs' probabilistic success rate expectations.[1] Eliciting expectations in surveys is, however, a controversial undertaking (e.g., Juster 1966, Manski 2004). The aim of this paper is to provide both a re-assessment of some of the available survey-based field evidence and new insights on what might drive the documented bias in entrepreneurs' judgment.

In their widely cited paper, Cooper at al. (1988) cleverly track and document the perceived chances for success of nearly 3000 new entrepreneurs in the U.S.. The entrepreneurs were asked, "What are the odds of your business succeeding?" and "What are the odds of any business like yours succeeding?". Based on these two questions the authors uncovered a number of striking findings: 33% of the interviewed entrepreneurs believed that their changes of success to be a certain 100%, and as many as 81% reported that their chances be no less than 70%. Given that the actual failure rates of new businesses is relatively high (about 40-50% newly created firms exit before the fourth year after the start-up), these numbers suggest that the entrepreneurs are optimistic in absolute terms. Around 68% of the

[1] A number of elegant theories have been put forward both in psychology and more recently in economics to explain the phenomenon and its implications. These theories include but are not limited to de Meza and Southey (1996), Bernardo and Welch (2001), Landier and Thesmar (2003), Brocas and Carillo (2004) and van den Steen (2004).

entrepreneurs also thought that they will do better than any business like theirs. Albeit this finding cannot literally be taken to mean that more than 50% believe that they will do better than the median (which would imply a logical error), it shows that the entrepreneurs are optimistic also in relative terms.

The questions used by Cooper at al. (1988) suffer, however, from certain defects. Perhaps the most obvious of them is that the time period over which the chances of success were supposed to be measured was not specified. This omission is potentially important, for in the long-run, all firms fail (and we are all dead). Pinfold's (2001) empirical note addresses this defect. In this study, a sample of new business founders in New Zealand were asked to estimate the probability that their business and similar ventures would still be operating five years after start-up. Less than 6% of the respondents believed that probability to be 50% or lower for their own venture. About 40% of the respondents believed the same probability to be 50% or lower for similar ventures. Again, the numbers suggest that new business founders suffer both from absolute and relative optimism.

Academic economists often assume that decision makers in general and entrepreneurs in particular have expectations that are rational and objectively correct. The profession is, in particular, quite skeptical about how useful probabilistic expectations data are and how accurately such expectations can be elicited in surveys (see the discussion in e.g. Dominitz and Manski 1997a, 1997b, and Manski 2004). This skepticism can be extended to the available field evidence that bears on entrepreneurial optimism, such as that provided by Cooper et al. and Pinfold. In particular, the available empirical evidence from the field does not seem to rule out certain obvious explanations for the optimistic bias in entrepreneurs' judgment: Is it an outcome of how the question is framed? Is the finding simply due to rash survey responses that the entrepreneurs (who obviously are busy in managing

their recently established businesses) sloppily give? Do biased failure rate expectations reflect a particular type of measurement error, because the elicited beliefs get obfuscated with the risk preferences of entrepreneurs? Does entrepreneurial optimism emerge from the difficulty entrepreneurs may have in understanding algebra of probability or because they do not make full use of the modern probability theory (cf. Camerer and Webber 1992)? Or is it related to the way entrepreneurs form and update their failure rate expectations? This paper addresses these apparent gaps in the empirical literature.

To that end, we document and characterize the failure rate expectations of owner-managers of a sample of four-month-old start-ups. This new field evidence comes from a recent Finnish survey (from 2004) and echoes the findings made in the previous literature: The founders of new businesses grossly underestimate the rate of failure of their own business both in absolute and in relative terms. What is more important is that these new field data allow us to rule out certain obvious explanations for entrepreneurial optimism: First, it is not due to a framing effect: Unlike the previous studies, we frame the question in terms of the probability that the respondent's new business *fails*. Despite this change of frame, the bias continues to exist. Second, it is not driven by rash and sloppy survey responses by busy entrepreneurs. We can find no evidence that absolute or relative optimism is related to the time it took for the respondents to provide their estimate of the failure probability. Further, the bias cannot be related to variables that proxy the risk preferences of entrepreneurs. Nor can we find a statistically significant relation between the estimated failure probabilities and managers' (in)ability to understand a simple law of probability. We do find, however, that people are clearly unwilling or unable to provide an estimate of the rate of failure for a typical start-up that is in the same line of business and founded at the same time as that of the respon-

dent. This question draws a large amount of non-response, as only 71% of the completed interviews included a response to this question. This finding indicates that people are not comfortable with presenting or forming estimates of base rates. We cannot relate this non-response to optimism, however.

After ruling out these apparently obvious explanations for entrepreneurial optimism, we consider the way entrepreneurs form and update their failure rate expectations. Entrepreneurs' risk beliefs are not rigid or immutable, as entrepreneurs do not simply stick to their priors when they are prompted to rethink the likelihood at which their new business will fail and when they are reminded of the base rate at which new firms in general fail. We find that 33% of the entrepreneurs update their risk beliefs, but we reject the formal restrictions of a Bayesian learning model. Optimists appear to update their beliefs in a way different from their more realistic (or pessimistic) counterparts.

Two final findings of ours are as follows: First, education matters: University education increases the likelihood that an entrepreneur believes that his newly started business will fail at a higher rate than the actual average failure rate. It also decreases the likelihood that he believes that his own business will fail at a lower rate than similar businesses. Second, 22% of the new entrepreneurs seem to suffer from the gambler's fallacy. This error in inference accounts for 58% percent of the incorrect answers to the survey question that measures entrepreneurs' ability to understand a simple law of probability.

To be sure, there also exists other empirical evidence (besides the referenced field evidence) which suggests that entrepreneurs' expectations may be biased. These other accounts do not, however, directly focus on success (or failure) rate expectations: Using British Household Panel Study, Arabsheibani, de Meza, Maloney and Pearson (2000) find, for example, that the self-employed predict better

financial outcomes than employees, but suffer from worse realizations. Landier and Thesmar (2003) document that when predicting the growth rates of their business, French entrepreneurs make optimistic expectation errors. Some interesting experimental evidence is reported in the famous study of Camerer and Lovallo (1999) and more recently in Coelho, de Meza and Reyneirs (2004).[2] Our paper augments this wider literature, for none of these studies consider the sources of the entrepreneurs' bias from the perspective we do, nor do they provide any empirical evidence on how entrepreneurs update their beliefs.

The plan of the paper is as follows. Section 1 describes our empirical research design and motives in more detail certain obvious explanations for entrepreneurial optimism that the previous empirical literature have not to our best knowledge addressed. Section 2 describes the field data and definition of variables. Section 3 presents the main results. Section 4 offers some concluding remarks.

2 Empirical research design

Our empirical research design consists of three parts:

Part I: In the first part we check whether our new field evidence from Finland echoes the findings made in the previous literature. A first robustness check of the previous field evidence is already incorporated into this check: Unlike the previous studies, we frame the question of the likelihood that the respondent's business succeeds in terms of the probability that the new business

[2] There also exists an extensive psychology and behavioral economics literature on unrealistic optimism and overconfidence (see, e.g., Debondt and Thaler 1995 and Rabin 1998 for reviews). This literature suggests that these biases in judgment are pervasive and not just an idiosyncratic characteristic of entrepreneurs.

fails. The entrepreneur were asked, “Please estimate the probability that your venture fails and is not in business three years from now?” (IFAIL) and “How many percent of the ventures that operate in the same line of business and were established at the same time as yours will fail and is not in business three years from now?” (OTHERSFAIL). The frame is thus opposite to that used in the previous studies, Cooper et al. and Pinfold in particular. This manoeuvre may sound trivial to some, but we would argue that it is not: There exists extensive amount of empirical evidence documenting that how people respond (in surveys and experiments) depends drastically on the way questions are presented to them (i.e., whether the questions are framed in a certain way; see, e.g., Tversky and Kahneman 1981 and also Diamond and Hausman 1994).

