A NETWORK SAMPLING PROCEDURE FOR ESTIMATING THE PREVALENCE OF NASCENT ENTREPRENEURS

Charles Pallit, Professor, University of Wisconsin Survey Research Laboratory
Paul Reynolds, Coleman Foundation Chair in Entrepreneurial Studies, Marquette University

Introduction

Conventional sampling procedures require that population elements have only one pathway into the sample. For example, a conventional Random Digit Dial sample gains access to the adult population by linking every adult to a housing unit (HU). Every adult in the population is thought of as linked to one and only one HU. An adult can enter the sample if and only if the adult is linked to a sample housing unit. This restriction, "... every adult is linked to one and only one HU," simplifies the theory somewhat, and helps to make the determination of the sample selection probabilities a little easier. However, other sampling procedures which relax this restriction and use rules allowing population elements multiple pathways into the sample are possible. These procedures have been referred to as network sampling procedures. The estimation procedures associated with these relaxed rules are called multiplicity estimators (Sirken 1974).

Levy (1977) defines the network sample survey as:

"A survey with multiplicity is one in which an element (e.g., birth, death, . . . , etc.) may be linked to more than one enumeration unit by an algorithm or counting rule."

Sirken (1974) points out that the class of multiplicity rules "... which have the property of supplementing the condition of a conventional rule with other conditions for linking elements to enumeration units" are of special interest because they permit the use of both conventional and multiplicity estimators.

As an example of such rules, Sirken presents the rule which links persons to their own residence, as well as the residences of their siblings and children. He points out that this rule permits four sets of estimates based on the following rules:

3. persons are linked to their own residence and to the residences of their children;
4. persons are linked to their own residence and to the residences of their siblings and children.

The first rule is a conventional rule. The other rules are multiplicity rules.

The multiplicity rules, i.e., rules 2, 3, and 4 casts nets which produce a greater number of adults in the final sample than the conventional rule 1, even though the number of residences in the sample is the same for all rules.

When the multiplicity rules are used, the selection probability for each adult is a function of the number of residences to which the adult is linked and the selection probabilities attached to those residences.

We use similar multiplicity rules.

Our objective was to obtain a probability sample of Wisconsin residents who are nascent entrepreneurs. The general procedure selected for this was as follows:

1. Select an RDD sample of phone numbers with replacement.
2. From each sample household contained in the RDD sample select a random adult respondent.
3. Ask the random respondent to list a well-defined set of individuals resident in Wisconsin. This list defines the 'net' from which the nascent entrepreneurs will be drawn.
4. Ask the random respondent which of the persons listed, if any, is a nascent entrepreneur. This identifies a pool of potential nascent entrepreneurs.
5. Interview the resulting sample of potential nascent entrepreneurs to determine if they are in fact entrepreneurs.
6. Ask the potential nascent entrepreneur to list the set of individuals resident in Wisconsin equivalent to the set defined in item 3. The count on this list is the size of the net used to catch the respondent. (Subject perhaps to some reporting error.)

We call a net whose size is obtainable as in step 6 a nice net. Not all nets are nice. We can define nets for which the specific net size is unavailable. Subject to one important assumption, and the use of a nice net the above rules allow us to specify the probability of selection for each nascent entrepreneur identified. Armed with this information we can produce valid estimates of the prevalence of nascent entrepreneurs in the population. The assumption is that the random respondent in the RDD sample identifies all nascent entrepreneurs in the net listed by the respondent. If this assumption fails then the prevalence estimate is an estimated lower bound for the true prevalence.

In this study five nice nets were used:

- a. Adult siblings by blood or marriage
- b. Parents or step-parents
- c. Grand-parents
- d. Adult children
- e. Spouse or Partner

These nets are defined by the response to the following questions. Two questions are used for each net. The question of residency is settled later in the process.

**SIBLING NET QUESTIONS**

**QUESTION 6**
My next question is, how many living brothers and sisters, INCLUDING step-brothers, step-sisters, half-brothers and sisters, adopted brothers and sisters do you have, if any?

**QUESTION 7**
I’d like to ask you some questions about them. Please give me just their FIRST name so that we may refer to them specifically later. What is the first name of one of your brothers or sisters? (BE SURE TO INCLUDE STEP, HALF AND ADOPTED SIBLINGS ON THIS TABLE) (ENTER THE NAME AND SEX, GO TO NEXT ONE)

**PARENT NET QUESTIONS**

**QUESTION 9**
Next, how many of your parents or step-parents are still living, if any?

**QUESTION 10**
I’d like to ask you some questions about your living parents or step-parents. Please give me just their FIRST name so we may refer to them later.

**GRAND-PARENT NET QUESTIONS**

**QUESTION 12**
Next, how many of your grand-parents are still living, if any?

**QUESTION 13**
I’d like to ask you some questions about these people (this person). Please give me just their FIRST name so that we may refer to them specifically later in the interview. What is the first name of one of your grandparents? (ENTER THE NAME AND SEX, GO TO NEXT ONE)

**CHILDREN’S NET QUESTIONS**

**QUESTION 15**
How many children, if any, do you have who are age 18 or older, including adopted children and step-children?

**QUESTION 16**
I’d like to ask you some questions about these people. Please give me just their FIRST name so that we may refer to them specifically later in the interview. What is the first name of one of your children, age 18 or older, including step and adopted children? (ENTER THE NAME AND SEX)
SIGNIFICANT OTHER NET QUESTIONS

QUESTION 18
Are you currently married or living with a partner?
1. MARRIED OR LIVING WITH A PARTNER
2. NOT MARRIED AND NOT LIVING WITH A PARTNER
7. DON'T KNOW
9. NOT ASCERTAINED / REFUSED

QUESTION 19
I'd like to ask you some questions about your spouse or partner. What is your spouse or partner's first name. (ENTER THE NAME AND SEX)

Counting Rules and Weights

For nice nets the selection probability for the entry of a nascent entrepreneur into the sample is determined in good measure by the number of ways that the entrepreneur could enter the sample. For example if the net is defined to be adult siblings, and a particular entrepreneur has no siblings then the net size is one and the probability of selecting the entrepreneur into the sample is

\[ \frac{p}{h_i} \]

where \( p \) = The selection probability for the entrepreneur's HU
and \( h_i \) = The number of adults in the entrepreneur's HU

With two sibling, there are two selection path-ways, the selection probability becomes:

\[ \frac{p}{h_i} + \frac{p}{h_2} \]

where \( p \) and \( h_i \) are as before.

The first term is for the nascent entrepreneur and the second is for the sibling.

