# Analyzing the 2004–2011 KFS Multiply Imputed Data

Prepared By:
Joseph Farhat
Alicia Robb

December 2013

The Kauffman Firm Survey
2004 2005 2006 2007 2008 2009 2010 2011

Ewing Marion KAUFFMAN Foundation

# The Kauffman Firm Survey

2004 2005 2006 2007 2008 2009 2010 2011

# Contents




Electronic copy available at: http://ssrn.com/abstract=2367300

# 1. Introduction

The purpose of this document is to provide instructions to the KFS users regarding the proper use of the KFS multiply imputed data to draw statistically valid inferences in their works. Also, it describes the general framework of the imputation process. [1]

Since data editing is first done prior to imputation, this document describes the general framework for data editing first. Data editing is a process that includes procedures for detecting and correcting errors in data. Errors occur due to misunderstanding the answers/questions, or item values for one item contradict with another item (inconsistency) and incorrect flow through prescribed skip patterns.

Understanding the underlying structure of the KFS questionnaire helps to distinguish data missing due to item non-response or due to unit non-response. Item non-response occurs when certain questions in a survey are not answered by a respondent. Unit non-response takes place when a business cannot be contacted or refuses to participate in a survey.

Legitimate missing (hard missing) value occurs when a business is sold, merged, temporarily stopped operations, permanently out of businesses, due to unit non-response or due to skip logic.

The following table shows the possible type of missing values by business status and year; it also shows the number of legitimate missing (hard missing) values in each year:

Cross-Section – Wide format

| Year | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | Missing values in year t |
|---|---|---|---|---|---|---|---|---|---|
| Refusal | | 561 | 743 | 825 | 816 | 743 | 776 | 676 | Hard missing |
| Stopped operation or sold or merged in any of previous follow-ups | | | 303 | 671 | 1,015 | 1,399 | 1,685 | 1,941 | Hard missing |
| Out of Business | | 260 | 321 | 299 | 344 | 250 | 218 | 209 | Hard missing |
| Merged Or Sold | | 43 | 47 | 45 | 40 | 36 | 38 | 40 | Hard missing |
| Temporarily Stopped | | 66 | 124 | 98 | 58 | 41 | 45 | 30 | Hard missing |
| Located : No data was collected | | | | 75 | 49 | 51 | 40 | 25 | Hard missing |
| Complete | 4,928 | 3,998 | 3,390 | 2,915 | 2,606 | 2,408 | 2,126 | 2,007 | Soft missing/Hard missing [a] |

[1] Data users who are interested in better understanding data imputation methods, its statistical properties, derivations and theoretical underpinnings of these methods, can see the references section for recommended readings.

### Longitudinal – Wide format

| Year | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | Missing values in year t |
|---|---|---|---|---|---|---|---|---|---|
| Stopped operation or sold or merged in any of previous follow-ups | | | 303 | 586 | 810 | 1,048 | 1,212 | 1,365 | Hard missing |
| Out of Business | | 260 | 247 | 188 | 213 | 141 | 133 | 114 | Hard missing |
| Merged Or Sold | | 43 | 36 | 36 | 25 | 23 | 20 | 17 | Hard missing |
| Temporarily Stopped | | 53 | 86 | 61 | 31 | 21 | 25 | 14 | Hard missing |
| Located : No data was collected | | | | 65 | 30 | 31 | 16 | 13 | Hard missing |
| Complete | 3,140 | 2,784 | 2,468 | 2,204 | 2,031 | 1,876 | 1,734 | 1,617 | Soft missing/Hard missing [a] |

[a] Hard missing: due to skip logic

### Longitudinal – Long format

| Status | n | Duration | Obs. |
|---|---|---|---|
| Permanently stopped operations in the first follow-up | 260 | 1 | 260 |
| Sold or Merged in first follow-up | 43 | 1 | 43 |
| Permanently stopped operations in second follow-up | 247 | 2 | 494 |
| Sold or Merged in second follow-up | 36 | 2 | 72 |
| Permanently stopped operations in third follow-up | 188 | 3 | 564 |
| Sold or Merged in third follow-up | 36 | 3 | 108 |
| Permanently stopped operations in fourth follow-up | 213 | 4 | 852 |
| Sold or Merged in fourth follow-up | 25 | 4 | 100 |
| Permanently stopped operations in fifth follow-up | 141 | 5 | 705 |
| Sold or Merged in fifth follow-up | 23 | 5 | 115 |
| Permanently stopped operations in sixth follow-up | 133 | 6 | 798 |
| Sold or Merged in sixth follow-up | 20 | 6 | 120 |
| Permanently stopped operations in seventh follow-up | 114 | 7 | 798 |
| Sold or Merged in seventh follow-up | 17 | 7 | 119 |
| Temporarily stopped operations in seventh follow-up | 14 | 7 | 98 |
| Responded to first follow-up to seventh follow-up | 1,630 | 8 | 13,040 |
| Total | 3,140 | | 18,286 |

To determined legitimate non-response, the KFS data has a variable - final_status_code - that allows us to determine if the business is still in operation (complete the survey), dropped out, permanently stopped operations, temporarily stopped operations, merged, or was sold. Based on the final_status_code variable, we created a new variable – class_f - for every follow-up; the class_f variable has the following values and labels:

| Value | Label |
|---|---|
| 0 | Located : No data was collected ( due to skip logic ) |
| 1 | Dropout (unit non-response) during follow-up t. |
| 2 | Missing because the business closed, sold or merged in the previous follow-ups. |
| 3 | Permanently stopped operations during follow-up t. |
| 4 | Merged, or sold during follow-up t. |
| 5 | Temporarily stopped operations during follow-up t. |
| 6 | Survival business(complete the survey)during follow-up t. |

The skip pattern in the KFS comes into two major types: question skip logic and a section skip logic. Question skip logic involves conditionally asking/skipping questions based upon responses to prior question(s). Meanwhile, section skip logic involves asking/skipping a whole section in the survey questionnaire, based upon responses to prior question(s).

The primary goal of understanding the skip patterns and the underlying structure of the questionnaire is to recognize non-applicable questions responses (hard missing or legitimate missing) and differentiate them from responses that are missing (soft missing), due to refusal or didn't know response. Having the missing values clearly defined (soft vs. hard missing) helps us to determine the correct number of observations for each variable; it also explains why the number of observations varying from variable to variable, correctly constructing aggregate variables, and correctly imputing soft missing data.

## 2. Logical Imputation (Data Editing)

Logical (or deductive) imputation refers to any method that uniquely identified the true value of the missing value with certainty from within the data set.

There are two major types of missing data in KFS: unit non-response and item non-response. Unit nonresponse occurs when a business refuses to participate in the survey. In the KFS, the unit nonresponses are dealt with through weighting adjustments; thus, it is not subject for logical imputation or any other kind of imputation. Meanwhile, item nonresponse occurs when certain questions in a survey are not answered by a respondent.

The single-cohort panel structure of the KFS offers many possibilities for logical imputation. To better understand what a deductive imputation procedure does, **some examples** are discussed:

## 2.1 Section C: Business Characteristics

As of the first follow-up, the KFS questionnaire asks the respondent to confirm legal status of the business, and then the legal status of the business is recorded under the “c1z2_legal_status” variable. In the baseline survey legal status of the business is recorded under “b2a_legal_status_0”. If the respondent confirms the legal status to be the same as in the previous year, the legal status of the business is copied from the previous year. Meanwhile, if the respondent confirms that the legal status is not the same as the previous year the legal status, he was asked to provide the new legal status of the business.

These types of questions do not have missing values due to skip logic but it has missing values due to legitimate missing values (business is sold, merged, temporarily stopped operations or permanently out of businesses, etc. ), as well as missing values due to item non-response.

In section C, the KFS questionnaire asks about the total number of employees excluding owner(s) who are paid employees of the business. As a result of the skip logic, data for “c6_num_ft_employees,” and “c7_num_pt_employees,” has to be recoded to zero if “c5_num_employees” is zero.

```
*  Example
global  suffix "_0 _1 _2 _3 _4 _5 _6 _7"
foreach fup in $suffix{
replace c6_num_ft_employees`fup'=0 if c5_num_employees`fup'==0
replace c7_num_pt_employees`fup'=0 if c5_num_employees`fup'==0

replace c5_num_employees`fup'   =.a if classf`fup'<6
replace c6_num_ft_employees`fup'=.a if classf`fup'<6
replace c7_num_pt_employees`fup'=.a if classf`fup'<6
}
```

## 2.2 Section D: Strategy and Innovation

All variables in section D should be recorded for legitimate missing values. An examination of the questions skip logics indicate that the following recoding is needed:

1. “d3_a_num_patent,” “d3_b_num_copyright,” and “d3_c_num_trademark,” need to be recoded to zero if the answers for “d3_a_have_patent,””d3_b_have_copyright,” or

"d3_c_have_trademark," is no, respectively.

2. "d4_a_lic_out_patent," "d4_b_lic_out_copyright," and "d4_c_lic_out_trademark," need to be recoded to zero if the answer for "d3_a_have_patent," "d3_b_have_copyright," or "d3_c_have_trademark," is no, respectively.

3. "d7_perc_sales_indiv," "d7_perc_sales_bus," and "d7_perc_sales_govt," need to be recoded to zero if the answer for "d6_have_sales," is no.

```
*  Example
foreach fup in $suffix{
replace d3_a_have_patent`fup'       =.a if classf`fup'<6
replace d3_a_num_patent`fup'        =.a if classf`fup'<6
replace d4_a_lic_out_patent`fup'    =.a if classf`fup'<6
replace d5_a_lic_in_patent`fup'     =.a if classf`fup'<6
}
foreach fup in $suffix{
replace d3_a_num_patent`fup'            = 0 if d3_a_have_patent`fup'  ==0
replace d4_a_lic_out_patent`fup'        = 0 if d3_a_have_patent`fup'  ==0
}
```

## 2.3 Section E: Business Organization and Human Resource Benefits

All variables in section E should be re-coded for legitimate missing values. Section E has both question skip logic and section skip logic. The entire section was skipped for businesses that has one owner (c2_owners) and reported zero employees (c5_num_employees). Meanwhile, the part-time employee benefits questions were skipped for businesses that reported the number of part-time employees is zero.

```
*  Example
foreach fup in $suffix{
gen skip_e`fup'=1 if c2_owners`fup'== 1  & c5_num_employees`fup'== 0
replace e1_a_num_human_res`fup'    =.a     if skip_e`fup'==1
replace e2a_ft_emp_bonus_plan`fup' =.a     if skip_e`fup'==1
replace e2b_pt_emp_bonus_plan`fup' =.a     if c7_num_pt_employees`fup'   ==
0
replace e1_a_num_human_res`fup'    =.a     if classf`fup'   <6
replace e2a_ft_emp_bonus_plan`fup' =.a if classf`fup'   <6
replace e2b_pt_emp_bonus_plan`fup' =.a if classf`fup'   <6
}
```

## 2.4 Section F: Business Finances

Section F deals with the major sources of financing, namely equity, debt and other financial information of the business. Since that the KFS collect data for up to 10 active-owner-operators, each owner was assigned a number. For all variables that are related to the active-owner-operators, the number prior to the "0, 1, 2, 3, 4, 5, 6, 7" suffixes indicates the number assigned to the owner. For example, “f2_owner_amt_eq_invest_02_1,” refers the equity injection by owner number two in the first follow-up. While “f2_owner_amt_eq_invest_09_1,” refers the equity injections by owner number nine in the first follow-up. Unless indicated, the variables in this document are listed without a suffix, if the variable name is the same across all rounds.

In the baseline survey, the respondent was always owner number one. Since that the respondent could change from one follow-up to the next one, starting from the first follow-up the variable “respondent” contains the number of the owners who responded for the business in a particular follow-up.

While KFS collect data for up to 10 active-owner-operators, the number assigned to the owner can be more than 10. Since some owners who use to be active (non-active) in one follow-up could be non-active (active) in another, thus the number assigned to the owner can be more than 10.

Starting from the first follow-up and to identify active-owner-operators in each follow-up surveys, a variable called “owner-active-owner-number,” were created to ensure users could see which owner was still an active-owner-operator in the business:

For all the financial variables in the KFS, if the respondent did not provided the exact amount of the variable in dollars, the respondent  was asked to provide a range of the amount instead. The range interval classes were standard across all the financial variables in the KFS. The interval classes are:

$0 .......... 00
$500 or less, .......... 01
$501 to $1,000, .......... 02
$1,001 to $3,000, .......... 03
$3,001 to $5,000, .......... 04
$5,001 to $10,000, .......... 05
$10,001 to $25,000, .......... 06
$25,001 to $100,000, .......... 07
$100,001 to $1,000,000, .......... 08
$1,000,001 or more? .......... 09
Don’t Know .......... "."
Refused .......... "."

Based on the exact amount of the variable or the range of the amount provided by the respondent, new variables were constructed for ease of analysis. The constructed variables represent the financial variables, in terms of range interval classes, by translating the exact amount into a range. To distinguish these variables from the range variables, the term "r" was included into the variables name.

### 2.4.1 Equity Injections by the Active-Owner-Operators

In every survey, the respondents were asked about their equity injections into the business in that year (indicator question), and the amount that was injected, if any. Starting from the first follow-up, respondents were asked to provide how much equity they injected into the business in all years.

For businesses that have more than one owner, equity injections by other active-owner-operators, up to nine of them, were collected through the respondents. The respondents were asked about the equity injections into the business by each of other active-owner-operators in that year (yearly inflow), and the amount that was injected, if any. Starting from the first follow-up, respondents were asked to provide how much equity each of other active-owner-operators injected, if they obtained equity financing during follow-up t, into the business in all years.

The amount of equity injections needs to be re-coded to zero, if the active-owner-operators states that he/they did not inject equity into the business in that follow-up.

Recoding hard missing values are required for two type of missing values. First, all variables in section F should be recorded for legitimate missing values. Second, the variables for non-active-owner-operator should be recoded to hard missing values.

Since that equity injection by other owners is not applicable for the businesses that reported having one active-owner-operator, the variables of equity injections for other owners needs to recode to hard missing values.

```
*  Example
global owners_1_15 "01 02 03 04 05 06 07 08 09 10 11 12 13 14 15"

forvalues i = 0/7 {
foreach ow in $owners_1_15  {
replace f2_owner_eq_invest_`ow'_`i'       =.a   if   owner_active_`ow'_`i'
!=1
}
}
```

```
forvalues i = 0/7 {
foreach ow in $owners_1_15  {
replace        f2_owner_amt_eq_invest_`ow'_`i'        =0                if
f2_owner_eq_invest_`ow'_`i'==0
}
}

forvalues i = 0/7 {
foreach ow in $owners_1_15  {
replace f2_owner_eq_invest_`ow'_`i'         =.a      if classf_`i'<6
replace f2_owner_amt_eq_invest_`ow'_`i'     =.a      if classf_`i'<6
}
}
```

### 2.4.2 Equity Injections by Other Owners

In every survey, the respondents were asked if the business obtain equity financing from owners who are not actively involved in operating the business, non-operator-owners, and the amount that was obtained, if any. The balance of each source of funding that was used during follow-up t was collected in every follow-up survey. Data collections for equity financing obtained from non-operator-owners were at the aggregate level.

The amount of equity injections needs to be recoded to zero if the non-active-owner-operators state that he/they did not inject equity into the business in that follow-up.

Since that equity injection by other owners is not applicable for the businesses that reported the legal status being sole proprietorship, the variables for sole proprietorship needs to recode to hard missing values.

```
*  Example
global List1 "spouse parents angels companies govt vent_cap other"

forvalues i = 0/7 {
foreach name in $List1  {
replace f3_eq_invest_`name'_`i'  =.a   if       c1z2_legal_status_`i'==1
replace f4_eq_amt_`name'_`i'     =.a   if       c1z2_legal_status_`i'==1
replace f4_eq_amt_`name'_`i'     = 0   if       f3_eq_invest_`name'_`i'==0
replace f3_eq_invest_`name'_`i'  =.a   if       classf_`i'<6
replace f4_eq_amt_`name'_`i'     =.a   if       classf_`i'<6
}
}
```

### 2.4.3 Personal Debt Obtained by the Respondent

Respondents were asked about all types of personal debt that was obtained in their names on behalf of the business, and how much of this debt was obtained, if any, during follow-up t. In addition to the amount obtained every year for each type of personal debt, the amount owed for each type of personal debt used in follow-up t was collected.

The amount of personal debt needs to be recoded to zero, if the respondent states that he/she did not use that source of funding. Also, the number of personal debt used needs to be recoded to zero, if the respondent states that they did not use that source of funding.

```
*  Example
forvalues i = 0/7 {

replace f7b_pers_loan_bank_numused_`i' =0  if f7a_pers_loan_bank_`i' ==0
replace f8c_pers_loan_bank_amt_`i'     =0  if f7a_pers_loan_bank_`i' ==0
replace f7a_pers_loan_bank_`i'         =.a if classf_`i'<6
replace f7b_pers_loan_bank_numused_`i' =.a if classf_`i'<6
}
```

### 2.4.4 Personal Debt Obtained by the Other Owners

For businesses that have more than one active-owner-operator, the respondents were asked to report all type of personal debt that was obtained by all other owners on behalf of the business, and how much was obtained, if any. Unlike equity financing by other owners, personal debt by other owners was collected at the aggregate level for active-owner-operators only. In addition to the amount of personal debt obtained by active-owner-operators in every year for each type of personal debt, the amount owed for each type of personal debt used in follow-up t was collected.

