Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Entrepreneurial Risk and Market Entry

Brian Wu, Anne Marie Knott,

MANAGEMENT SCIENCE
Vol. 52, No. 9, September 2006, pp. 1315–1330
ISSN 0025-1909 | EISSN 1526-5501 | 06 | 5209 | 1315

informs®
DOI 10.1287/mnsc.1050.0543

Entrepreneurial Risk and Market Entry

Brian Wu
The Wharton School, University of Pennsylvania, 2000 SH-DH, 3620 Locust Walk,
Philadelphia, Pennsylvania 19104, wux@wharton.upenn.edu

Anne Marie Knott
Olin School of Business, Washington University, Campus Box 1133, One Brookings Drive,
St. Louis, Missouri 63130-4899, knott@wustl.edu

This paper attempts to reconcile the risk-bearing characterization of entrepreneurs with the stylized fact that entrepreneurs exhibit conventional risk-aversion profiles. We propose that the disparity arises from confounding two distinct dimensions of uncertainty: demand uncertainty and ability uncertainty. We further propose that entrepreneurs will be risk averse with respect to demand uncertainty, yet "apparent risk seeking" (or overconfident) with respect to ability uncertainty. To examine this view, we construct a reduced-form model of the entrepreneur's entry decision, which we aggregate to the market level, then test empirically. We find that entrepreneurs in aggregate behave as we predict. Accordingly, risk-averse entrepreneurs are willing to bear market risk when the degree of ability uncertainty is comparable to the degree of demand uncertainty. Potential market failures exist in instances where there is a high demand uncertainty but low performance dispersion (insufficient entry), or low demand uncertainty but high performance dispersion (excess entry).

Key words: entry; risk; entrepreneur; uncertainty; overconfidence
History: Accepted by Pankaj Ghemawat, business strategy; received September 2, 2004. This paper was with the authors 3 months for 2 revisions.

1. Introduction

Throughout its history the literature on entrepreneurship has asserted that a critical economic role of the entrepreneur is risk bearing. One consequence of that perspective is that the theoretical and practitioner literature has assumed that entrepreneurs are risk seeking. To date, however, the empirical record indicates that entrepreneurs' risk profiles are indistinguishable from those of wage earners.

We believe that the disparity between intuition and theory versus the stylized facts lies in the dimensionality of uncertainty. In particular, we propose that there are two distinct sources of uncertainty in entrepreneurial ventures: (1) uncertainty regarding market demand, and (2) uncertainty regarding one's own entrepreneurial ability. We further propose that entrepreneurs display risk aversion with respect to demand uncertainty but exhibit overconfidence or "apparent risk seeking" with respect to ability uncertainty. Accordingly, while entrepreneurs are risk averse in the classic sense of preferring a certain payment to an uncertain payment with equivalent expected value, their overconfidence predisposes them to bear economic risk under a given set of circumstances.

Both dimensions of uncertainty have been proposed previously. Indeed, March and Shapira (1987) draw the contrast between "managerial gambling" (undertaking exogenous risks) and "risk taking" (undertaking risks over which managers believe they have some control). The contribution of this paper is to consider the two dimensions jointly in modeling and testing entrepreneurs' entry behavior. To conduct the empirical test, we characterize both dimensions of uncertainty across a set of equivalent markets over time, then estimate the degree to which aggregate entry responds to each dimension.

It is important to point out that while we model the individual's decision to enter markets, the primary purpose of this study is not to identify *who* in the employment pool will enter markets. Rather, we want to understand how individuals' decisions combined with market conditions affect the amount of entry. In this sense, our study follows in Blau's (1987) steps, who models the individual's self-employment decision, but then utilizes it to examine the impact of tax and social security policies on the amount of self-employment. A separate but related strand of literature examines the *who* question. That literature uses longitudinal data to understand characteristics affecting self-employment, e.g., Evans and Leighton (1989) and Dunn and Holtz-Eakin (2000).

Our test is novel in the sense that it jointly tests personality traits affecting individuals' entry decisions while also testing the economic implications of the trait-entry relationship. We examine whether entrepreneurs in aggregate exhibit particular decision biases and what market conditions activate those biases. We do not, however, assess whether the biases (if any) of entrepreneurs differ from those of wage earners.

2. Entrepreneurship and Demand Uncertainty

A fairly well-established theme in the entrepreneurial literature is that a key economic role of entrepreneurs is risk bearing. This view dates back to Cantillon (1755) who characterized the economy as consisting of two classes of inhabitants (aside from the prince and landowners): "hired people" on fixed wages and "undertakers" who purchase inputs (including labor) at fixed prices without assurance of profits. The key distinguishing feature of the second class is that it *undertakes* the risk of demand and price uncertainty (which at the time of his writing must have been quite high because one of the factors Cantillon considered was the number of deaths of local inhabitants). Included in the undertaker (entrepreneur) class were farmers, merchants, shopkeepers, and master craftsmen (even robbers).

This view was expounded by Knight (1921) and, in fact, the view of entrepreneur as risk taker is probably most associated with Knight. Knight's contribution was to draw a distinction between *risk*, which involves recurring events whose relative frequency can be known from past experience, and *uncertainty*, arising from unique events which can only be subjectively estimated. Risk is considered to be a relatively insignificant problem in that it can be accommodated through pooling and insurance. In contrast, uncertainty requires an economic functionary, the entrepreneur, whose job it is to decide what to do and how to do it in the face of uncertainties. Knight proposed that there is diversity among individuals with regard to confidence in one's judgment and the disposition to act on those judgments. Those who are "confident and venturesome 'assume the risk' or 'insure' the doubtful and timid by guaranteeing to the latter a specified income in return for an assignment of the actual results" (p. 269).

Anticipating later work by Camerer and Lovallo (1999), Knight asserts that "if men are poor judges of their own powers as well as ignorant of those of others (entrepreneurs), the size of the profit share depends on whether they tend on the whole to overestimate or underestimate the prospects of business operations" (p. 285).

Indeed, entrepreneurs do bear greater risk than wage earners. First, they bear income risk in that the stream of income from new ventures is uncertain. In the worst case, the firm fails and the income stream ceases. This risk of failure is considerable. Data from the U.S. Census Business Information Tracking Series (Headd 2003) indicate that 51% of firms exit within their first four years. Thus, failure risk looms large for entrants and is relatively nonexistent for established firms. An additional but related risk pertains to invested capital. While shareholders can minimize risk by diversifying their holdings precisely as Knight (1921) suggests, entrepreneurs typically must invest the bulk of their wealth in a single asset, the venture.

Given that entrepreneurs perform a risk-bearing role, most theoretical literature has assumed that entrepreneurs have greater risk tolerance than wage earners (McClelland 1961, Lucas 1978, Kanbur 1979, Kihlstrom and Laffont 1979). Empirical literature has emerged to test this inference. The surprising result has been that entrepreneurs do not appear to differ from wage earners on this trait. In fact, where there are differences they tend to indicate that entrepreneurs exhibit greater risk aversion than wage earners (Brockhaus 1980, Masters and Meier 1988, Miner and Raju 2004).

One of the reasons studies may fail to find risk tolerance is that the instruments used to test risk vary across studies and each instrument operationalizes risk quite differently. Brockhaus (1980) and Masters and Meier (1988) use a Kogan-Wallach Choice Dilemma Questionnaire (CDQ) which asks respondents what success threshold a given action would require before they would recommend the action to someone else. Both studies found no significant difference between entrepreneurs and managers on the recommended thresholds. Of course, one possible reason for this result is that people may be willing to undertake risk themselves while not recommending it for others. However, results from 12 studies using a different instrument, the Miner Sentence Completion Scale (MSCS), all found entrepreneurs to be risk avoiding relative to managers (Miner and Raju 2004). In another study, Sarasvathy et al. (1998) found that entrepreneurs are risk averse relative to bankers in that they trade higher expected value projects for ones with narrower variance (particularly avoiding negative outcomes).

Results from economic tests are similarly equivocal. Cramer et al. (2002) compare individuals' valuations for a lottery ticket and find that subjects who had ever been self-employed exhibit greater risk tolerance than wage earners even after controlling for wealth effects (the self-employed tend to have greater wealth and therefore bear less relative risk than wage earners). In contrast, O'Brien et al. (2003) examine the impact of demand uncertainty on entrepreneurial entry across 57 industries and find that demand uncertainty has a significant negative impact on entry. Mazzeo (2004) compares an entrepreneur's choice between operating an independent establishment (sole ownership) and becoming a franchisee (sharing risk with the franchisor) in the same market. He finds that franchising increases with the local demand uncertainty, thus suggesting risk aversion.

In summary, while there is some evidence to the contrary, the weight of the empirical evidence tends to indicate that entrepreneurs have comparable risk profiles to those of wage earners. This leaves the question of what accounts for their willingness to bear the high risk of failure. We believe the answer to that question lies in a second dimension of uncertainty.

3. Entrepreneurship and Performance Dispersion

While the most prevalent view of entrepreneurial risk pertains to classic risk aversion in that an individual is more interested in a certain payoff than an equivalent expected value from an uncertain payoff, a second view emerges from the work of Knight (1921), Kanbur (1979), Jovanovic (1982), Lippman and Rumelt (1982), March and Shapira (1987), Camerer and Lovallo (1999), and Norton and Moore (2002). In this view, a critical source of economic uncertainty pertains to entrepreneurial ability. For example, entrepreneurs know the performance distribution within a market but are uncertain about where they lie within that distribution.

Performance dispersion has two potential effects on entry: an options effect and an overconfidence effect. We discuss each of these in turn. High degrees of performance dispersion will lead to apparent risk seeking even if entrepreneurs are risk neutral (Lippman and Rumelt 1982). This occurs because entrepreneurial ventures have an options structure. The upside returns are uncertain, but the downside risk is limited by the entry fee. Accordingly, greater profit dispersion leads to higher expected values net of the entry fee. If entrepreneurs base decisions on net present values, risk-neutral entrepreneurs will appear to be risk seeking in that they will rationally prefer more disperse performance distributions.

