



Start-up financing, owner characteristics, and survival[☆]

Thomas Åstebro^{*,1}, Irwin Bernhardt

Department of Management Sciences, University of Waterloo, Waterloo, Ont., Canada N2L 3G1

Received 16 August 2001; received in revised form 12 December 2002; accepted 12 February 2003

Abstract

We investigated the relation between the survival of new small businesses and bank loans. This was done using a model that included other loan sources, human capital variables, and company and industry descriptors. We found there is a negative correlation between having a bank loan and business survival, and a positive correlation between having a non-bank loan and survival. However, having a bank loan was a *ceteris paribus* positive predictor of the survival of start-up companies. Our findings enabled some inferences about the process of loan source selection by start-up business owners, and about the banks' loan granting process.

© 2003 Elsevier Science Inc. All rights reserved.

JEL classification: G21; M20

Keywords: Bank loans; Start-up business survival; Forecast

1. Introduction

We investigated the relation between the survival of new small businesses (start-ups) and their having a bank loan. This was done in the context of a model that included other loan sources, human capital variables, and industry descriptors. The investigation was based on data referring

[☆] Data contained in this study were produced on-site at the Carnegie-Mellon Census Research Data Center. Research results and conclusions are those of the authors and do not necessarily indicate concurrence by the Bureau of the Census or the Carnegie-Mellon Census Research Data Center.

* Corresponding author. Tel.: +1-416-978-5703; fax: +1-519-746-7252.

E-mail address: tastebro@engmail.uwaterloo.ca (T. Åstebro).

¹ Current address: Rotman School of Management, University of Toronto.

to a set of small businesses launched in 1987 in the United States. The data source contains certain background data for 1987 and business survival data for 1991 for these businesses.

We found the following novel results: (1) There was a significant negative correlation between having a bank loan and the survival of the business across the sample. This is in spite of those firms in the sample with a bank loan having higher than average sales revenues at start-up. In contrast, there was a positive and significant correlation between having a loan from either friends, family, former owner, or from obtaining a house mortgage and the survival of the business. (2) Both bank loans and other loans were significantly positive indicators of survival, *ceteris paribus*. However, the coefficient for non-bank loans was more than twice as large as that for bank loans in the equation that estimated the chance of business survival. This difference was not statistically significant, with the *t*-statistic only slightly greater than one, however. This equation contained a large set of observable proxies for human capital. That is, there is weak evidence of adverse selection against banks. (3) We found, among companies that had some form of loan, that the chance that the loan was a bank loan to be negatively related to the education and other observable proxies for human capital of the owners. That is, there appeared to be self-selection by highly qualified owners in favor of sources other than commercial loans. (4) Our model of the determinants of survival permitted us to identify a large group (22%) of start-up companies with highly qualified owners that had higher survival rates than start-ups with bank loans.

We know of only one previously published work, Cressy (1996), concerning bank loans as predictors of start-up company survival that included a broad range of other predictors. Our results concerning the effect of human capital are consistent with Cressy's.

The remainder of this paper is divided into four parts: First, we describe the database and the sampling criteria we used. Second, we discuss the set of predictors. Third, we discuss our methods of analysis and present our findings. Finally, we have a brief conclusion.

2. The database

The US Census conducted the survey, "Characteristics of Business Owners" (CBO), in 3 years, 1982, 1987 and 1992. The 1982 and 1987 surveys covered a large set of human capital variables for multi-owner companies and sole proprietorships. The 1992 survey covered human capital variables only for sole proprietorships.¹ We relied on the 1987 CBO database, with its wider range of human capital variables for multi-owner as well as single-owner companies, for our study. The 1987 database is described in Nucci (1992) and Boden and Nucci (1997).² The observations in the 1987 database were derived from a stratified random sample. The sampling frame was the set of 1987 tax filings for various forms of small businesses, sole proprietorships, partnerships, and small (sub-chapter S) corporations (U.S. Bureau of the Census, 1991, iv). The survey contains information about businesses operating during 1987. The businesses included in the survey were started in 1987 or earlier. CBO owner and business information was assembled from questionnaire responses provided in 1991 by owners of the sampled businesses. The IRS supplied other firm-level information.

The CBO contains some unique features that make it useful for investigating the determinants of small business survival. In particular, the survey has relatively detailed data on owner characteristics. It has been shown that owner characteristics have a significant impact on the

business survival of start-ups (Cressy, 1996).³ The survey also contains questions regarding start-up financing conditions. That the data were collected in 1991 makes it possible to observe 4-year survival. As with other surveys, the CBO suffers from non-response bias (Boden & Nucci, 1997; Holmes & Schmitz, 1996).

A difficult question to answer when working with the CBO or similar data sets is how to deal with responses from multiple owners of the same firm. Most users of the CBO (see Nucci, 1992), as well as most other researchers in the area (except Cressy, 1996), analyze records only of single-owner firms, or simply ignore the fact that there may be more than one owner per firm. We believe multi-owner firms should be included because of their importance. Ignoring the fact of multiple owners is incorrect since the number of owners may be an important determinant of the probability of survival. Cressy (1996) found that multi-owner firms had higher probabilities of survival than single-owner firms. There is the potential, then, for bias in estimating the effects of financing conditions, and human capital on business survival if one fails to include data from all owners in the firm. We address non-response bias, as well as the issue of analysis of multi-owner firms, below.