Since that personal debt obtained by the other owners is not applicable for the businesses that reported having one active-owner-operator, these variables for businesses that reported having one active-owner-operators need to recoded to hard missing values.

The amount of personal debt obtained by the other owners needs to be recoded to zero, if the respondent states that they did not use that source of funding. Also, the number of personal debt used by the other owners needs to be recoded to zero, if the respondent states that they did not use that source of funding.

```
*  Example
forvalues i = 0/7 {
replace f9a_pers_loan_bank_`i'        =.a if c4_numowners_confirm_`i'<2
replace f9b_pers_loan_bank_numused_`i'=.a if c4_numowners_confirm_`i'<2
replace f10c_pers_loan_bank_amt_`i'   =.a if c4_numowners_confirm_`i'<2
replace f10c_pers_loan_bank_amt_`i'   =0  if f9a_pers_loan_bank_`i' ==0
replace f9b_pers_loan_bank_numused_`i'=0  if f9a_pers_loan_bank_`i' ==0
replace f9a_pers_loan_bank_`i'        =.a if classf_`i'<6
replace f9b_pers_loan_bank_numused_`i'=.a if classf_`i'<6
replace f10c_pers_loan_bank_amt_`i'   =.a if classf_`i'<6
}
```

### 2.4.5 Debt Obtained by the Business

In addition to personal debt financing, the KFS collects data about different types of debt financing that were obtained in the name of the business during baseline and each follow-up survey.

The amount of debt obtained by the business needs to be recoded to zero, if the respondent states that the business did not use that source of funding. Also, the number of business debt used needs to be recoded to zero if the respondent states that the business did not use that source of funding.

```
*  Example
forvalues i = 0/7 {
replace f11a_bus_loans_bank_`i'        =.a if classf_`i'<6
replace f12c_bus_loans_bank_amt_`i'    =.a if classf_`i'<6
replace f11b_bus_loans_bank_numused_`i'=.a if classf_`i'<6
replace f11b_bus_loans_bank_numused_`i'=0  if f11a_bus_loans_bank_`i'==0
replace f12c_bus_loans_bank_amt_`i'    =0  if f11a_bus_loans_bank_`i'==0
replace f11a_bus_loans_emp_`i'         =.a if c5_num_employees_`i'==0
replace f11a_bus_loans_owner_`i'       =.a if c2_owners_`i'==1
}
```

### 2.4.6 Other Financial Information

In addition to the sources of financing, the KFS collects much other financial information from the balance sheet and income statement, as well as financial information regarding the existence of R&D and rental or lease.

### 2.4.7 Section G: Work Behaviors and Demographics of Active-Owner-Operators

Information regarding work behaviors by active-owner-operators was collected in the baseline, as well as in every follow-up survey. Meanwhile demographics information was collected once. If the demographics information of an active-owner-operator was collected during follow-up t-1, then no demographics information will be collected during follow-up t. If the demographics information of an active-owner-operator was missing in follow-up t-1, then we keep asking about this missing information in the following surveys until we have a valid response. During each follow-up, the work behaviors and demographics of newly active-owner-operators were collected.

For example, if the race information of an active-owner-operator was missing in follow-up t-1, then we keep asking about this missing information in the following surveys until we have a valid response. Thus, missing data for follow-up t-1 can be directly filled from other portions of an individual's record.

The work behaviors and demographics variables for non-active-owner-operator should be recoded to hard missing values.

For the race categories, questions respondents were allowed to report multiracial or mixed-race. Thus, there is a question for each race. Since it is more practical for analysis purposes to have one variable with race coded as categorical variable, we created a new race variable "g6_race_group," having the following codes:

American Indian Or Alaska Native ........................................................ 01

Native Hawaiian Or Other Pacific Islander ........................................ 02

Asian .................................................................................................................... 03

Black Or African American ......................................................................... 04

White .................................................................................................................. 05

Other Races Or Mixed Race ......................................................................... 06

Owners reported multiracial or mixed-race were recorded as Other Races.

```
*  Example
forvalues i = 0/7 {
foreach ow in $owners_1_15  {
replace g1a_emp_owner_`ow'_`i'      =.a if owner_active_`ow'_`i'!=1
replace g1a_emp_owner_`ow'_`i'      =.a if classf_`i'<6
}
}
foreach ow in $owners_1_15{
/*Last observation carried backward  */
replace             g4_age_owner_`ow'_0=g4_age_owner_`ow'_1-1            if
g4_age_owner_`ow'_0==.
replace             g4_age_owner_`ow'_0=g4_age_owner_`ow'_2-2            if
g4_age_owner_`ow'_0==.
replace             g4_age_owner_`ow'_0=g4_age_owner_`ow'_3-3            if
g4_age_owner_`ow'_0==.
replace             g4_age_owner_`ow'_0=g4_age_owner_`ow'_4-4            if
g4_age_owner_`ow'_0==.
replace             g4_age_owner_`ow'_0=g4_age_owner_`ow'_5-5            if
```

```
g4_age_owner_`ow'_0==.
replace            g4_age_owner_`ow'_0=g4_age_owner_`ow'_6-6            if
g4_age_owner_`ow'_0==.
replace g4_age_owner_`ow'_0=g4_age_owner_`ow'_7-7 if
g4_age_owner_`ow'_0==.  }
foreach ow in $owners_1_15{
replace        g4_age_owner_`ow'_1        =g4_age_owner_`ow'_0+1        if
g4_age_owner_`ow'_1==.
replace        g4_age_owner_`ow'_2        =g4_age_owner_`ow'_0+2        if
g4_age_owner_`ow'_2==.
replace        g4_age_owner_`ow'_3        =g4_age_owner_`ow'_0+3        if
g4_age_owner_`ow'_3==.
replace        g4_age_owner_`ow'_4        =g4_age_owner_`ow'_0+4        if
g4_age_owner_`ow'_4==.
replace        g4_age_owner_`ow'_5        =g4_age_owner_`ow'_0+5        if
g4_age_owner_`ow'_5==.
replace        g4_age_owner_`ow'_6        =g4_age_owner_`ow'_0+6        if
g4_age_owner_`ow'_6==.
replace        g4_age_owner_`ow'_7        =g4_age_owner_`ow'_0+7        if
g4_age_owner_`ow'_7==.
}
forvalues i = 0/7 {
foreach ow in $owners_1_15  {
/* Recode legitimate missing values */
replace     g4_age_owner_`ow'_`i'=.a if owner_active_`ow'_`i'!=1
replace     g4_age_owner_`ow'_`i'=.a if classf_`i'<6
}
}
```

### 2.5 Questions Added / Dropped From the KFS Questionnaires during the Survey Period

In addition to the fixed core set of questions asked by all businesses in every follow-up survey, over the years some new questions were added to the survey questionnaires and some of these questions were dropped later on.

Appendix B summarizes all the questions that were added to KFS questionnaires during the survey period, as well as in which year the questions were dropped, if they were dropped.

## 3. Single imputation

### 3.1 Last observation carried forward (LOCF) and last observation carried backward (LOCB).

Both LOCF and LOCB methods can be used in longitudinal research designs, but they require the strong assumption of stability. LOCF takes into account the individual's previous observed value on a given variable. If an observation at a certain data collection wave is missing, the last observed value is then used as an estimate for this missing observation.

A related method, last observation carried backward (LOCB), works according to the same approach, but imputes a newer observation in the case of a missing earlier observation of the same individual.

- LOCF was used to impute the legal status variable and business location.
- To determine the missing values for d1a_provide_service and d1b_provide_product variables, we use the NAICS description to determine if the business is operating under the services or manufacturing industries. For the observations that we couldn't determine if they provide a service or product, we use LOCF method of imputation.
- For the Longitudinal file, LOCF is used to fill in all the fixed core set of questions for businesses reported being temporarily stopped or located ( no data was collected) ;LOCF implemented after the MI.

### 3.2 Internal consistency: using information from related observations

- For c2_owners: missing and zero values were replaced by total owners.
- The "c4_numowners_confirm," was replaced by the sum of "owner-active," if "c4_numowners_confirm," is not equal to sum of the owner-active.
- Replace "c1z2_legal_status=7," if "c1z2_legal_status=1," and "c4_numowners_confirm_1>1."

### 3.3 Other single imputations

- For d7_perc_sales_xxxx variables, if one variable has missing value we impute the missing value using: 100-sum (non-missing d7_perc_sales _xxxx).
- For f8xxx_line_y, f8xxx_ bal_y, f8xxx_ owed _y, f10xxx_line_y, f10xxx_ bal_y, f10xxx_ owed _y ,f12xxx_line_y, f12xxx_ bal_y f12xxx_ owed _y, , and f4_eq_amt_xxx_allyrs the missing values set to zero if the business never use these sources of funding. For example:

```
*  Example
Replace f10d_pers_loan_fam_owed_2 =0    if    f9a_pers_loan_fam_0 ==0 & ///
f9a_pers_loan_fam_1 ==0 & f9a_pers_loan_fam_2 ==0
```

## 4. Missing Data

The details of the percentage of missing values for each variable are provided in Appendix A.

The following table shows that for the fixed core set of questions (asked by all businesses in every survey) in about 90% of the variables the missing values are less than 5%, for 8% of the variables the missing values are within 5%-15%, and only about 2% have the missing values greater than 15%.

| Count / Wave | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | All |
|---|---|---|---|---|---|---|---|---|---|
| Proportion of missings <= 0.05 | 88.71% | 91.94% | 91.94% | 91.94% | 87.10% | 87.74% | 88.71% | 88.06% | 89.52% |
| Proportion of missings 0.05 - 0.15 | 9.03% | 8.06% | 6.13% | 3.23% | 10.97% | 10.32% | 9.35% | 10.00% | 8.39% |
| Proportion of missings > 0.15 | 2.26% | 0.00% | 1.94% | 4.84% | 1.94% | 1.94% | 1.94% | 1.94% | 2.10% |

The above classification is based on Harrell crude guidelines (2001).

## 5. The KFS multiply imputed data

Multiple imputations involve generating "m" substitute sets for the missing values, which allows for the uncertainty due to imputation to be reflected in the analysis (Rubin, 1978, 1987). The imputed values are ideally independent draws from the predictive distribution of the missing values conditional on the observed values. For the KFS multiply imputed data there are five completed data sets (m=5).

The multiple imputations for the KFS fixed core set of questions (asked by all businesses in every survey) were created using sequential regression multivariate imputation (SRMI) (Raghunathan et al., 2001), as implemented by the module mi impute chained in STATA software.

"... mi imputes chained fills in missing values in multiple variables iteratively by using chained equations, a sequence of univariate imputation methods with fully conditional specification (FCS) of prediction equations."

**Mi imputes chained** works as follows:

Let " $Z$" denote the fully-observed variables, and let $X_1, X_2, X_k$ denote $k$ variables with missing values, ordered by the amount missing, from least to most.

The imputation process for $X_1, X_2, X_k$ proceeds in10iterations.

- In the first iteration: the regression of $X_1$ on $Z$ is fitted, and the missing values of $X_1$ are imputed (randomly from an approximate predictive distribution based on the fitted regression).

- Then the regression of $X_2$ on $Z$ and $X_1$ (including the imputed values of $X_1$) is fitted, and the missing values of $X_2$ are imputed.
- and so on, until the regression of $X_k$ on $Z$ , $X_1$ , $X_2$..., $X_{k-1}$ is fitted, and the missing values of $X_k$ are imputed.

In iterations 2 through 10:

- The regression of $X_1$ on $Z$, $X_2$,$X_k$ is fitted, and the missing values of $X_1$ are imputed.
- Then the regression of $X_2$ on $Z$ , $X_1$, $X_3$, $X_k$ is fitted, and the missing values of $X_2$ are imputed.
- And so on, until the regression of $X_k$ on Z, $X_1$, $X_2$,$X_{k-1}$ is fitted, and the missing values of $X_k$ are imputed.

After 10 iterations, the final imputations of the missing values in $X_1$, $X_2$,$X_k$ are used. The entire procedure is repeated independently M times, yielding M imputed datasets.

The use of every variable other than the variable being imputed as a covariate in each regression model; and by sequencing through the regression models, the mi impute chained maintained relationships among the variables are included in the imputation models.

For each regression in the mi impute chained procedure, Stata allows the use of following regression models: regress, pmm, truncreg, intreg, logit, ologit, mlogit, poisson, and nbreg.

All the continuous variables in the KFS are non-negative and for the ones that have missing values, we have the range (contains the lower limit and upper limit) of the missing values, thus using intreg insures that the imputed values will be within the range provide by the respondents. Once a range response is used to impute a variable, the imputed value is used as a conditioning variable in all subsequent imputations

In the cases where there is no range provided or convergence issues raised, the alternative is Predictive Mean Matching (PMM). Since the PMM draws its imputed values from the observed data, the imputed values will never be outside the range of the observed values (honors any bounds that exist in the observed data).

For non-continuous variables an appropriate regression models (pmm, logit, truncreg, intreg) was used based on the nature of the variable and convergence (e.g, logit for indictor variables).

The imputation process is progressive in the sense that we first impute owners' level variables. Given owners level variables (converted to business level characteristics), we impute the rest of variables in series of steps and rounds. The other sequence of imputation was to impute data longitudinally first and then cross--sectional. The number of model specifications in the imputations is very large. In general, we have the same set of explanatory covariates for each imputed variable, which insures consistency across models. Consistency of covariate across models caused problems with variables where only small samples were available. For those variables we select explanatory covariates that fit the models best.

For the fixed core set of variables (asked by all businesses in every survey), we included the lag of the variables. We found that the best predictors of a missing value in one period are the values of that variable in the previous period. In all of model specifications, we control for other elements of the complex sample design of the KFS (strata and weights).

## 5.1 Wide vs. long format (Original KFS Data)

Both the confidential version of the KFS and the public version data come in a wide format.

Data from a repeated-measures design can be set up in two different data formats: wide and long formats. In the wide format all multi-wave variables from the same business and associated owners form just one record. For example, consider the following data: mprid is the businesses ID and var1_0, var1_1 and var1_2 are the sales for the baseline, first and second follow-up, and var2_0 is the gender of the owner, which was collected in the baseline survey only (time-constant).

| mprid | var1_0 | var1_1 | var1_2 | var2_0 |
|---|---|---|---|---|
| 1 | 485 | 2542 | 4095 | 1 |
| 2 | 2724 | 9292 | . | 1 |
| 3 | 9924 | 8049 | 2966 | 0 |

In the long (panel) format data, all variables from each wave and for the same business and associated owners form one record. For example:

| mprid | suffix | var1 | var2_0 |
|---|---|---|---|
| 1 | _0 | 485 | 1 |
| 1 | _1 | 2542 | 1 |
| 1 | _2 | 4095 | 1 |
| 2 | _0 | 2724 | 1 |
| 2 | _1 | 9292 | 1 |
| 2 | _2 | . | 1 |
| 3 | _0 | 9924 | 0 |
| 3 | _1 | 8049 | 0 |
| 3 | _2 | 2966 | 0 |

Two identification variables are needed for the long (panel) format data, in addition to the business ID, we need to have time variable (suffix), but only one variable is needed for the measurements “Var1” and “Var2.”

The long (panel) format is very useful for any longitudinal data analysis of the KFS panel (3,140). Longitudinal businesses that closed in a particular follow-up, sold or merged (records with no information) can be dropped in the long format, but not in the wide format. For example, if the business with ID two was sold in time two, then we can drop that record in time two, and as a result we will be dealing with unbalanced panel data.

| mprid | suffix | var1 | var2_0 |
|---|---|---|---|
| 1 | _0 | 485 | 1 |
| 1 | _1 | 2542 | 1 |
| 1 | _2 | 4095 | 1 |
| 2 | _0 | 2724 | 1 |
| 2 | _1 | 9292 | 1 |
| 3 | _0 | 9924 | 0 |
| 3 | _1 | 8049 | 0 |
| 3 | _2 | 2966 | 0 |

The advantage of the wide format data is that it is more convenient for the analysis of transitions and sequences, cross-tab (e.g. wave 1 vs. wave 2), lagged regression ($y_t = \alpha + \beta y_{t-1} + \gamma x_{t-1}$), recoding data into soft and hard missing value, logical imputation, defining subpopulation based on time varying variables, survival analysis, cross-sectional analysis and some data manipulation.

## 5.2 Wide vs. long format (KFS Multiply Imputed Data)

For MI data, the wide vs. long terminology is borrowed from reshape and the structures are similar but are not equivalent. All the KFS Multiply Imputed Data files are formatted using the flong style, which means, in addition to the original KFS data (regardless of the original data format, m=0), we have another five imputed dataset of the KFS (m=1,2,3,4,5).