A more provocative view finds that under certain circumstances, entrepreneurs exhibit "apparent risk seeking" with respect to ability uncertainty. This arises from overestimating their capability. In this view, entrepreneurs know the performance distribution and expected value of entry within a given market but they believe their own ability is drawn from a narrower and biased distribution (March and Shapira 1987, Zenger 1994, Busenitz 1999, Camerer and Lovallo 1999, Norton and Moore 2002). Indeed, this tendency was anticipated by Knight's (1921) insight that the distinguishing feature of entrepreneurs is their level of confidence in being able to handle unforeseeable events. Accordingly, it is not that entrepreneurs have greater risk tolerance; it is that they do not view the business situations as being risky (March and Shapira 1987, Palich and Bagby 1995, Sarasvathy et al. 1998). Instead, they believe that through skill and information they can condition the odds they face.

This bias is related to the "Lake Wobegon" or "better-than-average" effect (Alicke et al. 1995), where Lake Wobegon refers to Garrison Keillor's fictional hometown in Minnesota where "all the children are above average." A frequently cited example of the effect is that 60% of high-school seniors believe they are in the top 10% in ability to get along with others, while 25% believe they are in the top 1% (Myers 1982). If we extrapolate these numbers, it appears that the entire population of seniors believes it is in the top 20%.

A more pertinent example is engineers' assessment of their own performance relative to peers (Zenger 1994). That study found that in established companies, 32.4% of engineers believed they were in the top 5% of peer performance; while 61.8% believed they were in the top 10% and fully 89.7% believed they were in the top 25%. These biases were even more pronounced for engineers in entrepreneurial firms where self-assessments within the top 5%, 10%, and 25% of peers were, respectively, 42.0%, 73.3%, and 92.5%. Thus, while all engineers are prone to overconfidence, those drawn to start-ups are particularly overconfident.

Although the better-than-average effect has been demonstrated for both traits and behaviors, the bias is more likely for traits that are perceived to be controllable, e.g., fairness rather than intelligence (Alicke 1985, Allison et al. 1989); or in cases of greater ambiguity. For example, the bias is more likely when individuals compare themselves to an anonymous distribution rather than to a known set of individuals (Alicke 1985, Alicke et al. 1995, Allison et al. 1989). This qualitative characterization suggests a bias tied to dispersion of the trait in the referent population. Thus, if μ and σ capture, respectively, the mean and standard deviation of ability, unbiased entrepreneurs without private information should expect an ability draw equal to μ. If, however, entrepreneurs exhibit overconfidence (the better-than-average effect), their expected ability will have the form $\mu + A\sigma$, where the bias $A\sigma$ will depend on perceived controllability captured by A, and the degree of ambiguity captured by σ.

Considering the two types of uncertainty (demand uncertainty and ability uncertainty) jointly suggests entrepreneurs could appear to be risk seeking when in fact they are not. In particular, they may be risk averse with respect to demand uncertainty but overconfident with respect to ability uncertainty. If so, there may be settings where they will bear economic risk associated with uncertain demand so long as there is sufficient ability uncertainty.

It is worth noting that we do not postulate that entrepreneurs differ from wage earners on either dimension. Indeed, the empirical record discussed previously indicates that entrepreneurs are indistinguishable from wage earners with regard to risk aversion. In addition, we have just shown that high-school seniors (the pool from which most entrepreneurs and wage earners are drawn) are overconfident. All we require to reconcile entrepreneurs' risk aversion with their willingness to bear risk is that these dimensions affect entry behavior and that there is heterogeneity along some dimension such that not all individuals jump from wage employment to entrepreneurship at the same time.

4. A Model of Entrepreneurial Entry

4.1. Qualitative Description of the Entrepreneurial Entry Decision

There is substantial literature within management and economics regarding the decision to become an entrepreneur. Taken together, the literature suggests that individuals can be characterized as having underlying propensities to become entrepreneurs. These propensities are driven by human capital such as wealth and education (Dunn and Holtz-Eakin 2000, Hamilton 2000), demographic characteristics such as age, marital status, and whether your parents were themselves entrepreneurs (Dunn and Holtz-Eakin 2000), and the quality of alternative options (Lofstrom 2002). In addition to individual propensities, there are exogenous shocks that often trigger entrepreneurship such as divorce, graduation, termination from a prior job (Wooten et al. 1999), or working for a firm that becomes acquired (Stuart and Sorenson 2003). Finally, economic factors affect the level of entrepreneurship, e.g., tax rate, the unemployment rate, and the level of technology (Blau 1987).

The net effect of these factors is that at any given time typically 12%–13% of the nonagricultural work force is self-employed. However, there is considerable movement in and out of self-employment. Typically 2% to 3% of the workforce becomes entrepreneur each year and, as a result, over their entire career approximately 30% of the workforce tests the entrepreneurial waters at some point. These transitions occur both because individuals' human capital and demographic status are changing over time, and because their alternatives are changing over time.

Accordingly, a reasonable way to envision the self-employment decision is that individuals continually assess whether they have higher utility in conventional wage employment or self-employment. This utility comparison is meaningless, however, without the context of an actual market to enter. In most cases, this is precisely the market (industry plus geographic location) of the individual's prior employer. This makes sense. Individuals accrue experience, market knowledge, and affiliations during the course of their employment. The returns of this human and social capital are probably highest in the market in which they were accrued. Over time individuals observe market trends (the performance of all firms in the market) as well as their own performance relative to peers (both within the firm and within the industry). These observations offer noisy signals about underlying demand and entrepreneurial ability. At any given time, an individual remains in wage employment if his or her wage is higher than the expected entrepreneurial profits in that market, given the current estimates about the market and firms' abilities.

4.2. Reduced-Form Model of the Entry Decision

We can formalize the decision process just described using the basic entry decision model from the self-employment literature. The model examines an individual's choice between the expected value of entrepreneurial profits Π, and that of wage income W, given uncertainty about own ability (captured as marginal cost).

4.2.1. Risk Neutrality. We begin with a model of risk-neutral and unbiased behavior. The functional form for profits Π follows Lippman and Rumelt (1982):

$$E(\Pi) = -K + \frac{1}{r}\int_0^{c_0} q(p-c)\,dF(c), \quad \text{where} \tag{1}$$

q = firm output,
p = market price,
K = investment required to produce quantity q,
r = discount rate,
c = firm's marginal cost, drawn from distribution F, and
c_0 = largest solution to $p - c = 0$.

Prospective entrepreneurs compare the expected profit from their proposed venture to the alternative income stream from wage employment. Entry occurs when the expected profit stream exceeds the entrepreneur's opportunity cost (which comprises the sunk cost of entry K, plus the foregone wage stream from the best employment alternative W), as follows:

$$\int_0^{c_0} q(p-c)\,dF(c) > r(W+K). \tag{2}$$

We assume that $F(c)$ is normally distributed and, thus, can be fully characterized by its mean μ and dispersion σ. Accordingly, we can write a simple expression for the expected value of annual profits $E(\Pi)_t$, given the mean cost μ_s of surviving firms (those for whom $c \leq c_0$) and the failure rate λ, representing the entrepreneur's ex ante probability of drawing a value for $c > c_0$. This expression characterizes the truncated distribution of long-run profits, where λ firms are forced to exit ($c > c_0$) and obtain zero long-run profits, and $(1-\lambda)$ firms survive ($c \leq c_0$) and obtain $q(p-c)$ in perpetuity:

$$E(\Pi)_t = \lambda * 0 + (1-\lambda) * q(p-\mu_s) = q(p - \lambda p - \mu_s + \lambda * \mu_s). \quad (3)$$

Equation (3) implicitly captures the options structure of payoffs, which is a competing explanation to overconfidence for a positive relationship between entry and performance dispersion. What we mean by "options structure" is that with limited downside risk (losses are limited to the entry fee, K), expected net profits increase with dispersion in performance (marginal cost).

A potential empirical challenge is that the options effect and the overconfidence effect are both driven by $F(c)$. Thus, we may be unable to tease apart the two effects. Fortunately, Equation (3) fully captures the options effect associated with cost dispersion using only failure rate, λ, and the mean cost of surviving firms, μ_s. In other words, even though cost dispersion affects the entry decision of a risk-neutral entrepreneur, we do not need a direct measure of cost dispersion to capture the options effects. This allows us to use cost dispersion σ to capture overconfidence effects without their being confounded by options effects. Accordingly, we can distinguish between the two factors (options effects and overconfidence) that produce "apparent risk seeking." Replacing the expression for the left-hand side of Equation (2) yields the following entrepreneurial decision rule:

$$p(\text{entry}) = \Pr[q(p - \lambda p - \mu_s + \lambda * \mu_s) - r(W+K) > 0]. \quad (4)$$

Equation (4) captures individual propensity to enter, p(entry). We want to examine aggregate entry at the market level, however. This is for two reasons. First, we do not have data on individual characteristics. While there are longitudinal databases that include individual characteristics, such as the National Longitudinal Survey (NLS), these do not have industry-specific measures of overconfidence. Industry-specific measures are necessary because although I may be predisposed to overconfidence, I am not overconfident where I lack competence, e.g., medicine. More important than the data limitation is the fact that we are fundamentally interested in the aggregate economic impact of individual biases: How do individual biases interact with market conditions to affect the level of entry? Fortunately, our market-level test of aggregate entry allows us to jointly test both individuals' biases and their economic impact.