When no siblings exist the selection probability can be determined exactly. However as the number of siblings increase this becomes more problematic. For \( k \) siblings the selection probability can be written as the sum of the respondent selection probabilities from each adult in the net, i.e.,

\[ \frac{p}{h_i} + \frac{p}{h_2} + \frac{p}{h_3} + \ldots + \frac{p}{h_k} \]

Exact calculation of the probabilities requires a knowledge of the household size for each potential respondent, i.e., \( h_i \). Most of the values for \( h_i \) will be unknown to us. However we should always know the size of at least one household in the set, i.e., the value of \( h_i \) for the nominators household. So even though we can not compute the exact probability for the net, we can compute an estimate by using the nominator's household size \( h_i \).

Let \( d_i = \frac{1}{h_i} \).

Then the net selection probability is

\[ \Sigma d_i = kd \]

where

\[ d = \frac{1}{h} \Sigma d_i \]

Using \( d_i = \frac{1}{h_i} \), where \( h_i \) is the nominator's household size, as an estimate for \( d \), we can estimate the selection probability with

\[ kd_i \]

This probability can be used to estimate the total number of nascent entrepreneurs in the population. For a sample of one entrepreneur the estimate is

\[ e = (1/kd_i) \]

Note that this is a biased estimator, since \( d_i \) is a random variable. If \( m \) entrepreneurs are found in the sample then the estimated total is

\[ e = \Sigma_i = (e_i/m) \]

where

\[ e_i = (1/kd_{i0}) \]

The actual estimator used was adjusted for response rate, and was

\[ E = e/R \]

where \( 100R \) = the response rate.

We were fortunate enough to be able to include one example of a net which is not a 'nice net' in this study. In this net the entrepreneur is reached through
a network made up of friends and work confidants. As with the nets previously discussed the elements of this net are defined by the answers to questions asked of the first phase respondents. The questions are:

**WORK CONFIDANT NET QUESTIONS**

**QUESTION 3**
In this part of the study, we are exploring how informal sharing of help among friends and family actually works. We would like to ask you a few questions about those you talk with about work and career issues. First, we will explore how people might rely on each other.

If you were thinking about a major career change, like looking for a job or starting a new business, among the people that you know how many would you talk to about your career change?

**QUESTION 4**
I'd like to ask you some questions about these people. Please give me just their FIRST name and gender so that we may refer to them specifically later in the interview. What is the first name of one of the people you think you would talk to about your career change? (ENTER THE NAME AND PRESS <ENTER>)

Determining the size of this net is difficult. People who consider themselves part of this net for a given entrepreneur may not be regarded in the same light by the entrepreneur, etc. However in our investigation we decided to use the average size of an entrepreneur's network as reported by the entrepreneur as a value for k. Clearly this is an estimate and its use effects the error distribution of our estimator. It's use has the potential for adding both additional bias and additional random error to the estimate. Nevertheless we thought that this study gave us an opportunity to empirically test the potential effectiveness of such a procedure. We do this by comparing the estimates from each net as well as the variances of these estimates.

**Estimating The Variance**
We estimate the variance using a replication procedure with,

\[ \sum_{i=1}^{a} (E_i - E)^2 \]

\[ \frac{a(a-1)}{\sum_{i=1}^{a} E_i^2} \]

where a= Number of replicates,

\[ E_i = \text{The estimate of the number of entrepreneurs from the } i^{th} \text{ replicate, and} \]

\[ E = \frac{\sum_{i=1}^{a} E_i}{a} \]

In practice we used a disproportionate stratified RDD sample for the first phase and two replicates in the variance computation.

**Results**
Estimates of the number of nascent entrepreneurs in Wisconsin were computed for each type of net. These estimates and their associated standard deviations are displayed in Table 1.

**TABLE 1**
Estimates of The Number of Nascent Entrepreneurs in Wisconsin Using Different Types of Nets

<table>
<thead>
<tr>
<th>Net</th>
<th>Estimate</th>
<th>Est. S.D.</th>
<th>Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Siblings</td>
<td>81,980</td>
<td>31,088</td>
<td>13</td>
</tr>
<tr>
<td>Parents</td>
<td>100,212</td>
<td>114,057</td>
<td>10</td>
</tr>
<tr>
<td>Grand-parents</td>
<td>7,689</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Adult Children</td>
<td>71,957</td>
<td>26,892</td>
<td>9</td>
</tr>
<tr>
<td>Signify other</td>
<td>164,915</td>
<td>27,297</td>
<td>16</td>
</tr>
<tr>
<td>Work Confidant</td>
<td>103,404</td>
<td>67,754</td>
<td>31</td>
</tr>
<tr>
<td>No Net</td>
<td>146,870</td>
<td>40,688</td>
<td>8</td>
</tr>
</tbody>
</table>

The "No Net" entry shows the estimates and standard deviation which results if no network is used, i.e., only first phase data is used. This estimate is based only on the randomly selected respondents from the RDD sample used in the first phase.
The grand-parent net clearly fails because the frequency with which grand-parents are nascent entrepreneurs is very low.

The best of the remaining nets are the significant other net, the sibling net, and the adult children net. All these have about the same level of precision. There is a large difference between the significant other estimate of 164925 and the other two estimates. However, these differences are not large enough to prove malice, but large enough to raise questions and to suggest that the significant other may be a more reliable informant.

The work confidant net, that is the 'not nice net' seemed to work reasonably well. It's standard deviation is larger than most of the other nets, but that is expected.

Summary And Conclusions

The best single net is the significant other. This suggests that for nascent entrepreneurs the appropriate procedure in the future is to select a sample of households and to screen all adults in these households for nascent entrepreneurs. Such a procedure has the additional advantage of avoiding the tracking loss associated with the network procedure, which is quite high.

References


AN EFFICIENT NEW DUAL FRAME SAMPLE DESIGN FOR SURVEYS OF SMALL BUSINESS

David A. Hirschberg, U.S. Small Business Administration and Harold Nisselson, Westat, Inc.
Harold Nisselson, Westat, Inc., 1650 Research Blvd., Rockville, MD 20850

KEYWORDS: Dual Frame, Survey Design, Sampling, Area Sampling, List Sampling

1. Introduction

The use of commercial lists of businesses as a sampling frame for establishment surveys is widespread and relatively attractive from a cost standpoint. It provides flexibility for sampling purposes since businesses in the frame can be stratified by selected characteristics and sampled at varying rates as appropriate. However, relying solely on commercial lists has serious drawbacks for surveys of small businesses, the foremost being the substantial coverage bias due to the fact that these lists tend to miss many of the new businesses and the smallest small businesses. This is true even for lists created by government activities (Birely, 1984; Aldrich, et al., 1988), which in any event, are not generally available outside the agency itself.

Annual net changes in the number of firms do not adequately reflect the large gross flows that affect the business population. In 1992, the number of businesses with employees covered by reports to the U.S. Employment and Training Administration increased by 1 percent. However, new or successor firms over the course of the year represented 15 percent of the initial frame while terminated firms represented 14 percent.