Example: Wide original KFS - mi flong

| m | mprid | var1_0 | var1_1 | var1_2 | var2_0 |
|---|---|---|---|---|---|
| 0 | 1 | 485 | 2542 | 4095 | 1 |
| 0 | 2 | 2724 | 9292 | . | 1 |
| 0 | 3 | 9924 | 8049 | 2966 | 0 |
| 1 | 1 | 485 | 2542 | 4095 | 1 |
| 1 | 2 | 2724 | 9292 | Imputed 1 | 1 |
| 1 | 3 | 9924 | 8049 | 2966 | 0 |
| 2 | 1 | 485 | 2542 | 4095 | 1 |
| 2 | 2 | 2724 | 9292 | Imputed 2 | 1 |
| 2 | 3 | 9924 | 8049 | 2966 | 0 |
| 3 | 1 | 485 | 2542 | 4095 | 1 |
| 3 | 2 | 2724 | 9292 | Imputed 3 | 1 |
| 3 | 3 | 9924 | 8049 | 2966 | 0 |
| 4 | 1 | 485 | 2542 | 4095 | 1 |
| 4 | 2 | 2724 | 9292 | Imputed 4 | 1 |
| 4 | 3 | 9924 | 8049 | 2966 | 0 |
| 5 | 1 | 485 | 2542 | 4095 | 1 |
| 5 | 2 | 2724 | 9292 | Imputed 5 | 1 |
| 5 | 3 | 9924 | 8049 | 2966 | 0 |

Example: long original KFS -mi flong

| m | mprid | suffix | var1 | var2_0 |
|---|---|---|---|---|
| 0 | 1 | _0 | 485 | 1 |
| 0 | 1 | _1 | 2542 | 1 |
| 0 | 1 | _2 | 4095 | 1 |
| 0 | 2 | _0 | 2724 | 1 |
| 0 | 2 | _1 | 9292 | 1 |
| 0 | 2 | _2 | . | 1 |
| 0 | 3 | _0 | 9924 | 0 |
| 0 | 3 | _1 | 8049 | 0 |
| 0 | 3 | _2 | 2966 | 0 |
| 1 | 1 | _0 | 485 | 1 |
| 1 | 1 | _1 | 2542 | 1 |
| 1 | 1 | _2 | 4095 | 1 |
| 1 | 2 | _0 | 2724 | 1 |
| 1 | 2 | _1 | 9292 | 1 |
| 1 | 2 | _2 | Imputed 1 | 1 |
| 1 | 3 | _0 | 9924 | 0 |
| 1 | 3 | _1 | 8049 | 0 |
| 1 | 3 | _2 | 2966 | 0 |
| 2 | 1 | _0 | 485 | 1 |
| 2 | 1 | _1 | 2542 | 1 |
| 2 | 1 | _2 | 4095 | 1 |
| 2 | 2 | _0 | 2724 | 1 |
| 2 | 2 | _1 | 9292 | 1 |
| 2 | 2 | _2 | Imputed 2 | 1 |
| 2 | 3 | _0 | 9924 | 0 |
| 2 | 3 | _1 | 8049 | 0 |
| 2 | 3 | _2 | 2966 | 0 |
| 3 | 1 | _0 | 485 | 1 |
| 3 | 1 | _1 | 2542 | 1 |
| 3 | 1 | _2 | 4095 | 1 |
| . | . | . | . | . |
| 5 | 3 | _0 | 9924 | 0 |
| 5 | 3 | _1 | 8049 | 0 |
| 5 | 3 | _2 | 2966 | 0 |

The KFS Multiply Imputed Data files are available in Stata format. The following table shows the names of the files, the number of observations in each file, and the file's format:

| File name | Original KFS data | KFS Multiply Imputed Data - format | Number of variables |
|---|---|---|---|
| **MI_Flags_wide** | Wide (4,928 obs.) : Indictor variables for missing values | NA | 7,043 |
| **KFS8_LI** | Wide: KFS8 after logical imputation. | NA | 7,066 |
| **KFS8_Cross_Sectional_wide_MI_Long** | Wide (4,928 obs.) | Flong (4928*6=29,568 obs.) | 5,742 |
| **KFS8_Cross_Sectional_Long_MI_Long** | Long (4,928*8=39,424 obs.) | Flong (39,424*6=236,544 obs.) | 730 |
| **KFS8_Longitudinal_wide_MI_Long** | Wide (3,140 obs.) | Flong (3,140*6=18,840 obs.) | 5,728 |
| **KFS8_Longitudinal_Long_MI_Long** | Long | Flong (18,286*6=109,716 obs.) | 723 |

While Stata allows other formats (wide, mlong), we do not recommend converting the data to another format. This is due to the fact that when you construct a new variable is could be a super-varying variables in the KFS. You must use the flong format because in the wide and mlong formats, there is simply no place to store super-varying values.

A variable is said to be super varying if its values in the complete observations differ across m, while the existence of super-varying variables is usually an indication of error, it is not the case for the KFS's super-varying variables. The KFS has a complex skip logic that will produce super-varying variables

SPSS Users: while you can import the KFS Multiply Imputed Stata files to SPSS you should be aware that (as of SPSS19) Complex Sampling procedures in SPSS currently do not automatically analyze multiply imputed datasets.

SAS Users: you can import the KFS Multiply Imputed Stata files to SAS. SAS requires only the imputed data sets to be in the file, thus after importing the data to SAS, make sure to drop the original data (m=0). You can do this by using the variable named master (Keep if master>0). Use MIANALYZE procedures to analyze a multiply imputed dataset.

The following proc import statement will read the xxxx.dta data file and create a temporary data set called mydata.

```
Proc import datafile="Drive:\xxxx.dta" out=mydata dbms = dta replace;
run;
```

Use MIANALYZE procedures to analyze a multiply imputed data set. Also, SAS-callable SUDAAN includes a built-in option for analyzing multiply imputed data.

Super-varying variables are not an issue that you should be worried about in both SAS and SPSS.

### 5.3 Renamed variables/ Value recoded variables /Newly created variables

To use loops and reshape data efficiently, we need to have the names of the variables to be the same across all the follow-ups.

- Renamed variables

To insure consistency of the variables names across all year, the following variables were renamed:

| Old | New |
|---|---|
| fstatus_f2_2 | fstatus_2 |
| fstatus_f3_3 | fstatus_3 |
| fstatus_f4_4 | fstatus_4 |
| fstatus_f5_5 | fstatus_5 |
| fstatus_f6_6 | fstatus_6 |
| fstatus_f7_7 | fstatus_7 |
| b2a_legal_status_0 | c1z2_legal_status_0 |
| f2_owner_amt_eq_invest_allyrs_15 | f2_ownr_amt_eqinvest_allyrs_15_5 |
| f3a_xxxxxxxx | f3_xxxxxxxx |
| f3b_xxxxxxxx | f3_xxxxxxxx |
| f3c_xxxxxxxx | f3_xxxxxxxx |
| f3d_xxxxxxxx | f3_xxxxxxxx |
| f3e_xxxxxxxx | f3_xxxxxxxx |
| f3f_xxxxxxxx | f3_xxxxxxxx |
| f3g_xxxxxxxx | f3_xxxxxxxx |
| xx_2004_xx | xx_xx |
| xx_2005_xx | xx_xx |
| xx_2006_xx | xx_xx |
| xx_2007_xx | xx_xx |
| xx_2008_xx | xx_xx |
| xx_2009_xx | xx_xx |
| xx_2010_xx | xx_xx |

| Old | New |
|---|---|
| xx_2011_xx | xx_xx |
| cswgt_final_0 | wgt_final_0 |
| cswgt_final_1 | wgt_final_1 |
| cswgt_final_2 | wgt_final_f2_2 |
| cswgt_final_3 | wgt_final_f3_3 |
| cswgt_final_4 | wgt_final_f4_4 |
| cswgt_final_5 | wgt_final_f5_5 |
| cswgt_final_6 | wgt_final_f6_6 |
| cswgt_final_7 | wgt_final_f7_7 |
| wgt_1_long | wgt_final_1 |
| wgt_2_long | wgt_final_f12_long_2 |
| wgt_3_long | wgt_final_f123_long_3 |
| wgt_4_long | wgt_final_f1234_long_4 |
| wgt_5_long | wgt_final_f5_long_5 |
| wgt_6_long | wgt_final_f6_long_6 |
| wgt_7_long | wgt_final_f7_long_7 |

- Value recoded

For the race categories questions respondents were allowed to report multiracial or mixed-race. Thus, there is a question for each race. Since it is more practical for analysis purposes to have one variable with race coded as categorical variable, we created a new race variable "g6_race_group_xx_y" having the following codes/values:

American Indian Or Alaska Native .......................................................... 01

Native Hawaiian Or Other Pacific Islander .......................................... 02

Asian ............................................................................................................ 03

Black Or African American .................................................................. 04

White ............................................................................................................ 05

Other Races Or Mixed Race ..................................................................... 06

The g6 questions about race are still in the data, but they have the values 1/0 (yes/no). For owners reported multiracial or mixed-race we recoded them as *other.*

The following table shows the variables that were subject to values recoding, as well as the old and new values:

| Variable | New Values | Old Values |
|---|---|---|
| g3a_oth_bus_owner | 5 | 6 to ∞ |
| g6_race_amind_owner | 0,1 | 1 |
| g6_race_nathaw_owner | 0,1 | 2 |
| g6_race_asian_owner | 0,1 | 3 |
| g6_race_black_owner | 0,1 | 4 |
| g6_race_white_owner | 0,1 | 5 |
| g6_race_other_owner | 0,1 | 6 |

| Variable | New Values | Old Values |
|---|---|---|
| g10_gender_owner | 0 | 2 |
| f23_profit_or_loss | 0 | 2 |
| f7b_pers_other_numused | 5 | 6 to ∞ |
| f9b_pers_other_numused | 5 | 6 to ∞ |
| f11b_bus_other_numused | 5 | 6 to ∞ |
| f7b_bus_credcard_numused | 5 | 6 to ∞ |
| f9b_bus_credcard_numused | 5 | 6 to ∞ |
| f7b_pers_credcard_numused | 5 | 6 to ∞ |
| f7b_pers_loan_fam_numused | 5 | 6 to ∞ |
| f9b_pers_credcard_numused | 5 | 6 to ∞ |
| f9b_pers_loan_fam_numused | 5 | 6 to ∞ |
| f11b_bus_credcard_numused | 5 | 6 to ∞ |
| f7b_pers_loan_bank_numused | 5 | 6 to ∞ |
| f9b_pers_loan_bank_numused | 5 | 6 to ∞ |
| f11b_bus_cred_line_numused | 5 | 6 to ∞ |
| f11b_bus_loans_emp_numused | 5 | 6 to ∞ |
| f11b_bus_loans_fam_numused | 5 | 6 to ∞ |
| f7b_pers_loan_other_numused | 5 | 6 to ∞ |
| f9b_pers_loan_other_numused | 5 | 6 to ∞ |
| f11b_bus_loans_bank_numused | 5 | 6 to ∞ |
| f11b_bus_loans_govt_numused | 5 | 6 to ∞ |
| f11b_bus_loans_owner_numused | 5 | 6 to ∞ |
| f11b_bus_loans_nonbank_numused | 5 | 6 to ∞ |
| f11b_busloans_otherind_numused | 5 | 6 to ∞ |
| f11a_busloans_otherbus_numused | 5 | 6 to ∞ |

- Newly created variables

To use loops efficiently, we created the following variables and we set their values to hard missing:

| Variable Name | Added to survey | Variable Name | Added to survey |
|---|---|---|---|
| c10_morelocations | 0,1 | f12f_business_equip_veh | 0,1,2,3,4 |
| c11_num_locations | 0,1 | f12f_business_sec_dep | 0,1,2,3,4 |
| c12a_sba | 0,1,2,3,5,6,7 | f12f_intellectual_prop | 0,1,2,3,4 |
| c12b_fed_gov | 0,1,2,3,5,6,7 | f12f_inventory_acctrec | 0,1,2,3,4 |
| c12c_statelocal_gov | 0,1,2,3,5,6,7 | f12f_other | 0,1,2,3,4 |
| c12d_non_profit | 0,1,2,3,5,6,7 | f12f_other_pers_assets | 0,1,2,3,4 |
| c12e_college_univ | 0,1,2,3,5,6,7 | f12f_pers_real_estate | 0,1,2,3,4 |
| c12f_chamber_of_comm | 0,1,2,3,5,6,7 | f14d_new_loans | 0,1,2 |
| c12g_for_profit_org | 0,1,2,3,5,6,7 | f14e_approved_denied | 0,1,2 |
| c12h_other | 0,1,2,3,5,6,7 | f14f_bus_credit_hist | 0,1,2 |
| c9_loc_change_reason | 0 | f14f_inadeq_doc | 0,1,2 |
| d1_a_new_product | 0,1,2,3,4 | f14f_insuff_coll | 0,1,2 |
| d1_b_new_to_market | 0,1,2,3,4 | f14f_loan_toolarge | 0,1,2 |
| d1c_a_regional | 0,1,2,3,4 | f14f_new_bus | 0,1,2 |
| d1c_b_national | 0,1,2,3,4 | f14f_other | 0,1,2 |

| Variable Name | Added to survey | Variable Name | Added to survey |
|---|---|---|---|
| d1c_c_international | 0,1,2,3,4 | f14f_pers_credit_hist | 0,1,2 |
| d1d_new_processes | 0,1,2,3,4 | f14f_restr_on_lending | 0,1,2,3 |
| d2a_compadv_comp_reason | 0,1,2 | f14g_didnotapply | 0,1,2 |
| d2a_compadv_govlab_reason | 0,1,2 | f14h_loan_guarantees | 0,1,2,3 |
| d2a_compadv_patents_reason | 0,1,2 | f14i_economy_effect | 0,1,2,3,5,6,7 |
| d2a_compadv_univ_reason | 0,1,2 | f14j_most_challenging | 0,1,2,3 |
| d2b_compadv_comp_strength | 0,1,2 | f19a_res_dev_amt | 0,1,2 |
| d2b_compadv_govlab_strength | 0,1,2 | f19b_a_design | 0,1,2,3 |
| d2b_compadv_patents_strength | 0,1,2 | f19b_b_investments | 0,1,2,3 |
| d2b_compadv_univ_strength | 0,1,2 | f19b_c_brand_dev | 0,1,2,3 |
| d2c_compadv_cost_reason | 0,1,2,3,4,6,7 | f19b_d_org_dev | 0,1,2,3 |
| d2c_compadv_design_reason | 0,1,2,3,4,6,7 | f19b_e_worker_training | 0,1,2,3 |
| d2c_compadv_expertise_reason | 0,1,2,3,4,6,7 | f19b_f_other | 0,1,2,3 |
| d2c_compadv_marketing_reason | 0,1,2,3,4,6,7 | f19c_a_design_amt | 0,1,2,3,4 |
| d2c_compadv_price_reason | 0,1,2,3,4,6,7 | f19c_b_investments_amt | 0,1,2,3,4 |
| d2c_compadv_reputation_reason | 0,1,2,3,4,6,7 | f19c_c_brand_dev_amt | 0,1,2,3,4 |
| d2c_compadv_speed_reason | 0,1,2,3,4,6,7 | f19c_d_org_dev_amt | 0,1,2,3,4 |
| d2d_compadv_cost_strength | 0,1,2,3,4,6,7 | f19c_e_worker_training_amt | 0,1,2,3,4 |
| d2d_compadv_design_strength | 0,1,2,3,4,6,7 | f19c_f_other_amt | 0,1,2,3,4 |
| d2d_compadv_expertise_strength | 0,1,2,3,4,6,7 | f19c_intangassets_amt | 0,1,2,3,5,6,7 |
| d2d_compadv_marketing_strength | 0,1,2,3,4,6,7 | f32_chap11_bankruptcy | 0,1,2,3 |
| d2d_compadv_price_strength | 0,1,2,3,4,6,7 | f33_expected_growth | 0,1,2,3,5,6,7 |
| d2d_compadv_reput_strength | 0,1,2,3,4,6,7 | f34_future_revenue | 0,1,2,3,5,6,7 |
| d2d_compadv_speed_strength | 0,1,2,3,4,6,7 | f5a_seek_equity | 0,1,2,3,4 |
| d5a_founded_newprod | 0,1,2,3,4,6,7 | f6z_family_owned | 0,1,2,3,5,6,7 |
| d5b_a_personaluse | 0,1,2,3,4,6,7 | g10b_marital_status | 0,1,2,3 |
| d5b_b_previousjob | 0,1,2,3,4,6,7 | g1b2_reasonfor_business | 0,1,2,3,4,5,6 |
| d5b_c_startingbus | 0,1,2,3,4,6,7 | g10d_personal_outlook | 0,1,2,3,5,6,7 |
| d8_customer_locations | 0,1,2 | g10c_net_worth | 0,1,2,3 |
| d8a_international_sales | 0,1,2 | d9a_perc_internet_sales | 0,1,2 |
| d8b_perc_international_sales | 0,1,2 | f12e_collateral | 0,1,2,3,4 |
| d9_internet_sales | 0,1,2 | | |

For ease of running loops at owner level variables, all survey rounds have owner level variables for 15 owners.