2.1. Sampling

Three concerns guided our decisions regarding sampling: (1) We sought to avoid complicating factors that would make the interpretation of our results more difficult. To this end, we adopted the white male sub-sample. This limitation was established to avoid issues related to racial and sex discrimination. (2) We were concerned with possible sample attrition bias. The CBO survey was based on 1987 tax filings. The dropping out of weaker companies probably reduced the pre-1987 cohorts. Given this, we limited our sample to those companies that were started in 1987. This includes businesses that existed before 1987, but that had completely new ownership in 1987. (3) We wanted to exclude companies that were not started for business reasons and companies that would not be reasonable candidates for commercial bank loans under any circumstances. Some “companies” may actually have been individuals who filed business tax forms to reduce tax liability rather than for business reasons. To remove such individuals, Bates (1990), in using the 1982 CBO survey, deleted companies that had zero capital at start-up. We deleted such companies also. After these deletions, our sample had 923 observations of companies. Bates also required that sample firms had sales of at least \$5,000 in the sample year. This is problematic because a company’s 1987 sales will have depended on whether or not it received a bank loan. Companies that would have had \$5,000 in sales, had they received a loan, may have been eliminated from the sample if, in fact, they did not receive a loan. The effect of this selection criterion would be to tilt the sample in favor of those firms holding loans. We did reduce our initial sample, however, by eliminating from it all companies with *predicted* sales of less than \$5,000 in 1987. Very small firms would not have been candidates for a commercial loan under any circumstances because the cost of evaluating and administering loans to them would make such loans unprofitable for banks. The time required for evaluating a small business loan by traditional methods was estimated to be over 12 h (Mester, 1997). Predicted sales were estimated from a set of exogenous variables that did not include any variables that would be affected by a company receiving a loan.⁴ We used weighted least squares to generate this estimate.⁵ After this selection, our sample had 738 observations.

Table 1
Sources of debt^a

Source of debt ^b (1)	Percent of owners ^c (2) (%)	Percent of owners ^d (3) (%)	Percent of owners ^e (4) (%)
Spouse	0.9	1.4	1.5
Other family members	7.2	10.9	11.5
Friends	1.7	2.5	2.8
Personal credit card	1.2	1.9	2.5
Refinanced home	2.0	4.7	6.5
Former owner	1.9	2.9	4.2
Commercial bank loan	16.0	20.1	22.4
Other type loan	2.0	1.9	2.4
Federal, state, or local government	0.5	0.1	0.0
Other	0.6	1.3	2.1

^a All data are weighted to reflect sampling plan and survey non-response. Missing data are not imputed for comparability across sources. Frequency counts therefore differ from those presented in Table 3.

^b There can be more than one source per owner.

^c Firms operating in 1987 and started by predominantly white male owners. Source: U.S. Census Bureau (1992).

^d Firms with capital >0, started in 1987 by predominantly white male owners.

^e As in col. (3) plus actual firm sales in 1987 >\$5,000.

Table 1 reports weighted statistics regarding the use of bank loans as a source of debt at start-up. As indicated, the percentage of owners with bank loans increases from 16.0% for all start-ups owned by white males to 20.1% for those with some capital, and to 22.4% for those with some capital and actual sales of more than \$5,000 during the start-up year.⁶ This finding corroborates our assumption that the imposed selection criteria removed firms that were not candidates for bank loans from the sample. Another indication that non-candidates for bank loans were excluded by the selection criteria is that the percentage of owners holding state, federal, or local government loans declined from 0.5 to 0%. State, federal, or local government loans are intended to provide start-up capital for those firms that are not candidates for regular bank loans.

2.2. Non-response adjustments

The CBO, in common with most other questionnaire-based surveys, suffers from two types of non-response bias, unit non-response and item non-response. Unit non-response occurs when an owner fails to return a questionnaire. The unit response rate among all white males in the 1987 CBO was approximately 74% (Nucci, 1992, Table 1). This relatively low response rate for Census surveys is attributable, in part, to a difference of approximately 3 years between the year of business tax filing and receipt of a CBO questionnaire.⁷ Item non-response occurs when an owner opts not to answer a particular question on the questionnaire even though the question is applicable.

The survival bias among responses in the CBO is estimated at approximately 3% (Holmes & Schmitz, 1996). Unit non-response and business survival are likely to be correlated. Survival bias is reduced, in the case of multiple owners, because we were able to use information about the survival of the firm supplied by responding owners where data were missing from

non-respondents in multi-owner firms. As well, the survey contains weights that adjust for unit non-response in both single-owner and multiple-owner firms. That is, researchers at the Census determined the incidence of non-response according to firm size, location and industry. The inverses of these response frequencies by stratum are employed as weights throughout the analysis and tabulations. This method is supported by [Holt, Smith, and Winter \(1980\)](#). The weights affect results considerably because there is a clear tendency for owners of smaller firms not to respond.

Item non-response varies by survey item. An appendix (available from the corresponding author on request) shows the percentage response for selected questionnaire items. The appendix shows that item response is above 86% for all variables except college concentration, which has an item response of only 59%.⁸ We imputed values for item non-responses using the Bayesian multinomial multiple imputation method described by [Rubin \(1987\)](#). In short, we generated five complete data sets where missing data were randomly replaced conditional on observed data and survey structure (contact the authors for a complete description). We report results with these imputations.

3. The predictors

We constructed a reduced form model of the probability of survival of start-up businesses. We utilized three sets of predictors of small business survival. They are: (1) indicators for whether or not the company held bank loans and/or other loans at time of start-up; (2) human capital proxies; and (3) industry and company descriptors. The names and definitions of the predictors are shown in [Table 2](#). All of the predictors were observable at the time of start-up. They were, therefore, pre-determined. The implications of these predictors for survival depend on the definition of survival. Accordingly, we state our definition of survival before our discussion of the various predictors.