We build on the reported failure probabilities elicited using the above described questions to form two indicators of optimism. The first measures the likelihood that an entrepreneur believes that his newly started business will fail at a lower rate than the actual average three-year failure rate of new Finnish businesses. To measure this, we code a dummy variable, AOPT, which equals one if the respondent thinks that the probability that his venture fails in three years from now is lower than the actual average three-year failure rate of new Finnish businesses. We use this dummy as a proxy for absolute optimism. While very crude, the primary use of this proxy is that it allows us to identify the set of entrepreneurs who, *on average*, hold optimistic expectations and to study their observable characteristics and behavior. The cut-off point we use is 40% and hence AOPT = 1 if IFAIL < 40% and zero otherwise. The cut-off is taken from Eurostat (2004, p. 47)

and it refers to the actual average three-year failure rate of new Finnish businesses. It is the best point estimate we are aware of.[3]

The second indicator of optimism measures an entrepreneur's estimate of how much more (less) likely it is that similar businesses fail than it is that his own business will fail in three years time. We use the difference between the two stated probabilities as a proxy for this type of relative optimism: The difference is computed ROPT = OTHERSFAIL – IFAIL and is therefore increasing in the degree of relative optimism.

Part II: Having established that the new field evidence available to us indeed echoes the findings made in the previous studies, we evaluate the robustness of this evidence. We consider, in particular, the following assertions often made about the meaningfulness of the (probabilistic) survey responses:

- Assertion 1: "Founders of new businesses have neither incentive nor time to answer carefully in surveys and thus to provide an accurate estimate of the failure rate of their businesses." If lack of care is pervasive, it might render the responses next to meaningless. We evaluate this assertion using our indicators of optimism in a number of ways: We regress the indicators on the time it took for the respondents to answer to the survey questions used to elicit the expectations. Our hypothesis is that if optimism is due to rash and sloppy responses, we should find a relation between AOPT (or ROPT) and the response times because it takes, on average, less time to give a sloppy response than a well-thought-out one. Another pattern of response that might indicate a lack of care is the frequency at which the responses are rounded to {0, 50, 100} (see Dominitz and Manski 1997a, 1997b, Manski 2004). We check for this possibility, too. Finally, we com-

[3] Use of other cut-offs (for example, 30% or 50%) yielded similar results.

pare the average durations it took for the respondents to reply to the probabilistic failure rate questions to the average durations of the other questions that were presented in the survey.

- Assertion 2: “Probabilistic expectations get in surveys obfuscated with the risk preferences of entrepreneurs.” A view underlying this type of criticism is that that the method used to elicit expectations does not allow for a proper distinction between the risk preferences of entrepreneurs and their probabilistic risk assessments. If that is the case or if the expected utility theory “fails” and entrepreneurs let their preferences toward risk intervene with their probabilistic assessments of the risk, spurious evidence for optimism might follow. In particular, if the respondents’ risk aversion correlates positively with their (stated) assessments of risk, the previous estimates of entrepreneurial optimism would be biased upwards (as suggested e.g. by Kahneman and Tversky 1979). We regress the indicators on a number of proxies of risk aversion to check for this possible source of bias in the field evidence. Our hypothesis is that if probabilistic expectations get in surveys obfuscated with the risk preferences, we should find a statistically significant relation between the proxies of optimism and risk aversion.
- Assertion 3: “Entrepreneurs have difficulties in understanding algebra of probability and hold therefore less than rational expectations on the failure rate of their business”. The analysis of Camerer and Webber (1992), for example, supports an assertion like this. Moreover, if respondents find it difficult to think in terms of probabilities, they may be unwilling to respond to probabilistic questions (Dominitz and Manski 1997a, 1997b, Manski 2004). To examine the effects of this type of bias in the field evi-

dence, we regress the indicators of optimism on a variable that describes how well the respondents understand a simple law of probability. We also check whether the questions eliciting probabilistic failure rates draw more non-response than other questions. Finally, we report the frequency at which the founders of the new ventures in our data seem to suffer from the "gambler's fallacy" (i.e., rely on a form of "law of small numbers"; see Rabin 2002).

It is important to emphasize that our aim is *not* to take a stance on whether these assertions or more generally any type of hostility to subjective expectations data are warranted. Being cautious is prudent. The aim of this part of our empirical design is just to apply this standard of prudence to the specific question at hand by checking whether the field evidence on entrepreneurial optimism is robust to the criticism.

Part III: In the third part we develop and estimate a simple model of Bayesian updating to bear some new evidence on the way entrepreneurs update their risk beliefs. Our field evidence allows us to say something about the updating of risk beliefs, because interviewers provided the respondents with a piece of base rate information. This information was about the average likelihood of a new firm failing in Finland during its first three years. Based on a random assignment, about 60% of the respondents in our estimating sample were (just before the last questions of the survey) told "It is often argued that only one out of five new firms fails and is not in business three years after the start-up". The remaining 40% were told "It is often argued that as many as four out of five new firms fails and is not in business three years after the start-up". Right after this base rate information had been presented, the entrepreneurs were asked "In light of this information, please estimate the probability that your venture fails and is not in business

three years from now?" (IFAIL_UD). The estimates of the Bayesian learning model (that we are about to formulate) will be based on this updating opportunity.

Why was the survey designed in this way? Before answering the question, it is worth noting a couple of problems that the designer of a survey faces when planning this type of set-up: In general, she/he (as an econometrician) does not have any information about the likelihood of failure of a respondent's *own* firm. Firms cannot therefore be provided any new information about it. Moreover, the actual *average* three-year default rate is not observed either, as the survey designer only has statistical estimates for it. At the beginning of 2004 when the survey was completed, no one had bankruptcy data to estimate the rate for the relevant period from 2004 to 2007. The best point estimate (based on historical data) available at the time was 40-41%, which literally only applies to firms that were founded in 1998 (see Eurostat, 2004, p. 47). While informing the respondents about this best point estimate was of course an option, it would not have not induced variation in the piece of information that was to be provided to the cross-section of respondents. Moreover, to induce variation (and thereby to make it possible to empirically identify learning/updating), one would have had to choose at least two points from the range over which the actual three-year default rate is likely to be distributed.

Entrepreneurs were *not* provided with two close-by point estimates of the base rate around its mean. Instead, the two randomly chosen sub-samples were reminded of the two somewhat arbitrarily chosen base rates: 20% and 80%. For the purposes of this study, these base rates can be reconciled at least in part with the following reasoning: Based on the available evidence, one can expect that entrepreneurs would, on average, provide quite optimistic failure probabilities, both for their own business and for their fellows' businesses. A reasonable conjecture

is that the average of the latter would be somewhere around 20-25%. Given the prior subjective beliefs, informing the respondents about this base rate provides the average respondent with *no* new information. In this paper, we call the group of entrepreneurs to whom this base rate information was presented "the low base rate group". The second number, 80%, is relatively large, implying a very high average risk of failure. We call the group of entrepreneurs to whom this higher number was presented "the high base rate group". This base rate is clearly higher than the average of the stated subjective probabilities. It is probably regarded relatively – if not entirely – implausible by most respondents. These two pieces of information impose hence no (or only a negligible) need to update beliefs, because if both of them are deemed imprecise (implausible) or are not new, they can and will be ignored at no cost. It thus appears that the survey design in other words created a bias *towards* finding evidence that the prior dominates. Note, finally, that neither the survey framework nor the Bayesian learning model forces the respondents to use in any way the piece of base rate information provided to them.