# 6. Analytic Examples Using Stata® 12.0

## 6.1 The mi suite of commands: info from the Stata user guide

" The mi suite of commands deals with multiple-imputation data, abbreviated as mi data. In summary,

1. mi data may be stored in one of four formats—flongsep, flong, mlong, and wide—known as styles.

2. mi data contain M imputations numbered m = 1, 2, : : : , M, and contain m = 0, the original data with missing values.

3. Each variable in mi data is registered as imputed, passive, or regular, or it is unregistered.

a. Unregistered variables are mostly treated like regular variables.

b. Regular variables usually do not contain missing, or if they do, the missing values are not imputed in m > 0.

c. Imputed variables contain missing in m = 0, and those values are imputed in m > 0.

d. Passive variables are algebraic combinations of imputed, regular, or other passive variables.

4. If an imputed variable contains a value greater than . in m = 0—it contains .a, .b, : : : , .z—then that value is considered a hard missing and the missing value persists in m > 0."

### 6.1.1 Data Management Commands

- `Reporting Commands`

| Command | Function |
|---|---|
| `mi query` | Reports whether the data in memory are mi data. |
| `mi describe,noupdate` | Provides a more detailed report on mi data. |
| `mi set` | Declare multiple-imputation data |
| `mi xeq` | `Execute command(s) on individual imputations or on selected ones` |

```
use KFS8_Cross_Sectional_wide_MI_Long,clear
(KFS8 Cross Sectional in wide format & Multiply Imputed Data in long
format)

mi query

data mi set flong, M = 5

mi describe,noupdate

  Style:  flong

  Obs.:   complete        0
          incomplete      4,928     (M = 5 imputations)
          ---------------------
          total           4,928

  Vars.:  imputed:  0

          passive:  0

          regular:  1; mprid

           system:   3; _mi_m _mi_id _mi_miss

         (there are 5738 unregistered variables)


mi misstable summarize c5_*

                                                            Obs<.
                                              +------------------------------
                                                Unique
Variable         Obs=.     Obs>.     Obs<.   values        Min          Max
-------------+--------------------------------+------------------------------
c5_num_emp~0       105               4,823       49          0          165
c5_num_emp~1        46       930     3,952       63          0         1100
c5_num_emp~2        37     1,538     3,353       62          0          160
c5_num_emp~3        25     2,013     2,890       60          0          350
c5_num_emp~4         4     2,322     2,602       63          0          320
c5_num_emp~5        10     2,520     2,398       63          0          265
c5_num_emp~6         5     2,802     2,121       58          0          400
c5_num_emp~7         7     2,921     2,000       71          0          900
-----------------------------------------------------------------------------
```

All KFS multiply imputed data files already declared to be multiple-imputation data, thus do not use `mi set`.

➢ Setting mi Data

Commands like `svyset`, `stset`, and `xtset` also have mi versions: use `mi svyset` to declare survey data, use `mi stset` to declare survival data, and use `mi xtset` to declare panel data.

```
use KFS8_Cross_Sectional_wide_MI_Long,clear
(KFS8 Cross Sectional in wide format & Multiply Imputed Data in long format)

mi svyset  [pweight=cswgt_final_0],  strata(sampleinfo_samplestrata_0)

      pweight: cswgt_final_0
          VCE: linearized
  Single unit: missing
     Strata 1: sampleinfo_samplestrata_0
         SU 1: <observations
        FPC 1: <zero

* sts is not supported by svy nor mi

mi stset Duration [pweight=cswgt_final_0] , failure(event==1)

      pweight: cswgt_final_0
          VCE: linearized
  Single unit: missing
     Strata 1: sampleinfo_samplestrata_0
         SU 1: <observations>
        FPC 1: <zero>

. mi stset Duration [pweight=cswgt_final_0] , failure(event==1)

     failure event:  event == 1
obs. time interval:  (0, Duration]
 exit on or before:  failure
            weight:  [pweight=cswgt_final_0]

------------------------------------------------------------------------------
     4928  total obs.
        0  exclusions
------------------------------------------------------------------------------
     4928  obs. remaining, representing
     2190  failures in single record/single failure data
    27568  total analysis time at risk, at risk from t =         0
                             earliest observed entry t =         0
                                  last observed exit t =         8
use KFS8_Longitudinal_Long_MI_Long,clear
(KFS8 Longitudinal in long format & Multiply Imputed Data in long format)

mi xtset  mprid year
       panel variable:  mprid (unbalanced)
        time variable:  year, 2004 to 2011
                delta:  1 unit

mi xeq 0:   xtdescribe

   mprid:  10000016, 10000090, ..., 10324611                n =       3140
    year:  2004, 2005, ..., 2011                            T =          8
           Delta(year) = 1 unit
           Span(year)  = 8 periods
```

```
           (mprid*year uniquely identifies each observation)

Distribution of T_i:   min      5%     25%       50%       75%     95%     max
                         1       1       3         8         8       8       8

     Freq.  Percent    Cum. |  Pattern
 ---------------------------+----------
     1630     51.91   51.91 |  11111111
      303      9.65   61.56 |  1.......
      283      9.01   70.57 |  11......
      238      7.58   78.15 |  1111....
      224      7.13   85.29 |  111.....
      164      5.22   90.51 |  11111...
      153      4.87   95.38 |  111111..
      145      4.62  100.00 |  1111111.
 ---------------------------+----------
     3140    100.00         |  XXXXXXXX

mi svyset  [pweight=wgt_7_long],  strata(sampleinfo_samplestrata)

      pweight: wgt_7_long
          VCE: linearized
  Single unit: missing
     Strata 1: sampleinfo_samplestrata
         SU 1: <observations>
        FPC 1: <zero>
```

### 6.1.2 Creating or Changing Variables

In Stata , the definitions of MI variables are:

1. A regular variable is a variable that is neither imputed nor passive and that has the same values, whether missing or not, in all m. (e.g., c4_numowners_confirm)
2. An imputed variable is a variable that has missing values and for which you have imputations. An imputed variable will have missing values in m = 0 and varying values for observations in m > 0.
3. A passive variable is a varying variable that is a function of imputed variables or of other passive variables. A passive variable will have missing values in m = 0 and varying values for observations in m > 0.

Two other definitions that they use in the manual, the definitions for varying and super varying.

4. Varying: a variable is said to be varying if its values in the incomplete observations (missing) differ across imputations. Imputed and passive variables are varying. Regular variables are nonvarying. Unregistered variables can be either.

5. Super varying: a variable is said to be super varying if its values in the complete observations (no missing in m=0) differ across imputations.

The distinction between varying variables and super-varying variables allows -mi- to detect inconsistencies among complete observations across imputations and fix such inconsistencies. Variables that are functions of the values of other imputed variables are likely to be super-varying (e.g., skip logic)

#### 6.1.2.1 Creating or Changing Variables - Regular Variables

New or changed variable that is functions of existing regular variables are also regular variable. Use `mi xeq` to create new regular variable or change the value of existing regular variable.

**Example of regular variables in KFS**

| |
|---|
| mprid |
| b1_bus_start |
| c1z2_legal_status |
| c2_owners |
| c3a_owner_operators |
| c4_numowners_confirm |
| d1a_provide_service |
| d1b_provide_product |
| c8_primary_loc |
| hightech |
| techempl |
| techgenr |
| naics_code |
| sampleinfo_samplestrata |
| website |
| email |

```
use KFS8_Cross_Sectional_wide_MI_Long,clear
(KFS8 Cross Sectional in wide format & Multiply Imputed Data in long format)

mi xeq: recode  c8_primary_loc_0 (2/5=0)
m=0 data:
-> recode  c8_primary_loc_0 (2/5=0)
(c8_primary_loc_0: 2437 changes made)

m=1 data:
-> recode  c8_primary_loc_0 (2/5=0)
(c8_primary_loc_0: 2437 changes made)

m=2 data:
-> recode  c8_primary_loc_0 (2/5=0)
(c8_primary_loc_0: 2437 changes made)

m=3 data:
-> recode  c8_primary_loc_0 (2/5=0)
(c8_primary_loc_0: 2437 changes made)

m=4 data:
-> recode  c8_primary_loc_0 (2/5=0)
(c8_primary_loc_0: 2437 changes made)

m=5 data:
-> recode  c8_primary_loc_0 (2/5=0)
(c8_primary_loc_0: 2437 changes made)

mi xeq: gen provide_service_product_1=(d1a_provide_service_1==  ///
d1b_provide_product_1 ==1) if d1a_provide_service_1<. &          ///
d1a_provide_service_1<.

m=0 data:
-> gen
provide_service_product_1=(d1a_provide_service_1==d1b_provide_product_1==1) if
d1a_provide_service_1<. & d1a_provide_service_1<.
(930 missing values generated)

m=1 data:
-> gen
provide_service_product_1=(d1a_provide_service_1==d1b_provide_product_1==1) if
d1a_provide_service_1<. & d1a_provide_service_1<.
(930 missing values generated)

m=2 data:
-> gen
provide_service_product_1=(d1a_provide_service_1==d1b_provide_product_1==1) if
d1a_provide_service_1<. & d1a_provide_service_1<.
(930 missing values generated)

m=3 data:
-> gen
provide_service_product_1=(d1a_provide_service_1==d1b_provide_product_1==1) if
d1a_provide_service_1<. & d1a_provide_service_1<.
(930 missing values generated)

m=4 data:
-> gen
provide_service_product_1=(d1a_provide_service_1==d1b_provide_product_1==1) if
d1a_provide_service_1<. & d1a_provide_service_1<.
(930 missing values generated)
```

```
m=5 data:
-> gen
provide_service_product_1=(d1a_provide_service_1==d1b_provide_product_1==1) if
d1a_provide_service_1<. & d1a_provide_service_1<.
(930 missing values generated)


mi xeq: replace provide_service_product_1=.a if classf_1<6

m=0 data:
-> replace provide_service_product_1=.a if classf_1<6
(930 real changes made, 930 to missing)

m=1 data:
-> replace provide_service_product_1=.a if classf_1<6
(930 real changes made, 930 to missing)

m=2 data:
-> replace provide_service_product_1=.a if classf_1<6
(930 real changes made, 930 to missing)

m=3 data:
-> replace provide_service_product_1=.a if classf_1<6
(930 real changes made, 930 to missing)

m=4 data:
-> replace provide_service_product_1=.a if classf_1<6
(930 real changes made, 930 to missing)

m=5 data:
-> replace provide_service_product_1=.a if classf_1<6
(930 real changes made, 930 to missing)
```

#### 6.1.2.2 Creating or Changing Variables - Passive variables

A passive variable is a function of imputed variables. Use `mi passive` followed by `gen` or `replace` commands to generate/replace and register passive variables. You can use mi passive with any function that produces values that solely depend on values within the observation. In general, you cannot use mi passive with functions that produce values that depend on groups of observations.

You can use `mi passive` only with the variables that already registered. You should be careful not to mistakenly use `mi passive` to create super-varying variables.

We already register the varying variables in the KFS MI files. All variables were checked for consistency across imputation and were registered correctly. You can use the command "`mi varying`" to see the registered / unregistered variables.

```
use KFS8_Cross_Sectional_wide_MI_Long,clear
(KFS8 Cross Sectional in wide format & Multiply Imputed Data in long format)
```

```
mi passive : gen tot_assets_0 = f29_assetval_acctrec_0 + f29_assetval_cash_0 + f29_assetval_equip_0 + f29_assetval_inv_0 + f29_assetval_landbuild_0 + f29_assetval_othbusprop_0 + f29_assetval_other_0 + f29_assetval_veh_0
m=0:(1046 missing values generated)
m=1:
m=2:
m=3:
m=4:
m=5:
mi passive : egen tot_assets_0= rowtotal(f29_assetval_acctrec_0 f29_assetval_cash_0 f29_assetval_equip_0 f29_assetval_inv_0 f29_assetval_landbuild_0 f29_assetval_othbusprop_0 f29_assetval_other_0 f29_assetval_veh_0),missing
m=0: (21 missing values generated)
m=1:
m=2:
m=3:
m=4:
m=5:

mi passive : egen pr_race_w_0 =rowmean( g6_race_white_owner_*_0)
m=0:(11 missing values generated)
m=1:
m=2:
m=3:
m=4:
m=5:
```

#### 6.1.2.3 Creating or Changing Variables - Super-varying Variables

A super-varying variable is generated when the variable value is a function of the values of other imputed variables. Thus, its values in the complete observations (m=0) differ across m.

A super-varying variable could be a result of a skip logic where the sample varies across imputations, incorrect flow through prescribed skip patterns, or/and inconsistency in values after editing the raw data.

In the long format, a super-varying variable is generated when the variable value is a function of the values of imputed variables for other observations.

**<u>Super-varying variables must not be registered</u>**

Use `mi xeq` followed by `gen`, `egen` or `replace` commands to generate/replace super-varying variables.

```
use KFS8_Cross_Sectional_wide_MI_Long,clear
(KFS8 Cross Sectional in wide format & Multiply Imputed Data in long format)

mi xeq :egen  tot_ex_equity_0= rowtotal(f4_eq_amt_spouse_allyrs_0
f4_eq_amt_parents_allyrs_0 f4_eq_amt_angels_allyrs_0
f4_eq_amt_companies_allyrs_0 f4_eq_amt_govt_allyrs_0
```

```
f4_eq_amt_vent_cap_allyrs_0 f4_eq_amt_other_allyrs_0),missing


m=0 data:
-> egen  tot_ex_equity_0= rowtotal(f4_eq_amt_spouse_allyrs_0
f4_eq_amt_parents_allyrs_0 f4_eq_amt_angels_allyrs_0 f4_eq_amt_companies_allyrs_0
f4_eq_amt_govt_allyrs_0 f4_eq_amt_vent_cap_allyrs_0
f4_eq_amt_other_allyrs_0),missing
(2588 missing values generated)

m=1 data:
-> egen  tot_ex_equity_0= rowtotal(f4_eq_amt_spouse_allyrs_0
f4_eq_amt_parents_allyrs_0 f4_eq_amt_angels_allyrs_0 f4_eq_amt_companies_allyrs_0
f4_eq_amt_govt_allyrs_0 f4_eq_amt_vent_cap_allyrs_0
f4_eq_amt_other_allyrs_0),missing
(2576 missing values generated)

m=2 data:
-> egen  tot_ex_equity_0= rowtotal(f4_eq_amt_spouse_allyrs_0
f4_eq_amt_parents_allyrs_0 f4_eq_amt_angels_allyrs_0 f4_eq_amt_companies_allyrs_0
f4_eq_amt_govt_allyrs_0 f4_eq_amt_vent_cap_allyrs_0
f4_eq_amt_other_allyrs_0),missing
(2576 missing values generated)

m=3 data:
-> egen  tot_ex_equity_0= rowtotal(f4_eq_amt_spouse_allyrs_0
f4_eq_amt_parents_allyrs_0 f4_eq_amt_angels_allyrs_0 f4_eq_amt_companies_allyrs_0
f4_eq_amt_govt_allyrs_0 f4_eq_amt_vent_cap_allyrs_0
f4_eq_amt_other_allyrs_0),missing
(2576 missing values generated)

m=4 data:
-> egen  tot_ex_equity_0= rowtotal(f4_eq_amt_spouse_allyrs_0
f4_eq_amt_parents_allyrs_0 f4_eq_amt_angels_allyrs_0 f4_eq_amt_companies_allyrs_0
f4_eq_amt_govt_allyrs_0 f4_eq_amt_vent_cap_allyrs_0
f4_eq_amt_other_allyrs_0),missing
(2576 missing values generated)

m=5 data:
-> egen  tot_ex_equity_0= rowtotal(f4_eq_amt_spouse_allyrs_0
f4_eq_amt_parents_allyrs_0 f4_eq_amt_angels_allyrs_0 f4_eq_amt_companies_allyrs_0
f4_eq_amt_govt_allyrs_0 f4_eq_amt_vent_cap_allyrs_0
f4_eq_amt_other_allyrs_0),missing
(2576 missing values generated)

mi xeq : mean tot_ex_equity_0

m=0 data:
-> mean tot_ex_equity_0

Mean estimation                     Number of obs    =    2340

-----------------------------------------------------------------
                |       Mean   Std. Err.     [95% Conf. Interval]
----------------+------------------------------------------------
tot_ex_equity_0 |     126749    35987.7      56177.88    197320.1
-----------------------------------------------------------------

m=1 data:
-> mean tot_ex_equity_0

Mean estimation                     Number of obs    =    2352
```

```
-----------------------------------------------------------------
                 |       Mean   Std. Err.     [95% Conf. Interval]
-----------------+-----------------------------------------------
tot_ex_equity_0 |   131588.3   35971.22      61049.66    202126.9
-----------------------------------------------------------------

m=2 data:
-> mean tot_ex_equity_0

Mean estimation                      Number of obs    =    2352

-----------------------------------------------------------------
                 |       Mean   Std. Err.     [95% Conf. Interval]
-----------------+-----------------------------------------------
tot_ex_equity_0 |   131219.8   36010.29      60604.58      201835
-----------------------------------------------------------------

m=3 data:
-> mean tot_ex_equity_0

Mean estimation                      Number of obs    =    2352

-----------------------------------------------------------------
                 |       Mean   Std. Err.     [95% Conf. Interval]
-----------------+-----------------------------------------------
tot_ex_equity_0 |   131090.2   35951.89      60589.51    201590.9
-----------------------------------------------------------------

m=4 data:
-> mean tot_ex_equity_0

Mean estimation                      Number of obs    =    2352

-----------------------------------------------------------------
                 |       Mean   Std. Err.     [95% Conf. Interval]
-----------------+-----------------------------------------------
tot_ex_equity_0 |   131320.4   35967.15      60789.76      201851
-----------------------------------------------------------------

m=5 data:
-> mean tot_ex_equity_0

Mean estimation                      Number of obs    =    2352

-----------------------------------------------------------------
                 |       Mean   Std. Err.     [95% Conf. Interval]
-----------------+-----------------------------------------------
tot_ex_equity_0 |     132369   35973.14       61826.6    202911.3
-----------------------------------------------------------------
```