It is necessary to trace sample businesses that cannot be located at the address given in the frame, in order to determine whether they are currently out-of-business. If still in business but at a different location, followup would be needed to obtain interviews from them. These facts entail both costs for followup and bias due to failure to resolve all such cases. It is common experience to find a substantial proportion of the sample out-of-business or moved-and-not-locatable even after followup. Scheirer (1993), for example, reported an experience with a list two years old for which a third of the nonrespondents could not be found by long distance operators. Although the costs of these activities can be controlled by techniques such as subsampling, they may still be relatively large.

Area samples to identify and survey businesses minimize these problems of list sampling but are quite expensive and are not efficient for sampling different kinds of businesses at varying rates.

2. A Dual Frame Approach

At the request of the U.S. Small Business Administration, Westat developed, designed and piloted a dual frame approach which minimizes the coverage bias, can be implemented efficiently, and which meets the requirements for coverage and flexibility. The design integrates a commercial list with an area sample to supplement the list. An essential feature is to treat the list as representing a set of business addresses rather than the businesses themselves. This eliminates the need to find the new location of businesses that may have moved from their original sites, as well as giving the businesses that replaced the movers a chance of selection from the list sample. It is supplemented by an area sample of (nonresidential) buildings to help assure completeness of coverage. This is more efficient than supplementing a sample of named businesses. The dual approach avoids many of the problems created by treating the businesses while in a large part, retaining its statistical and cost advantages. It represents the first attempt known to us, to develop a (national) sample of businesses by developing a (national) sample of business locations.

3. Some Design Considerations

Figure A attached provides an outline of the stages that might be used in a multi-stage design under the dual frame design we explored and an independent area sample design we used as a standard for comparison. In a multi-stage design, the dual frame approach is implemented in the next-to-last stage, the final stage being the list/area samples. The design of the dual frames will vary from country to country.

We suggest that the next-to-last stage be a sample of mail-delivery areas such as ZIP code areas in the U.S.A. The reason for this is that commercial lists show the ZIP code location for almost every establishment they list. In addition, in the U.S.A. tabulations of economic census data are available by ZIP code -- an advantage for stratification and sampling for the list sample component. Finally, maps and boundary descriptions are available for ZIP code areas -- an advantage for the area sample supplement. If ZIP code areas cover a large geographic area, they may be segmented for further sampling.

Segmenting can be automated if a suitable geographic information system is available. In their paper at this conference, Petrucci and Pratesi (1993) describe methods of building an area frame for business establishment surveys using their Geographical Information System (GIS) for geocoding lists of establishments and defining area segments. In the U.S.A., the Census Bureau's TIGER system can be used to perform these functions efficiently. Area segments can also be defined in terms of road patterns which are useful for businesses that do not have an address location or are mobile.
The second stage is the sampling of area segments

(1) The contribution to the relvariance

are not subject

achieved by optimizing the sampling at each of the
case

assume that the primary and second-stage sampling

within PSUs; Var2 represents a term that reflects the variability in size

and subsampling rates of segments

The factor in square brackets represents the design
effect due to clustering of the sample within PSUs.

(2) The following simple cost function reflects the

variable cost of a three-stage design such as that

within PSU

\[
c = c_1 + c_2 \overline{T} + c_3 \overline{Tn}
\]

where

\[
c_1 = \text{variable cost per ZIP code in sample}
\]

\[
c_2 = \text{variable cost per segment selected in a}
\]

ZIP code

\[
c_3 = \text{variable cost per interview for a small}
\]

business selected in a segment

c1 includes the cost of sampling and of obtaining maps

of the boundaries of the selected ZIP code areas. c2 includes the cost of coding the commercial list source

addresses in the selected ZIP codes to Census
geography, establishing the segment boundaries,

selecting the sample segments and listing businesses in

the sample segments for the sampling of businesses for

interview. c3 includes the cost of interviewing to

achieve a completed interview, as well as the necessary

callbacks to achieve a satisfactory response rate.

(3) Following standard theory, the optimum value of

the design parameter \( \bar{n} \) is given by the equation

\[
\bar{n} = \left( \frac{c_2}{c_3} \right)^{1/2} \left( \frac{c_1 c_2 - c_3}{c_2} \right)^{1/2}
\]

and the number of segments per ZIP code by the equation

\[
\frac{A}{T} = \left( \frac{c_1 c_2}{c_2} \right)^{1/2}
\]

Sampling Variances and Cost - Dual Frame
Sample

The dual frame sample within a ZIP code may be
viewed as a stratified sample with two strata: (1) the
list frame and (2) the area sample frame.

(1) Accordingly, the contribution to the relvariance
of a survey estimate arising from the sampling
within a ZIP code area can be expressed as
The following example illustrates the sample design analysis:

**Independent Area Sample**

We assume segments from which the \( \bar{n} \) will be sampled that will average 60 listings. To be conservative, we take \( \delta_1 = 0.10 \) and \( \delta_2 = 0.0125 \), and based on experience in other surveys we take \( a = 0.5 \).

For cost factors we assume 3 minutes per listing for the sample segments, consistent with the experience in the pilot study, and a 45 minute interview. With the other cost factors involved we take

\[
c_2/c_3 = 210/45 = 4.67
\]

and

\[
c_1/c_2 = 120/210 = 0.57
\]

Then, referring to equations (1) and (2)

\[
\hat{n} = \sqrt{(4.67)(14)} = \sqrt{65.38} \approx 8
\]

\[
\hat{t} = \sqrt{(0.57)(8)} = 4.56 \approx 2
\]

Accordingly, the independent area sample design would sample 8 small businesses per segment and 2 segments per ZIP code area.

**Dual Frame Sample Optimization**

Referring to equation (3), it is reasonable to estimate that the list frame might cover 75 percent of small businesses and the area frame 25 percent. Then

\[ P_L/P_A = 3 \]

The design effect can be expressed approximately in the form

\[ 1 + \delta_3 \left( \frac{\bar{n}}{\bar{n}^*} - 1 \right) \]

where \( \bar{n}^* \) = average number of businesses per segment not represented in the list frame. If the same design is used as described for the independent area sample, segments of 8 businesses would average \( 8/4 = 2 \) businesses not represented in the list frame. Also, it is reasonable to estimate that in a list frame sample perhaps 20 percent might be lost due to post office addresses, business sites no longer in existence, and other reasons. Therefore, the list frame sample should designate 25 percent more businesses than the target to provide for losses.