The data for our sample refer to two time periods. The first time period is the time of start-up for our sample, 1987. The second time period is 4 years later, 1991. We define survival as occurring if a company still existed in 1991 whether or not there was a change of ownership. Failure of the business is synonymous with closing of operations. We do not define a firm that is sold or changes ownership as failing. If a business changed ownership (either through acquisition, merger, or other reason) during the period 1987 (year of start-up) to 1991, but remained in operation it is defined as surviving. If we had defined survival as the continuance of the 1987 owners in the business, as was done by [Bates \(1990\)](#), the implications of high values of human capital indicators would have been unclear. Most of the indicators are non-specific in terms of relevance for starting an enterprise, managing an enterprise, or for paid work. The choice of roles depends on comparative advantage ([Evans & Jovanovic, 1989](#); [Holmes & Schmitz, 1990](#)). While a transfer of ownership is consistent with continued operation, it is however the case that most (approximately 76%) of the firms in our sample that transferred ownership eventually ceased operations within 4 years from establishment. A more restrictive definition considers survival as occurring if a company still existed in 1991 and included at least one of the original 1987 owners as an owner in 1991. Regression results under this alternative definition are reported in an appendix available from the corresponding author on request. The sensitivity analysis shows that the estimated model is fairly robust to this alternative definition of survival.

Table 2
Definitions of variables

Name	Definition
survive	=1 if any of the original owners remained owners for 4 years, or if the firm was still operating after 4 years but under new owners, else 0
edu3	Proportion of owners who were high school graduates
edu4	Proportion of owners who were college drop-outs
edu5	Proportion of owners who were college graduates
edu6	Proportion of owners who did post-graduate studies
scieng	Proportion of owners with science or engineering degree
business	Proportion of owners with business degree
wrkfam	Proportion of owners who previously have worked in family business
owner	Proportion of owners who had previous ownership experience
workexp4	Proportion of owners with 2–9 years of work experience
workexp5	Proportion of owners with 10–19 years of work experience
workexp6	Proportion of owners with at least 20 years of work experience
respno	Number of owners
denovo	=1 if newly formed business in 1987, else 0
toteq3	=1 if total equity from all owners between \$10,000 and \$25,000, else 0
toteq4	=1 if total equity from all owners at least \$25,000, else 0
inc	=1 if firm's legal form is partnership or sub-chapter-S corporation, else 0
scale20	Proportion of firms in 2-digit industry with 1–19 employees
scale50	Proportion of firms in 2-digit industry with 1–49 employees
loan	=1 if any owner in the firm had a start-up commercial bank loan, else 0
othloan	=1 if any owner financed the start-up with a loan from either friends, family spouse, former owner, or home mortgage, else 0

The CBO distinguishes among the following sources of loans: commercial bank loans, loans from spouses, other family members, friends, former owners, personal credit cards, home mortgages, federal/state or local government loans, and other types. We define a firm as holding a given type of loan if any of the owners reported using that type in financing the start-up. Table 1 reports weighted statistics regarding the use of these sources of debt at start-up for non-imputed data. As indicated, commercial bank loans were the most frequently cited source of debt capital (22.4%), followed by other family members (11.5%), and a refinanced home (6.5%). When the firm is a sole proprietorship there is some ambiguity about what a commercial bank loan means since there is no legal separation between personal assets and business assets. We assume, though, that respondents interpret the category “commercial bank loan” to mean that a start-up loan is secured mainly against business assets.⁹

There are three reasons to expect that holding a bank loan will be a predictor of company survival. First, securing a loan from any source relaxes financial constraints on investment. Several researchers have found evidence that suggests that financial constraints may be binding (Blanchflower & Oswald, 1998; Evans & Jovanovic, 1989; Holtz-Eakin, Joulfaian, & Rosen, 1994a, 1994b). Second, securing a bank loan may increase a company's credibility with potential suppliers and customers (Best & Zhang, 1993). Third, the indicator for a loan may proxy for qualities of the owners and the company that were observable by the lender but are not observable by the researcher. On the other hand, banks may be unable to assess the risk inherent in lending

to small new companies accurately. Adverse selection may then occur, where the most able entrepreneurs select not to apply for bank loans but turn to other sources of debt where prices and other loan conditions better reflect their abilities (Stiglitz & Weiss, 1981).

The predictive value of other sources of loans is less clear, *a priori*. Such loans may improve the survivability of the firm for the same reasons as those associated with a bank loan. However, a loan obtained from friends, family members and other related investors may be a poor signal of business credibility. On the other hand, friends and family may have a better knowledge of personal characteristics that are unobservable to a bank. In addition, the signal value of a house mortgage is less than that of a commercial bank loan because there is no assessment of the quality of the business when a business owner obtains a house mortgage to finance the business. On the other hand, an owner might be more committed to continue operations and service debt if a house mortgage is involved than if there is only limited liability (Bester, 1985). Also, securing a house mortgage may imply further access to greater financial wealth that, in turn, implies relaxed financial constraints. We cannot say, *a priori*, therefore, whether holding other loans or obtaining a house mortgage will have a greater or smaller effect on business survival than the effect created by holding a commercial bank loan. For the purposes of this paper, we collapsed obtaining loans from spouse, other family members, friends, and from the former owner, as well as obtaining a house mortgage into the indicator “other loans.” If any of the owners in the firm obtained debt capital from any of these sources, we say the firm had other loans.¹⁰ We do not analyze the remaining sources of debt due to their paucity of use.