What if we use `mi passive`?

```
use KFS8_Cross_Sectional_wide_MI_Long,clear
(KFS8 Cross Sectional in wide format & Multiply Imputed Data in long format)

mi passive : egen  tot_ex_equity_0= rowtotal(f4_eq_amt_spouse_allyrs_0
f4_eq_amt_parents_allyrs_0 f4_eq_amt_angels_allyrs_0
f4_eq_amt_companies_allyrs_0 f4_eq_amt_govt_allyrs_0
f4_eq_amt_vent_cap_allyrs_0 f4_eq_amt_other_allyrs_0),missing
m=0: (2588 missing values generated)
m=1:(2576 missing values generated)
```

```
m=2:(2576 missing values generated)
m=3:(2576 missing values generated)
m=4:(2576 missing values generated)
m=5:(2576 missing values generated)
(451 values of passive variable tot_ex_equity_0 in m>0 updated to match values in
m=0)


mi xeq : mean tot_ex_equity_0

m=0 data:
-> mean tot_ex_equity_0

Mean estimation                     Number of obs    =    2340

-----------------------------------------------------------------
                |       Mean   Std. Err.     [95% Conf. Interval]
----------------+------------------------------------------------
tot_ex_equity_0 |     126749    35987.7      56177.88    197320.1
-----------------------------------------------------------------

m=1 data:
-> mean tot_ex_equity_0

Mean estimation                     Number of obs    =    2340

-----------------------------------------------------------------
                |       Mean   Std. Err.     [95% Conf. Interval]
----------------+------------------------------------------------
tot_ex_equity_0 |     126749    35987.7      56177.88    197320.1
-----------------------------------------------------------------

m=2 data:
-> mean tot_ex_equity_0

Mean estimation                     Number of obs    =    2340

-----------------------------------------------------------------
                |       Mean   Std. Err.     [95% Conf. Interval]
----------------+------------------------------------------------
tot_ex_equity_0 |     126749    35987.7      56177.88    197320.1
-----------------------------------------------------------------

m=3 data:
-> mean tot_ex_equity_0


Mean estimation                     Number of obs    =    2340

-----------------------------------------------------------------
                |       Mean   Std. Err.     [95% Conf. Interval]
----------------+------------------------------------------------
tot_ex_equity_0 |     126749    35987.7      56177.88    197320.1
-----------------------------------------------------------------

m=4 data:
-> mean tot_ex_equity_0

Mean estimation                     Number of obs    =    2340

-----------------------------------------------------------------
                |       Mean   Std. Err.     [95% Conf. Interval]
```

```
-----------------+------------------------------------------------
tot_ex_equity_0 |     126749    35987.7      56177.88    197320.1
------------------------------------------------------------------

m=5 data:
-> mean tot_ex_equity_0

Mean estimation                    Number of obs    =    2340

------------------------------------------------------------------
                |       Mean   Std. Err.     [95% Conf. Interval]
-----------------+------------------------------------------------
tot_ex_equity_0 |     126749    35987.7      56177.88    197320.1
------------------------------------------------------------------
```

Since that the `f4_eq_amt_xxxx_allyrs` variables are not registered, mistakenly use `mi passive` to create super-varying variables will trigger an automatic update by Stata. The results of such update is that all the data for this variable in the imputed data (m=1, 2, 3, 4, 5) will be replaced by the data in the original data (m=0).

Leaving newly created variables unregistered makes `mi update` leave them alone. So, it is always safer to use `mi xeq` when you create any new variable.

### 6.1.3 Estimating

Multiple-imputation data analysis in Stata is similar to standard data analysis. All what we have to do is to prefix the estimation commands with `mi estimate`:

The following estimation commands support the `mi estimate` prefix.

**Linear regression models**

| | |
|---|---|
| regress | Linear regression |
| cnsreg | Constrained linear regression |
| mvreg | Multivariate regression |

**Binary-response regression models**

| | |
|---|---|
| logistic | Logistic regression, reporting odds ratios |
| logit | Logistic regression, reporting coefficients |
| probit | Probit regression |
| cloglog | Complementary log-log regression |
| binreg | GLM for the binomial family |

**Count-response regression models**

| | |
|---|---|
| poisson | Poisson regression |
| nbreg | Negative binomial regression |
| gnbreg | Generalized negative binomial regression |

**Ordinal-response regression models**

| | |
|---|---|
| ologit | Ordered logistic regression |

| | |
|---|---|
| oprobit | Ordered probit regression |
| **Categorical-response regression models** | |
| mlogit | Multinomial (polytomous) logistic regression |
| mprobit | Multinomial probit regression |
| clogit | Conditional (fixed-effects) logistic regression |
| **Quantile regression models** | |
| qreg | Quantile regression |
| iqreg | Interquantile range regression |
| sqreg | Simultaneous-quantile regression |
| bsqreg | Bootstrapped quantile regression |
| **Survival regression models** | |
| stcox | Cox proportional hazards model |
| streg | Parametric survival models |
| stcrreg | Competing-risks regression |
| **Other regression models** | |
| glm | Generalized linear models |
| areg | Linear regression with a large dummy-variable set |
| rreg | Robust regression |
| truncreg | Truncated regression |
| **Descriptive statistics** | |
| mean | Estimate means |
| proportion | Estimate proportions |
| ratio | Estimate ratios |
| total | Estimate totals |
| **Panel-data models** | |
| xtreg | Fixed-, between- and random-effects, and population-averaged linear models |
| xtmixed | Multilevel mixed-effects linear regression |
| xtrc | Random-coefficients regression |
| xtlogit | Fixed-effects, random-effects, and population-averaged logit models |
| xtprobit | Random-effects and population-averaged probit models |
| xtcloglog | Random-effects and population-averaged cloglog models |
| xtpoisson | Fixed-effects, random-effects, and population-averaged Poisson models |
| xtnbreg | Fixed-effects, random-effects, and PA negative binomial models |
| xtmelogit | Multilevel mixed-effects logistic regression |
| xtmepoisson | Multilevel mixed-effects Poisson regression |
| xtgee | Fit population-averaged panel-data models by using GEE |
| **Survey regression models** | |
| svy | Estimation commands for survey data |

### 6.1.4 Examples – KFS wide MI style long

```
use KFS8_Cross_Sectional_wide_MI_Long,clear
(KFS8 Cross Sectional in wide format & Multiply Imputed Data in long format)

   *Calculating Company-Level Characteristics - equally weighted - 2004 only
```

```
forvalues i = 0/0 {
mi xeq :egen s_owner_amt_eq_invest_`i'=rowtotal
(f2_owner_amt_eq_invest_*_`i')
mi xeq :egen s_hours_owner_`i'       =    rowtotal
(g1b1_hours_owner_*_`i')
mi xeq :egen m_work_exp_owner_`i'     =   rowmean
(g2_work_exp_owner_*_`i')
mi xeq :egen m_age_owner_`i' =      rowmean (   g4_age_owner_*_`i')
mi xeq :egen pr_emp_owner_`i' =rowmean     (    g1a_emp_owner_*_`i')
mi xeq :egen pr_hisp_origin_owner_`i'  =  rowmean      (
      g5_hisp_origin_owner_*_`i')
mi xeq :egen pr_native_born_owner_`i'  =  rowmean      (
      g7_native_born_owner_*_`i')
mi xeq :egen pr_gender_owner_`i'       =  rowmean      (
      g10_gender_owner_*_`i')
mi xeq :egen     pr_race_w_`i'            =   rowmean     (
g6_race_white_owner_*_`i')
mi xeq :egen     pr_race_a_`i'            =   rowmean     (
g6_race_asian_owner_*_`i')
mi xeq :egen     pr_race_b_`i'            =   rowmean     (
g6_race_black_owner_*_`i')
mi xeq :egen     pr_race_o_`i'            =   rowmean     (
g6_race_other_owner_*_`i')
mi xeq :egen     pr_race_na_`i'           =   rowmean     (
g6_race_nathaw_owner_*_`i')
mi xeq :egen     pr_race_am_`i'           =   rowmean     (
g6_race_amind_owner_*_`i')
mi xeq :gen     homebased_`i' =(c8_primary_loc_`i'==1)
mi   xeq   :egen      tot_assets_`i'=   rowtotal(f29_assetval_acctrec_`i'
f29_assetval_cash_`i'     f29_assetval_equip_`i'     f29_assetval_inv_`i'
f29_assetval_landbuild_`i'                     f29_assetval_othbusprop_`i'
f29_assetval_other_`i' f29_assetval_veh_`i')
}



forvalues i = 0/0 {
mi xeq :replace  s_owner_amt_eq_invest_`i'  =.a if classf_`i'<6
mi xeq :replace  s_hours_owner_`i'          =.a if classf_`i'<6
mi xeq :replace  m_work_exp_owner_`i'       =.a if classf_`i'<6
mi xeq :replace  m_age_owner_`i'            =.a if classf_`i'<6
mi xeq :replace  pr_emp_owner_`i'           =.a if classf_`i'<6
mi xeq :replace  pr_hisp_origin_owner_`i'   =.a if classf_`i'<6
mi xeq :replace  pr_native_born_owner_`i'   =.a if classf_`i'<6
mi xeq :replace  pr_gender_owner_`i'        =.a if classf_`i'<6
mi xeq :replace   pr_race_w_`i'             =.a if classf_`i'<6
mi xeq :replace   pr_race_a_`i'             =.a if classf_`i'<6
mi xeq :replace   pr_race_b_`i'             =.a if classf_`i'<6
mi xeq :replace   pr_race_o_`i'             =.a if classf_`i'<6
mi xeq :replace   pr_race_na_`i'            =.a if classf_`i'<6
mi xeq :replace   pr_race_am_`i'            =.a if classf_`i'<6
mi xeq :replace     homebased_`i'           =.a if classf_`i'<6
mi xeq :replace   tot_assets_`i'            = .a if classf_`i'<6
}

mi svyset  [pweight=cswgt_final_0],  strata(sampleinfo_samplestrata_0)
```

## ➢ Descriptive statistics

```
mi estimate: svy: mean   pr_*_0

Multiple-imputation estimates      Imputations     =          5
Survey: Mean estimation            Number of obs   =       4928

Number of strata  =          6      Population size = 73278.441
Number of PSUs    =       4928
                                   Average RVI     =     0.0042
                                   Largest FMI     =     0.0159
                                   Complete DF     =       4922
DF adjustment:   Small sample      DF:     min     =    3720.43
                                           avg     =    4717.88
Within VCE type:   Linearized              max     =    4919.96

------------------------------------------------------------------------
                       |       Mean   Std. Err.     [95% Conf. Interval]
-----------------------+------------------------------------------------
        pr_emp_owner_0 |   .4697238   .0079859      .4540679    .4853797
pr_hisp_origin_owner_0 |   .0656113   .0040787      .0576153    .0736074
pr_native_born_owner_0 |   .8880152   .0050114      .8781907    .8978397
     pr_gender_owner_0 |   .6775468   .0053296      .6670985    .6879951
           pr_race_w_0 |   .8089889   .0064614      .7963216    .8216562
           pr_race_a_0 |   .0379152   .0030877      .0318618    .0439685
           pr_race_b_0 |   .0854109   .0046647      .0762661    .0945558
           pr_race_o_0 |   .0511647   .0036271      .0440533    .0582761
          pr_race_na_0 |   .0058139     .00133      .0032065    .0084213
          pr_race_am_0 |   .0107064   .0016509      .0074697    .0139431
------------------------------------------------------------------------
```

*Based on available-cases - not complete cases

```
mi xeq 0: svy:    mean   pr_*_0

m=0 data:
-> svy:  mean   pr_*_0
(running mean on estimation sample)

Survey: Mean estimation

Number of strata =       6          Number of obs    =    4887
Number of PSUs   =    4887          Population size  =   72617
                                    Design df        =    4881

------------------------------------------------------------------------
                       |             Linearized
                       |       Mean   Std. Err.     [95% Conf. Interval]
-----------------------+------------------------------------------------
        pr_emp_owner_0 |   .4691385   .0080128      .4534298    .4848471
pr_hisp_origin_owner_0 |   .0654262   .0040914      .0574052    .0734472
pr_native_born_owner_0 |   .8886723   .0050182      .8788343    .8985103
     pr_gender_owner_0 |   .6774446   .0053524      .6669515    .6879377
           pr_race_w_0 |    .809692   .0064732      .7970017    .8223824
           pr_race_a_0 |   .0373706   .0030706      .0313508    .0433904
           pr_race_b_0 |   .0858202   .0046952      .0766155    .0950249
           pr_race_o_0 |   .0508361   .0036048      .0437691    .0579031
          pr_race_na_0 |   .0058605   .0013419      .0032297    .0084913
          pr_race_am_0 |   .0104205    .001627      .0072309    .0136102
------------------------------------------------------------------------
```

```
mi estimate: svy:       proportion g6b_race_group_01_0

Multiple-imputation estimates      Imputations     =          5
Survey: Proportion estimation      Number of obs   =       4928

Number of strata  =          6     Population size = 73278.441
Number of PSUs    =       4928
                                   Average RVI     =     0.0087
                                   Largest FMI     =     0.0200
                                   Complete DF     =       4922
DF adjustment:   Small sample      DF:     min     =    3275.22
                                           avg     =    4393.06
Within VCE type:   Linearized              max     =    4920.00

      _prop_1: g6b_race_group_01_0 = American Indian Or Alaska Native
      _prop_2: g6b_race_group_01_0 = Native Hawaiian Or Other Pacific
      _prop_4: g6b_race_group_01_0 = Black Or African American
      _prop_6: g6b_race_group_01_0 = Other Races Or Mixed Race

--------------------------------------------------------------------
                    | Proportion   Std. Err.     [95% Conf. Interval]
--------------------+-----------------------------------------------
            _prop_1 |   .0118408   .0018489      .0082159    .0154658
            _prop_2 |   .0056011   .0013188      .0030156    .0081865
              Asian |   .0362738   .0031042      .0301881    .0423596
            _prop_4 |   .0857726   .0047379      .0764842     .095061
              White |   .8100227   .0066022      .7970793    .8229661
            _prop_6 |    .050489   .0037277      .0431801     .057798
--------------------------------------------------------------------

mi estimate: svy:       total   c5_num_employees_0

Multiple-imputation estimates      Imputations     =          5
Survey: Total estimation           Number of obs   =       4928

Number of strata  =          6     Population size = 73278.441
Number of PSUs    =       4928
                                   Average RVI     =     0.0064
                                   Largest FMI     =     0.0064
                                   Complete DF     =       4922
DF adjustment:   Small sample      DF:     min     =    4658.43
                                           avg     =    4658.43
Within VCE type:   Linearized              max     =    4658.43

-------------------------------------------------------------------
                   |      Total   Std. Err.     [95% Conf. Interval]
-------------------+-----------------------------------------------
c5_num_employees_0 |   136511.5   6687.389      123401.1      149622
-------------------------------------------------------------------
```

```
mi estimate: svy: ratio ft_0: c6_num_ft_employees_0/c5_num_employees_0

Multiple-imputation estimates      Imputations     =          5
Survey: Ratio estimation           Number of obs   =       4928

Number of strata  =         6      Population size = 73278.441
Number of PSUs    =      4928
                                   Average RVI     =     0.0093
                                   Largest FMI     =     0.0093
                                   Complete DF     =       4922
DF adjustment:   Small sample      DF:     min     =    4413.69
                                           avg     =    4413.69
Within VCE type:   Linearized              max     =    4413.69

         ft_0: c6_num_ft_employees_0/c5_num_employees_0

--------------------------------------------------------------
             |      Ratio   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
        ft_0 |   .6506912   .0182751      .6148629    .6865195
--------------------------------------------------------------
```

```
mi estimate: svy: ratio ft_0: c6_num_ft_employees_0/c5_num_employees_0
,over( homebased_0)

Multiple-imputation estimates      Imputations     =          5
Survey: Ratio estimation           Number of obs   =       4928

Number of strata  =         6      Population size = 73278.441
Number of PSUs    =      4928
                                   Average RVI     =     0.0089
                                   Largest FMI     =     0.0093
                                   Complete DF     =       4922
DF adjustment:   Small sample      DF:     min     =    4410.21
                                           avg     =    4455.66
Within VCE type:   Linearized              max     =    4501.12

         ft_0: c6_num_ft_employees_0/c5_num_employees_0

            0: homebased_0 = 0
            1: homebased_0 = 1

--------------------------------------------------------------
        Over |      Ratio   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
           0 |   .6587501   .0214913      .6166163    .7008838
           1 |   .6117805   .0242893      .5641615    .6593994
--------------------------------------------------------------
```