Then with the cost factors used we take

\[ C_A/C_L = 1.6 \]

\[ \sqrt{C_A/C_L} = 1.3 \]

Also, with only an average of 2 businesses added per segment \((\text{DEF})_A\) would be 1.10 and \( \sqrt{(\text{DEF})_A} \) would be 1.05. So, referring to equation (3)

\[
\frac{\hat{n}_L}{\hat{n}_A} = (3) \left( \frac{1}{1.05} \right) (1.3) = 3.7
\]

Then the area sample proportion of a total allocation would be

\[
\frac{\hat{n}_A}{\hat{n}_A + \hat{n}_L} = \frac{1}{1 + 3.7} = \frac{1}{4.7} = 0.21
\]

or 20 percent, say, with the list frame accounting for 80 percent.

**6. The Pilot Study**

Six ZIP code areas were selected in which to conduct a pilot study:
1. An urban area of high-rise (multiple occupancy) office buildings
2. An area of dense retail and service establishments
3. A suburban area simulating a mixture of business clusters and strip-mall development along major streets
4. An area simulating a smaller town with surrounding residential neighborhoods
5. An area of small industry
6. A rural area.

We excluded from the pilot study government operated activities such as schools, hospitals, and the like.

Field staff were given:

- Area segment maps with boundaries marked
- Random systematic samples of business locations in each of the six study areas, taken independently from two major commercial list sources.

Businesses shown in the list frame with only a Post Office Box mail address were excluded from the list frame and assumed to be covered by the area sample. However, there are other ways of dealing with such cases.

Field staff had two tasks:

I. Attempt to find the list sample business locations and report the outcomes in terms of criteria which they were given

II. Independently canvas their assigned area segments. List all business locations they encounter and their current business (or other) occupants. Identify any establishments that could not be matched to listings in the commercial lists.

The definition of business location (BL) involves considerations similar to those for defining "listing units" in a demographic survey, in particular a coverage evaluation survey. To be efficient BLs should be addresses which according to the commercial list have fewer than a designated number of establishments. In larger multi-occupant buildings BLs should be described by specific location (e.g., floor, suite number). Directories displayed in multi-occupant buildings should not be used as the list for the structure. Rather, the building manager should be interviewed, and similarly for shopping malls.

Specific rules to help assure unbiased survey estimates are needed to deal with problems such as variations in the name by which an establishment may be shown in the list frame and the name of the same establishment as observed by an interviewer in the field, and with businesses found to have moved within, say, a given building or ZIP code area, to deal with problems such as possible multiple chances of selection. The field instructions must be readily implemented by the interviewers. The training of field staff in these respects should be as extensive as in demographic surveys, unlike the usual economic survey.

7. Pilot Study Findings

Use of a List Sample

Table 1 summarizes, by list source, the outcomes of the attempts to locate the sample list BLs in the field. The differences between the two list sources are not statistically significant. The percent of listings matched (or with a possibility of match) indicates the percent of the target universe expected to be covered by the list component. The complementary percent indicates how much of the list frame will be lost and, therefore, needed to be covered through the area sample. Our conclusion is that commercial business lists may usefully be considered as a sampling frame in the sense that a high proportion of the listings in the frame can be successfully located in the field. Survey effort may be conserved by limiting, under strict rules, the effort to find listings in individual cases. The effect of failure to find listings is to transfer the coverage of the BLs to the area sample supplement but not necessarily to introduce coverage biases.

Use of an Area Sample

Table 2 summarizes, by list source, the outcomes of the attempts to find occupied area sample listings anywhere in the list source entries for the entire ZIP code area. The area sample cases that could not be found in a list presumably represent undercoverage of the list frame, although in some cases they represent instances of incorrect address locations or incorrect ZIP codes for list source BLs.

Table 3 summarizes, by list source and both sources combined, the percent of cases in each ZIP code area that it was possible in the field to match to a list entry. This table shows that the use of a combination of lists may be cost effective, despite any problems in unduplicating the multiple lists for the list sampling.

Businesses not readily visible are likely to be missed in the area sample supplement. In particular, businesses operated out of private homes with no outward sign of business activity should not be taken to be within the scope of this dual frame approach. Such businesses are best covered by a household survey or, possibly, a list sample from business-income tax records.

Businesses not operated from fixed locations, such as street or road-side vendors, can be accommodated within the scope of the area sample. Special rules should be established to ensure unique probabilities of inclusion.

8. Discussion

The pilot study indicated that a procedure to identify businesses missed in list sources by comparing, on an address basis, listings from an area sample canvass with business locations from list sources is
feasible and practical. The process is not perfect. However, survey statistics from the dual frame approach described are not likely to be substantially biased by failure of the area sample component to detect all undercoverage in the list sample component.

Unit costs for individual survey operations will vary from country to country and, even within a country, from one survey organization to another. Variance components will also vary, particularly in the area sample. Based on our experience, we would expect that the list sample component would account for on the order of 80 percent of the businesses interviewed, and the area sample 20 percent. Thus, about 80 percent of the sample would reflect the efficiencies of list sampling from which an independent area sample would not benefit. Considering variance and cost factors jointly, we conclude that the dual frame approach described can be expected to show cost savings of 20 to 50 percent compared to an area sample approach used alone.

Bibliography


* The views expressed are those of the authors and do not necessarily reflect official positions of the SBA.

Table 1. Success in finding list cases in the field

<table>
<thead>
<tr>
<th>Outcome of attempt to locate list case</th>
<th>Percent of listings attempted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>List Source 1</td>
</tr>
<tr>
<td>Found by match criterion:</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>83.3</td>
</tr>
<tr>
<td>Matched on name and address</td>
<td>28.9</td>
</tr>
<tr>
<td>Matched on name</td>
<td>17.8</td>
</tr>
<tr>
<td>Matched on address</td>
<td>11.1</td>
</tr>
<tr>
<td>Residential address:</td>
<td></td>
</tr>
<tr>
<td>Business activity</td>
<td>5.6</td>
</tr>
<tr>
<td>No business activity</td>
<td>18.8</td>
</tr>
<tr>
<td>Other (out-of-scope)</td>
<td>1.1</td>
</tr>
<tr>
<td>Not found but possibility to match with followup</td>
<td>10.0</td>
</tr>
<tr>
<td>Not possible to locate</td>
<td>6.7</td>
</tr>
<tr>
<td>All cases</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Table 2. Success in finding area sample cases in any list

<table>
<thead>
<tr>
<th>Outcome of attempt to find area sample case in list source</th>
<th>Percent of area sample cases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>List Source 1</td>
</tr>
<tr>
<td>Matched by criterion:</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>50.2</td>
</tr>
<tr>
<td>Matched on name and address</td>
<td>30.1</td>
</tr>
<tr>
<td>Matched on name</td>
<td>9.2</td>
</tr>
<tr>
<td>Matched on address</td>
<td>10.9</td>
</tr>
<tr>
<td>No match on criterion but possibility with followup</td>
<td>26.8</td>
</tr>
<tr>
<td>Not possible to find (match)</td>
<td>23.0</td>
</tr>
<tr>
<td>All cases</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Table 3. Success in finding area sample cases in any list by type of locality