Increased human capital provides owner-managers greater ability to create and manage viable enterprises. Therefore, we expect high values of human capital proxies to be predictors of company survival. We used variables that represent formal as well as informal human capital as well as variables that reflect general and specific human capital. We included the level of formal education, the field of any post-secondary study, family business background, previous ownership experience, years of prior work experience, and the number of owners. Some of these variables, like the level of education, are closely related to general human capital, while others, like family business background are probably more closely related to the specific skills required for running a small business.

Except for the number of owners, these human capital proxies need no explanation. A greater number of owners will increase start-up company viability for three reasons. First, multiple owners implies greater availability of managerial labor to accomplish the required tasks of launching and maintaining a new enterprise. Second, multiple owners may imply a greater variety of complementary skills. Third, multiple owners may proxy for a deeper commitment to a successful enterprise. All of our human capital variables, except for the number of owners, were categorical variables. We measured the variables as the proportions of owners in the various categories.

We included two indicators of whether or not the company had total equity between \$10,000 and \$25,000 or more than \$25,000 at start-up.¹¹ Firm size is commonly used as a variable to analyze financing conditions (e.g., Avery, Bostik, & Samolyk, 1998) and firm survival (e.g., Dunne & Hughes, 1994). As discussed in Section 2.1, firm size is endogenous to finance conditions at time of start-up. One can, however, consider the total amount of equity contributed by the owners as exogenous. Equity is a component of firm size. High equity also contributes to the viability of a company by reducing vulnerability to fluctuations in revenue.

We included two additional company descriptors, an indicator of whether or not the company was new or purchased from a previous owner, and an indicator of the legal form of the company. We expect an ongoing company with new owners to have a greater chance of survival than a completely new company because uncertainties about markets and production methods will be reduced for an ongoing company. This is consistent with previous work (Bates, 1990; Cressy, 1996). The impact of the legal form of the company on its survival probability is not clear. Sole proprietorships may have higher survival rates than partnerships and corporations because the costs of failure are borne, in theory, entirely by the owner, and this influences owner behavior towards maximizing business survival. On the other hand, partnerships and, in particular, incorporated businesses, may incur substantial start-up organizational costs when forming. These costs may therefore serve to sort firms by their expected survival probabilities. The legal form's net effect on the survival probability is unknown. We found no large empirical differences in the coefficients for whether or not the firm is a partnership as opposed to a corporation. We, consequently, collapsed the two types of firms into one category and compared the category to sole proprietorships. (Regression results that separate the three legal entities are available from the corresponding author on request.)

The industry descriptors we adopted were proxies for scale economies and historic survival rates. The descriptors reflected the 2-digit SIC level of aggregation. We expect that start-ups in industries that offer apparent advantages for large companies or plants would have lower chances of survival than companies in industries where small companies appear to do well. We proxied scale economies with the percentages of plants in the industries with either 1–19 employees or with 1–49 employees. Data for these proxies were obtained from *County Business Patterns* (U.S. Bureau of the Census, 1987). We also tried to proxy unobservable industry effects on business survival with the industry average survival rate of businesses launched between 1980 and 1982 by non-minority males and females. These data were obtained for the 1982 CBO. We found no significant additional industry effects using 1980–1982 survival rates, and do not report results with it.¹²

4. Statistical procedures and results

Our statistical procedures had several steps. First, we implemented weighted least squares to estimate predicted sales during 1987 for the set of 923 observations meeting the selection criteria discussed in Section 2.1.¹³ We applied the criterion of predicted sales greater than \$5,000 to reduce the sample from 923 observations to 738 observations. Weighted univariate statistics for 738 observations in one of five complete data sets are in Table 3. The table also reveals sample characteristics broken down by whether there was a bank loan or not. Most striking is the difference in education across the two sub-samples. Firms with bank loans had owners with less education, less previous business experience, and less work experience. Firms with bank loans more often had more than one owner, were less often started from scratch, and were less likely to have loans from other sources. We also computed sales in the first year of operations and found that while average sales revenues were \$93,341, those with bank loans had average sales more than twice of those without bank loans: \$167,751 versus \$71,700 ($t = 1.97$, $p < 0.05$).

Table 3

Total sample weighted univariate statistics^a

	Mean	SD	Bank loan		<i>t</i> -value
			Loan = 0	Loan = 1	
survive	0.692	0.390	0.695	0.679	0.626
edu3	0.251	0.366	0.186	0.468	−9.545
edu4	0.309	0.392	0.359	0.137	9.817
edu5	0.205	0.336	0.219	0.155	3.174
edu6	0.093	0.240	0.112	0.030	5.256
scieng	0.112	0.268	0.121	0.084	1.563
business	0.314	0.392	0.347	0.201	5.528
wrkfam	0.279	0.380	0.284	0.262	0.821
owner	0.248	0.360	0.263	0.195	2.745
workexp4	0.303	0.390	0.309	0.284	0.927
workexp5	0.247	0.364	0.241	0.269	−1.075
workexp6	0.268	0.369	0.286	0.208	3.036
respno	1.107	0.373	1.088	1.168	−2.872
denovo	0.693	0.394	0.720	0.601	4.151
toteq3	0.092	0.247	0.092	0.094	−0.118
toteq4	0.134	0.291	0.134	0.133	0.013
inc	0.182	0.330	0.173	0.215	−1.727
scale20	0.883	0.068	0.884	0.880	0.921
seale50	0.956	0.037	0.956	0.955	0.157
loan	0.230	0.360	0.000	1.000	
othloan	0.232	0.361	0.256	0.138	4.832

^a All data are weighted to reflect sampling plan and survey non-response. Data are from one of five complete data sets.