To examine entry in market j, we aggregate individual propensity in Equation (4) over the pool of prospective entrants, n_j. The pool includes those with relevant industry experience together with sufficient wealth to cover the entry cost K:

$$\text{entry}_j = p(\text{entry}) * n_j = \Pr[q(p - \lambda p - \mu_s + \lambda * \mu_s) - r(W+K) > 0] * n_j. \quad (5)$$

To consider economic risk, we further assume that r is decomposed into a risk-free component r_f, and a risk premium r_m, reflecting market volatility (demand uncertainty). Note that if entrepreneurs are risk neutral, there should be no risk premium. Equation (5) together with the risk decomposition generates the following propositions regarding a risk-neutral entrepreneur:

Proposition 1. *Entry is decreasing in the failure rate* λ.

Proposition 2. *Entry is decreasing in the mean survivor cost* μ_s.

Proposition 3. *Entry is increasing in cross-term* $\lambda * \mu_s$.

Proposition 4. *Entry is decreasing in opportunity cost* $(W+K)$.

Proposition 5A. *Entry is insensitive to cost dispersion* σ.[1]

Proposition 6A. *Entry is insensitive to demand uncertainty* r_m.

4.2.2. Risk Aversion and Overconfidence. Equation (5) and the related propositions characterize unbiased and risk-neutral entry. In essence, these form our null hypotheses. We now wish to consider the risk biases that are the central concern of this paper. We first take the most straightforward bias, risk aversion with respect to demand uncertainty. The null hypothesis, Proposition 6A, is that entrepreneurs will be insensitive to demand uncertainty r_m. If, however, entrepreneurs are risk averse, then entry should be decreasing in r_m. This bias has already been demonstrated for entrepreneurial entry in 57 industries (O'Brien et al. 2003) and for entrepreneur choice of franchise versus independent ownership (Mazzeo 2004).

[1] Note that entry will be insensitive to σ *only* after controlling for the options effect through λ and μ_s.

The second dimension, overconfidence with respect to ability uncertainty, requires us to define a functional form for overconfidence. Following the discussion in §3, we assume that overconfidence takes the form $A\sigma$, where σ is ability (cost) dispersion, and A is the degree of overestimation with respect to the dispersion.[2] Prior studies suggest that A is higher for traits that are perceived to be controllable (Alicke 1985, Allison et al. 1989), and that entrepreneur ability is a trait that is believed to be controllable (March and Shapira 1987).

If entrepreneurs' decisions are biased in the manner discussed previously (risk averse with respect to market uncertainty and overconfident with respect to ability uncertainty), we replace Propositions 5A and 6A as follows:

Proposition 5B. *Entry is increasing in cost dispersion* σ *(overconfidence).*

Proposition 6B. *Entry is decreasing in demand uncertainty* r_m *(risk aversion).*

Remember that we already capture the contributions of cost dispersion associated with the options effect through the mean cost of surviving firms μ_s and the failure rate λ.

5. Empirical Approach

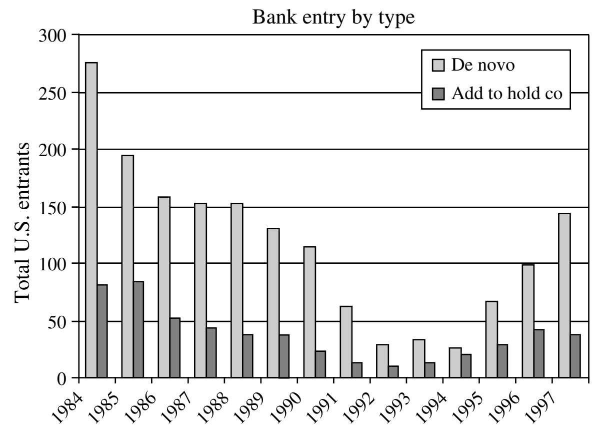
5.1. Industry

To test the propositions, we need a setting that (1) allows us to characterize the cost distribution over the full set of firms, (2) has substantial entry and provides reliable counts of that entry for firms of all sizes, and (3) comprises a large number of comparable markets that share common technology and common demand functions. We could find only one such industry—commercial banking post deregulation. The banking industry is ideal because it is fragmented with localized competition and is marked by significant de novo entry; also, the Federal Deposit Insurance Corporation (FDIC) collects complete cost data on all firms. Fragmentation allows us to compare discrete markets with a common demand function and common technology. Thus, we can compare variance in market conditions (demand volatility and cost dispersion) while controlling for other factors that differ across industries. We can also control for differences in the level of market demand through time-varying differences in market conditions.

In addition to the quasi-experimental advantages of banking for our purposes, banking is one of the most important industries in the economy. Financial services' output is roughly $2.1 trillion (7% of U.S. total). Furthermore, even though banking is one of the oldest industries, it has been growing at 6.5% annual rate over the past 10 years. Accordingly, a study of entry in banking is potentially important in its own right.

[2] In the empirical test, A will be captured by the coefficient on σ.

Figure 1 Trends in New Bank Charters by Bank Type



5.2. De Novo Entry in Banking

Banking does not typically come to mind when naming entrepreneurial industries. However, there is substantial entry in the banking industry over the period we examine (1984–1997), as shown in Figure 1. The figure also demonstrates that the dominant mode of entry is de novo start-ups by entrepreneurs, rather than expansion of existing bank holding companies. We are interested exclusively in the de novo start-ups.

One of the primary drivers behind de novo entry is the Community Reinvestment Act (U.S. Congress 1977), designed to encourage banks to meet the credit needs of local communities. The Act has spawned a growing industry of consultants, service providers, and websites dedicated to small and start-up banking (e.g., http://www.denovobanks.com). This "de novo bank support industry" facilitates new banks both with the start-up process and with ultimate operations. The consultants allow banks to compile necessary documentation (business plans and bank charter applications) and resources (human capital, physical capital, and financial capital) for start-up more quickly. The service providers allow community banks to compete more effectively with large banks. These providers offer online banking, check clearing, account maintenance, and statement processing on a fee basis. Accordingly, community banks can emulate large banks without costly investments in technology. This allows them to reach profitability more quickly, both through higher demand (from improved service offerings) and lower fixed cost (outsourced technology).

Perhaps one reason entrepreneurship seems incongruous with banking is because of industry regulation. One important aspect of that regulation is approval of new charters by either the respective state banking department or its federal counterpart, the Office of the Comptroller of Currency. In either case, banks also require approval from the FDIC for deposit insurance. The two major components of the application to the chartering agencies as well as the FDIC are the business plan and the summary of the management team. The business plan provides details for the bank's first three years of operations and must demonstrate that operations are consistent with market demand, customer base, competition, and general economic conditions, and that the bank will achieve acceptable leverage ratios by the end of three years. The management team summary must demonstrate that the team has substantial expertise in bank administration, commercial and consumer lending, bank operations, and investment/funds management. In particular, at least two members should have previous financial institution experience and the majority of directors must have close ties to the community (Federal Reserve Bank of Atlanta 2005).

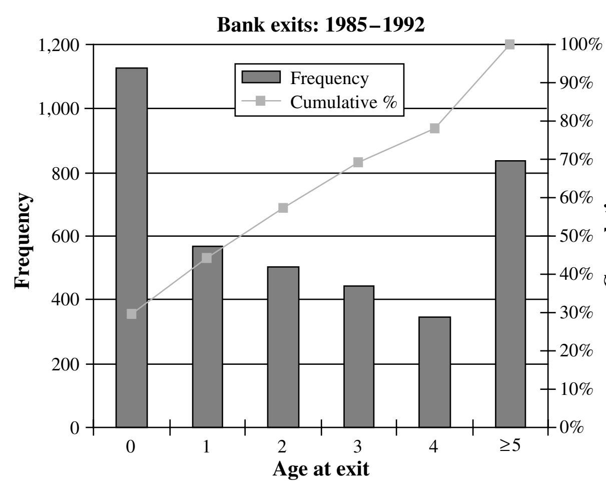
An interesting question is why there is any de novo entry in an industry that is otherwise experiencing dramatic consolidation. Interviews with industry experts reveal that consolidation actually provides the impetus for entry. Merger activity has three effects. First, it creates the "market void" that is required for approval of a new charter, which is typically small business lending (Goldberg and White 1998). Second, mergers create "liquidity events" (Stuart and Sorenson 2003), wherein displaced bank executives are both in search of new employment and flush with proceeds from the merger.[3] This displacement would appear as a supply shock in Equation (4) in that W drops to zero. The experience of these displaced executives and their financial assets supply the two other main criteria required for charter approval. Third, the expectation of future mergers provides a liquidation mechanism for the start-up banks, similar to the exit strategies sought by venture capitalists in the high-tech sectors. This liquidation potential appears as a demand shock in Equation (4) in that the expected returns include an acquisition premium above and beyond the present value of the remaining profit stream.

5.3. Bank Failure

The flip side of bank start-up is bank failure. While we are not interested in failure per se, it offers face validity for the overconfidence effect. An ideal measure of overconfidence would compare an entrepreneur's expected cost to the firm's realized cost. We do not, however, know the entrepreneur's expected cost. All we can infer is that all entrepreneurs anticipate having cost below the survivor threshold—they expect the new firm to survive. Thus, if we see failure, we can assume some form of error in the entry decision. Figure 2a demonstrates that there is substantial failure in banking.

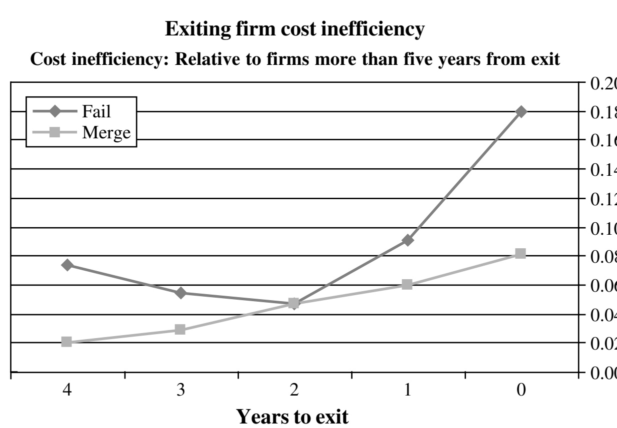
To further demonstrate that the error is overconfidence, we need to show that failure is associated with high cost (inefficiency). While this has been documented in the literature on banking (e.g., Berger and Humphrey 1992), we can also demonstrate it for the set of firms examined here. To do so, we take the cost efficiencies we derive in Stage 1 analysis (described in §§5.4 and 6). We then create distributed lag models for exiting firms to track their efficiency preceding failure. Figure 2b presents graphical results of

Figure 2a Substantial Failure in Banking



Note. Banks founded between 1985 and 1992 that exit by end of 1996. Exits include mergers and failures.