➢ **Regression**

```
mi estimate, post: svy: reg f24_profitloss_amt_0 pr_gender_owner_0
pr_race_w_0 pr_hisp_origin_owner_0 m_age_owner_0 pr_native_born_owner_0
i.homebased_0
```

```
Multiple-imputation estimates                    Imputations       =          5
Survey: Linear regression                        Number of obs     =       4928

Number of strata  =         6                    Population size   = 73278.441
Number of PSUs    =      4928
                                                 Average RVI       =     0.1515
                                                 Largest FMI       =     0.3027
                                                 Complete DF       =       4922
DF adjustment:   Small sample                    DF:     min       =      51.57
                                                         avg       =    1155.88
                                                         max       =    4640.70
Model F test:       Equal FMI                    F(   6,  704.3)   =       4.39
Within VCE type:   Linearized                    Prob > F          =     0.0002

----------------------------------------------------------------------------------------
  f24_profitloss_amt_0 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-----------------------+----------------------------------------------------------------
     pr_gender_owner_0 |  -1215.076   4182.292    -0.29   0.772    -9430.876    7000.724
           pr_race_w_0 |  -3162.242   4065.806    -0.78   0.437    -11157.76    4833.277
pr_hisp_origin_owner_0 |  -1511.337   5834.695    -0.26   0.797    -13221.82    10199.14
         m_age_owner_0 |  -389.5985   284.4032    -1.37   0.171    -947.5724    168.3754
pr_native_born_owner_0 |   6998.205   5471.778     1.28   0.201    -3746.241    17742.65
         1.homebased_0 |   13849.18   4833.538     2.87   0.004     4373.153    23325.21
                 _cons |   4017.858   12905.33     0.31   0.756    -21324.63    29360.35
----------------------------------------------------------------------------------------
```

```
estimates store mi

mi xeq 0:        svy:   reg f24_profitloss_amt_0  pr_gender_owner_0
pr_race_w_0 pr_hisp_origin_owner_0 m_age_owner_0 pr_native_born_owner_0
i.homebased_0
```

```
m=0 data:
-> svy:    reg  f24_profitloss_amt_0   pr_gender_owner_0  pr_race_w_0  pr_hisp_origin_owner_0
m_age_owner_0 pr_native_born_owner_0 i.homebased_0
(running regress on estimation sample)

Survey: Linear regression

Number of strata   =         6                   Number of obs     =       4130
Number of PSUs     =      4130                   Population size   = 61202.686
                                                 Design df         =       4124
                                                 F(   6,   4119)   =       5.47
                                                 Prob > F          =     0.0000
                                                 R-squared         =     0.0015

----------------------------------------------------------------------------------------
                       |             Linearized
  f24_profitloss_amt_0 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-----------------------+----------------------------------------------------------------
     pr_gender_owner_0 |  -1629.636    4457.54    -0.37   0.715    -10368.82    7109.547
           pr_race_w_0 |  -2732.914     4280.3    -0.64   0.523    -11124.61    5658.783
pr_hisp_origin_owner_0 |  -2681.984   5151.077    -0.52   0.603    -12780.87    7416.906
         m_age_owner_0 |  -439.5909   324.9976    -1.35   0.176    -1076.761    197.5796
pr_native_born_owner_0 |   5401.048   5771.056     0.94   0.349    -5913.335    16715.43
         1.homebased_0 |   14759.29   5691.326     2.59   0.010     3601.217    25917.35
                 _cons |   6741.159   14366.55     0.47   0.639    -21425.03    34907.35
----------------------------------------------------------------------------------------
```

```
estimates store cc
```

```
estimates table cc mi, b se p

----------------------------------------
    Variable |     cc           mi       
-------------+--------------------------
pr_gender_~0 | -1629.6358   -1215.0758  
             |  4457.5399    4182.2918  
             |     0.7147       0.7714  
 pr_race_w_0 | -2732.9137   -3162.2421  
             |     4280.3    4065.8063  
             |     0.5232       0.4367  
pr_hisp_or~0 | -2681.9838   -1511.3371  
             |   5151.077    5834.6951  
             |     0.6026       0.7956  
m_age_owne~0 | -439.59093   -389.59851  
             |  324.99756    284.40319  
             |     0.1763       0.1708  
pr_native_~0 |  5401.0483    6998.2053  
             |  5771.0565    5471.7782  
             |     0.3494       0.2010  
             |
 homebased_0 |
          1  |  14759.286    13849.184  
             |   5691.326    4833.5375  
             |     0.0095       0.0042  
             |
       _cons |  6741.1588     4017.858  
             |   14366.55    12905.333  
             |     0.6389       0.7556  
----------------------------------------
                          legend: b/se/p

mi estimate: svy:       logit f2_owner_eq_invest_01_0
i.g10_gender_owner_01_0  g1b1_hours_owner_01_0 i.g5_hisp_origin_owner_01_0
g4_age_owner_01_0  i.g6_race_white_owner_01_0  tot_assets_0 i.homebased_0

Multiple-imputation estimates                  Imputations       =          5
Survey: Logistic regression                    Number of obs     =       4928

Number of strata  =         6                  Population size   = 73278.441
Number of PSUs    =      4928
                                               Average RVI       =     0.0049
                                               Largest FMI       =     0.0149
                                               Complete DF       =       4922
DF adjustment:   Small sample                  DF:     min       =    3836.47
                                                       avg       =    4574.53
                                                       max       =    4919.50
Model F test:       Equal FMI                  F(   7, 4884.5)   =       3.29
Within VCE type:   Linearized                  Prob > F          =     0.0017

-------------------------------------------------------------------------------------------
  f2_owner_eq_invest_01_0 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
--------------------------+----------------------------------------------------------------
  1.g10_gender_owner_01_0 |  -.0510739   .0921519    -0.55   0.579    -.2317328     .129585
    g1b1_hours_owner_01_0 |   .0078883   .0018613     4.24   0.000     .0042394    .0115372
1.g5_hisp_origin_own~01_0 |  -.2070301   .1671076    -1.24   0.215    -.5346356    .1205753
        g4_age_owner_01_0 |   .0053884   .0039348     1.37   0.171    -.0023261     .013103
1.g6_race_white_owne~01_0 |   .0209044   .1076682     0.19   0.846    -.1901759    .2319847
             tot_assets_0 |   1.66e-08   1.43e-08     1.17   0.243    -1.13e-08    4.46e-08
            1.homebased_0 |    .032439   .0867136     0.37   0.708    -.1375593    .2024372
                    _cons |   .8194928    .225925     3.63   0.000     .3765719    1.262414
-------------------------------------------------------------------------------------------
```

```
mi xeq 0: svy: logit f2_owner_eq_invest_01_0 i.g10_gender_owner_01_0 g1b1_hours_owner_01_0 i.g5_hisp_origin_owner_01_0 g4_age_owner_01_0 i.g6_race_white_owner_01_0 tot_assets_0 i.homebased_0

m=0 data:
-> svy: logit f2_owner_eq_invest_01_0 i.g10_gender_owner_01_0 g1b1_hours_owner_01_0 i.g5_hisp_origin_owner_01_0 g4_age_owner_01_0 i.g6_race_white_owner_01_0 tot_assets_0 i.homebased_0
(running logit on estimation sample)

Survey: Logistic regression

Number of strata   =         6                  Number of obs      =      4791
Number of PSUs     =      4791                  Population size    = 71131.726
                                                Design df          =      4785
                                                F(   7,   4779)    =      3.19
                                                Prob > F           =    0.0023

------------------------------------------------------------------------------------------
                          |             Linearized
  f2_owner_eq_invest_01_0 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
--------------------------+---------------------------------------------------------------
  1.g10_gender_owner_01_0 |  -.0473134   .0933929    -0.51   0.612    -.2304065    .1357796
    g1b1_hours_owner_01_0 |   .0077903   .0018915     4.12   0.000     .0040822    .0114985
1.g5_hisp_origin_own~01_0 |  -.2241187   .1685798    -1.33   0.184    -.5546126    .1063752
        g4_age_owner_01_0 |   .0055839   .0039507     1.41   0.158    -.0021614    .0133291
1.g6_race_white_owne~01_0 |   .0370744   .1090216     0.34   0.734    -.1766581     .250807
             tot_assets_0 |   1.39e-08   1.20e-08     1.16   0.247    -9.61e-09    3.74e-08
            1.homebased_0 |   .0379414   .0878312     0.43   0.666    -.1342481     .210131
                    _cons |   .7956481   .2282219     3.49   0.000     .3482282    1.243068
```

## ➢ Survival Analysis

```
mi stset Duration [pweight=cswgt_final_0] , failure(event==1)

     failure event:  event == 1
obs. time interval:  (0, Duration]
 exit on or before:  failure
            weight:  [pweight=cswgt_final_0]

------------------------------------------------------------------------------
     4928  total obs.
        0  exclusions
------------------------------------------------------------------------------
     4928  obs. remaining, representing
     2190  failures in single record/single failure data
    27568  total analysis time at risk, at risk from t =         0
                             earliest observed entry t =         0
                                  last observed exit t =         8



sts list if _mi_m==1,  survival

         failure _d:  event == 1
   analysis time _t:  Duration
             weight:  [pweight=cswgt_final_0]

           Beg.          Net          Survivor
  Time    Total   Fail   Lost           Function
------------------------------------------------
     1  73278.4   4697   1902             0.9359
     2  66679.3   5982    965             0.8519
     3  59732.1   5347   1259             0.7757
     4  53125.8   5759   1375             0.6916
     5  45991.6   4350   1113             0.6262
     6    40528   3900   1525             0.5659
     7  35103.3   3985   2615             0.5017
     8  28503.2      0  2.9e+04           0.5017
------------------------------------------------

 sts list if _mi_m==0,  survival

         failure _d:  event == 1
   analysis time _t:  Duration
             weight:  [pweight=cswgt_final_0]

           Beg.          Net          Survivor
  Time    Total   Fail   Lost           Function
------------------------------------------------
     1  73278.4   4697   1902             0.9359
     2  66679.3   5982    965             0.8519
     3  59732.1   5347   1259             0.7757
     4  53125.8   5759   1375             0.6916
     5  45991.6   4350   1113             0.6262
     6    40528   3900   1525             0.5659
     7  35103.3   3985   2615             0.5017
     8  28503.2      0  2.9e+04           0.5017
------------------------------------------------
```

```
mi estimate: svy: stcox pr_gender_owner_0 pr_race_w_0 pr_hisp_origin_owner_0 m_age_owner_0 pr_native_born_owner_0 tot_assets_0 homebased_0


Multiple-imputation estimates                   Imputations       =          5
Survey: Cox regression                          Number of obs     =       4928

Number of strata  =         6                   Population size   = 73278.441
Number of PSUs    =      4928
                                                Average RVI       =     0.0025
                                                Largest FMI       =     0.0066
                                                Complete DF       =       4922
DF adjustment:   Small sample                   DF:     min       =    4639.83
                                                        avg       =    4822.04
                                                        max       =    4919.89
Model F test:       Equal FMI                   F(   7, 4912.8)   =       0.67
Within VCE type:   Linearized                   Prob > F          =     0.7000

------------------------------------------------------------------------------------------
                    _t |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-----------------------+------------------------------------------------------------------
     pr_gender_owner_0 |  -.0482578   .0565372    -0.85   0.393     -.159096    .0625803
           pr_race_w_0 |  -.0847179    .066121    -1.28   0.200    -.2143457    .0449099
pr_hisp_origin_owner_0 |   .0852772   .1017117     0.84   0.402    -.1141232    .2846776
         m_age_owner_0 |   .0003172   .0023025     0.14   0.890    -.0041969    .0048313
pr_native_born_owner_0 |   .0345032   .0863364     0.40   0.689    -.1347551    .2037614
          tot_assets_0 |  -1.38e-09   6.55e-09    -0.21   0.833    -1.42e-08    1.15e-08
           homebased_0 |  -.0470351   .0466294    -1.01   0.313    -.1384496    .0443794
------------------------------------------------------------------------------------------
```

### 6.1.5 Examples – KFS long MI style long

```
use KFS8_Longitudinal_Long_MI_Long,clear
(KFS8 Longitudinal in long format & Multiply Imputed Data in long format)

mi xtset  mprid year
       panel variable:  mprid (unbalanced)
        time variable:  year, 2004 to 2011
                delta:  1 unit

mi xeq 0:       xtdescribe

   mprid:  10000016, 10000090, ..., 10324611                 n =       3140
    year:  2004, 2005, ..., 2011                             T =          8
           Delta(year) = 1 unit
           Span(year)  = 8 periods
           (mprid*year uniquely identifies each observation)

Distribution of T_i:   min      5%     25%       50%       75%     95%     max
                         1       1       3         8         8       8       8

 Freq.  Percent    Cum. |  Pattern
 ---------------------------+----------
    1630     51.91   51.91 |  11111111
     303      9.65   61.56 |  1.......
     283      9.01   70.57 |  11......
     238      7.58   78.15 |  1111....
     224      7.13   85.29 |  111.....
     164      5.22   90.51 |  11111...
     153      4.87   95.38 |  111111..
     145      4.62  100.00 |  1111111.
 ---------------------------+----------
    3140    100.00         |  XXXXXXXX
```

```
mi svyset  [pweight=wgt_7_long],  strata(sampleinfo_samplestrata)

      pweight: wgt_7_long
          VCE: linearized
  Single unit: missing
     Strata 1: sampleinfo_samplestrata
         SU 1: <observations>
        FPC 1: <zero>
```