We conducted a preliminary investigation of the value of having a bank loan on survival. There is a negative correlation between the bank loan indicator and the survival indicator, with $r = -0.160$ ($p < 0.01$). In contrast, there is a positive correlation between the indicator for other loans and the survival indicator, with $r = 0.212$ ($p < 0.001$).¹⁴ Analyzed in another manner, we found the probability of survival for those with bank loans to be 0.686, while it was 0.696 for those without bank loans. These preliminary results indicate either that banks are poor at selecting among start-up applicants, or that more of the poor start-ups apply for bank loans.

To investigate the marginal value of having a bank loan, we applied weighted maximum likelihood to estimate the probit prediction model separately for each of the five complete data sets using Huber robust standard error estimators (White, 1982). Estimation results are reported using the formulae in Little and Rubin (1987, equations 12.17–12.20).¹⁵

The results of the maximum likelihood estimate of the probit survival model are in Table 4. The most interesting results indicate that holding a commercial bank loan or other loans were both significantly positive predictors of survival, *ceteris paribus*. Note that the coefficient for other loans is more than twice the magnitude of that for a commercial bank loan. This difference is not statistically significant, with the *t*-statistic being only slightly greater than one, however. The difference in coefficient values suggests that there may be sorting by borrowers. Those with high unobservable business qualities may choose a source of debt with less risk sharing

Table 4

Probit MLE of predictors of start-up business survival^a

Predictor	Coefficient estimate ^b	SE	<i>t</i> -value	<i>p</i> -value	Degrees of freedom
edu3	0.475	0.356	1.333		11
edu4	0.760	0.332	2.292	*	15
edu5	0.558	0.335	1.665	*	68
edu6	1.555	0.411	3.763	**	79
scieng	0.214	0.324	0.661		29
business	0.024	0.239	0.101		19
wrkfam	0.119	0.253	0.470		10
owner	0.508	0.264	1.926	*	9
workexp4	0.244	0.379	0.643		7
workexp5	0.138	0.412	0.335		6
workexp6	0.459	0.416	1.102		6
respno	0.140	0.216	0.646		50
denovo	−0.158	0.227	−0.696		194
toteq3	0.279	0.359	0.777		18
toteq4	0.355	0.308	1.150		11
inc	0.549	0.298	1.842	*	17
scale20	5.238	3.548	1.477	+	10
scale50	−10.088	6.137	−1.644	+	12
loan	0.347	0.240	1.445	+	72
othloan	0.730	0.361	2.021	*	6
_cons	4.198	3.171	1.324		13

Average *n* = 727.4Average pseudo- R^2 = 0.183

Average chi-square

(d.f. = 20) = 62.55, $p < 0.001$ ^a All data are weighted to reflect sampling plan and survey non-response. Standard errors are robust.^b Coefficient values, standard errors, and degrees of freedom computed over five samples according to [Little and Rubin \(1987, equations 12.17–12.20\)](#).+ $p < 0.10$.* $p < 0.05$.** $p < 0.01$.

and, by extension, a lower interest rate. That is, there is weak evidence of adverse selection against bank loans.

The human capital and company coefficients have the expected signs. The positive sign for the coefficient of the proportion of firms in the 2-digit industry with fewer than 20 employees was expected. The negative sign for the proportion of firms with fewer than 50 employees is somewhat surprising. It suggests a very low threshold for scale economies.

That there was a negative correlation between having a bank loan and the probability of survival, while having a bank loan had a positive coefficient in the equation that included other determinants of survival, suggests that start-ups with strong values for the other predictors did not get bank loans. We examined the relation between having a commercial bank loan and applicant characteristics. We estimated a simple probit equation that explained the probability

Table 5

Probit MLE of predictors of bank loans^a

Predictor	Coefficient estimate ^b	SE	<i>t</i> -value	<i>p</i> -value	Degrees of freedom
edu3	−0.085	0.446	−0.19		12
edu4	−1.069	0.468	−2.28	*	11
edu5	−0.554	0.439	−1.26		264
edu6	−1.087	0.551	−1.97	*	43
scieng	0.131	0.437	0.30		38
business	−0.453	0.299	−1.51	+	28
wrkfam	0.165	0.261	0.63		18
owner	0.403	0.339	1.19		33
workexp4	−0.229	0.414	−0.55		13
workexp5	−0.611	0.426	−1.44	+	9
workexp6	−0.886	0.476	−1.86	+	7
respno	0.904	0.354	2.56	*	17
denovo	0.162	0.250	0.65		63
toteq3	−0.065	0.471	−0.14		6
toteq4	−0.173	0.396	−0.44		8
inc	0.062	0.362	0.17		9
scale20	2.356	4.089	0.58		3,329
scale50	0.104	7.757	0.01		265
_cons	−2.104	4.391	−0.48		57

Average $n = 392.8$ Average pseudo- $R^2 = 0.201$

Chi-square (d.f. = 19) = 42.16,

 $p < 0.0015$

^a Coefficient values, standard errors, and degrees of freedom computed over five samples according to [Little and Rubin \(1987, equations 12.17–12.20\)](#).

⁺ $p < 0.10$.

^{*} $p < 0.05$.

of having a bank loan, *given that the company had a loan of some type*.¹⁶ That is, we estimate the probability of obtaining a bank loan among those owners that actively sought debt capital. We found that the probability of having a bank loan decreases with education, work experience, and also with equity. The results appear in [Table 5](#).