To formulate the Bayesian model, we follow Viscusi and O'Connor (1984), Viscusi (1997) and Hakes and Viscusi (1997), and Viscusi and Evans (1998) who assume that individuals adopt a Bayesian learning approach. We also adopt this assumption and focus on a single equation in which the dependent variable is entrepreneurs' *updated* estimate of the probability that their venture fails (i.e., IFAIL_UD). In this framework, the events of failing are regarded as independent Bernoulli trials and the associated risk probabilities are characterized using a beta distribution. The parameters of the prior distribution are p, the prior risk assessment (of the likelihood of failure, i.e. IFAIL), and γ, the precision that the entrepreneurs attach to their prior. We regard the receipt of new information about the

average failure rate as roughly equivalent to observing "additional Bernoulli trials" concerning the riskiness of new ventures. The informational content of the numbers that was provided to the entrepreneurs in the survey depend on two parameters: ξ, which is the precision of the new information and *s*, which is assumed to be the risk implied by the new information (e.g., the fraction of the additional Bernoulli trials that reflect failure.)

For the beta family of distributions, the posterior probability of failure after receiving the new information is given by (see for example Viscusi and O'Connor 1984 and Viscusi 1997)

$$p^{posterior} = \frac{\gamma p + \xi s}{\gamma + \xi} = \frac{\gamma p}{\gamma + \xi} + \frac{\xi s}{\gamma + \xi} \qquad (1)$$

To derive our estimating equation, we let *i* index the respondents and j denote the base rate group to which respondent *i* belongs. There is cross-sectional variation in *p* (i.e., it varies across the respondents), but conditional on the group respondent *i* was assigned to, *s* does not. The reason for this is that each respondent *i* was provided with only one piece of base rate information.

The empirical counterpart of (1) is

$$IFAIL\_UD_{ji} = \alpha_j + \beta_j IFAIL_{ji} + \varepsilon_{ji} \qquad (2)$$

where $\alpha_j = s_j \xi_j / (\xi_j + \gamma)$, $\beta_j = \gamma / (\gamma + \xi_j)$, and ε_{ji} is the error term. The nature of the learning process is captured by coefficients α_j and β_j. If the respondents completely disregard the information provided to them, $\alpha_j = 0$ and $\beta_j = 1$. On the other hand, if the new information is taken at face value, we should observe $\alpha_j > 0$ and $\beta_j = 0$. The risk level implied by the new information is given by $s_j = \alpha_j / (1 - \beta_j)$, but this estimate is not easy to interpret in the present context. It reflects, at least to some extent, how the default rate of respondent *i*'s new busi-

ness maps to changes in the average default rate. The informational content of the stated average default rate of new businesses relative to the prior, Ψ_j, is easier to interpret, because higher values of Ψ_j mean greater precision of the stated average default probability relative to the respondents' initial judgments. It is given by $\Psi_j = \xi_j / \gamma = 1/\beta_j - 1$.

3 Data

3.1 Survey description

The field evidence for this study comes from a Finnish cross-sectional survey of a sample of (approximately) four-month-old start-ups conducted by The Research Institute of the Finnish Economy (ETLA). The central features of the survey are as follows:

The survey was nation-wide and it covered initially *all* start-ups that were founded in October 2003. The survey was completed between January 22 – March 11, 2004. The survey was conducted by computer assisted telephone interviews. Interviews were carried out by Tietoykkönen Ltd, which specializes in research and marketing information services, fieldwork, and statistical data analysis and has about 15 years experience from the field.

The initial population consisted of 2207 firms which were granted a new business identity code by one of the many possible registration offices, such as the Finnish Trade Register and tax authorities. Of this initial population 870 firms were eventually reached for an interview. The primary reason for not reaching a start-up was "technical" in nature: no response was received from the registered

phone number despite numerous attempts (758 firms).[4] Entrepreneurs' contact information was not available due to confidentiality reasons in 579 cases.[5] From those that were successfully contacted, 202 responded that their firms were not yet active in business. Of those firms which were active and available for an interview, 64 percent (426/668 firms) agreed to be interviewed. A small number of interviews could not be fully completed, primarily due to linguistic problems (33 firms) which reduced the response rate to 59 percent (393/668).

The number of completed interviews available to us is 393. Most of the analysis in this paper focuses, however, on the 272 businesses which responded to all three failure-expectations questions and which provided some basic demographic and other data we decided to use for this study. We occasionally use a larger sample that consists of 366 firms to illustrate the robustness of our main findings, and to consider possible biases due to non-response to certain questions.

The core questions in the survey asked respondents about their (and other founders if more than one) personal background, various details of the start-up and problems related to starting the business. There were also a set of general attitude questions related to entrepreneurship and start-ups. In this study we focus on the subset of the questions elicited entrepreneurs' views about the chances of success of their new business.

---

[4] The number of calls per phone number varied from 1 to 11. Busy lines were redialed at least 5 times.

[5] According to the Finnish legislation very small businesses have right to keep their contact information secret. A criterion for secrecy is that the annual turnover is less than 8500 euros.

3.2 Control variables

We have already described and defined IFAIL, OTHERSFAIL, AOPT, and ROPT, which are the key variables of this study. The other variables that we use are as follows:

RETIME1 measures the time (in seconds) it took for the respondent to come up with an answer to IFAIL and RETIME2 to OTHERSFAIL, respectively. Risk preferences are proxied by two variables, the first of which is called LOTTERY. This variable is defined as the sum of money that the respondent was willing to pay for a single ticket in a (hypothetical) lottery that had 10% chance to win 1000 euros and 90% change to win nothing. The second proxy for risk preferences is called RISKLOVE. This variable simply reflects the answers to question "I love risk-taking". The respondent had to use a scale from 1 to 4 (in which 1 stood for complete disagreement and 4 for complete agreement). RISKLOVE equals 1 if the answer was 3 or 4, and 0 otherwise. Understanding algebra of probability is captured by PROB, which is a dummy variable that equals 1 if the correct answer was given to the following question (and 0 otherwise): "We toss a regular coin four times and get every time tails. What is the probability for heads in the next toss?"

Variables that characterize respondents' background include the following: i) ACADEMIC = 1 if respondent has an academic degree and 0 otherwise; and ii) COLLEGE = 1 if respondent has a college level degree. The omitted category in is other than academic or college level education. Previous work and entrepreneurship experience is measured by variable EXPERIENCE, which sums respondent's work and entrepreneurial experience in the field of his/her start-up in years. Another measure of experience is WORKED, which equals 1 if respondent was employed prior to starting his/her new business and 0 otherwise. In addition, we

have dummies for marital status and sex: SINGLE = 1 if the respondent was single at the time of interview and 0 otherwise; and FEMALE = 1 for females and 0 for males. NCHILD measures the number of children in the respondent's household. Finally, AGE is the age of the respondent in years.

4 Results

4.1 Part I: New field evidence

Table 1 reports the distribution of IFAIL and OTHERSFAIL. It shows that the average of IFAIL is 13% and the median 5%. As many as 89% of the respondents think that the likelihood of failure of their new venture is 40% or less in three years time. Given that the actual three-year default rate is on average about 40%, the entrepreneurs seem to hold beliefs that are optimistic in absolute terms. Most entrepreneurs consider themselves better managers than average, for 76% of the respondents think that the likelihood of failure of the new ventures run by other entrepreneurs is 40% or less in three years time. The average of OTHERSFAIL is 26% and the median 20%. These numbers indicate relative optimism. Albeit we have framed the question in terms of the probability that the business *fails*, the bias exist in our data and is similar in magnitude to that documented in the earlier analyses using the opposite frame (cf. Cooper et al. 1988, and Pinfold 2001).