Figure 2b Failure Associated with Inefficiency (Excess Cost)



[3] The gains from liquidated ventures raise the issue of a "house money effect" (Thaler and Johnson 1990), wherein prior gains lead to risk-taking behavior. To date, this effect has been examined in the context of pure gambles rather than contests of skill, although a similar effect surfaces in the Camerer and Lovallo (1999) experiment, where confidence of players increases as they progress through a tournament despite the fact that the expected skill of players they will be facing is increasing across rounds. Accordingly, we would expect liquidation events to increase risk taking along both dimensions of uncertainty. Without controls for these events, results will understate the level of risk aversion along both dimensions.

this analysis.[4] The figure indicates that exits come from high-cost firms. Failing firms not only have higher cost than the industry overall (26% higher than an average bank of comparable scale),[5] but their cost deteriorates as they approach their exit year. While these effects hold for both failures and mergers, they are more pronounced for failures. Taken together, the failure rates and the cost data suggest that entrepreneurs are overconfident about their likely performance.

5.4. Empirical Model

Our empirical approach compares aggregate entry across markets to characterize entrepreneurs' behavior along each dimension of uncertainty. This allows us to examine if "apparent risk seeking" (overconfidence) with respect to ability uncertainty might compensate for risk aversion with respect to demand uncertainty. If so, risk-averse entrepreneurs may be willing to bear economic risk in certain markets. Note that our approach examines aggregate behavior in each market; thus, it ignores characteristics of individual entrepreneurs. Accordingly, we are unable to answer the question of whether entrepreneurs' biases differ from wage earners. We can only answer the question of whether entrepreneurs on average appear to be risk averse and/or overconfident.

Analysis proceeds in two stages. In the first stage, we model an industry cost frontier to collect measures of cost efficiency for each firm in each year. This allows us to characterize μ_s and σ for each market in each year. In the second stage, we model Equation (5) to test Propositions 5B and 6B that entrepreneurs are risk averse with respect to demand uncertainty and overconfident with respect to ability (cost) uncertainty.

5.4.1. Stage 1: Characterizing Firm Cost Efficiency. We follow convention in studies of bank efficiency by modeling a stochastic cost frontier using a translog cost function (Cebenoyan et al. 1992, Hermalin and Wallace 1994, Berger et al. 1993, Mester 1993). Stochastic frontier analysis, developed by Aigner et al. (1977), is based on the econometric specification of a cost frontier. The stochastic frontier model assumes that the log of firm i's cost in year t, c_{it}, differs from the cost frontier $c^{\min}$ by an amount that consists of two distinct components: a standard normally distributed error term e_{it}, and a cost inefficiency term modeled as a nonnegative random variable u_{it}, with a truncated normal distribution.[6]

We use the translog cost function to accommodate the complex array of bank inputs and outputs. In addition, the translog form accommodates trade-offs in both market strategies (product mixes and prices) and operational strategies (input mixes). The basic translog cost function mimics the linear programming problem for firm i in year t, that is minimizing total cost c_{it} by choice of output levels y_{it}, taking input prices w_{it} as given:

$$\begin{aligned} c_{it} = {} & \beta_0 + \sum_j \beta_{1j} y_{it}^j + \sum_k \beta_{2k} w_{it}^k + \frac{1}{2}\sum_j \sum_j \beta_{3ij} y_{it}^j y_{it}^j \\ & + \frac{1}{2}\sum_k \sum_k \beta_{4kk} w_{it}^k w_{it}^k + \sum_j \sum_k \beta_{5jk} y_{it}^j w_{it}^k \\ & + u_{it} + e_{it}, \quad \text{where} \end{aligned} \tag{6}$$

c_{it} = log observed firm cost,
y_{it}^j = vector of log output levels; j indexes output elements,
w_{it}^k = vector of log input prices; k indexes input elements,
u_{it} = cost inefficiency with truncated normal distribution, and
e_{it} = error term with normal distribution.

We pool data for all firms over 14 years using the stochastic frontier model to capture firm-year measures of cost inefficiency relative to a fixed global cost frontier. We collect the estimates of the expected value of firm-year cost inefficiency in Stage 1, $E(u_{it} \mid e_{it})$, which for convenience we continue to label as u_{it}, and use the estimates to form the cost distribution for each market in each year. The frontier analysis generates values that have a truncated normal distribution, whereas the model in Equation (3) assumes that cost is normally distributed. Accordingly, we transform values of u_{it} by taking their natural log, $\ln(u_{it})$.

5.4.2. Stage 2: Variables. To test the propositions derived from Equation (5), we need to make assumptions about how entrepreneurs assess demand uncertainty r_{jt}, the cost distributions μ_{sjt} and σ_{jt}, and the failure rate λ_{jt}, as well as opportunity costs for the entry fee K_{jt} and foregone wage W_{jt} in their respective markets j.

Demand uncertainty. Our measure of demand uncertainty follows the conventional approach (Stock and Watson 1998, Morgan et al. 2004), where risk is approximated by variability in demand. Demand uncertainty is measured as the root mean squared error (RMSE) from the regression $\ln[(\text{market demand}_t/\text{market demand}_{t-1}) - 1] = \alpha_0 + \alpha_1 * (\text{trend})$ over a moving 10-year window. We measure demand in two alternative ways. Our primary measure, *state personal income*, uses economic conditions as a proxy for demand. While *gross state product* (*GSP*) sounds

[4] Coefficient estimates are available from the authors.

[5] The cost inefficiency (in ln form) increases from the mean of 0.18 to 0.35. This corresponds to a 26% cost increase.

[6] Other distributional assumptions are also possible, the most common of which are the half-normal and exponential distributions. All results are robust to these alternative distributions.

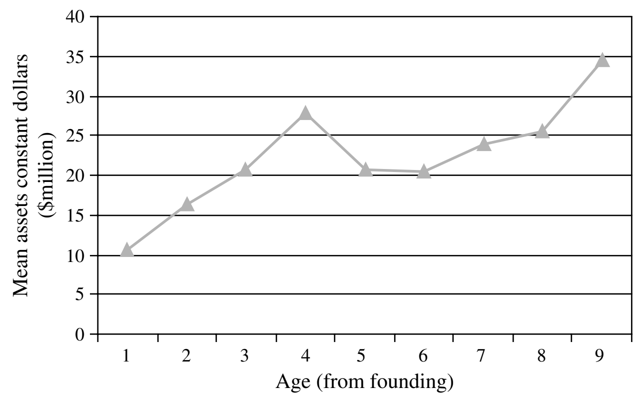
like a better proxy for market conditions, the two are defined to be equal. By familiar identity, output must be equal to the sum of payments to inputs (wages, rent, and profits). We use state income rather than GSP because GSP is not available for all years. The correlation coefficient between state income and GSP for the years both were available was 99.7%, indicating that the identity holds for our data. In addition to the demand volatility, RMSE(state income) or RMSE(output), we also capture the trend (growth) of the respective variable as a control for market opportunity. An alternative measure, *aggregate output*, captures demand directly. We construct the measure for each market year by summing the output over all firms in a given market and year.

Cost distribution. A central issue in characterizing the cost distributions is the relevant set of incumbents assessed by potential entrants. There are two logical candidates. The first is members of bank holding companies (the larger firms engaged in the merger activity), and the second is independent banks. We characterize the cost distributions for each group in each market j for each year t. We then test Equation (5) separately for the two groups. We construct two measures of the cost distribution for each group μ_{sjt} and σ_{jt}, where the set of survivor firms s is defined as those who have remained in the industry for three years. We use three years as the cutoff because this matches expert opinion about the time to reach profitability as well as results in past studies regarding the time to fail (Knott and Posen 2005), but we also test robustness to alternative age thresholds.

Failure rate. The third measure entrepreneurs use to make their decision is the failure rate λ_{jt}. *Failure* is less straightforward in banking than in other settings because the FDIC has an interest in the resolution of failing banks. "Forced mergers" between a failed institution and a healthier bank are the dominant form of FDIC resolution. The less common resolution is "paid outs," where the FDIC pays depositors for their lost deposits under FDIC insurance provisions. We adopt the FDIC definition of failure as the sum of "forced mergers" and FDIC "paid outs," and we measure failure rate λ_{jt} as the ratio of the number of these failures over the number of incumbents.

A third mechanism of bank exit is a "voluntary merger" with an existing bank. It is unclear how voluntary mergers affect entry. Because they, like forced mergers, are more likely to occur for a failing bank, it is possible that their effect combines with failure to reduce entry. However, the discussion in §5.2 suggests that mergers might stimulate entry. If so, merger rate will have a positive effect on entry, whereas Proposition 1 anticipates failure rate (paid outs and forced mergers) to have a negative effect on entry. Given the ambiguity about how voluntary mergers affect entry, we examine them separately from failure.