Studies of bank practices in making commercial loan decisions show that they stress financial evidence of creditability.¹⁷ However, it is not reasonable to assume that banks discriminate against applicants with high levels of education or work experience. They certainly do not discriminate against loan applicants with high levels of equity. It is likely that start-ups with high levels of these characteristics instead self-selected against bank loans. It appears that start-ups with owners that have high human capital and that have sufficient wealth to have high levels of equity are able to get other, more favorable, non-bank loans. This substitution between bank loans and other loans shows more directly in the negative correlation, over the entire sample, between the bank loan indicator and the indicator for other loans, where $r = -0.622$ ($p < 0.001$).

Table 6
Pooled results for various groups in hold-out sample^a

Variable ^b	Hold-out sample (1) ^c	$p(\text{survive}) = 0.9$ (2) ^c	Bank loan (3) ^c	Bank loan and $p = 0.9$ (4) ^c	No bank loan and $p = 0.9$ (5) ^c	No bank loan ^c
survive	0.694 (0.019)	0.858 (0.063)	0.686 (0.067)	0.816 (0.092)	0.872 (0.087)	0.696 (0.020)
loan	0.240 (0.023)	0.278 (0.113)	1.000 (0.000)	1.000 (0.000)	0.000 (0.000)	0.000 (0.000)
othloan	0.215 (0.026)	0.428 (0.081)	0.143 (0.106)	0.256 (0.035)	0.504 (0.127)	0.238 (0.034)
Number	279 (17.7)	65 (10.2)	67 (6.6)	18 (7.1)	47 (11.7)	212 (16.2)
Percent of hold-out sample (%)	100	23	24	6	17	76

Difference in survival rates, col. 1 versus col. 2: $t = -11.203^{***}$; col. 1 versus col. 3: $t = 0.542$; col. 2 versus col. 3: $t = 8.382^{***}$; col. 3 versus col. 4: $t = -5.089^{***}$; col. 3 versus col. 6: $t = -0.63$; col. 2 versus col. 4: $t = 1.701$; col. 5 versus col. 3: $t = 7.575^{***}$.

* $p < 0.05$, d.f. = 18, two-tailed t -test.

** $p < 0.01$, d.f. = 18, two-tailed t -test.

*** $p < 0.001$, d.f. = 18, two-tailed t -test.

^a All data are weighted to reflect sampling plan and survey non-response.

^b Results are pooled from estimations on 10 different subsets applied to 10 hold-out samples.

^c Standard deviation over 10 estimations in parentheses.

To analyze the relative predictive value of the survival equation, we re-estimated the model for 10 random subsets of one of the five complete data sets. We subsequently applied the coefficients from the re-estimation to a holdout sample (subset #2). Pooled weighted univariate statistics for subset #2 across 10 repeated forecasts are shown in Table 6. We divided subset #2 according to two factors. The first factor used for dividing the holdout sample was whether or not the companies held bank loans at the time of start-up. The second factor was whether or not the predicted survival probability, based on the coefficients from the subset #1 re-estimation, was at least 0.9.

There is evidence that the model is reliable. The average survival rate for the entire subset #2 was 69.4%, while the average rate for those companies with predicted probability of survival of at least 0.9 was 85.8%. The difference is significant at better than 0.1% ($t = -11.2$, d.f. = 18). Notice that the average survival rate for all those firms receiving bank loans, 68.6%, was not significantly different from the rate for entire subset #2 ($t = 0.54$). Notice, as well, that the average survival rate of companies with $p(\text{survive}) \geq 0.9$, 85.8%, was better than that of companies holding commercial bank loans, 68.6% ($t = 8.38$, $p < 0.001$). In addition, the average size of the sample with $p(\text{survive}) > 0.9$, 65%, was not significantly different from the sample holding bank loans, 67% ($t = -0.77$). Even the survival rate of companies with $p(\text{survive}) \geq 0.9$ that did *not* hold a commercial bank loan at start-up, 87.2%, was significantly better than that of companies securing commercial bank loans, 68.6% ($t = 7.58$, $p < 0.001$). And the group that had a high probability of survival but did not hold bank loans is almost as large (70%) as the group with start-up loans. These results imply that banks could have fared better in terms of exposure to borrower's business failure risk with an increase in the number of loans granted, if their loan portfolios had included start-ups with high values of the reported predictors of survival.

5. Conclusion

We investigated the relation between the survival of new small businesses (start-ups) and their having a bank loan. As well, we were able to make inferences concerning the selection of capital sources by start-up businesses. A bank loan is likely to improve the survival chances of a new small business for many reasons. A probit regression model of start-up company survival based on owner human capital, loan sources, and wealth as well as industry and company characteristics showed indeed that having a bank loan is, *ceteris paribus*, a significant positive predictor of the survival of start-up companies. However, the unconditional correlation between having a bank loan and survival was negative. In contrast, an unconditional positive correlation between having other loans and survival was established. There appears to have been a substantial number of start-ups with high survival rates that did not receive bank loans. These companies made significantly more use of other sources of borrowed capital than did those companies receiving bank loans. There was evidence of self-selection against commercial bank loans by owners with high levels of human capital and wealth. There was also weak evidence of adverse selection against banks.