[Insert Table 1 here]

Table 2 reports descriptive statistics for AOPT and ROPT. The table provides further evidence for entrepreneurial optimism. Of particular interest is to note that 63% of the respondents think that their new business has a lower prob-

ability of failure than other similar ventures in the same line of business that were founded at the same time. Only 14% of the respondents think the opposite. These numbers speak for relative optimism: Almost all entrepreneurs think they are better than average. The mean of ROPT tells us that the respondents think on average that their own business is 12 percentage points less likely to fail than that of the others in three years time.

[Insert Table 2 here]

Our field evidence provides us with some additional evidence for entrepreneurial optimism on and above that presented in Table 1 and 2. First, the entrepreneurs were asked to evaluate their own ability to be a successful entrepreneur on a Likert scale from 1 to 10. The median and average score to this question is 8, and no one in our estimating sample reported a score below 5. As many as 76% of the entrepreneurs think that their ability is 8 or higher. Second, the entrepreneurs were asked to evaluate their own ability to be a successful entrepreneur relative to the entrepreneurs who are in the same line of business and who established their firms at the same time. On a five-point scale, only 2% (51%) of the respondents thought that their ability was lower (higher) than that of their fellow entrants. Interestingly, when the entrepreneurs were asked to evaluate their own ability to be a successful entrepreneur relative to incumbent entrepreneurs who are in the same line of business, relative optimism disappears. On a five-point scale, 40% of the respondents thought that their ability is *lower* than that of the incumbents and only 14% thought that it is higher.

4.2 Part II: Ruling out obvious explanations

Table 3 and 4 present a number of estimations that try to link AOPT and ROPT to {RETIME1 (or RETIME2), LOTTERY, RISKLOVE, PROB, ACADEMIC, COLLEGE, AGE, FEMALE, SINGLE, NCHILD, WORKED, EXPERIENCE}. In table 3 the dependent variable is AOPT, and the method of estimation Logit. In table 4 the dependent variable is ROPT and the method of estimation OLS with a heteroskedastic-robust variance-covariance matrix.

The results show that we cannot find a relation between AOPT (or ROPT) and the response time, RETIME1 (or RETIME2). Nor is there a statistically significant relation between the proxies of optimism and risk aversion (LOTTERY or RISKLOVE). PROB has no effect on ROPT, but it does have a marginally significant positive effect on AOPT (at the 10% level) when the demographic controls are included in the regression. The positive effect means, however, that understanding algebra of probability *increases* absolute optimism, which is in contrast to the hypothesis that optimism is due to problems in understanding the laws of probability.

[Insert Table 3 here]

[Insert Table 4 here]

A final finding from Table 3 and 4 is that education matters: University education (ACADEMIC), for example, increases the likelihood that an entrepreneur believes that his newly started business will fail at a higher rate than the actual average failure rate. It also decreases the likelihood that he believes that his own

business will fail at a lower rate than similar businesses. The estimates suggest, moreover, that EXPERIENCE increases absolute but not relative optimism.

Are there any other signs that might indicate a lack of care or unwillingness to respond to the probabilistic questions? While not reported in the table, a pattern worth noting is that while some respondents round their probabilities to the values {0, 50}[6], it also is very frequent in our data that the probabilities are round to the nearest multiple of five. This finding is similar to what, e.g., Dominitz and Manski (1997a, pp. 270) report in their analysis of the perceptions of economic insecurity. When we regress a dummy indicating whether the respondent *did* round his probability to the values {0, 50} on the time it took for him to reply to the question, we find that in the case of IFAIL, there is a negative and statistically significant (at the 5% level) relation. For OTHERSFAIL, we find no such statistically significant effect (albeit the estimated coefficient is negative). This suggests that rounding to {0, 50} may indeed signal a care of lack. Note, however, that we find no evidence that the response times correlate with optimism (see Table 3). We can also compare the average time it took for the respondents to reply to the probabilistic failure rate questions to the average time of the other survey questions. Albeit this comparison is not without its problems, it provides no indication that people have responded to them more sloppily than for example to the other questions involving probabilities. The average response time to IFAIL (OTHERSFAIL) in the estimating sample is 16 seconds (18 seconds), which compares to 25 seconds to LOTTERY and 21 seconds to PROB. Note, however, that these response times include the time it took by the interviewer to read the question and that the word-

[6] No one reported that his or her firm would fail in three years with probability one.

ing of the LOTTERY and PROB -questions is somewhat longer than that of IFAIL and OTHERSFAIL.[7]

We can also check whether the questions eliciting probabilistic failure rates draw more non-response than other questions. The number of observations in our basic estimating sample is 272, but the number of fully completed interviews is 393. This larger sample indicates that 97% of the respondents gave probability responses to IFAIL. This compares to PROB (96%), LOTTERY (99%), AGE (100%) and to the response rates to the probabilistic questions concerning labor force participants' perceptions of economic security in the U.S. (response rate = 98%; see Dominitz and Manski 1997a, p. 270). A dramatic finding is, however, that the response rate to OTHERSFAIL is as low as 71%. People are clearly unwilling or unable to provide an estimate of the rate of failure for a typical start-up that is in the same line of business and founded at the same time as that of the respondent. This non-response indicates that people are not comfortable with presenting or forming estimates of base rates. It also is the primary reason why our basic estimating sample is smaller than the sample consisting of completed interviews. We cannot use the larger sample throughout the paper, because we would not be able to compute ROPT. However, the main results presented so far for absolute optimism (based on AOPT) are not changed if we use the larger sample. For example, if we re-run the regressions of Table 3 using the largest possible sample of 381 observations, essentially nothing changes.[8]

---

[7] The LOTTERY-question consists, for example, of 20 words (140 characters with not spaces) and the IFAIL-question of 15 words (131 characters with not spaces). Of course, asking and responding to easier questions, such as the year of birth (AGE; 5 seconds), is faster.

[8] The only minor changes relate to significance levels: in the models (4) and (8) the significance of PROB slightly increases and is now significant at 10% and 5% levels. In addition, significance of EXPERIENCE increases and ACADEMIC and COLLEGE decrease: in the models (5)-(8) EX-

To conclude this part of our empirical research set-up, we report the frequency at which the founders of the new ventures in our data seem to suffer from the "gambler's fallacy" (i.e., rely on a form of "law of small numbers"). In our estimating sample, 22% of the respondents think that after observing four heads in row, the likelihood that the next toss of the coin results in tails is strictly higher than 50%. This error in inference accounts for 58% percent of the incorrect answers to PROB. When a dummy equaling one for those who suffer from this error is included as a new regressor to the regression models (4) and (8) of Table 3 and 4, we find no evidence that "gambler's fallacy" is related to entrepreneurial optimism.

4.3 Part III: How are risk beliefs updated?

To evaluate how entrepreneurs update their risk beliefs, we first compare IFAIL_UD to IFAIL. On average entrepreneurs revise only slightly their failure rate expectations in the light of the base rate information provided to them: The overall mean of IFAIL_UD is 12.8%, which is nearly identical to the mean of IFAIL. The difference is -0.6 percentage points, which is not statistically significant. The medians of IFAIL_UD and IFAIL are also similar, and we fail to reject the null hypothesis of equality of the medians. As many as 67% of the respondents (i.e., 184 out of 272) stick exactly to their prior, but the remaining 33% update their beliefs in the expected direction: *Conditional on IFAIL_UD* $\neq$ *IFAIL*, the mean decrease in the low base rate group is -2.68 percentage points and the mean increase in the high base rate group is 1.64 percentage points.

PERIENCE is significant at 1% level, ACADEMIC is significant at 10 percent level in the models (5)-(7) and 5% level in the model (8) and COLLEGE is not significant in any models even at 10 percent level. The mean of AOPT in this larger sample of 381 observations is 0.882 and IFAIL 13.320. These values are quite identical to the means calculated from our estimating sample.