## ➢ Regression

```
mi estimate: svy:       reg f24_profitloss_amt  tot_assets
pr_gender_owner pr_race_w pr_hisp_origin_owner  m_age_owner
       pr_native_born_owner    i.homebased i.d_education_owner     i.year

Multiple-imputation estimates                 Imputations       =          5
Survey: Linear regression                     Number of obs     =      18286

Number of strata  =         6                 Population size   = 408495.43
Number of PSUs    =     18286
                                              Average RVI       =     0.0253
                                              Largest FMI       =     0.0252
                                              Complete DF       =      18280
DF adjustment:   Small sample                 DF:     min       =    4725.50
                                                      avg       =   14876.49
                                                      max       =   18277.99
Model F test:       Equal FMI                 F(  14,15295.8)   =       4.17
Within VCE type:   Linearized                 Prob > F          =     0.0000

--------------------------------------------------------------------------------------
  f24_profitloss_amt |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
---------------------+----------------------------------------------------------------
          tot_assets |   .0189758   .0169062     1.12   0.262    -.0141619    .0521135
     pr_gender_owner |  -35698.17   50521.82    -0.71   0.480    -134725.7    63329.38
           pr_race_w |   47286.26    46706.4     1.01   0.311    -44262.72    138835.2
pr_hisp_origin_owner |  -60596.78   35829.83    -1.69   0.091      -130830    9636.444
         m_age_owner |  -1723.759   1027.937    -1.68   0.094    -3738.612    291.0948
pr_native_born_owner |   69215.13    50117.4     1.38   0.167    -29019.69    167449.9
          1.homebased |  -94632.89   88725.17    -1.07   0.286    -268542.6    79276.77
 1.d_education_owner |      71605   70967.41     1.01   0.313    -67497.79    210707.8
                     |
                year |
               2005  |   9401.701   10058.23     0.93   0.350    -10314.16    29117.56
               2006  |   40747.76   13168.15     3.09   0.002     14935.55    66559.96
               2007  |    46025.7   16883.61     2.73   0.006     12932.19     79119.2
               2008  |   34029.68   10851.88     3.14   0.002     12757.12    55302.23
               2009  |   26803.69   11328.97     2.37   0.018     4593.622    49013.76
               2010  |   36590.08   13245.08     2.76   0.006     10627.52    62552.63
               2011  |   554536.9   517519.5     1.07   0.284    -459849.9     1568924
                     |
               _cons |    1446.26    57628.8     0.03   0.980    -111511.7    114404.2
--------------------------------------------------------------------------------------
```

```
mi estimate: svy: reg  f24_profitloss_amt  tot_assets     pr_gender_owner
pr_race_w     pr_hisp_origin_owner     m_age_owner     pr_native_born_owner
i.homebased i.d_education_owner i.year , vce(cluster mprid)

  option vce() of regress is not allowed with the svy prefix
an error occurred when mi estimate executed svy:regress on m=1
```

```
mi estimate: xtgee f24_profitloss_amt tot_assets pr_gender_owner
pr_race_w pr_hisp_origin_owner m_age_owner pr_native_born_owner
i.homebased i.d_education_owner i.year i.sampleinfo_samplestrata
[pweight=wgt_7_long]
```

```
Multiple-imputation estimates                  Imputations       =          5
GEE population-averaged model                  Number of obs     =      18286

Group variable:                      mprid     Number of groups  =       3140
Link:                             identity     Obs per group: min =          1
Family:                           Gaussian                    avg =        5.6
Correlation:                  exchangeable                    max =          8
Scale parameter:                        x2
                                               Average RVI       =     0.0284
                                               Largest FMI       =     0.0304
                                               DF:     min       =    4459.56
                                                       avg       =   1.73e+12
DF adjustment:                Large sample             max       =   3.64e+13
                                               F(  18,      .)   =          .
Within VCE type:                    Robust     Prob > F          =          .

                                         (Within VCE adjusted for clustering on mprid)
------------------------------------------------------------------------------------------
     f24_profitloss_amt |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
------------------------+-----------------------------------------------------------------
             tot_assets |   .0190176   .0172165     1.10   0.269    -.0147262    .0527613
        pr_gender_owner |  -130814.2   135291.8    -0.97   0.334    -395981.2    134352.9
              pr_race_w |   43322.97   42586.63     1.02   0.309    -40145.34    126791.3
   pr_hisp_origin_owner |  -56046.05   40340.62    -1.39   0.165      -135115    23022.87
            m_age_owner |  -1574.015   1352.307    -1.16   0.244    -4224.489    1076.459
   pr_native_born_owner |   77475.11   56675.55     1.37   0.172    -33606.93    188557.1
             1.homebased |  -93161.77   81828.98    -1.14   0.255    -253543.6    67220.07
    1.d_education_owner |   70558.25   78073.33     0.90   0.366    -82462.66    223579.2
                        |
                   year |
                  2005  |     8349.6   7817.362     1.07   0.286    -6973.857    23673.06
                  2006  |   38754.95   10313.37     3.76   0.000     18538.23    58971.66
                  2007  |   44068.51   14812.11     2.98   0.003     15037.26    73099.76
                  2008  |   32099.43   10403.34     3.09   0.002     11707.15     52491.7
                  2009  |   24326.13   10249.35     2.37   0.018     4232.326    44419.94
                  2010  |   34260.88   12593.69     2.72   0.007     9576.631    58945.13
                  2011  |   551415.6   516541.1     1.07   0.286    -460986.3     1563817
                        |
sampleinfo_samplestrata |
                   102  |   22023.43   107438.8     0.20   0.838    -188553.3    232600.1
                   201  |  -3334.977   83221.16    -0.04   0.968    -166445.7    159775.7
                   202  |   152655.7     119241     1.28   0.200    -81052.49    386363.8
                   301  |  -21942.74   81839.49    -0.27   0.789    -182345.2    138459.8
                   302  |   129366.9   162482.7     0.80   0.426    -189093.5    447827.3
                        |
                  _cons |  -41546.75   138386.3    -0.30   0.764      -312779    229685.5
------------------------------------------------------------------------------------------
```

```
mi estimate: xtgee f24_profitloss_amt tot_assets pr_gender_owner
pr_race_w pr_hisp_origin_owner m_age_owner pr_native_born_owner
i.homebased i.d_education_owner i.year i.sampleinfo_samplestrata
[pweight=wgt_7_long] ,vce(robust)

Multiple-imputation estimates                   Imputations       =          5
GEE population-averaged model                   Number of obs     =      18286

Group variable:                      mprid      Number of groups  =       3140
Link:                             identity      Obs per group: min =          1
Family:                           Gaussian                     avg =        5.6
Correlation:                  exchangeable                     max =          8
Scale parameter:                        x2
                                                Average RVI       =     0.0284
                                                Largest FMI       =     0.0304
                                                DF:     min       =    4459.56
                                                        avg       =   1.73e+12
DF adjustment:                Large sample              max       =   3.64e+13
                                                F(  18,      .)   =          .
Within VCE type:                    Robust      Prob > F          =          .

                                             (Within VCE adjusted for clustering on mprid)
------------------------------------------------------------------------------------------
     f24_profitloss_amt |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
------------------------+-----------------------------------------------------------------
             tot_assets |   .0190176   .0172165     1.10   0.269    -.0147262    .0527613
        pr_gender_owner |  -130814.2   135291.8    -0.97   0.334    -395981.2    134352.9
              pr_race_w |   43322.97   42586.63     1.02   0.309    -40145.34    126791.3
   pr_hisp_origin_owner |  -56046.05   40340.62    -1.39   0.165      -135115    23022.87
            m_age_owner |  -1574.015   1352.307    -1.16   0.244    -4224.489    1076.459
   pr_native_born_owner |   77475.11   56675.55     1.37   0.172    -33606.93    188557.1
           1.homebased |  -93161.77   81828.98    -1.14   0.255    -253543.6    67220.07
    1.d_education_owner |   70558.25   78073.33     0.90   0.366    -82462.66    223579.2
                        |
                   year |
                  2005  |     8349.6   7817.362     1.07   0.286    -6973.857    23673.06
                  2006  |   38754.95   10313.37     3.76   0.000     18538.23    58971.66
                  2007  |   44068.51   14812.11     2.98   0.003     15037.26    73099.76
                  2008  |   32099.43   10403.34     3.09   0.002     11707.15     52491.7
                  2009  |   24326.13   10249.35     2.37   0.018     4232.326    44419.94
                  2010  |   34260.88   12593.69     2.72   0.007     9576.631    58945.13
                  2011  |   551415.6   516541.1     1.07   0.286    -460986.3     1563817
                        |
sampleinfo_samplestrata |
                   102  |   22023.43   107438.8     0.20   0.838    -188553.3    232600.1
                   201  |  -3334.977   83221.16    -0.04   0.968    -166445.7    159775.7
                   202  |   152655.7     119241     1.28   0.200    -81052.49    386363.8
                   301  |  -21942.74   81839.49    -0.27   0.789    -182345.2    138459.8
                   302  |   129366.9   162482.7     0.80   0.426    -189093.5    447827.3
                        |
                  _cons |  -41546.75   138386.3    -0.30   0.764      -312779    229685.5
------------------------------------------------------------------------------------------
```

```
mi   estimate:   xtgee   f24_profitloss_amt   tot_assets     pr_gender_owner
pr_race_w     pr_hisp_origin_owner     m_age_owner     pr_native_born_owner
i.homebased     i.d_education_owner     i.year   i.sampleinfo_samplestrata
[pweight=wgt_7_long]     ,vce(robust) corr(ar2)

Multiple-imputation estimates                   Imputations       =          5
GEE population-averaged model                   Number of obs     =      17417

Group and time vars:            mprid year      Number of groups  =       2554
Link:                             identity      Obs per group: min =         3
Family:                           Gaussian                     avg =       6.7
Correlation:                         AR(2)                     max =         8
Scale parameter:                        x2
                                                Average RVI       =     0.0293
                                                Largest FMI       =     0.0311
                                                DF:     min       =    4250.25
                                                        avg       =   1.49e+12
DF adjustment:                Large sample              max       =   3.14e+13
                                                F(  18,      .)   =          .
Within VCE type:                    Robust      Prob > F          =          .

                                         (Within VCE adjusted for clustering on mprid)
------------------------------------------------------------------------------------------
     f24_profitloss_amt |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
------------------------+-----------------------------------------------------------------
             tot_assets |   .0190593   .0172152     1.11   0.268    -.0146818    .0528005
        pr_gender_owner |  -140363.8   144985.1    -0.97   0.333    -424529.4    143801.8
              pr_race_w |   47302.65   44925.77     1.05   0.292    -40750.27    135355.6
   pr_hisp_origin_owner |   -67546.9   47845.61    -1.41   0.158      -161325    26231.19
            m_age_owner |  -1651.657    1440.89    -1.15   0.252    -4475.749    1172.435
   pr_native_born_owner |   81033.46   61549.75     1.32   0.188    -39601.82    201668.7
             1.homebased |  -104555.9   88726.17    -1.18   0.239      -278456    69344.19
    1.d_education_owner |   78165.25   84746.76     0.92   0.356    -87935.35    244265.8
                        |
                   year |
                  2005  |   14693.31   7216.369     2.04   0.042     548.1031    28838.51
                  2006  |    37866.2   10034.37     3.77   0.000     18196.43    57535.96
                  2007  |   43337.13   15158.19     2.86   0.004     13627.61    73046.65
                  2008  |   31323.24   10279.84     3.05   0.002     11173.06    51473.43
                  2009  |   23462.58   9987.374     2.35   0.019     3882.112    43043.05
                  2010  |   33475.91   12369.95     2.71   0.007     9230.315    57721.51
                  2011  |   550612.5   516669.3     1.07   0.287    -462040.7     1563266
                        |
sampleinfo_samplestrata |
                   102  |   28948.57   114731.8     0.25   0.801    -195922.2    253819.3
                   201  |  -1745.666   87501.31    -0.02   0.984    -173245.3      169754
                   202  |     165904   127408.5     1.30   0.193    -83812.11    415620.2
                   301  |  -18729.67   86258.42    -0.22   0.828    -187793.1    150333.8
                   302  |   140351.3   174571.1     0.80   0.421    -201801.8    482504.3
                        |
                  _cons |  -44675.91   147857.3    -0.30   0.763    -334470.9    245119.1
------------------------------------------------------------------------------------------
```

```
mi estimate: xtreg f24_profitloss_amt tot_assets pr_gender_owner pr_race_w pr_hisp_origin_owner m_age_owner pr_native_born_owner i.homebased i.d_education_owner i.year [pweight=wgt_7_long] , fe i(mprid)
```

```
Multiple-imputation estimates                   Imputations       =          5
Fixed-effects (within) regression               Number of obs     =      18286

Group variable: mprid                           Number of groups  =       3140
                                                Obs per group: min =         1
                                                               avg =       5.8
                                                               max =         8

                                                Average RVI       =     3.5153
                                                Largest FMI       =     0.3841
                                                Complete DF       =       3139
                                                DF:     min       =      32.55
                                                        avg       =    2515.01
DF adjustment:   Small sample                           max       =    3137.00
                                                F(  13,      .)   =          .
Within VCE type:       Robust                   Prob > F          =          .

                                     (Within VCE adjusted for 3140 clusters in mprid)
-------------------------------------------------------------------------------------
  f24_profitloss_amt |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
---------------------+---------------------------------------------------------------
          tot_assets |   .0206624   .0178933     1.15   0.248    -.0144213    .0557461
     pr_gender_owner |  -324047.9   261310.7    -1.24   0.215    -836405.6    188309.8
           pr_race_w |   24317.75   79214.76     0.31   0.759    -131103.4    179738.9
pr_hisp_origin_owner |    69976.4   146461.4     0.48   0.636    -228159.1    368111.9
         m_age_owner |   17675.14   17416.05     1.01   0.310    -16472.86    51823.14
pr_native_born_owner |  -117354.7   135703.2    -0.86   0.387    -383588.8    148879.5
          1.homebased |  -35335.12   21210.76    -1.67   0.097    -77060.77    6390.533
 1.d_education_owner |   22384.41   68396.98     0.33   0.743    -111723.1    156491.9
                     |
                year |
               2005  |  -10425.74   18343.58    -0.57   0.570    -46392.59    25541.11
               2006  |  -3225.158   35904.37    -0.09   0.928     -73623.7    67173.38
               2007  |  -15172.64   52031.09    -0.29   0.771    -117191.1    86845.81
               2008  |   -49400.2    69060.9    -0.72   0.474    -184809.3    86008.95
               2009  |  -76011.94   84237.11    -0.90   0.367    -241177.4    89153.51
               2010  |  -87995.17   100528.7    -0.88   0.381      -285104    109113.6
               2011  |   400228.2     412385     0.97   0.332    -408343.7     1208800
                     |
               _cons |  -486096.4     726080    -0.67   0.503     -1909749    937556.2
---------------------+---------------------------------------------------------------
             sigma_u |  1223907.7
             sigma_e |  4887936.4
                 rho |  .05899794   (fraction of variance due to u_i)
-------------------------------------------------------------------------------------
Note: sigma_u and sigma_e are combined in the original metric.
```

```
mi estimate: xtreg f24_profitloss_amt tot_assets pr_gender_owner
pr_race_w pr_hisp_origin_owner m_age_owner pr_native_born_owner
i.homebased i.d_education_owner i.year [pweight=wgt_7_long] ,
fe i(mprid) vce(robust)
```

```
Multiple-imputation estimates                   Imputations       =          5
Fixed-effects (within) regression               Number of obs     =      18286

Group variable: mprid                           Number of groups  =       3140
                                                Obs per group: min =         1
                                                               avg =       5.8
                                                               max =         8

                                                Average RVI       =     3.5153
                                                Largest FMI       =     0.3841
                                                Complete DF       =       3139
                                                DF:     min       =      32.55
                                                        avg       =    2515.01
DF adjustment:   Small sample                           max       =    3137.00
                                                F(  13,      .)   =          .
Within VCE type:       Robust                   Prob > F          =          .

                                     (Within VCE adjusted for 3140 clusters in mprid)
-------------------------------------------------------------------------------------
  f24_profitloss_amt |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
---------------------+---------------------------------------------------------------
          tot_assets |   .0206624   .0178933     1.15   0.248    -.0144213    .0557461
     pr_gender_owner |  -324047.9   261310.7    -1.24   0.215    -836405.6    188309.8
           pr_race_w |   24317.75   79214.76     0.31   0.759    -131103.4    179738.9
pr_hisp_origin_owner |    69976.4   146461.4     0.48   0.636    -228159.1    368111.9
         m_age_owner |   17675.14   17416.05     1.01   0.310    -16472.86    51823.14
pr_native_born_owner |  -117354.7   135703.2    -0.86   0.387    -383588.8    148879.5
         1.homebased |  -35335.12   21210.76    -1.67   0.097    -77060.77    6390.533
 1.d_education_owner |   22384.41   68396.98     0.33   0.743    -111723.1    156491.9
                     |
                year |
               2005  |  -10425.74   18343.58    -0.57   0.570    -46392.59    25541.11
               2006  |  -3225.158   35904.37    -0.09   0.928     -73623.7    67173.38
               2007  |  -15172.64   52031.09    -0.29   0.771    -117191.1    86845.81
               2008  |   -49400.2    69060.9    -0.72   0.474    -184809.3    86008.95
               2009  |  -76011.94   84237.11    -0.90   0.367    -241177.4    89153.51
               2010  |  -87995.17   100528.7    -0.88   0.381      -285104    109113.6
               2011  |   400228.2     412385     0.97   0.332    -408343.7     1208800
                     |
               _cons |  -486096.4     726080    -0.67   0.503     -1909749    937556.2
---------------------+---------------------------------------------------------------
             sigma_u |  1223907.7
             sigma_e |  4887936.4
                 rho |  .05899794   (fraction of variance due to u_i)
-------------------------------------------------------------------------------------
Note: sigma_u and sigma_e are combined in the original metric.
```

References

Allison, P.D. (2001) Missing Data. Thousand Oaks, CA:Sage

Little, R.J. and Rubin, D.B. (2000) Statistical Analysis with Missing Data, New York: Wiley.

Raghunathan, T., Lepkowski, J., van Hoewyk, J. and Solenberger, P. (2001) 'A Multivariate Technique for Multiply Imputing Missing Values Using a Sequence of Regression Models, Survey Methodology, 27(1), pp. 85-95.

Rubin, D. B. (1987), Multiple Imputation for Nonresponse in Surveys, New York: John Wiley & Sons, Inc.

Schafer, J. (1999). Multiple imputation: A primer. Statistical Methods in Medical Research, 8(1), 3–15.