We have reported evidence of self-selection of highly qualified start-up owners away from commercial bank loans. The other side of this phenomenon is that banks may have ceded the

business with highly qualified applicants to the informal market. Banks may have deemed it more profitable to compete in a niche with higher default rates and correspondingly higher interest rates. A possible explanation for this concerns the difference in information costs between banks and informal lenders. We have noted that banks focus on financial evidence of creditability, as opposed to information about owners' human capital, in making loan decisions. This focus makes it difficult for banks to seek business from highly qualified owners. The information that we used in constructing our model would not be expensive to collect. (One might expect such information to be required routinely from managerial job applicants.) However, it may be that verification costs are too high to be justified by a loan to a start-up. The problem may be most acute for informal elements of human capital like the extent of work experience and family background. Verification might be much easier for lenders in the informal loan market.

Notes

1. The CBO is available on-site at several Census Research Data Centers.
2. The 1982 CBO is documented for example by [Bates \(1990\)](#) and [Holmes and Schmitz \(1996\)](#). But there currently exists little documentation on the 1992 survey.
3. While most researchers use business closure to indicate business failure (e.g., [Brüderl, Preisendörfer, & Ziegler, 1992](#); [Holtz-Eakin et al., 1994a](#)), Cressy use closing of a bank account as an indicator of business failure.
4. The model used to predict sales contained three groups of variables: owner characteristics, firm characteristics, and industry characteristics. They were: owner age, owner education, type of graduate degree (if applicable), household income, whether the owner was single or not, whether the owner was born in a country other than the US, prior work experience, previous business ownership experience, if close relatives owned a business, if the owner had worked for these relatives, the legal form of the business, whether the firm was a franchise or not, whether the firm was started *de novo* or not, the amount of time spent on the business, number of owners, total equity, 2-digit industry mean assets per employee for sole proprietorships and partnerships (from *Statistics of Income*), whether a knowledge-based industry or not as defined by [Baldwin, Hinchley, and Johnson \(1997\)](#), labour scale economies (from *County Business Patterns*), 2-digit industry average survival rate between 1982 and 1986 of new small start-ups (from 1982 CBO), and average asset size of small start-ups surviving operations between 1982 and 1986 (from 1982 CBO). The model had an R^2 of approximately 0.60.
5. The weights reflected the unequal probability of being selected in the stratified sampling for the CBO and the probability of survey response.
6. Missing data are not imputed in cols. (3) and (4) of [Table 1](#) for the purpose of comparability across data sources. Frequency counts therefore differ, and are generally lower than with imputed data. Because data are not imputed we cannot use predicted sales to display data. We instead use actual sales to present data in col. (4) of [Table 1](#). This choice was made purely for illustrative purposes.
7. Other surveys of small business finance, such as the National Survey of Small Business Finance (NSSBF), experience even lower response rates. The 1993 NSSBF experienced a 59% unit non-response rate ([Price Waterhouse LLP, 1996](#)).

8. The reason for this substantial drop in item response is a questionnaire design error: respondents to the preceding question regarding highest level of education were not asked to skip the college concentration question if they did not attend college. After classifying as ineligible for response all those owners reporting their highest level of education as below college attendance, the item response on college concentration increases to 93.1%.
9. This assumption is corroborated by [Avery et al. \(1998\)](#) who show that 77.0% of all small business loans extended in 1987 were secured, in part, by business collateral. However, many smaller business loans are typically unsecured as the relative cost of valuing assets becomes too high.
10. Regression results were unfortunately sensitive to low cell counts when using more than one indicator variable to distinguish these various sources of debt from each other. We include the effect of obtaining a mortgage on a house to finance the start-up as “other loan” because a bank need not evaluate the business’ prospects when an entrepreneur applies for a home mortgage. Discussions with lenders confirmed this view.
11. There was an apparent ambiguity in the questionnaire from which the data were derived. It appeared that some persons who utilized personal loans to enter the business did not count the proceeds of the loans as part of their equity. Other owners did count the loans as part of equity. We noticed this ambiguity because there was substantial numbers of companies with positive capital that claimed to have no equity. This is why we were forced to use these broad categories—essentially “low equity” and “high equity.”
12. We tried other variables that did not have explanatory power. There were three human capital proxies. These were indicators for prior managerial experience, the age of the owner, and whether or not the owners had ever been married. A proxy for the wealth of the owners did not have an appreciable effect on business survival. There was also one additional industry descriptor, the median assets for survivors in the same 2-digit industry in the 1982 CBO.
13. The explanatory variables included human capital variables, age, education, work experience, citizenship status, marital status, and employment status in the business. The company descriptors included the following: total equity, franchise, new company as opposed to an acquired existing company, and the legal form of the company. Industry descriptors were as follows: the survival rate of companies, and median assets for survivors in the same 2-digit industry in the 1982 CBO, and scale economies. There was also a proxy for personal wealth of the owners.
14. These results are consistent with [Avery et al.’s \(1998\)](#) finding that personal commitments appear to be substitutes for business collateral.
15. It should be noted that the standard errors increase significantly when using these formulae, compared to standard errors estimated for any of the five complete data sets. It should also be noted that standard errors are larger for any of the complete data sets where missing data have been imputed using the Bayesian multinomial model than when using (traditional) mode (or mean) imputation ([Chen & Åstebro, 2003](#)).
16. The reason for not estimating a bivariate probit of bank loan selection and business survival ([Boyes, Hoffman, & Low, 1989](#); [Greene, 1998](#); [Jacobson & Roszbach, 2000](#) use bivariate probits for credit card selection and default), was that we were limited in the number of explanatory variables that would uniquely identify the bank loan decision.

We also detected a significant correlation between the error terms. The combination of these two facts leads to convergence problems for the bivariate probit.