<table>
<thead>
<tr>
<th>ZIP Code area</th>
<th>Locality</th>
<th>Percent matched*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>List Source 1</td>
<td>List Source 2</td>
</tr>
<tr>
<td>1</td>
<td>Highrise office</td>
<td>98</td>
</tr>
<tr>
<td>2</td>
<td>Dense retail/service</td>
<td>65</td>
</tr>
<tr>
<td>3</td>
<td>Mixed clusters/strip</td>
<td>70</td>
</tr>
<tr>
<td>4</td>
<td>Mixed residential/business</td>
<td>84</td>
</tr>
<tr>
<td>5</td>
<td>Small industry</td>
<td>69</td>
</tr>
<tr>
<td>6</td>
<td>Rural area</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>All area sample cases</td>
<td>77</td>
</tr>
</tbody>
</table>

* Includes cases not matched on criteria but possibility of a match with followup.
Figure A
Outline of Stages in Alternative Sample Designs

**FIRST-STAGE SAMPLE**
A stratified sample of PSUs (primary sampling units) appropriate for a national survey. The PSUs would be large metropolitan areas, and counties or groups of counties outside those areas. The first-stage sample would be the same under either design.

**SECOND-STAGE SAMPLE**
A (stratified) sample of SSUs (secondary sampling units) within the PSUs selected in the first-stage sample. The suggested SSUs for the U.S.A. would be five-digit ZIP code areas. A single sample of ZIP codes would be used to develop further components of each of the alternative designs.

**THIRD-STAGE SAMPLE**
Within the selected ZIP codes recode the addresses of businesses from the business list sources to (1990) Census geography: Blocks/Block Groups for the U.S.A.

Within the selected ZIP codes select a sample of businesses from the list source whose addresses are to be used to establish (by field check) a current list of business addresses.

Within a subsample of the selected ZIP codes define "area segments" on the basis of Census geography to be field listed to supplement the sample of addresses from the list source.

Within the selected ZIP codes define "area segments" on the basis of Census geography to be field listed.

Samples of business addresses for surveys of small businesses.
CAPITAL STRUCTURE LIFE CYCLE: STATIC PROCESS OR DYNAMIC EVOLUTION?

Joel Shulman, Babson College; Rick Cooper, Wayne State University; David Brophy, University of Michigan
Joel Shulman, 229 Tomasso, Babson College, Wellesley, MA 02157

KEY WORDS: Capital Structure, Life Cycle

INTRODUCTION

Small firm owner/managers (entrepreneurs) often encounter financial obstacles in the early stages of growth when the delicate tradeoffs of wealth and control must be faced. While rational investors attempt to maximize wealth through growth and external financing, some entrepreneurs may not be willing to surrender associated control of their firm by soliciting outside funds. Alternatively, some firms may be able to manage growth through internally generated funds, informal sources of funds, or venture capital. Whatever the outcome, it seems clear that small firm capital structures may not remain static. Rather, small firm capital structures may experience a dynamic evolution as the firm grows. This paper examines small firm capital structures using a newly-released, comprehensive data set of small private firms and reports developments based on age, industry and cash-flow considerations. Results are analyzed in the context of traditional capital structure theory and contrasted with prior empirical work which has thus far been limited to studies of publicly-traded, large firms or small samples of private firms.

PECKING ORDER, STATIC PREFERENCE, AGENCY COSTS, INFORMATION ASYMMETRIES AND OPTIMAL CAPITAL STRUCTURE

Capital structure theories, beginning with Modigliani and Miller (1958) have addressed issues related to corporate optimal capital structure, tax benefits and bankruptcy costs related to debt. Donaldson (1965) discussed how corporations might opt for a target debt-to-total assets ratio. These theories, while useful in describing rational manager behavior within the large organization, have been less useful in describing financing difficulties for the emerging small firm owner/manager. However, agency costs and information asymmetries theories as described by Jensen and Meckling (1976), Fama (1980), Barnea, Haugen and Senbet (1981), Petit and Singer (1985) and Myers and Majluf (1984) may be very important in explaining differing capital structures between small and large firms.

These articles describe how agents of the firm may attempt to consume corporate perquisites at the expense of other stakeholders. As a result of agency costs, potential external sources of equity and debt including financial institutions, venture capitalists and common shareholders attempt to limit their risk exposure by imposing high monitoring standards, collateral requirements, or restrictive covenants on the firm raising the funds. These actions are designed to limit managerial freedom. Additional theories posited by Myers (1984) and Myers and Majluf (1984) describe pecking order and static trade-off approaches to corporate financing. The pecking order theory suggests that firms utilize funds in a sequential fashion beginning with internal funds and then later switching to external sources once all internal sources are relinquished. The static trade-off theory on the other hand, implies that firms will borrow funds up to the point where the debt tax write-off equals the bankruptcy tax shield. Although none of the theories have been specifically targeted toward small firms, studies by Churchill and Lewis (1986) provide a discussion on bank lending practices and the problems experienced by small borrowing firms, while James (1986) describes information disclosures required of small firm owner/managers in bank related financing transactions.

With the emergence of expanded small firm data bases, researchers are now afforded unprecedented opportunities to explore small firm capital structure differences and similarities with large firms. Moreover, recent surveys have included questions on disclosure requirements, contracting arrangements, types of financing, rates, fees, term, and sex/race of borrower. Each of these questions enable more refined analyses of agency costs and information asymmetries issues, as well as, tests of discrimination. This study explores the evolving
nature of emerging small firm capital structures. We offer a life cycle hypothesis and compare it with existing theories.

CAPITAL STRUCTURE DIFFERENCES: RECENT EVIDENCE

Early tests of capital structure differences suggest small firms have less debt than large firms (Martin and Petty 1978). Other studies which examined differences in small firm and large firm capital structures have provided inconclusive results due to the paucity of small firm capital structure data. For instance, Day, Stoll, and Whaley (1985) and Kester (1986) found no capital structure differences between large and small firms using a COMPUSTAT data base of 7000 public firms. However, as several researchers later pointed out, even the smallest, publicly-traded firms in the COMPUSTAT data base are much larger and, accordingly, much different than the average, small private firm.