Figure 3 Bank Assets (Physical Plus Financial Less Deposits) vs. Age



Opportunity costs. The final element in Equation (5) is *opportunity cost*. We use input data from the Stage 1 model to construct measures for both foregone wage W_{jt} and entry cost K_{jt} for each market year. Foregone wage W_{jt} is the mean value for the labor price w_1 for all firms in market j in year t. Entry cost K_{jt} is more complex. Ideally, with sufficient entrants in each market in each year, we could merely take the mean value for physical assets plus financial assets for those entering. Unfortunately, there are several market years with no entrants. One approach to circumvent this problem is to use mean entry cost across all markets $\mu(K)_t$. The problem with this approach is that our entire empirical methodology is to exploit differences across markets. Accordingly, we use a hybrid approach. We pool data across all firms to define an asset growth path. We regress firm assets (physical plus financial) on age, year, and market dummies. This generates a set of coefficients for each of the age dummies controlling for year and market (Figure 3). We use these coefficients to de-age the assets of all firms in market j in year t to define the time-varying entry cost for each market in each year K_{jt}.

Other issues with variables. One issue with respect to the full set of independent variables is the delay between observation of market conditions by the prospective entrepreneur and the actual entry. This delay will depend on the length of the setup process. Discussions with industry experts indicate that the setup process takes from 12 to 18 months. We therefore lag explanatory variables by one year but check robustness using alternative lags.

5.4.3. Model. We model entry rate (with entry counts on the left-hand side and the number of incumbents on the right-hand side to allow the rate to vary with the number of incumbents[7]) as a function of the two types of market uncertainty, RMSE_{jt-1} and σ_{jt-1}, the failure rate λ_{jt-1}, the mean cost of the

[7] All results are robust when excluding the number of incumbents.

survivor pool μ_{sjt-1}, and opportunity cost W_{jt-1} and K_{jt-1}, while controlling for time-varying market factors (growth and concentration) as well as market fixed effects γ_j and industrywide year effects δ_t:

$$\begin{aligned}\text{entry}_{jt} &= \beta_0+\beta_1*\text{incum}_{jt-1}+\beta_2*\lambda_{jt-1}+\beta_3*\mu_{sjt-1}\\ &\quad+\beta_4*\sigma_{jt-1}+\beta_5*(\lambda_{jt-1}*\mu_{sjt-1})+\beta_6*\text{growth}\\ &\quad+\beta_7*\text{RMSE}_{jt-1}+\beta_8 W_{jt-1}+\beta_9 K_{jt-1}\\ &\quad+\delta_t+\gamma_j+e_{jt}. \end{aligned} \tag{7}$$

If entrepreneurs are rational, we expect β_2, β_3, β_8, and β_9 to be negative, and β_5 to be positive (Propositions 1–4). If, in addition, they are unbiased and risk neutral, we expect β_4 and β_7 to be insignificant (Propositions 5A and 6A). If, however, entrepreneurs are biased in the manner discussed previously, then they are risk averse with respect to demand uncertainty and β_7 should be negative (Proposition 6B); and they are overconfident with respect to cost uncertainty and β_4 should be positive (Proposition 5B). Note that the coefficient β_4 captures the effects of cost uncertainty above and beyond the effects associated with the options structure of entrepreneurial returns. This is true because the options effects, derived in Equation (3), are captured by β_2, β_3, and β_5.

Estimation of Equation (7) requires count data models because the dependent variable, the number of entries in a market year, is a nonnegative integer. Given that, the use of ordinary least squares (OLS) violates the assumptions of homoskedasticity and normality (Greene 1997). Accordingly, we employ a fixed effects negative binomial model (Cameron and Trivedi 1998, Greene 1997). However, we also run robustness checks using generalized estimating equations (GEE) and fixed effects specifications.

5.5. Data

The data for the study comes from the FDIC research database which contains quarterly financial data for all commercial banks filing the "Report of Condition and Income" (the so-called Call Report). We examine each of the 50 states plus the District of Columbia for the period 1984–1997. This initial data set contains 694,587 firm-quarter observations. Following convention in the banking literature, we aggregate to annual data by averaging the quarterly data (Mester 1993). The final first-stage data set comprises 170,859 firm-year observations.

While there is considerable debate as to the choice of inputs and outputs in the banking sector, a review of the literature suggests that there is some convergence around a model that sees capital and labor as inputs to the production process and various forms of loans as outputs (Wheelock and Wilson 1995). We collect data to construct seven variables related to banking efficiency in log thousands of constant 1996 dollars. The dependent variable is total cost c—total interest and noninterest expenses. The six independent variables are divided between input prices and output quantities. Input prices are (a) labor price w_1 (salary divided by the number of full-time equivalent employees), (b) physical capital price w_2 (occupancy and other noninterest expenses divided by the value of physical premises and equipment), and (c) capital price w_3 (total interest expense divided by the sum of total deposits, other borrowed funds, subordinated notes, and other liabilities). Output quantities are stocks of (d) mortgage loans y_1, (e) nonmortgage loans y_2, and (f) investment securities y_3.

Our operational definition of a market in the analysis is a state. In part, this definition arises from a data limitation. The unit of observation in the FDIC data is an insurance certificate, which is the unique number assigned to a bank upon entry into a given state. A separate certificate is required for each state in which a bank operates, but covers all branches for that bank operating within the state. Ignoring for a moment the data limitation, there are two discrete definitions of market: the state, representing certificate/headquarters level competition, and municipality, representing branch level competition. A reasonable argument for not doing branch-level analysis, even if data were available, is that it is difficult to determine a relevant radius for competition. Consumers might choose a branch close to their home or one close to their office, but they may also choose a bank based on the fact that it has branches near both, suggesting aggregation to a metropolitan area. Continuing that logic, a state is merely further aggregation, representing on average 7.1 metropolitan statistical areas (MSA), 1.3 primary metropolitan statistical areas (PMSA), or 0.4 consolidated metropolitan statistical areas (CMSA). Given the difficulty of choosing a level of aggregation for branch-level competition, and given the fact that the state captures headquarters competition, we define a market as a state.[8]

To test the entry decision, we gather aggregate market-year data on entry from the FDIC database. We define entry as a new commercial banking institution that comes into existence by way of a new charter. This definition restricts the sample to de novo entry by entrepreneurs, i.e., it excludes conversions, recharters, and the addition of new banks to existing bank holding companies. It is interesting to note that

[8] As an additional test of reasonableness, Petersen and Rajan (2002) examine the distance between small firms and the bank branch they use most frequently. They find that the distance capturing 75% of firms in 1990–1993 is 68 miles and growing rapidly due to information technology. This implies a circumscribed area of 14,524 miles, which is greater than the land area of 10 states and is equal to 26.3% of the mean land area of all states excluding Texas and Alaska.

Table 1 Stage 1 Data Summary

Variable	Description	Obs.	Mean	Std. dev.	Min	Max
c	Cost	174,869	7.98	1.27	−0.82	16.76
w_1	Price_labor	174,673	2.97	0.28	−4.85	9.39
w_2	Price_physical capital	173,999	−1.64	0.73	−9.93	5.63
w_3	Price_capital	173,789	−3.54	0.41	−12.58	3.32
y_1	Mortgage	172,503	9.47	1.58	−1.37	17.86
y_2	Other loans	173,373	9.63	1.40	−0.70	18.55
y_3	Securities	173,828	9.58	1.47	−1.27	17.51

Variable	c	w_1	w_2	w_3	y_1	y_2	y_3
c	1						
w_1	0.1929*	1					
w_2	−0.0418*	0.2792*	1				
w_3	0.1464*	0.0921*	0.0452*	1			
y_1	0.8640*	0.1151*	−0.1322*	−0.0463*	1		
y_2	0.9146*	0.1493*	−0.0594*	0.1201*	0.7711*	1	
y_3	0.7603*	0.0821*	−0.0922*	0.0259*	0.6883*	0.6934*	1

Note. Units: ln(thousand—1996 dollars). *Significant at 5%.

such de novo entry comprises 75.8% of new charters over the period we examine (see Figure 1). Table 1 provides variable descriptions and summary statistics of the data used in Stage 1.

6. Results

6.1. Characterizing Firm Cost Efficiency

We estimate the Stage 1 stochastic frontier model assuming a truncated normal distribution for the inefficiency term and a normally distributed error term. Results from the Stage 1 analysis using Equation (6) are given in Table 2.

The most important result of the Stage 1 frontier estimation is the value of the inefficiency terms u_{it}. The distribution of firm cost inefficiency over all market years is given in Figure 4a, and the mean value over time is depicted in Figure 4b. The mean u_{it} over the entire period is 0.171, which indicates that the mean firm has cost 18.6% above that of a firm on the cost frontier. The data also indicate that while the mean cost inefficiency changes over time in response to changing technologies and demand conditions, the general trend (particularly over the mid-1990s) is toward increasing efficiency (decreasing cost). This is consistent with prior studies of the industry.

While a discussion of the estimated coefficients of the frontier model is outside the scope of this paper, the coefficient estimates are consistent with expectations as (a) total costs appear to rise with output and increases in the price of capital, and (b) firms substitute labor and physical capital in response to changing prices for these inputs. These results reinforce confidence in the frontier technique.

6.2. Test of Propositions

The firm cost efficiencies from Stage 1 were transformed (by taking natural logs) to convert their values from the truncated normal distribution in frontier analysis to the normal distribution assumed by Equation (3). The transformed cost efficiencies were then pooled by market year and used to create estimates of the cost distribution for each market j in each year t. We characterized each ex ante cost distribution by its dispersion σ_{jt}. We then excluded firms less than three years old to characterize the means of the

Table 2 Results from Stage 1 Regression

Dependent variable: ln(cost)
170,859 observations

	Coefficient	Std. error		Coefficient	Std. error
w_1	−8.691e−01**	(30.644)	w_1w_3	−2.312e−01**	(61.988)
w_2	−2.085e−01**	(17.544)	$(w_2)^2/2$	−4.964e−03**	(4.675)
w_3	2.078e+00**	(85.535)	w_2w_3	−2.519e−02**	(15.625)
y_1	1.942e−02*	(2.073)	$(w_3)^2/2$	2.564e−01**	(72.979)
y_2	4.262e−01**	(41.009)	y_1w_1	3.230e−02**	(22.174)
y_3	2.784e−01**	(30.382)	y_1w_2	−7.015e−04	(0.969)
$(y_1)^2/2$	9.331e−02**	(165.733)	y_1w_3	−3.159e−02**	(24.569)
y_1y_2	−6.028e−02**	(120.751)	y_2w_1	−2.019e−02**	(12.595)
y_1y_3	−2.049e−02**	(39.425)	y_2w_2	−9.330e−03**	(10.312)
$(y_2)^2/2$	1.225e−01**	(165.107)	y_2w_3	2.952e−02**	(21.413)
y_2y_3	−5.541e−02**	(95.188)	y_3w_1	−3.038e−02**	(21.670)
$(y_3)^2/2$	8.827e−02**	(190.369)	y_3w_2	1.418e−02**	(19.267)
$(w_1)^2/2$	2.011e−01**	(43.998)	y_3w_3	1.620e−02**	(12.903)
w_1w_2	3.016e−02**	(16.267)	Constant	5.320e+00**	(54.442)
			$E(u_{it} \mid e_{it})$	1.707e−01**	0.172

Notes. Absolute value of t statistics are in parentheses. +Significant at 10%. *Significant at 5%. **Significant at 1%.