These preliminary observations (and the fact that the design of the field experiment is somewhat biased towards finding that the prior dominates) lead us to conjecture that the Bayesian learning model might fit the data well: If, for example, prior information dominates, we should find that $\alpha_j = 0$, $\beta_j = 1$, and that the coefficient of determination is near to unity. This is not quite the case, as we will show next.

The OLS estimates of the Bayesian learning model (and the associated heteroscedasticity-robust standard errors) are presented in Table 5, separately for the low and high base rate groups. The first two columns show that the estimate of β_j is 0.620 and 0.791 and that of α_j 0.029 and 0.042 in the low and high base rate groups, respectively. Albeit plausible as such, these estimates allow us to easily reject the joint hypothesis of $\alpha_j = 0$ and $\beta_j = 1$ at better than 5% level for both groups. This finding implies that the respondents do not simply stick to their priors. Neither the estimates nor this rejection hints of any fundamental misspecification of the learning model. However, Ramsey's RESET tests indicate that the functional form can be questioned and that the models may miss some potentially important variables or non-linearities. As shown in the table, the tests reject the null of no misspecification at better than 1% significance level.

One way to try to salvage the linear learning model is to evaluate whether the signs of statistical misspecification are driven by those whose prior default probability was exactly zero. These probabilities cannot be updated in the Bayesian framework, which cannot handle certain events. Moreover, it is possible that the signs of misspecification are due to peculiar updating by those who are optimistic either in absolute or in relative terms. To allow for these possibilities, we include ZERO (= 1 if IFAIL = 0 and 0 otherwise), AOPT and KAOPT (= AOPT

times IFAIL), as well as DROPT (= 1 if ROPT > 0 and 0 otherwise) and KDROPT (= DROPT times IFAIL).

Columns (3)-(6) report the results. Three findings are of particular interest: First, Ramsey's RESET tests no longer reject the null hypothesis of no problems with the functional form in columns (3) and (4), where AOPT and KAOPT are included. These variables obtain in three out of the four cases statistically significant coefficients. However, the tests do reject in columns (5) and (6), where neither DROPT nor KDROPT obtain significant coefficients. These findings suggest that allowing for absolute optimism in the linear learning model salvages the model in statistical terms at least to some extent. The estimates also indicate that those who are optimistic in absolute terms update their beliefs in a way different from their (realistic) counterparts. The second finding of interest is that "pessimists" (i.e., those with AOPT = 0) give clearly a much larger weight on the low base rate than optimisms (AOPT = 1), who let their prior dominate. We cannot reject the null hypothesis that $\beta_j = 0$ when AOPT = 0 in this low base rate group. However, $\beta_j = 0$ can easily be rejected at better than 1% significance level when AOPT = 1 (Coefficient = -0.225 + 0.989 ≈ 0.764, p-value < 0.01). For the high base rate group, we see no such difference, as the coefficient of KAOPT is not significantly different from zero. Finally, ZERO obtains a positive and significant coefficient in the high base rate group, but not in the low base rate group. This finding indicates that some entrepreneurs responded in an "alarmist" fashion to the high base rate reported to them, which interestingly compares to the results reported in Viscusi (1997).[9]

[9] A similar finding puts the negative (but insignificant) coefficient of IFAIL in column (3) into a perspective: see the next footnote.

Columns (7) and (8) reports the most parsimonious versions of the Bayesian learning model that pass the RESET specification tests. For the low base rate case (column (7)), we reject the joint hypothesis that $\beta_j = 1$ and $\alpha_j = 0$ for the optimists (i.e. for those with AOPT = 1). In this specification, the constant term ($\hat{\alpha}_j = 0.367$) has a simple interpretation, as it equals the average of IFAIL_UD, conditional on being a realist (i.e. AOPT = 0) in the low base rate group. This conditional average is quite a bit lower than the corresponding conditional average of the prior (IFAIL), which is 0.539.[10] For the high base rate group (column (8)), we cannot reject at the 5% level the joint hypothesis that $\beta_j = 1$ and $\alpha_j = 0$ for the optimists, as the Wald-test obtains a value of 2.87 and has p-value of 0.061. For realists, the corresponding Wald-test obtains a value of 3.81 and has a p-value of 0.025.

As measured by the coefficient of determination, R^2, the overall fit of the basic learning models in columns (1) and (2) is somewhat worse than that of the re-specified learning models reported in column (7) and (8). These models also have the desirable property that the predicted values from the OLS regressions fall within the unit interval. As Papke and Wooldrige (1996) have recently emphasized, this property cannot be taken for granted: A linear model for fractional data may well suffer from the same drawback (of failing to provide plausible predictions) as the linear probability models suffer when fitted to binary data. To further

[10] Conditional on not being optimistic in absolute terms, the entrepreneurs responded in an "alarmist" fashion to the relatively low base rate reported to them, which (again) interestingly compares to the results reported in Viscusi (1997). In this group it apparently holds that the higher was the prior, the bigger was the update downwards. In this sub-sample, the (conditional pair wise) correlation between IFAIL and IFAIL_UD is -0.13 (p-value 0.59) for those with AOPT = 0, whereas it is 0.73 (p-value = <0.01) for those with AOPT = 1.

study the statistical properties of the learning model, we repeat the estimations of Table 5 using the larger sample of 366 firms to the extent it is possible (i.e., columns (1)-(4), and (7)-(8)). These estimations show that we cannot find any evidence that those who did not respond to the base rate question (OTHERSFAIL) update their beliefs differently from others. Nor do the statistical results reported in Table 5 change when this large sample is used. For example, the models reported in columns (1)-(2) do not pass the RESET tests, while those of column (7)-(8) do. We can also look for omitted variables by re-estimating the models of Table 5 with various combinations of control variables available to us, such as demographics (cf. the control variables used in Table 3 and 4). However, we are able to detect no statistically significant effects in these experiments.

Despite being robust to the above checks, the estimated learning models reported in column (7) and (8) suffer from some form of heteroscedasticity. For example, White's (1980) general heteroscedasticity test rejects the null hypothesis of homoscedasticity at better than 1% level for the models reported in column (1) and (7), but not for the models reported in column (2) and (8). These findings do not render the reported standard errors in Table 5 invalid, because they are robust to an unspecified form of heteroscedasticity. However, the White test can also indicate a specification error. We investigate this potential problem by running the classical Park/Harvey/Glejser heteroscedasticity regressions that explore its nature. We do so in a couple of ways using the models reported in column (7) and (8): First, the logarithm of the squared residuals from these models is regressed on the logarithm of (IFAIL+1). This specification assumes that the heteroscedastic variance is proportional to the prior. Second, we do not take logarithms, but consider instead a simpler linear model. For this test, the squared residuals are regressed on IFAIL. Finally, the absolute value of the residuals is regressed on

IFAIL. These explorative regressions show that heteroscedasticity (rather than an unknown specification error) might indeed be driving the high value of the White test, for the error variance is typically increasing in the prior. The coefficient of IFAIL (or its transformation) is always positive. Moreover, it is statistically significant in each of the six explorative regressions we run for model (7), and in four out of the six regressions for model (8). This finding suggests that there is more variability in the posterior of those who initially thought that their venture is quite likely to fail. How to reconcile this pattern with updating is left for future research.

To sum up, entrepreneurs do not simply stick to their priors when prompted to rethink the rate at which their own new firms fails and reminded of the average likelihood at which new firms on average fail (i.e., of the base rate). We nevertheless reject the formal restrictions of the linear Bayesian model. We do so in part because optimists appear to update their beliefs in a way different from their more realistic (or pessimistic) counterparts.