Petersen and Shulman (1987) using a data set of 3600 small, private firms from 12 countries found distinct capital structure differences between smaller, younger firms and older, more established firms. They attributed the differences to a capital structure evolution during the firm's product life cycle. In their "life cycle" capital structure model, they discuss how emerging small firms initially rely on self-financing and relatives/friends. As the firm grows, self- and relative financing declines due to lack of personal diversification as well as inefficiencies in borrowing amounts, terms, independence and control (relatives and friends are poor intermediaries). They found weak evidence supporting the life cycle theory including differences in access to capital depending on age of the firm and geographic location. They associated the change in capital structure to less costly and more efficient sources of funds made available as the firm grew or matured. They hypothesized that information asymmetries and agency costs may be involved.

More recently, Mull and Winters (1991) used a data set of 34 IPOs in examining the life cycle hypothesis. They compared capital structures between public and private firms and find private firms use more debt compared to public firms. They infer from this result that as firms move from private to public ownership changes occur in the relationships between the determinants of capital structure and firm debt levels. They hypothesize in the same vein as Peterson and Shulman (1987), that as firms move through their life cycle of ownership, they gain access to more efficient capital markets and lower transaction costs.

Norton (1991), on the other hand, used survey responses from 117 private US firms and finds that factors dealing with bankruptcy costs, agency costs and information asymmetries play an insignificant role in affecting capital structure policy. He argues that financial officers follow a "pecking order" (i.e. use-up cash, then debt, then equity. etc.), and are principally concerned with market conditions and their own preferences. He found no data to support the static preference theory which balances a firm's debt-related tax benefits with a firm's potential bankruptcy costs. Moreover, Norton found little evidence to support the notion that firms have a target capital structure to which they adhere. The target capital structure theory, first proposed by Donaldson (1965), is frequently discussed in finance textbooks, but has had little support since its origins.

Ou (1991) uses a Small Business Administration (SBA) and Federal Reserve survey of 3400 private US firms and reports that the incidence and timing of outside equity support depends on such things as the rate of growth, the level of profits, the timing of operating cash flows and the initial infusion of equity capital. He also finds that a firm's debt-equity relationship is influenced by the age, company organization, and borrowing frequency. For example, he finds that younger firms have lower equity ratios, but that it varies by corporate classification. In addition, the rate of growth affects the equity ratio in an inverse manner.

EMPIRICAL TEST AND METHODOLOGY

This study uses the Small Business Administration and the Federal Reserve database consisting of 3400 observations. We use the database to test capital structure differences among small and large firms. We also test the SBA/Federal Reserve data for evidence of monitoring provisions, types of financing, borrower costs, potential bias against minorities or females and other borrower-specific
information. Finally, we examine the data for evidence supporting the "pecking order", "static preference", "agency cost", "asymmetric information" and "life cycle" theories. Standard statistical differences and ordinary least squares regression techniques are employed and shown in Tables 1-8.

Figure 1 shows the basic life cycle model. We hypothesize that firms borrow progressively more funds from traditional sources (i.e. bankers) as they mature and establish a financial track record. Borrowing may depend on such factors as: 1) whether or not the firm is in a capital intensive industry; 2) the timing of payments to suppliers and cash receipts (i.e. cash conversion cycle); 3) the level of sales and asset growth; 4) the profit level of the firm; 4) the geographical location of the firm; 5) initial capital infusions from the principals; 6) prior experience among principals and their net worth; 7) race or gender of borrowing parties; 8) organizational type; 9) asset size or number of employees, and 10) asset collateral or loan covenants.

Early sources of commercial bank debt are presumed to be short-term in nature due to the reduced risk to the lenders and the restrictive nature of loan covenants and lender guarantees. As the firm matures (i.e. increases number of employees) additional sources of funds are presumed available at a more attractive price. The greater access to funds is derived from the reduced risk to the lending parties (fewer loan covenants and provisions) and greater loan amount economies of scale (lower transactions cost on a percentage basis) as the firm develops. Consequently, the life cycle model which we propose offers a blend of agency cost, information asymmetry, static trade-off and pecking order theories. We believe firms follow a pattern of minimizing costs consistent with the pecking theory (Myers 1984). Greater efficiencies (lower transactions costs) accrue as the firm matures enabling firms greater access to new fund sources. Agency costs and information asymmetries presumably decrease with maturity since the corporation is establishing a profile by which to judge future performance. Moreover, we hypothesize that many firms pursue a public equity issuance in order to facilitate investment liquidity or continue asset growth. This would be consistent with the static trade-off theory, (i.e balanced bankruptcy/tax shield), as described by Myers (1984).

Table 1 shows the frequency of organization type based on the number of employees in the firm. In excess of 33% of all firms in the database are in the form of proprietorship. Most of the proprietorships exist among smaller firms (employees < 4), whereas Subchapter S and C corporations dominate whenever a firm has 5 or more employees. Table 2 shows minority and female representation in the database. Interestingly, 255 out of 3404 of the observations are minority controlled whereas only 421 firms are controlled by females. Clearly, minority representation in this sample more closely resembles the overall percentages in the U.S. compared to the female representation. Table 2 demonstrates how both minority and female control declines as firm size increases. Several explanations could explain this finding. For example, bias in lending behavior or discrimination in the financial markets might prevent minorities and women from growing their firms. This might be consistent with agency cost or information asymmetry arguments. Alternatively, minorities and women might cluster in certain industries which by their nature do not grow beyond a certain employee base. Table 3, for instance provides support for this argument. Minorities and women tend to operate in retailing and service industries (shown in Table 3 as SIC groups 3 and 4) and may plateau at an earlier level than say, a manufacturing firm (shown in Table 3 as SIC 2).

Table 4 shows the most recent type of financing broken down by corporate maturity (number of employees). Not surprisingly, more than one-half of the smallest firms (0-1 employees) have not yet used (or were not able to get) traditional forms of financing. Access to financing increases as firms mature. This should be obvious given the larger representation of different financing categories in the later stages of maturity. Moreover, access to longer-term or permanent sources of funds increases with maturity. Both of these findings are consistent with the pecking order theory as well as agency cost and information asymmetries theories. Firms may follow a path of financing which becomes more accessible and less expensive. Additionally, Table 4 supports the notion that firms pursue sources of financing in a manner consistent with the life cycle hypothesis. Close
examination of Table 4 suggests a financing pattern evolves as the emerging firm grows.

Tables 5-8 illustrate simple regressions which test relationships among employees, short-term debt, long-term debt, equity and principle owner debts. Table 5, demonstrates how short-term debt at first increases with firm maturity and then later declines (employees is significantly positive and the employees$^2$ term is significantly negative). In addition, Table 5 shows how short-term debt is negatively related to proprietorship ownership and positively related to profits. These findings are consistent with pecking order, asymmetric information, agency cost, static trade-off and life cycle hypotheses.