Figure 4a **Histogram of Firm-Year Efficiency Metrics**

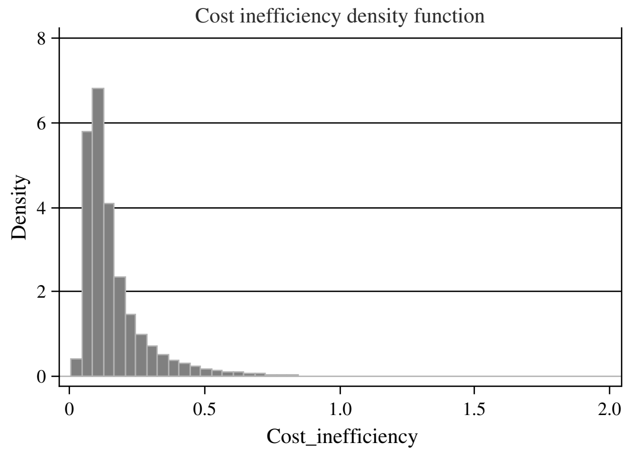
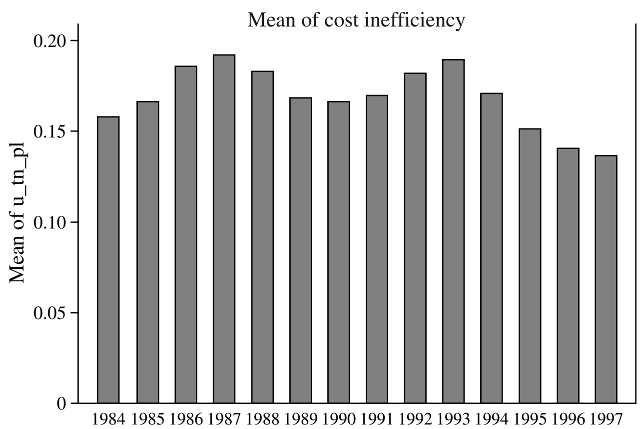


Figure 4b **Mean Cost of Inefficiency**



survivor populations μ_{sjt}.[9] These measures were combined with measures for demand uncertainty RMSE_{jt}, the failure rate λ_{jt}, and opportunity cost W_{jt} and K_{jt} to test the propositions derived from Equation (5). Summaries of these measures, as well as control variables for growth and concentration, are provided in Table 3.

Table 4 presents the results for the test of Equation (7). Model 1 is a simple model of controls for market opportunity; Model 2 is the baseline model for risk-neutral and unbiased entry given an options structure for entrepreneurial returns (testing Propositions 1–4); and Model 3 adds terms for both dimensions of uncertainty to test Propositions 5 and 6.

Results for Model 1 indicate that entry responds to conventional measures of market opportunity. De novo entry increases with demand growth and decreases with the number of incumbents and the degree of market concentration. Entry does not appear to vary with our measures of opportunity cost, entry cost K, and foregone wage W. This could occur for two reasons. First, both variables are correlated with state income, which is the basis for our measures for demand uncertainty and growth. Accordingly, market growth likely captures some opportunity cost and foregone wage effects. An alternative explanation for foregone wage is that equilibrium models of firm formation (Kihlstrom and Laffont 1979, Blau 1987) expect wages to increase as the level of entrepreneurship increases. In these models, wages are an adjustment mechanism: The more individuals become entrepreneurs, the fewer laborers there are for them to employ, and the more wages increase. A robustness check which uses aggregate output rather than state income as the basis for demand uncertainty and growth (Table 5, Model 1) indicates that W takes on the expected sign. This suggests that correlation between state income and foregone wage W offers a better explanation for the finding than does wage adjustment.

Adding terms for the options payoff structure captured in Equation (4) (Model 2) indicates that entry behavior responds rationally to the options structure of payoffs. Entry is decreasing in the failure rate λ_{jt} (Proposition 1) and mean survivor cost μ_{sjt} (Proposition 2), and is increasing in the cross-term $\lambda_{jt} * \mu_{sjt}$ (Proposition 3). The economic impact of a standard-deviation decrease in either the failure rate or survivor cost is to increase entry by one firm. All results are as expected.

Following discussions in §§5.2 and 5.4, we treated voluntary mergers separately from failure. The coefficient on mergers is positive but not significant. Greater significance would have provided support for the liquidity effect, where mergers add qualified bank executives with newly acquired financial assets to the pool of potential entrepreneurs. As results stand now, the main implication is that voluntary mergers should not be counted as failures.

Model 3 presents our main test. It examines entrepreneurial behavior along both dimensions of uncertainty. Results indicate that entry is increasing with cost uncertainty σ_{jt} (Proposition 5B) and decreasing with demand uncertainty $\text{RMSE(state income)}_{jt}$ (Proposition 6B). Accordingly, entrepreneurs in aggregate appear to be overconfident with respect to ability (cost) uncertainty, while risk averse with respect to demand uncertainty. The economic impact of a standard deviation increase in demand uncertainty is to decrease entry by 0.31 firms. The economic impact of a standard deviation increase in ability uncertainty is to increase entry by 0.32 firms. Again, these results are as expected.

Model 4 adds a simple test comparing two alternative reference sets that entrepreneurs could use to estimate the market cost distribution. The baseline in

[9] Our baseline model uses the set of independent banks to construct the cost distribution measures; however, we separately examine the cost distribution of bank holding companies.

Table 3 Summary of Data for Stage 2

Variable	Mean	Std. dev.	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Entry$_{jt}$	2.051	4.453	1.00												
2. ln(incumbents)$_{jt-1}$	4.866	1.219	0.28	1.00											
3. Growth(state income$_{jt-1 \text{ to } jt-11}$)	0.031	0.023	0.29	−0.10	1.00										
4. CR4$_{jt-1}$	0.503	0.231	−0.09	−0.75	0.19	1.00									
5. Mean wage$_{jt}$	20.728	4.236	−0.04	−0.22	−0.13	0.32	1.00								
6. Entry cost$_{jt}$	8,265.6	65,533.9	−0.05	−0.22	0.09	0.08	0.19	1.00							
7. Failure rate$_{jt}$	0.007	0.020	−0.08	−0.10	−0.29	0.10	0.07	−0.06	1.00						
8. Merger rate$_{jt}$	0.043	0.041	0.07	−0.07	0.02	0.10	0.06	0.05	0.02	1.00					
9. $\mu(F(c_{\text{ind}}))_{jt*}$	0.180	0.073	0.03	−0.38	0.05	0.43	0.16	0.24	0.21	0.05	1.00				
10. $\mu *$ failure rate$_{jt*}$	0.002	0.005	−0.07	−0.13	−0.27	0.15	0.08	−0.05	0.95	0.03	0.30	1.00			
11. RMSE(state income$_{jt-1 \text{ to } jt-11}$)	0.023	0.012	−0.14	−0.08	−0.22	−0.07	0.01	−0.13	0.16	−0.10	−0.09	0.15	1.00		
12. $\sigma(F(c_{\text{ind}})))_{jt}$	0.552	0.162	0.09	−0.25	0.17	0.29	0.30	0.43	0.01	0.07	0.66	0.06	−0.20	1.00	
13. $\sigma(F(c_{\text{hold co}}))_{jt}$	0.551	0.222	0.00	−0.27	−0.08	0.31	0.35	0.40	0.09	0.05	0.38	0.13	−0.02	0.34	1.00

Models 2 and 3 reflects the cost distribution of other independent banks; Model 4 adds the cost dispersion for members of bank holding companies. The coefficient for holding company dispersion is near zero, while that for independent banks remains positive and highly significant. This suggests that potential entrepreneurs compare themselves to other independent entrepreneurs rather than to the large bank holding companies.