5 Conclusions

In this paper we document and characterize the failure rate expectations of owner-managers of a sample of four-month-old start-ups. This new field evidence comes from a recent Finnish survey (from 2004) and echoes the findings made in the previous literature: The founders of new businesses grossly underestimate the rate of failure of their own business both in absolute and in relative terms. According to the survey results as many as 89% of the respondents think that the likelihood of failure of their new venture is below the actual three-year default rate (40%), the average estimate being 13%. In contrast, only 76% of the respondents think that the likelihood of failure of new ventures run by other entrepreneurs is below

40% in three years time. Respondents' average estimate of the failure likelihood of the new ventures run by other entrepreneurs is significantly higher (26%) than their estimate of their own venture's failure likelihood.

Our field data allow us to rule out certain obvious but previously overlooked explanations for entrepreneurial optimism, such as framing effect, rash and sloppy survey responses by busy entrepreneurs, respondents' risk preferences and their (in)ability to understand a simple law of probability. For instance, we can find no evidence that absolute or relative optimism is related to the time it took for the respondents to come up with their estimate of the failure probability for their own venture. We find, however, that high level education reduces optimism and that previous work and entrepreneurial experience increases absolute (but not relative) optimism.

This paper is the first to document that 22% of the new entrepreneurs seem to suffer from the "gambler's fallacy", i.e., rely on a form of "law of small numbers". This error in inference accounts for 58% percent of the incorrect answers to the survey question that measures entrepreneurs' ability to understand a simple law of probability. We are, however, not able to link the gambler's fallacy to entrepreneurial optimism. We moreover find that people are clearly unwilling or unable to provide an estimate of the rate of failure for a typical start-up that is in the same line of business and founded at the same time as that of the respondent. This question draws a large amount of non-response, as only 71% of the completed interviews included a response to this question. This finding indicates that people are not comfortable with presenting or forming estimates of base rates. We cannot relate this non-response to optimism, however.

We also studied the way entrepreneurs form and update their failure rate expectations. Entrepreneurs' risk beliefs are not rigid or immutable, as entrepre-

neurs do not simply stick to their priors when prompted to rethink the rate at which their new business fail and reminded of the base rate at which new firms in general fail. We find that 33% of the entrepreneurs update their risk beliefs, but we reject the formal restrictions of a Bayesian learning model. Optimists appear to update their beliefs in a way different from their more realistic (or pessimistic) counterparts.

References

Arabsheibani, Gholamreza, de Meza, David, Maloney, John and Pearson, Bernard, 2000, And a vision appeared into them of a great profit: Evidence of self-deception among the self-employed, *Economic Letters*, Vol. 67, pp. 35-41.

Bernardo, Antonio, E. and Welch, Ivo, 2001, On the evolution of overconfidence and entrepreneurs, *Journal of Economics and Management Strategy*, Vol. 10, No. 3, pp. 301 - 330.

Brocas, Isabelle and Carillo, Juan, D., 2004, Entrepreneurial boldness and excessive investment, *Journal of Economics and Management Strategy*, Vol. 13, No. 2, pp. 321-350.

Camerer, Colin and Lovallo, Dan, 1999, Overconfidence and excess entry: An experemential approach, *American Economic Review*, Vol. 89, No. 1, pp. 306-318.

Camerer, Colin and Webber, Martin, W., 1992, Recent developments in modelling preferences: Uncertainty and ambiguity, *Journal of Risk and Uncertainty*, Vol. 5, pp. 325-370.

Coelho, Marta, de Meza, David and Reyneirs, Diane, 2004, Irrational exuberance, entrepreneurial finance and public policy, *International Tax and Policy Finance*, Vol. 11, pp. 391-417.

Cooper, Arnold, C., Woo, Carolyn, Y. and Dunkelberg, William, C., 1988, Entrepreneurs' perceived chances for success, *Journal of Business Venturing*, 3, pp. 97-108.

Debondt, Werner, F., M. and Thaler, Richard, H., 1995, Financial decision-making in markets and firms: A behavioral perspective, in Jarrow, R., Maksimovic, V. and Ziemba, W.(eds.): *Finance, vol. 9 of Handbook in Operations Research and Management Science*, Elsevier, Amsterdam.

de Meza, David and Southey, Clive, 1996, The borrower's curse: Optimism, finance and entrepreneurship, *Economic Journal*, Vol. 106, pp. 375-386.

Diamond, Peter, A. and Hausman, Jerry, A., 1994, Contingent valuation: Is some number better than no number?, *Journal of Economic Perspectives*, Vol. 8, No. 4, pp. 45-64.

Dominitz, Jeff and Manski, Charles, F., 1997a, Perceptions of economic insecurity. Evidence from the survey of economic expectations, *Public Opinion Quarterly*, Vol. 61, pp. 261-287.

Dominitz, Jeff and Manski, Charles, F., 1997b, Using expectations data to study subjective income expectations, *Journal of the American Statistical Association*, Vol. 92, No. 439, pp. 855-867.

Eurostat, 2004, *Business Demography in Europe. Results for 10 Member States and Norway*, European Communities, Luxembourg.

Hakes, Jahn, K. and Viscusi, Kip, W., 1997, Mortality risk perceptions: A Bayesian reassessment, *Journal of Risk and Uncertainty*, Vol. 15, pp. 135-150.

Juster, T., 1966, Consumer buying intentions and purchase probability: An experiment in survey design, *Journal of the American Statistical Association*, Vol. 61, pp. 658-696.

Kahneman, Daniel, and Tversky, Ames, 1979, Prospect theory: An analysis of decision under risk, *Econometrica*, Vol. 47, pp. 263-291.

Landier, Augustin and Thesmar, David, 2003, *Financial contracting with optimistic entrepreneurs: Theory and evidence*, mimeo, University of Chicago.

Manski, Charles, F., 2004, Measuring expectations, *Econometrica*, Vol. 72, No. 5, pp. 1329-1376.

Papke, Leslie, E. and Wooldridge, Jeffrey, M., 1996, Econometric methods for fractional response variables with an application to 401(k) plan participation rates, *Journal of Applied Econometrics*, Vol. 11, pp. 619-632.

Pinfold, John, F., 2001, The expectations of new business founders: The New Zealand case, *Journal of Small Business Management*, Vol. 39, No. 3, pp. 279-285.

Rabin, Matthew, 1998, Psychology and economics, *Journal of Economic Literature*, Vol. 36, pp. 11-46.

Rabin, Matthew, 2002, Inference by believers in the law of small numbers, *Quarterly Journal of Economics*, Vol. 117, No. 3, August, pp. 775-816.

Tversky, Ames, and Kahneman, Daniel, 1981, The framing of decisions and the rationality of choice, *Science*, Vol. 211, pp. 453-458.

van den Steen, Eric, J., 2004, Rational overoptimism (and other biases), *American Economic Review*, Vol 94, No. 4, pp. 1141-1151.

Viscusi, Kip, W., 1997, Alarmist decision with devergent risk information, *Economic Journal*, Vol. 107, November, pp. 1657-1670.

Viscusi, Kip, W. and Evans, William, N., 1998, Estimation of revealed probabilities and utility functions for product safety decisions, *Review of Economics and Statistics*, Vol. 80, No. 1, pp. 28-33.

Viscusi, Kip, W. and O'Connor, Charles, J., 1984, Adaptive responses to chemical labeling: Are workers Bayesian decision makers? *American Economic Review*, Vol. 74, No. 5, pp. 942-956.

White, Halbert, L. Jr., 1980, A heteroscedasticity consistent covariance matrix estimator and a direct test of heteroscedasticity, *Econometrica*, Vol. 48, pp. 817-818.