Table 6 shows how principle owner debts are positively related to corporate organization (intercept t statistic = 4.5), negatively related to minority controlled firms and negatively related to public ownership. Moreover, the positive t statistic for employees ($t = 2.09$) and negative t statistic for employees$^2$ suggests that principle owners basically lend money to their organizations early in corporate life. The public variable is consistent with an explanation that any firm which has issued public stock need not borrow money from its principles. Both the corporate organization and the public variables are consistent with the life cycle hypothesis. They are also consistent with static trade-off and pecking order theories. A negative minority t statistic (-3.36) might be indicative of agency cost or information asymmetry theories and suggests that minority owned firms either do not need cash from principle owners or can not get cash from principle owners (i.e. they don't have it).

Table 7 shows how equity to total assets levels move positively with proprietorships and inversely with profits and Subchapter S corporations. This information suggests low-profit proprietorships in the early stages of development rely heavily on equity funds whereas subchapter S corporations don't. This regression is consistent with Tables 1 and 4 which showed how small firms (i.e. proprietorships) in the early stages of growth often had no traditional forms of bank capital.

Table 8 provides additional evidence that long-term debt does not exist with proprietorships or immature firms. The t-statistic for proprietorship is (-2.7) and employees and employees$^2$ are (-2.9 and 2.49 respectively. In addition, the t-statistic for Profit (1.62) shows how long-term debt is more pronounced with corporate profits. Each of these results is consistent with the overall life cycle hypothesis.

**CONCLUSIONS**

While prior studies have tested for differences between small and large firms, difficulties stemming from inadequate data on small, private firms have allowed only tentative conclusions. However, an extensive database, recently developed by the SBA and Federal Reserve, may enable researchers to address many questions which have been left unanswered. For example, capital structure issues involving agency costs, information asymmetries, static trade-off and pecking order theories can now be discussed in the context of a broader life cycle hypothesis.

Although the results in this paper are preliminary in nature, they nevertheless indicate that small firm capital structures undergo an evolutionary process. We believe maturing firms have increasing access to less expensive sources of financing. Thus, depending on the industry, growth, timing of payments and organizational type and maturity level, institutional sources and costs will vary. The life cycle hypothesis presented in this paper appears to be supported by the SBA data and consistent with other existing theories of capital structure.
### TABLE 1

Frequency of Organization Type

<table>
<thead>
<tr>
<th>Number of Employees</th>
<th>0-1</th>
<th>2-4</th>
<th>5-14</th>
<th>15-24</th>
<th>25-49</th>
<th>50-99</th>
<th>100-249</th>
<th>250+</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proprietorship</td>
<td>296</td>
<td>692</td>
<td>216</td>
<td>22</td>
<td>14</td>
<td>10</td>
<td>4</td>
<td>1</td>
<td>1,255</td>
</tr>
<tr>
<td>Partnership</td>
<td>30</td>
<td>113</td>
<td>76</td>
<td>19</td>
<td>12</td>
<td>12</td>
<td>8</td>
<td>4</td>
<td>274</td>
</tr>
<tr>
<td>Subchapter S</td>
<td>29</td>
<td>106</td>
<td>135</td>
<td>38</td>
<td>40</td>
<td>65</td>
<td>49</td>
<td>14</td>
<td>476</td>
</tr>
<tr>
<td>Corporation</td>
<td>60</td>
<td>181</td>
<td>526</td>
<td>125</td>
<td>146</td>
<td>187</td>
<td>115</td>
<td>39</td>
<td>1,379</td>
</tr>
<tr>
<td>Total</td>
<td>415</td>
<td>1192</td>
<td>853</td>
<td>204</td>
<td>222</td>
<td>284</td>
<td>176</td>
<td>58</td>
<td>3,404</td>
</tr>
</tbody>
</table>

### TABLE 2

Minority and Female Representation

<table>
<thead>
<tr>
<th>Number of Employees</th>
<th>0-1</th>
<th>2-4</th>
<th>5-14</th>
<th>15-24</th>
<th>25-49</th>
<th>50-99</th>
<th>100-249</th>
<th>250+</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minority</td>
<td>40</td>
<td>112</td>
<td>51</td>
<td>11</td>
<td>23</td>
<td>11</td>
<td>6</td>
<td>1</td>
<td>255</td>
</tr>
<tr>
<td>Female</td>
<td>60</td>
<td>199</td>
<td>95</td>
<td>21</td>
<td>14</td>
<td>22</td>
<td>4</td>
<td>6</td>
<td>421</td>
</tr>
<tr>
<td>Overall Total Sample</td>
<td>415</td>
<td>1192</td>
<td>853</td>
<td>204</td>
<td>222</td>
<td>284</td>
<td>176</td>
<td>58</td>
<td>3,404</td>
</tr>
</tbody>
</table>
### TABLE 3
Minority and Female Representation By SIC Code

<table>
<thead>
<tr>
<th>Number of Employees</th>
<th>0-1</th>
<th>2-4</th>
<th>5-14</th>
<th>15-24</th>
<th>25-49</th>
<th>50-99</th>
<th>100-249</th>
<th>250+</th>
<th>Total by SIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minority 1</td>
<td>8</td>
<td>19</td>
<td>7</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>36/473</td>
</tr>
<tr>
<td>Female 1</td>
<td>3</td>
<td>5</td>
<td>8</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>23/473</td>
</tr>
<tr>
<td>Minority 2</td>
<td>0</td>
<td>13</td>
<td>5</td>
<td>2</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>30/525</td>
</tr>
<tr>
<td>Female 2</td>
<td>1</td>
<td>12</td>
<td>13</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>1</td>
<td>2</td>
<td>45/525</td>
</tr>
<tr>
<td>Minority 3</td>
<td>18</td>
<td>44</td>
<td>22</td>
<td>5</td>
<td>10</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>106/1468</td>
</tr>
<tr>
<td>Female 3</td>
<td>31</td>
<td>120</td>
<td>47</td>
<td>9</td>
<td>6</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>223/1468</td>
</tr>
<tr>
<td>Minority 4</td>
<td>14</td>
<td>36</td>
<td>17</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>83/938</td>
</tr>
<tr>
<td>Female 4</td>
<td>25</td>
<td>62</td>
<td>27</td>
<td>3</td>
<td>2</td>
<td>8</td>
<td>0</td>
<td>3</td>
<td>130/938</td>
</tr>
</tbody>
</table>

Categories: 1 = Mining and Construction; 2 = Manufacturing; 3 = Retailing; 4 = Service