17. For a detailed review of industry practice in small business lending see Eisenbeis (1996) and Mester (1997). Credit scoring has only recently been applied to commercial loans, starting in the early 1990s (Berger, Frame, & Miller, 2001). The personal information used in commercial loan credit scores may include the owner's monthly income, outstanding debt, financial assets, employment tenure, home ownership, and previous loan defaults or delinquencies (Mester, 1997).

Acknowledgments

Åstebro acknowledges financial support from the Natural Sciences and Engineering Research Council of Canada and the Social Sciences and Humanities Research Council of Canada's joint program in Management of Technological Change as well as support from the Canadian Imperial Bank of Commerce. We thank Gongyue Chen for research assistance.

References

- Avery, R. B., Bostik, R. W., & Samolyk, K. A. (1998). The role of personal wealth in small business finance. *Journal of Banking and Finance*, 22, 1019–1061.
- Baldwin, J., Hinchley, C., & Johnson, J. (1997). *Successful entrants: Creating the capacity for survival and growth*. Catalogue No. 61-524-XPE, Statistics Canada.
- Bates, T. (1990, November). Entrepreneur human capital inputs and small business longevity. *Review of Economics and Statistics*, 72, 551–559.
- Berger, A. N., Frame, W. S., & Miller, N. H. (2001, September 5–7). Credit scoring and the price and availability of small business credit. In L. C. Thomas, J. N. Crook, & D. B. Edelman (Eds.), *Proceedings of conference on credit scoring and credit control* (Vol. VII). University of Edinburgh, Edinburgh, Scotland.
- Best, R., & Zhang, H. (1993). Alternative information sources and the information content of bank loans. *Journal of Finance*, 48, 1507–1522.
- Bester, H. (1985). Screening versus rationing in credit markets with imperfect information. *American Economic Review*, 75, 558–627.
- Blanchflower, D. G., & Oswald, A. J. (1998). What makes a young entrepreneur? *Journal of Labor Economics*, 16, 26–60.
- Boden, R. J., & Nucci, A. R. (1997). Counting the self-employed using household and business sample data. *Small Business Economics*, 9, 427–436.
- Boyes, W., Hoffman, D., & Low, S. (1989). An econometric analysis of the bank credit scoring problem. *Journal of Econometrics*, 40, 3–14.
- Brüderl, J., Preisendörfer, P., & Ziegler, R. (1992). Survival chances for newly founded business organizations. *American Sociological Review*, 57, 227–242.
- Chen, G., & Åstebro, T. (2003). *How to deal with missing categorical data: Test of a simple Bayesian method* (under review).
- Cressy, R. (1996, September). Are business startups debt rationed? *Economic Journal*, 106, 1253–1270.
- Dunne, P., & Hughes, A. (1994). Age, size growth and survival: UK companies in the 1980s. *Journal of Industrial Economics*, XLII, 115–140.
- Eisenbeis, R. (1996). Recent development in the application of credit scoring techniques to the evaluation of commercial loans. *IMA Journal of Mathematics Applied in Business and Industry*, 7, 271–290.
- Evans, D. S., & Jovanovic, B. (1989). An estimated model of entrepreneurial choice under liquidity constraints. *Journal of Political Economy*, 97, 808–827.

- Greene, W. (1998). Sample selection in credit-scoring models. *Japan and the World Economy*, 10, 299–316.
- Holmes, T. J., & Schmitz, J. A., Jr. (1990). A theory of entrepreneurship and its application to the study of business transfers. *Journal of Political Economy*, 98, 265–294.
- Holmes, T. J., & Schmitz, J. A. (1996). Nonresponse bias and business turnover rates: The case of the characteristics of business owners survey. *Journal of Business & Economic Statistics*, 14, 231–241.
- Holt, D., Smith, T. M. F., & Winter, P. D. (1980). Regression analysis of data from complex surveys. *Journal of the Royal Statistical Society*, 143(Series A, Part 4), 474–487.
- Holtz-Eakin, D., Joulfaian, D., & Rosen, H. S. (1994a). Sticking it out: Entrepreneurial survival and liquidity constraints. *Journal of Political Economy*, 112, 53–75.
- Holtz-Eakin, D., Joulfaian, D., & Rosen, H. S. (1994b). Entrepreneurial decisions and liquidity constraints. *RAND Journal of Economics*, 25, 334–347.
- Jacobson, T., & Roszbach, K. F. (2000). Evaluating bank lending policy and consumer credit risk. In Y. Abu-Mostafa, B. LeBaron, A. W. Lo, & A. S. Weigand (Eds.), *Computational finance 1999* (pp. 535–548). Cambridge: MIT Press.
- Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: Wiley.
- Mester, L. (1997). *What's the point of credit scoring?* *Federal Reserve Bank of Philadelphia business review* (pp. 3–16).
- Nucci, A. R. (1992). *The characteristics of business owners database*. Discussion Paper 92-7. Center for Economic Studies, U.S. Bureau of the Census.
- Price Waterhouse LLP. (1996). *National Survey of Small Business Finance: Methodology report*. Available from the Federal Reserve Board's web site: <http://www.bog.frb.fed.us/pubs/oss/>.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley.
- Stiglitz, J., & Weiss, A. (1981). Credit rationing in markets with imperfect information. *The American Economic Review*, 71, 393–410.
- U.S. Bureau of the Census. (1987). *County Business Patterns*. Government Printing Office, Washington, DC.
- U.S. Bureau of the Census. (1991). *Characteristics of business owners, government printing office*. Washington, DC.
- White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica*, 50, 1–25.