Table 1. Frequency of optimism

PANEL A: SUMMARY STATISTICS

	Mean	Median	S.D.	Min	Max	N
IFAIL	13.434	5	16.919	0	80	272
OTHERSFAIL	25.665	20	18.723	0	80	272

PANEL B: FREQUENCY DISTRIBUTIONS

	IFAIL			OTHERSFAIL		
	Freq.	Percent	Cum.	Freq.	Percent	Cum.
0-10	179	65.81	65.81	96	35.29	35.29
11-20	43	15.81	81.62	41	15.07	50.37
21-30	18	6.62	88.24	50	18.38	68.75
31-40	2	0.74	88.97	21	7.72	76.47
41-50	25	9.19	98.16	54	19.85	96.32
51-60	1	0.37	98.53	6	2.21	98.53
61-70	2	0.74	99.26	2	0.74	99.26
71-80	2	0.74	100.00	2	0.74	100.00
Total	272	100.00		272	100.00	

Notes: IFAIL is an entrepreneur's estimate of the probability that his/her venture fails during the next three years. OTHERSFAIL is an entrepreneur's estimate of the three-year failure probability of other entrepreneurs' ventures that operate in the same line of business and were established at the same time as his/hers.

Table 2. Descriptive statistics for the indicators of absolute and relative optimism

PANEL A: SUMMARY STATISTICS

	Mean	Median	S.D.	Min	Max	N
AOPT	0.886	1	0.318	0	1	272
ROPT	12.232	10	19.452	-40	60	272

PANEL B: FREQUENCY DISTRIBUTIONS

	Freq.	Percent	Cum.
AOPT = 0	31	11.40	11.40
AOPT = 1	241	88.60	100.00
ROPT < 0	39	14.34	14.34
ROPT = 0	63	23.16	37.50
ROPT > 0	170	62.50	100.00

Notes: AOPT = 1 if IFAIL < 40% and zero otherwise. ROPT = OTHERSFAIL – IFAIL.

Table 3. Covariates of absolute optimism

	Dependent variable AOPT Logit estimates							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
RETIME1	0.020 (0.029)				0.025 (0.033)			
LOTTERY		0.004 (0.007)				0.002 (0.007)		
RISKLOVE			0.093 (0.382)				0.066 (0.407)	
PROB				0.624 (0.384)				0.734 * (0.428)
ACADEMIC					-2.457 ** (1.134)	-2.488 ** (1.130)	-2.491 ** (1.131)	-2.691 ** (1.143)
COLLEGE					-2.081 ** (1.053)	-2.080 ** (1.049)	-2.092 ** (1.052)	-2.049 * (1.053)
AGE					-0.032 (0.020)	-0.029 (0.020)	-0.030 (0.020)	-0.025 (0.020)
FEMALE					-0.177 (0.423)	-0.126 (0.430)	-0.159 (0.424)	-0.068 (0.424)
SINGLE					0.668 (0.614)	0.692 (0.613)	0.673 (0.621)	0.870 (0.627)
NCHILD					0.176 (0.169)	0.158 (0.170)	0.162 (0.169)	0.147 (0.172)
WORKED					0.464 (0.450)	0.470 (0.450)	0.462 (0.453)	0.425 (0.457)
EXPERIENCE					0.070 ** (0.035)	0.070 ** (0.034)	0.070 ** (0.034)	0.068 * (0.035)
CONSTANT	1.740 *** (0.473)	1.992 *** (0.215)	2.002 *** (0.275)	1.721 *** (0.263)	3.826 *** (1.349)	4.078 *** (1.297)	4.129 *** (1.284)	3.587 *** (1.325)
Observations	272	272	272	272	272	272	272	272
Wald	0.549	0.359	0.060	2.650	17.209 **	16.700 *	16.630 *	19.578 **
Pseudo R^2	0.003	0.002	0.000	0.014	0.089	0.087	0.086	0.101
Log likelihood	-96.214	-96.309	-96.459	-95.163	-87.884	-88.139	-88.173	-86.700

Notes: Standard errors are presented in the parentheses. The superscript asterisks indicate the statistical significance of coefficients and test statistics: (***) denotes significance at 1 percent level, (**) at 5 percent level and (*) at 10 percent level.

Table 4. Covariates of relative optimism

	Dependent variable ROPT OLS estimates with robust standard errors							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
RETIME1	0.011 (0.130)				-0.006 (0.131)			
RETIME2	0.053 (0.133)				0.022 (0.136)			
LOTTERY		0.032 (0.029)				0.028 (0.033)		
RISKLOVE			2.738 (2.360)				3.329 (2.315)	
PROB				0.864 (2.468)				1.762 (2.675)
ACADEMIC					-8.989 ** (4.472)	-8.910 ** (4.465)	-9.115 ** (4.440)	-9.456 ** (4.522)
COLLEGE					-3.206 (2.994)	-3.165 (3.013)	-3.513 (3.009)	-3.176 (3.023)
AGE					0.173 (0.174)	0.182 (0.171)	0.178 (0.170)	0.184 (0.171)
FEMALE					-2.556 (2.530)	-2.080 (2.589)	-2.697 (2.501)	-2.384 (2.541)
SINGLE					0.169 (2.925)	-0.003 (2.913)	-0.669 (2.862)	0.573 (3.035)
NCHILD					0.030 (0.973)	-0.054 (0.950)	-0.112 (0.974)	1.73E-04 (0.975)
WORKED					-3.114 (2.992)	-3.199 (2.989)	-3.611 (2.973)	-3.146 (2.979)
EXPERIENCE					0.064 (0.190)	0.067 (0.191)	0.074 (0.191)	0.056 (0.189)
CONSTANT	11.081 *** (3.192)	11.663 *** (1.328)	10.762 *** (1.710)	11.723 *** (2.009)	12.027 * (6.369)	11.427 * (6.224)	11.277 * (6.167)	10.904 * (6.514)
Observations	272	272	272	272	272	272	272	272
Wald	0.095	1.168	1.346	0.123	0.801	0.946	1.222	0.919
R^2	0.001	0.004	0.005	0.000	0.032	0.036	0.039	0.034

Notes: Robust standard errors are presented in the parentheses. The superscript asterisks indicate the statistical significance of coefficients and test statistics: (***) denotes significance at 1 percent level, (**) at 5 percent level and (*) at 10 percent level. Wald test is based on heteroskedastic-robust variance matrix.

Table 5. Estimates of Bayesian learning model

	Dependent variable IFAIL_UD OLS estimates with robust standard errors Sample base rate							
	"20%"	"80%"	"20%"	"80%"	"20%"	"80%"	"20%"	"80%"
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IFAIL	0.620 ***	0.791 ***	-0.225	1.389 ***	0.596 ***	0.894 ***		1.155 ***
	(0.074)	(0.087)	(0.279)	(0.165)	(0.089)	(0.102)		(0.085)
AOPT			-0.473 ***	0.310 **			-0.347 ***	0.184 ***
			(0.166)	(0.132)			(0.039)	(0.063)
KAOPT			0.989 ***	-0.262			0.733 ***	
			(0.301)	(0.188)			(0.090)	
ZERO			0.009	0.065 *	-8.78E-04	0.049		0.069 *
			(0.014)	(0.037)	(0.016)	(0.037)		(0.037)
DROPT					-2.35E-04	0.040		
					(0.016)	(0.028)		
KDROPT					0.155	0.051		
					(0.195)	(0.149)		
CONSTANT	0.029 ***	0.042 ***	0.488 ***	-0.311 **	0.025 *	-0.011	0.367 ***	-0.189 ***
	(0.007)	(0.016)	(0.166)	(0.132)	(0.013)	(0.022)	(0.038)	(0.069)
Observations	158	114	158	114	158	114	158	114
Wald	70.115 ***	82.499 ***	33.655 ***	3431.22 ***	25.653 ***	41.404 ***	64.620 ***	92.159 ***
R^2	0.630	0.561	0.677	0.601	0.640	0.580	0.674	0.599
RESET	6.290 ***	3.570 **	0.950	0.470	6.230 ***	2.870 **	0.720	0.470

Notes: Robust standard errors are presented in the parentheses. The superscript asterisks indicate the statistical significance of coefficients and test statistics: (***) denotes significance at 1 percent level, (**) at 5 percent level and (*) at 10 percent level. RESET reports Ramsey's regression specification error test for omitted variables. Wald test is based on heteroskedastic-robust variance matrix. In column (4), the very large Wald test statistic is a bit spurious, for it reduces to 40.98 if the non-robust (conventional) variance matrix is used.