Table 4 Test of Propositions

	(1)	(2)	(3)	(4)
$\beta_2$failure$_{jt-1}$		−37.17**	−37.17**	−35.18**
		(2.77)	(2.82)	(2.69)
Mergers$_{jt-1}$		0.90	0.87	1.08
		(0.74)	(0.72)	(0.90)
$\beta_3\mu(\text{cost}_s)_{jt-1}$		−5.06**	−6.67**	−6.72**
		(3.23)	(4.25)	(4.25)
$\beta_5\mu(\text{cost}_s)_{jt-1} *$ failure$_{jt-1}$		150.3**	153.9**	150.9**
		(2.87)	(3.01)	(2.97)
β_7RMSE(state income$_{jt-1 \text{ to } jt-11}$)			−25.80*	−33.20**
			(2.34)	(2.94)
$\beta_4\sigma(\text{cost}_{\text{ind}})_{jt-1}$			1.95**	2.02**
			(3.75)	(3.86)
$\sigma(\text{cost}_{\text{hold co}})_{jt-1}$				0.10
				(0.31)
β_1 ln(incumbents)$_{jt-1}$	−0.94**	−0.88**	−1.16**	−1.39**
	(3.46)	(3.05)	(3.13)	(3.62)
$\beta_6$growth(state income$_{jt-1 \text{ to } jt-11}$)	25.64**	25.45**	20.43**	19.43**
	(10.73)	(9.63)	(7.08)	(6.76)
CR4$_{jt-1}$	−1.93*	−1.83*	−1.33^	−1.89*
	(2.54)	(2.36)	(1.68)	(2.26)
β_8(mean wage)$_{jt-1}$	0.01	0.01	0.01	0.01
	(0.76)	(0.82)	(0.47)	(0.51)
β_9(entry cost)$_{jt-1}$	0.00^	0.00	0.00	0.00
	(1.76)	(1.45)	(1.28)	(1.38)
Year effects				
Constant	6.70**	7.36**	9.39**	1.12**
	(3.57)	(3.72)	(3.58)	(4.04)
Log likelihood	−774.53	−766.06	−756.41	−742.61
Chi square	262.76	274.83	307.92	317.08
Observations	637	637	637	629
Number of markets	49	49	49	49

Notes. Dependent variable is entry$_{jt}$. Absolute value of z statistics are in parentheses. ^Significant at 10%. *Significant at 5%. **Significant at 1%.

6.3. Robustness Checks

We conducted a number of robustness checks of the main results reflecting concerns expressed in §§5.4.2 and 5.4.3. These checks are presented in Table 5 alongside the main model (Model 3 from Table 4). The checks include sensitivity to changes in each of the following: the measure used for the market demand variables, the functional form of the options effect, lags for the explanatory variables, definition of the survivor pool, the distributional assumption in Stage 1 frontier analysis, and the econometric specification of the main model. Results are robust to all these changes.

7. Discussion

The goal of this paper is to reconcile the risk-bearing role of entrepreneurs with the stylized fact that entrepreneurs exhibit conventional risk-aversion profiles. We proposed that the disparity arises from confounding two distinct dimensions of uncertainty: demand uncertainty and ability uncertainty. We further proposed that entrepreneurs will be risk averse with respect to demand uncertainty, while "apparent risk seeking" (overconfident) with respect to ability uncertainty. To examine this view, we constructed a reduced-form model of the entrepreneurial entry decision, which we then aggregated to the market level. In the model, entrepreneurs compare the expected value of an uncertain profit stream against the opportunity cost of continuing in wage employment. The baseline model anticipates that with rational options behavior, entry will be increasing in mean cost, and decreasing in the failure rate as well as the entrepreneur's opportunity cost. The model with risk preferences adds variables for demand uncertainty as well as ability (cost) uncertainty. We tested the aggregate model

Table 5 Results from Robustness Checks

	Main	(1) Demand = agg output	(2) Options form	(3) Lag = 2	(4) Survivor ≥ 4 yr	(5) Survivor ≥ 5 yr	(6) Stage 1 1/2 normal	(7) Stage 1 exponential	(8) Eqn 7: GEE	(9) Eqn 7: FE
$\beta_2$failure$_{jt-1}$	-37.17^{**}	-47.62^{**}	-45.05^{**}	-21.02	-39.81^{**}	-40.12^{**}	-15.48	-37.12^{**}	-11.63	-18.11
	(2.82)	(3.26)	(3.192)	(1.52)	(3.15)	(3.14)	(1.21)	(2.82)	(0.50)	(0.77)
(Failure$_{jt-1}$)^2			102.00							
			(1.565)							
Mergers$_{jt-1}$	0.87	1.93	0.96	$2.16^{\wedge}$	0.85	0.82	1.51	0.87	$6.81^{\wedge}$	3.96
	(0.72)	(1.48)	(0.793)	(1.67)	(0.71)	(0.69)	(1.23)	(0.72)	(1.72)	(0.94)
$\beta_3\mu(\text{cost}_s)_{jt-1}$	-6.67^{**}	-7.88^{**}	-11.32^{*}	-4.70^{**}	-6.72^{**}	-6.98^{**}	-6.65^{**}	-6.70^{**}	$-8.07^{\wedge}$	-19.90^{**}
	(4.25)	(4.83)	(2.533)	(2.80)	(4.21)	(4.40)	(3.27)	(4.25)	(1.90)	(3.95)
$(\mu(\text{cost}_s)_{jt-1})^2$			8.83							
			(1.121)							
$\beta_5\mu(\text{cost}_s)_{jt-1} * \text{failure}_{jt-1}$	153.9^{**}	159.5^{**}	146.70^{**}	99.99	163.1^{**}	164.1^{**}	64.81	154.4e^{**}	47.04	76.50
	(3.01)	(2.87)	(3.039)	(1.80)	(3.41)	(3.38)	(1.34)	(3.01)	(0.49)	(0.78)
β_7RMSE(state income$_{jt-1 \text{ to } jt-11}$)	-25.80^{*}		-24.08^{*}	-34.67^{**}	-25.60^{*}	-26.17^{*}	-27.70^{*}	-25.81^{*}	-43.24^{*}	$-50.52^{\wedge}$
	(2.34)		(2.161)	(2.92)	(2.32)	(2.38)	(2.46)	(2.34)	(2.21)	(1.94)
RMSE(agg bank output$_{jt-1 \text{ to } jt-11}$)		-1.25								
		(0.95)								
$\beta_4\sigma(\text{cost}_{\text{ind}})_{jt-1}$	1.95^{**}	2.87^{**}	1.87^{**}	1.76^{**}	1.89^{**}	1.88^{**}	1.08^{**}	1.94^{**}	4.59^{**}	7.59^{**}
	(3.75)	(5.49)	(3.578)	(3.01)	(3.68)	(3.69)	(2.39)	(3.74)	(2.82)	(4.03)
β_1 ln(incumbents)$_{jt-1}$	-1.16^{**}	-0.59^{*}	-1.08^{**}	-1.37^{**}	-1.13^{**}	-1.15^{**}	-0.95^{**}	-1.16^{**}	1.04^{**}	-4.32^{**}
	(3.13)	(2.28)	(2.775)	(3.44)	(3.12)	(3.15)	(2.83)	(3.13)	(2.93)	(2.76)
$\beta_6$growth(state income$_{jt-1 \text{ to } jt-11}$)	20.43^{**}		19.78^{**}	19.68^{**}	20.38^{**}	20.18^{**}	20.53^{**}	20.43^{**}	35.53^{**}	29.79^{**}
	(7.08)		(6.725)	(6.51)	(7.07)	(7.01)	(6.99)	(7.08)	(4.75)	(3.82)
Growth(agg bank output$_{jt-1 \text{ to } jt-11}$)		2.34^{**}								
		(3.94)								
CR4$_{jt-1}$	$-1.33^{\wedge}$	-1.55^{*}	$-1.36^{\wedge}$	-1.41	$-1.35^{\wedge}$	$-1.37^{\wedge}$	-1.21	$-1.33^{\wedge}$	0.98	-7.73^{**}
	(1.68)	(1.98)	(1.708)	(1.61)	(1.70)	(1.73)	(1.53)	(1.68)	(0.57)	(2.69)
β_8(mean wage)$_{jt-1}$	0.01	-0.03	0.00	-0.01	0.01	0.01	0.01	0.01	0.03	0.01
	(0.47)	(1.56)	(0.347)	(0.81)	(0.45)	(0.49)	(0.72)	(0.47)	(0.74)	(0.18)
β_9(entry cost)$_{jt-1}$	0.00	$0.00^{\wedge}$	0.00	0.00	0.00	0.00	$0.00^{\wedge}$	0.00	0.00	0.00^{**}
	(1.28)	(1.82)	(1.478)	(1.24)	(1.44)	(1.56)	(1.93)	(1.28)	(0.67)	(2.94)
Year effects										
Constant	9.39^{**}	6.05^{**}	9.52^{**}	11.08^{**}	9.27^{**}	9.50^{**}	8.20^{**}	9.38^{**}	-3.64	2.84^{**}
	(3.58)	(3.32)	(3.414)	(3.82)	(3.60)	(3.65)	(3.54)	(3.58)	(1.33)	(3.32)
Log likelihood	-756.41	-787.77	-754.691	-671.52	-756.02	-755.27	-761.97	-756.44		
Chi square	307.92	217.01	311.101	312.43	308.21	311.57	292.52	307.83	144.10	
Observations	637	637	637	588	637	637	637	637	663	663
Number of markets	49	49	49	49	49	49	49	49	51	

Notes. Dependent variable is entry$_{jt}$. Absolute value of z statistics are in parentheses. $^{\wedge}$Significant at 10%. *Significant at 5%. **Significant at 1%.

across markets and over time in the banking industry. We found that entrepreneurs behave rationally to the options structure of payoffs, but that in addition they appear to be risk averse with respect to demand uncertainty and "risk seeking" (overconfident) with respect to cost uncertainty.

The test was constructed in a manner that allowed us to draw inferences at both the individual and market levels. For the characterization of entrepreneurial personality, our results using large sample data (1,635 entrants) confirm observations made previously for small samples—namely, that entrepreneurs as a group appear to be both risk averse and overconfident. It is noteworthy that we obtain these results from real entry behavior rather than with personality instruments. Thus, this test is similar to lab experiments in that we observe how people actually behave rather than how they report they will behave. Moreover, this behavior occurs in real settings with substantial investments (approximately $11 million per entry), thereby offering greater external validity than lab experiments.

It is worth repeating that our results about individual biases pertain exclusively to entrepreneurs who actually enter a banking market. Thus, we are unable to say anything about how (if at all) entrepreneurs' biases differ from those of wage earners. Indeed, our earlier discussions suggest that risk aversion and overconfidence are common to both groups: Entrepreneurs' risk preferences are indistinguishable from wage earners, and at least 80% of high-school seniors (the pool from which both wage earners and entrepreneurs are drawn) are overconfident in at least one dimension. Accordingly, what may distinguish entrepreneurs from wage earners in this setting are simple differences in human capital and wealth.