### TABLE 4
Most Recent Financing Experience

<table>
<thead>
<tr>
<th>Number of Employees</th>
<th>0-1</th>
<th>2-4</th>
<th>5-14</th>
<th>15-24</th>
<th>25-49</th>
<th>50-99</th>
<th>100-249</th>
<th>250+</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commercial Bank</td>
<td>89</td>
<td>358</td>
<td>347</td>
<td>89</td>
<td>108</td>
<td>145</td>
<td>93</td>
<td>261</td>
<td>1255</td>
</tr>
<tr>
<td>S&amp;L</td>
<td>45</td>
<td>179</td>
<td>142</td>
<td>25</td>
<td>34</td>
<td>33</td>
<td>18</td>
<td>9</td>
<td>485</td>
</tr>
<tr>
<td>Savings Bank</td>
<td>16</td>
<td>51</td>
<td>57</td>
<td>16</td>
<td>20</td>
<td>24</td>
<td>12</td>
<td>10</td>
<td>206</td>
</tr>
<tr>
<td>Credit Union</td>
<td>2</td>
<td>21</td>
<td>17</td>
<td>4</td>
<td>8</td>
<td>16</td>
<td>9</td>
<td>2</td>
<td>79</td>
</tr>
<tr>
<td>Finance</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>21</td>
</tr>
<tr>
<td>Insurance</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>7</td>
<td>8</td>
<td>3</td>
<td>29</td>
</tr>
<tr>
<td>More Than 1</td>
<td>26</td>
<td>91</td>
<td>51</td>
<td>13</td>
<td>12</td>
<td>13</td>
<td>7</td>
<td>1</td>
<td>215</td>
</tr>
<tr>
<td>No Funding (or response)</td>
<td>236</td>
<td>488</td>
<td>236</td>
<td>48</td>
<td>34</td>
<td>41</td>
<td>28</td>
<td>3</td>
<td>1114</td>
</tr>
<tr>
<td>Total Sample</td>
<td>415</td>
<td>1192</td>
<td>853</td>
<td>204</td>
<td>222</td>
<td>284</td>
<td>176</td>
<td>58</td>
<td>3404</td>
</tr>
</tbody>
</table>
Table 5

Relationship Between Short-Term Debt, Employees and Organization Type

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>t-statistic</th>
<th>Adjusted R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.01532</td>
<td>0.28570</td>
<td>-0.054</td>
<td></td>
</tr>
<tr>
<td>Employees</td>
<td>0.00077</td>
<td>0.00033</td>
<td>2.355</td>
<td></td>
</tr>
<tr>
<td>Employees²</td>
<td>-0.00000</td>
<td>0.00000</td>
<td>-2.044</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>-0.02736</td>
<td>0.01905</td>
<td>-1.436</td>
<td></td>
</tr>
<tr>
<td>Proprietorship</td>
<td>-0.04457</td>
<td>0.01649</td>
<td>-2.703</td>
<td></td>
</tr>
<tr>
<td>Partnership</td>
<td>-0.02439</td>
<td>0.02555</td>
<td>-0.955</td>
<td></td>
</tr>
<tr>
<td>Subchapter S</td>
<td>-0.00911</td>
<td>0.02127</td>
<td>-0.428</td>
<td></td>
</tr>
<tr>
<td>Profit</td>
<td>0.03827</td>
<td>0.00739</td>
<td>5.177</td>
<td>0.028</td>
</tr>
</tbody>
</table>

Table 6

Principle Debt, Employees, Organization Type and Ownership

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>t-statistic</th>
<th>Adjusted R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>4005186</td>
<td>899017</td>
<td>4.50</td>
<td></td>
</tr>
<tr>
<td>Employees</td>
<td>2262</td>
<td>1084</td>
<td>2.09</td>
<td></td>
</tr>
<tr>
<td>Employees²</td>
<td>-6</td>
<td>3</td>
<td>-1.75</td>
<td></td>
</tr>
<tr>
<td>Minority</td>
<td>-465797</td>
<td>138683</td>
<td>-3.36</td>
<td></td>
</tr>
<tr>
<td>Proprietorship</td>
<td>-71260</td>
<td>327577</td>
<td>-0.22</td>
<td></td>
</tr>
<tr>
<td>Partnership</td>
<td>-65510</td>
<td>139596</td>
<td>-0.47</td>
<td></td>
</tr>
<tr>
<td>Subchapter S</td>
<td>11552</td>
<td>66037</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>Public</td>
<td>-1733828</td>
<td>449391</td>
<td>-3.86</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Adjusted R²
Table 7

Equity, Employees, Organization Type and Profits
Construction/Mining Industry

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>t-statistic</th>
<th>Adjusted R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.355</td>
<td>0.558</td>
<td>0.636</td>
<td></td>
</tr>
<tr>
<td>Employees</td>
<td>-0.001</td>
<td>0.000</td>
<td>-1.726</td>
<td></td>
</tr>
<tr>
<td>Employees²</td>
<td>0.000</td>
<td>0.000</td>
<td>1.231</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>-0.026</td>
<td>0.042</td>
<td>-0.607</td>
<td></td>
</tr>
<tr>
<td>Proprietorship</td>
<td>0.175</td>
<td>0.036</td>
<td>4.829</td>
<td></td>
</tr>
<tr>
<td>Partnership</td>
<td>0.079</td>
<td>0.056</td>
<td>1.407</td>
<td></td>
</tr>
<tr>
<td>Subchapter S</td>
<td>-0.089</td>
<td>0.047</td>
<td>-1.918</td>
<td></td>
</tr>
<tr>
<td>Profit</td>
<td>-0.041</td>
<td>0.016</td>
<td>-2.481</td>
<td></td>
</tr>
<tr>
<td>Adjusted R²</td>
<td></td>
<td></td>
<td>0.033</td>
<td></td>
</tr>
</tbody>
</table>

Table 8

Long-Term Debt, Employees, Organization Type and Profits
Manufacturing Industry

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>t-statistic</th>
<th>Adjusted R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.349</td>
<td>0.351</td>
<td>0.994</td>
<td></td>
</tr>
<tr>
<td>Employees</td>
<td>-0.001</td>
<td>0.000</td>
<td>-2.906</td>
<td></td>
</tr>
<tr>
<td>Employees²</td>
<td>0.000</td>
<td>0.000</td>
<td>2.489</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>-0.069</td>
<td>0.063</td>
<td>-1.108</td>
<td></td>
</tr>
<tr>
<td>Proprietorship</td>
<td>-0.142</td>
<td>0.052</td>
<td>-2.704</td>
<td></td>
</tr>
<tr>
<td>Partnership</td>
<td>-0.122</td>
<td>0.092</td>
<td>-1.326</td>
<td></td>
</tr>
<tr>
<td>Subchapter S</td>
<td>0.069</td>
<td>0.047</td>
<td>1.483</td>
<td></td>
</tr>
<tr>
<td>Profit</td>
<td>0.043</td>
<td>0.026</td>
<td>1.627</td>
<td></td>
</tr>
<tr>
<td>Adjusted R²</td>
<td></td>
<td></td>
<td>0.022</td>
<td></td>
</tr>
</tbody>
</table>
REFERENCES