E L I N K E I N O E L Ä M Ä N T U T K I M U S L A I T O S (ETLA)
THE RESEARCH INSTITUTE OF THE FINNISH ECONOMY
LÖNNROTINKATU 4 B, FIN-00120 HELSINKI

Puh./Tel. (09) 609 900 Telefax (09) 601753
Int. 358-9-609 900 Int. 358-9-601 753
http://www.etla.fi

KESKUSTELUAIHEITA - DISCUSSION PAPERS ISSN 0781-6847

Julkaisut ovat saatavissa elektronisessa muodossa internet-osoitteessa:
http://www.etla.fi/finnish/research/publications/searchengine

No 957 DEREK C. JONES – PANU KALMI – MIKKO MÄKINEN, The Determinants of Stock Option Compensation: Evidence from Finland. 01.12.2004. 32 p.

No 958 EDVARD JOHANSSON, Job Satisfaction in Finland – Some results from the European Community Household panel 1996-2001. 01.12.2004. 46 p.

No 959 HANNU PIEKKOLA – ANNI HEIKKILÄ, Active Ageing and Pension System: Finland. 07.12.2004. 35 p.

No 960 ANTTI KAUHANEN, Yrittäjien ansiot, työajat ja työkuormitus: Selvitys ekonomien ja insinöörien yrittäjyydestä. 09.12.2004. 22 s.

No 961 ANNI HEIKKILÄ, The Regional Distribution of Professional Competence in Finland. 16.12.2004. 20 p.

No 962 KARI E.O. ALHO, A Gravity Model under Monopolistic Competition. 31.12.2004. 15 p. Revised version 18.02-2005. 17 p.

No 963 KARI E.O. ALHO – VILLE KAITILA – MIKA WIDGRÉN, Speed of Convergence and Relocation: New EU Member Countries Catching up with the Old. Original version31.12.2004. 20 p., This version 23.05.2005. 21 p.

No 964 MAIJA GAO – ARI HYYTINEN – OTTO TOIVANEN, Demand for Mobile Internet: Evidence from a Real-World Pricing Experiment. 11.01.2005. 39 p.

No 965 MIKA MALIRANTA, Foreign-owned firms and productivity-enhancing restructuring in Finnish manufacturing industries. 19.01.2005. 21 p.

No 966 CHRISTOPHER PALMBERG – MIKA PAJARINEN, Determinants of Internationalisation through Strategic Alliances – Insights Based on New Data on Large Finnish Firms. 28.01.2005. 22 p.

No 967 OLLI-PEKKA RUUSKANEN, Ajankäytön muutosten vaikutus työllistymishalukkuuteen. 01.02.2055. 21 s.

No 968 SERGEY BOLTRAMOVICH – VLADISLAV YURKOVSKY – PAVEL FILIPPOV – HANNU HERNESNIEMI, Russian Infrastructure Clusters. A Preliminary Study. 01.02.2005. 67 p.

No 969 PEKKA SULAMAA – MIKA WIDGRÉN, Economic Effects of Free Trade between the EU and Russia. Original version 22.02.2005, this version 23.05.2005. 14 p.

No 970 HANNU HERNESNIEMI – KATI JÄRVI – JARI JUMPPONEN – GRIGORI DUDAREV – TAUNO TIUSANEN, Itäisen Suomen ja Venäjän liiketaloudellisen yhteistyön mahdollisuudet. 04.03.2005. 49 s.

No 971 JYRKI ALI-YRKKÖ – MONIKA JAIN, Offshoring Software Development – Case of Indian Firms in Finland. 07.03.2005. 14 p.

No 972 HANNU PIEKKOLA, Knowledge Capital as the Source of Growth. 17.03.2005. 35 p.

No 973 PEKKA YLÄ-ANTTILA – CHRISTOPHER PALMBERG, The Specificities of Finnish Industrial Policy – Challenges and Initiatives at the Turn of the Century. 29.03.2005. 25 p.

No 974 TUOMAS MÖTTÖNEN, Talouspoliittisen päätöksenteon tietoperustat. Esimerkkinä yritys- ja pääomaverouudistus. 29.03.2005. 90 s.

No 975 JYRKI LESSIG, Suhdannevaihteluiden symmetriaa kultakannan aikana. Ruotsin modernisoituminen, ulkomaankauppa ja taloudellinen integraatio 1800-luvun eurooppalaisten valuuttaliittojen aikana. 31.03.2005. 56 s.

No 976 SAMI NAPARI, Occupational Segregation during the 1980s and 1990s – The Case of Finnish Manufacturing. 18.04.2005. 54 p.

No 977 JYRKI ALI-YRKKÖ – ANTHONY DE CARVALHO – PAAVO SUNI, Intia maailmantaloudessa. 03.06.2005. 31 s.

No 978 RAINE HERMANS – MARTTI KULVIK – ANTTI-JUSSI TAHVANAINEN, ETLA 2004 Survey on the Finnish Biotechnology Industries – Background and Descriptive Statistics. 22.04.2005. 40 p.

No 979 ELIAS OIKARINEN, The Diffusion of Housing Price Movements from Centre to Surrounding Areas. 25.04.2005. 36 p.

No 980 JYRKI ALI-YRKKÖ, Impact of Public R&D Financing on Employment. 06.05.2005. 24 p.

No 981 MAARIT LINDSTRÖM, Onko luovilla aloilla taloudellista merkitystä? Luovat alat, kulttuurialat ja taidekoulutetut eri toimialoilla. 19.05.2005. 26 s.

No 982 MARTTI NYBERG – MAARIT LINDSTRÖM, Muotoilun taloudelliset vaikutukset. 20.05.2005. 25 s.

No 983 NIKU MÄÄTTÄNEN, Vapaaehtoiset eläkevakuutukset, verotus ja säästäminen. 24.05.2005. 31 s.

No 984 TUOMO NIKULAINEN – MIKA PAJARINEN – CHRISTOPHER PALMBERG, Patents and Technological Change – A Review with Focus on the Fepoci Database. 25.05.2005. 26 p.

No 985 PEKKA SULAMAA – MIKA WIDGRÉN, Asian Regionalism versus Global Free Trade: A Simulation Study on Economic Effects. 27.05.2005. 12 p.

No 986 EDVARD JOHANSSON – PETRI BÖCKERMAN – RITVA PRÄTTÄLÄ – ANTTI UUTELA, Alcohol Mortality, Drinking Behaviour, and Business Cycles: Are Slumps Really Dry Seasons? 16.06.2005. 10 p.

No 987 ARI HYYTINEN – MIKA PAJARINEN, Why Are All New Entrrepreneurs Better than Average? Evidence from Subjective Failure Rate Expectations. 23.06.2005. 34 p.

Elinkeinoelämän Tutkimuslaitoksen julkaisemat "Keskusteluaiheet" ovat raportteja alustavista tutkimustuloksista ja väliraportteja tekeillä olevista tutkimuksista. Tässä sarjassa julkaistuja monisteita on mahdollista ostaa Taloustieto Oy:stä kopiointi- ja toimituskuluja vastaavaan hintaan.

Papers in this series are reports on preliminary research results and on studies in progress. They are sold by Taloustieto Oy for a nominal fee covering copying and postage costs.