The test does, however, allow us to draw inferences about the market-level implications of entrepreneurs' biases. In particular, our results offer a possible reconciliation of observed risk aversion with the entrepreneur's role as economic risk bearer. Entrepreneurs are willing to bear economic risk when the degree of performance dispersion dominates the degree of demand uncertainty. In those instances, overconfidence can compensate for risk aversion to achieve sufficient entry. This appears to be the case in banking. Potential market failures exist, however, in instances where there is either a high degree of demand uncertainty but low performance dispersion (insufficient entry) or low demand uncertainty but high performance dispersion (excess entry).

There are, of course, caveats to our results. The test was conducted in the banking industry because of its high rate of entry and nice experimental properties. One question, then, is how results for banking might generalize to other settings. The most obvious distinction between entry in banking versus other industries is that it is regulated. A review of charter requirements suggests that this has three effects. First, bank founders are highly qualified. To gain charter approval, they must have both extensive banking experience and strong ties to the local community. Second, banks require substantial up front investment of approximately $11 million in physical and financial assets for our sample. Third, the chartering agencies "match" the number of entrants to market demand through the requirement that the new bank satisfies an unmet need. The implication of all three effects is that there should be less entry in banking than in other industries, both through the supply effect (fewer people will satisfy the high human capital and physical or financial capital requirements or both required by charters) and through the demand effect (new charters must satisfy unmet needs).

Acknowledgments

The authors thank Tim Folta, Andrew King, Dan Lovallo, Barry Rubens, Mirjam Van Praag, Govert Vroom, the associate editor, and two anonymous referees, as well as seminar participants at Dartmouth College, Purdue University, the University of Maryland, the United States Association for Small Business and Entrepreneurship, and The Wharton School for helpful comments.

References

Aigner, D., A. K. Lovell, P. Schmidt. 1977. Formulation and estimation of stochastic frontier production function models. *J. Econometrics* **6** 21–37.

Alicke, M. D. 1985. Global self-evaluation as determined by the desirability and controllability of trait adjectives. *J. Personality Soc. Psych.* **49** 1621–1630.

Alicke, M. D., M. L. Klotz, D. L. Breitenbecher, T. J. Yurak, D. S. Vredenburg. 1995. Personal contact, individuation and the better-than-average effect. *J. Personality Soc. Psych.* **68** 804–825.

Allison, S. T., D. M. Messick, G. R. Goethals. 1989. On being better but not smarter than others: The Muhammad Ali effect. *Soc. Cognition* **7** 275–296.

Berger, A. N., D. B. Humphrey. 1992. Measurement and efficiency issues in commercial banking. Z. Griliches, ed. *Output Measurement in the Service Sectors.* Vol. 56. University of Chicago Press, Chicago, IL, 245–279.

Berger, A. N., D. Hancock, D. Humphrey. 1993. Bank efficiency derived from the profit function. *J. Banking Finance* **17**(2, 3) 317–347.

Blau, D. M. 1987. A time-series analysis of self-employment in the United States. *J. Political Econom.* **95**(3) 445–467.

Brockhaus, R. 1980. Risk-taking propensity of entrepreneurs. *Acad. Management J.* **23**(3) 509–520.

Busenitz, L. W. 1999. Entrepreneurial risk and strategic decision making: It's a matter of perspective. *J. Appl. Behav. Sci.* **35**(3) 325–340.

Camerer, C., D. Lovallo. 1999. Overconfidence and excess entry: An experimental approach. *Amer. Econom. Rev.* **89**(1) 306–318.

Cameron, A. C., P. K. Trivedi. 1998. *Regression Analysis of Count Data*. Cambridge University Press, Cambridge, UK, and New York.

Cantillon, R. 1755. *Essay on the Nature of Commerce*. Higgs translation, 1931. Macmillan, London, UK.

Cebenoyan, A., G. Papaioannou, N. Travlos. 1992. Foreign takeover activity in the United States and wealth effects for target firm shareholders. *Financial Management* **21**(3) 58–68.

Cramer, J. S., J. Hartog, N. Jonker, C. M. van Praag. 2002. Low risk aversion encourages the choice for entrepreneurship: An empirical test of a truism. *J. Econom. Behav. Organ.* **48**(1) 29–36.

Dunn, T., D. Holtz-Eakin. 2000. Financial capital, human capital, and the transition to self-employment: Evidence from intergenerational links. *J. Labor Econom.* **18** 282–305.

Evans, D., L. Leighton. 1989. Some empirical aspects of entrepreneurship. *Amer. Econom. Rev.* **79**(3) 519–535.

Federal Reserve Bank of Atlanta. 2005. *Banking Applications*. Retrieved July 1, 2006, http://www.frbatlanta.org/ba.

Goldberg, L. G., L. J. White. 1998. De novo banks and lending to small businesses. *J. Banking Finance* **22** 851–867.

Greene, W. H. 1997. *Econometrics Analysis*, 3rd ed. Prentice-Hall, Upper Saddle River, NJ.

Hamilton, Barton H. 2000. Does entrepreneurship pay? An empirical analysis of the returns to self-employment. *J. Political Econom.* **108**(3) 604–631.

Headd, B. 2003. Redefining business success: Distinguishing between closure and failure. *J. Small Bus. Econom.* **21** 51–61.

Hermalin, B., N. Wallace. 1994. The determinants of efficiency and solvency in savings and loans. *RAND J. Econom.* **25**(3) 361–381.

Jovanovic, B. 1982. Selection and the evolution of industry. *Econometrica* **50**(3) 649–670.

Kanbur, S. M. 1979. On risk taking and the personal distribution of income. *J. Political Econom.* **87**(4) 760–797.

Kihlstrom, R. E., J. J. Laffont. 1979. A general equilibrium entrepreneurial theory of firm formation based on risk aversion. *J. Political Econom.* **87**(4) 719–748.

Knight, F. H. 1921. *Risk, Uncertainty and Profit*. Houghton Mifflin, New York.

Knott, A. M., H. Posen. 2005. Is failure good? *Strategic Management J.* **26**(7) 617–641.

Lippman, S. A., R. P. Rumelt. 1982. Uncertain imitability: An analysis of interfirm differences in efficiency under competition. *Bell J. Econom.* **13**(2) 418–438.

Lofstrom, M. 2002. Labor market assimilation and the self-employment decision of immigrant entrepreneurs. *J. Population Econom.* **15** 83–114.

Lucas, R. E. 1978. On the size distribution of business firms. *Bell J. Econom.* **9**(2) 508–523.

March, J. G., Z. Shapira. 1987. Managerial perspectives and risk taking. *Management Sci.* **33**(11) 1404–1418.

Masters, R., R. Meier. 1988. Sex differences and risk-taking propensity of entrepreneurs. *J. Small Bus. Management* **26**(1) 31–35.

Mazzeo, M. 2004. Retail contracting and organizational form: Alternatives to chain affiliation in the motel industry. *J. Econom. Management Strategy* **13**(4) 599–615.

McClelland, D. C. 1961. *The Achieving Society*. Van Nostrand, Princeton, NJ.

Mester, L. 1993. Efficiency in the savings and loan industry. *J. Banking Finance* **17** 267–286.

Miner, J. B., N. S. Raju. 2004. Risk propensity differences between managers and entrepreneurs and between low- and high-growth entrepreneurs: A reply. *J. Appl. Psych.* **89**(1) 3–13.

Morgan, D. P., B. Rime, P. E. Strahan. 2004. Bank integration and state business cycles. *Quart. J. Econom.* **119**(4) 1555–1584.

Myers, D. G. 1982. The inflated self: Human illusions and the biblical call to hope. *Christian Century.* (December 1) 1226.

Norton, W. I., Jr., W. T. Moore. 2002. Entrepreneurial risk: Have we been asking the wrong question? *Small Bus. Econom.* **18**(4) 281–287.

O'Brien, J. P., T. B. Folta, D. R. Johnson. 2003. A real options perspective of entrepreneurial entry in the face of uncertainty. *Managerial Decision Econom.* **24**(8) 515–533.

Palich, L. E., D. R. Bagby. 1995. Using cognitive theory to explain entrepreneurial risk-taking—challenging conventional wisdom. *J. Bus. Venturing* **10**(6) 425–438.

Petersen, M., R. Rajan. 2002. Does distance still matter? The information revolution in small business lending. *J. Finance* **57**(6) 2533–2570.

Sarasvathy, D., H. Simon, L. Lave. 1998. Perceiving and managing business risks: Differences between entrepreneurs and bankers. *J. Econom. Behav. Organ.* **33**(2) 207–225.

Stock, J. H., M. W. Watson. 1998. Business cycle fluctuations in U.S. macroeconomic time series. Working Paper 6528, National Bureau of Economic Research, Cambridge, MA.

Stuart, T. E., O. Sorenson. 2003. Liquidity events and the geographic distribution of entrepreneurial activity. *Admin. Sci. Quart.* **48** 175–201.

Thaler, R. H., E. J. Johnson. 1990. Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management Sci.* **36**(6) 643–660.

United States Congress (U.S. Congress). 1977. Community Reinvestment Act, law number 12 U.S.C. 2901. U.S. Congress, Washington, D.C.

Wheelock, D., P. Wilson. 1995. Evaluating the efficiency of commercial banks: Does our view of what banks do matter? *Rev. Federal Reserve Bank St. Louis* **77** 39–52.

Wooten K. C., T. A. Timmerman, R. Folger. 1999. The use of personality and the five-factor model to predict new business ventures: From outplacement to start-up. *J. Vocational Behav.* **54**(1) 82–101.

Zenger, T. R. 1994. Explaining organizational diseconomies of scale in R&D: Agency problems and allocation of engineering talent, ideas and effort by firm size. *Management Sci.* **40** 708–729.