Schafer, J. L. (1997), Analysis of Incomplete Multivariate Data, New York: Chapman and Hall

## Appendix A: Percentage of soft missing values

| Follow-Up | _0 | _1 | _2 | _3 | _4 | _5 | _6 | _7 |
|---|---|---|---|---|---|---|---|---|
| b1_bus_start | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| c1z2_legal_status | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| c2_owners | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| c3a_owner_operators | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| c4_numowners_confirm | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| c5_num_employees | 2.13 | 1.15 | 1.09 | 0.86 | 0.15 | 0.42 | 0.24 | 0.35 |
| c6_num_ft_employees | 2.41 | 3.63 | 3.60 | 3.81 | 3.03 | 2.57 | 2.40 | 2.74 |
| c7_num_pt_employees | 2.86 | 4.80 | 4.63 | 5.49 | 3.84 | 4.86 | 3.95 | 3.99 |
| c8_primary_loc | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| c9_loc_change_reason | | 0.00 | 0.82 | 18.71 | 0.00 | 0.00 | 4.55 | 0.00 |
| c10_morelocations | | | 0.12 | 0.17 | 0.15 | 0.29 | 0.19 | 0.35 |
| c11_num_locations | | | 0.27 | 0.21 | 0.23 | 0.50 | 0.24 | 0.40 |
| c12a_sba | | | | | 0.42 | | | |
| c12b_fed_gov | | | | | 0.65 | | | |
| c12c_statelocal_gov | | | | | 0.58 | | | |
| c12d_non_profit | | | | | 0.50 | | | |
| c12e_college_univ | | | | | 0.58 | | | |
| c12f_chamber_of_comm | | | | | 0.54 | | | |
| c12g_for_profit_org | | | | | 0.50 | | | |
| c12h_other | | | | | 1.69 | | | |
| d1a_provide_service | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| d1b_provide_product | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| d1_a_new_product | | | | | | 0.12 | 0.19 | 0.10 |
| d1d_new_processes | | | | | | 0.33 | 0.19 | 0.20 |
| d1_b_new_to_market | | | | | | 0.75 | 1.29 | 0.72 |
| d1c_a_regional | | | | | | 11.82 | 2.38 | 1.59 |
| d1c_b_national | | | | | | 8.18 | 2.38 | 1.59 |
| d1c_c_international | | | | | | 13.03 | 2.72 | 1.59 |
| d2_comp_advantage | 1.42 | 0.60 | 0.44 | 0.69 | 0.35 | 0.21 | 0.19 | 0.10 |
| d2a_compadv_comp_reason | | | | 5.44 | 1.00 | 0.93 | 0.81 | 0.91 |
| d2a_compadv_govlab_reason | | | | 5.55 | 1.20 | 0.84 | 1.11 | 1.14 |
| d2a_compadv_patents_reason | | | | 5.44 | 1.46 | 1.09 | 1.11 | 1.14 |
| d2a_compadv_univ_reason | | | | 5.44 | 1.13 | 0.76 | 1.01 | 0.80 |
| d2b_compadv_comp_strength | | | | 18.30 | 3.55 | 2.93 | 3.10 | 4.23 |
| d2b_compadv_govlab_strength | | | | 66.89 | 30.51 | 19.61 | 22.92 | 26.19 |
| d2b_compadv_patents_strength | | | | 38.96 | 15.49 | 13.79 | 12.77 | 12.36 |
| d2b_compadv_univ_strength | | | | 44.09 | 16.19 | 8.74 | 13.64 | 10.53 |
| d2c_compadv_cost_reason | | | | | | 1.94 | | |
| d2c_compadv_design_reason | | | | | | 1.94 | | |
| d2c_compadv_expertise_reason | | | | | | 1.68 | | |
| d2c_compadv_marketing_reason | | | | | | 2.19 | | |
| d2c_compadv_price_reason | | | | | | 1.43 | | |

| Follow-Up | _0 | _1 | _2 | _3 | _4 | _5 | _6 | _7 |
|---|---|---|---|---|---|---|---|---|
| d2c_compadv_reputation_reason | | | | | | 1.94 | | |
| d2c_compadv_speed_reason | | | | | | 2.19 | | |
| d2d_compadv_cost_strength | | | | | | 4.98 | | |
| d2d_compadv_design_strength | | | | | | 3.98 | | |
| d2d_compadv_expertise_strength | | | | | | 3.14 | | |
| d2d_compadv_marketing_strength | | | | | | 6.32 | | |
| d2d_compadv_price_strength | | | | | | 8.59 | | |
| d2d_compadv_reput_strength | | | | | | 3.72 | | |
| d2d_compadv_speed_strength | | | | | | 4.82 | | |
| d3_a_have_patent | 0.63 | 1.48 | 1.45 | 1.10 | 0.50 | 0.37 | 0.56 | 0.40 |
| d3_a_num_patent | 0.79 | 1.93 | 1.77 | 1.37 | 0.96 | 0.83 | 0.99 | 0.90 |
| d3_b_have_copyright | 0.71 | 1.40 | 1.39 | 1.23 | 0.58 | 0.66 | 0.85 | 0.95 |
| d3_b_num_copyright | 1.18 | 3.05 | 2.65 | 2.40 | 2.38 | 2.91 | 2.63 | 2.29 |
| d3_c_have_trademark | 0.53 | 1.03 | 0.94 | 1.06 | 0.46 | 0.71 | 0.52 | 0.55 |
| d3_c_num_trademark | 1.22 | 3.03 | 3.24 | 2.64 | 3.57 | 3.36 | 2.45 | 2.24 |
| d4_a_lic_out_patent | 0.69 | 1.50 | 1.45 | 1.10 | 0.50 | 0.37 | 0.56 | 0.45 |
| d4_b_lic_out_copyright | 0.77 | 1.43 | 1.42 | 1.27 | 0.65 | 0.66 | 0.85 | 0.95 |
| d4_c_lic_out_trademark | 0.63 | 1.03 | 1.00 | 1.10 | 0.54 | 0.75 | 0.56 | 0.55 |
| d5_a_lic_in_patent | 0.53 | 1.08 | 0.80 | 0.79 | 0.77 | 0.96 | 0.80 | 0.80 |
| d5_b_lic_in_copyright | 0.61 | 0.78 | 0.74 | 0.58 | 0.61 | 0.46 | 0.42 | 0.60 |
| d5_c_lic_in_trademark | 0.67 | 0.83 | 0.86 | 0.69 | 0.81 | 1.04 | 1.08 | 0.90 |
| d5a_founded_newprod | | | | | | 0.21 | | |
| d5b_a_personaluse | | | | | | 1.15 | | |
| d5b_b_previousjob | | | | | | 1.47 | | |
| d5b_c_startingbus | | | | | | 1.31 | | |
| d6_have_sales | 0.18 | 0.13 | 0.24 | 0.00 | 0.12 | 0.08 | 0.14 | 0.05 |
| d7_perc_sales_bus | 0.51 | 0.25 | 0.35 | 0.58 | 0.27 | 0.25 | 0.28 | 0.05 |
| d7_perc_sales_govt | 0.51 | 0.38 | 0.35 | 0.72 | 0.27 | 0.29 | 0.24 | 0.05 |
| d7_perc_sales_indiv | 0.49 | 0.28 | 0.32 | 0.10 | 0.27 | 0.25 | 0.24 | 0.05 |
| d8_customer_locations | | | | 0.41 | 0.21 | 0.14 | 0.21 | 0.05 |
| d8a_international_sales | | | | 0.48 | 0.33 | 0.18 | 0.42 | 0.05 |
| d8b_perc_international_sales | | | | 3.12 | 2.15 | 1.04 | 2.96 | 0.62 |
| d9_internet_sales | | | | 0.74 | 0.33 | 0.28 | 0.42 | 0.11 |
| d9a_perc_internet_sales | | | | 3.28 | 1.25 | 1.08 | 1.86 | 0.43 |
| e1_a_num_human_res | 1.39 | 0.86 | 0.83 | 0.65 | 0.68 | 0.70 | 0.74 | 0.71 |
| e1_b_num_sales | 1.39 | 0.79 | 0.75 | 0.60 | 0.68 | 0.70 | 0.74 | 0.86 |
| e1_c_num_exec_admin | 1.36 | 0.82 | 0.79 | 0.60 | 0.74 | 0.70 | 0.89 | 0.79 |
| e1_d_num_resdev | 1.36 | 0.86 | 0.87 | 0.60 | 0.91 | 0.70 | 0.89 | 0.79 |
| e1_e_num_prod_manu | 1.42 | 0.93 | 0.79 | 0.60 | 0.85 | 0.76 | 0.89 | 0.79 |
| e1_f_num_gen_admin | 1.36 | 0.75 | 0.75 | 0.60 | 0.85 | 0.82 | 0.89 | 0.86 |
| e1_g_num_fin_admin | 1.36 | 0.82 | 0.87 | 0.60 | 0.85 | 0.70 | 0.89 | 0.86 |
| e1_h_num_other | 1.85 | 1.20 | 1.03 | 0.65 | 1.47 | 1.20 | 1.19 | 1.02 |
| e2a_ft_emp_bonus_plan | 1.20 | 0.31 | 0.28 | 0.30 | 0.57 | 0.70 | 0.96 | 1.18 |

| Follow-Up | _0 | _1 | _2 | _3 | _4 | _5 | _6 | _7 |
|---|---|---|---|---|---|---|---|---|
| e2a_ft_emp_flex_time | 1.26 | 0.31 | 0.28 | 0.20 | 0.57 | 0.63 | 0.59 | 0.63 |
| e2a_ft_emp_hlth_plan | 1.33 | 0.14 | 0.12 | 0.20 | 0.51 | 0.44 | 0.37 | 0.63 |
| e2a_ft_emp_other | 1.13 | 0.41 | 0.20 | 0.40 | 0.68 | 1.65 | 1.85 | 2.12 |
| e2a_ft_emp_paid_sick | 1.36 | 0.21 | 0.12 | 0.40 | 0.85 | 0.57 | 0.67 | 0.71 |
| e2a_ft_emp_paid_vaca | 1.33 | 0.21 | 0.36 | 0.30 | 0.57 | 0.63 | 0.52 | 0.79 |
| e2a_ft_emp_retire_plan | 1.20 | 0.10 | 0.16 | 0.25 | 0.57 | 0.51 | 0.44 | 0.63 |
| e2a_ft_emp_stock_own | 1.30 | 0.14 | 0.32 | 0.30 | 0.62 | 1.14 | 1.19 | 1.89 |
| e2a_ft_emp_tuit_reim | 1.39 | 0.24 | 0.16 | 0.35 | 0.68 | 0.82 | 0.89 | 0.63 |
| e2b_pt_emp_bonus_plan | 14.09 | 14.32 | 13.05 | 15.41 | 10.14 | 13.61 | 11.38 | 12.93 |
| e2b_pt_emp_flex_time | 14.09 | 14.25 | 13.28 | 15.23 | 10.14 | 13.03 | 11.38 | 11.82 |
| e2b_pt_emp_hlth_plan | 14.00 | 14.25 | 12.98 | 14.97 | 10.04 | 13.15 | 11.13 | 11.96 |
| e2b_pt_emp_other | 13.90 | 14.25 | 13.13 | 15.23 | 10.04 | 13.84 | 11.88 | 14.05 |
| e2b_pt_emp_paid_sick | 14.09 | 14.38 | 13.05 | 15.32 | 10.14 | 13.03 | 11.25 | 12.38 |
| e2b_pt_emp_paid_vaca | 14.09 | 14.32 | 13.05 | 15.41 | 9.94 | 13.03 | 11.13 | 12.10 |
| e2b_pt_emp_retire_plan | 14.09 | 14.18 | 12.82 | 15.15 | 10.04 | 13.15 | 11.00 | 11.96 |
| e2b_pt_emp_stock_own | 14.18 | 14.11 | 13.21 | 15.15 | 10.14 | 13.61 | 11.38 | 12.52 |
| e2b_pt_emp_tuit_reim | 14.09 | 14.25 | 13.13 | 14.97 | 9.94 | 13.61 | 11.88 | 12.24 |
| owner_active | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| f2_owner_eq_invest | 0.76 | 0.84 | 0.79 | 0.33 | 0.35 | 0.32 | 0.47 | 0.32 |
| f2_owner_amt_eq_invest | 12.12 | 3.69 | 1.88 | 1.22 | 0.99 | 0.94 | 0.74 | 0.82 |
| tot_equity_owner_r | 3.96 | 1.59 | 1.27 | 0.48 | 0.46 | 0.41 | 0.61 | 0.61 |
| f2_ownr_amt_eqinvest_allyrs | 12.12 | 7.51 | 5.30 | 3.17 | 2.95 | 3.36 | 2.64 | 3.14 |
| tot_equity_allyrs_owner_r | 3.96 | 2.19 | 1.74 | 0.69 | 0.64 | 1.20 | 0.61 | 1.04 |
| f2_owner_perc_own | 1.11 | 1.41 | 0.55 | 0.24 | 0.48 | 0.61 | 0.17 | 0.11 |
| f6_perc_owned_owner | 1.02 | 1.27 | 0.50 | 0.21 | 0.43 | 0.53 | 0.15 | 0.10 |
| g1a_emp_owner | 0.63 | 1.44 | 0.73 | 0.29 | 0.35 | 0.47 | 0.10 | 0.18 |
| g1b1_hours_owner | 2.02 | 3.51 | 1.70 | 0.91 | 1.10 | 1.20 | 0.91 | 0.46 |
| total_hours_owner_r | 0.57 | 1.53 | 0.61 | 0.12 | 0.40 | 0.29 | 0.07 | 0.18 |
| g2_work_exp_owner | 0.33 | 0.48 | 0.34 | 0.21 | 0.19 | 0.18 | 0.00 | 0.18 |
| g3a_oth_bus_owner | 0.63 | 0.63 | 0.42 | 0.26 | 0.21 | 0.20 | 0.03 | 0.29 |
| g3b_bus_same_ind_owner | 1.46 | 1.49 | 1.01 | 0.62 | 0.52 | 0.49 | 0.08 | 0.68 |
| g4_age_owner | 0.82 | 0.77 | 0.57 | 0.57 | 0.54 | 0.58 | 0.14 | 0.21 |
| age_owner_r | 0.38 | 0.60 | 0.45 | 0.31 | 0.29 | 0.26 | 0.07 | 0.14 |
| g5_hisp_origin_owner | 0.36 | 0.50 | 0.36 | 0.26 | 0.24 | 0.20 | 0.07 | 0.14 |
| g6_race_amind_owner | 0.41 | 0.48 | 0.34 | 0.29 | 0.24 | 0.18 | 0.07 | 0.18 |
| g6_race_asian_owner | 0.41 | 0.48 | 0.34 | 0.29 | 0.24 | 0.18 | 0.07 | 0.18 |
| g6_race_black_owner | 0.41 | 0.48 | 0.34 | 0.29 | 0.24 | 0.18 | 0.07 | 0.18 |
| g6_race_nathaw_owner | 0.41 | 0.48 | 0.34 | 0.29 | 0.24 | 0.18 | 0.07 | 0.18 |
| g6_race_other_owner | 0.41 | 0.48 | 0.34 | 0.29 | 0.24 | 0.18 | 0.07 | 0.18 |
| g6_race_white_owner | 0.41 | 0.45 | 0.34 | 0.29 | 0.24 | 0.18 | 0.07 | 0.18 |
| g6b_race_group | 0.41 | 0.45 | 0.34 | 0.29 | 0.24 | 0.18 | 0.07 | 0.18 |
| g7_native_born_owner | 0.28 | 0.55 | 0.40 | 0.31 | 0.24 | 0.18 | 0.07 | 0.18 |
| g8_us_cit_owner | 0.25 | 0.46 | 0.36 | 0.21 | 0.16 | 0.12 | 0.03 | 0.14 |

| Follow-Up | _0 | _1 | _2 | _3 | _4 | _5 | _6 | _7 |
|---|---|---|---|---|---|---|---|---|
| g9_education_owner | 0.64 | 0.67 | 0.38 | 0.17 | 0.19 | 0.20 | 0.07 | 0.18 |
| g10_gender_owner | 0.19 | 0.55 | 0.38 | 0.38 | 0.38 | 0.32 | 0.07 | 0.07 |
| g10b_marital_status | | | | | 0.04 | 0.08 | 0.09 | 0.00 |
| g10c_net_worth | | | | | 0.50 | 0.42 | 0.38 | 0.35 |
| g10d_personal_outlook | | | | | 0.50 | | | |
| g1b2_reasonfor_business | | | | | | | | 0.35 |
| respondent | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| f3_eq_invest_angels | 1.02 | 0.67 | 0.37 | 0.21 | 0.08 | 0.39 | 0.09 | 1.16 |
| f4_eq_amt_angels | 2.25 | 1.06 | 0.50 | 0.41 | 0.30 | 0.39 | 0.18 | 1.16 |
| tot_equity_angels_r | 1.62 | 0.73 | 0.43 | 0.34 | 0.23 | 0.39 | 0.18 | 1.16 |
| f4_eq_amt_angels_allyrs | 2.25 | 1.48 | 0.63 | 0.40 | 0.12 | 0.68 | 0.32 | 2.29 |
| tot_equity_angels_allyrs_r | 1.62 | 1.30 | 0.47 | 0.30 | 0.12 | 0.68 | 0.32 | 2.29 |
| f5_perc_owned_angels | 1.23 | 0.86 | 0.47 | 0.40 | 0.12 | 0.68 | 0.16 | 2.29 |
| f3_eq_invest_companies | 0.77 | 0.67 | 0.37 | 0.28 | 0.38 | 0.85 | 0.70 | 1.07 |
| f4_eq_amt_companies | 1.45 | 0.95 | 0.56 | 0.69 | 0.83 | 1.09 | 0.70 | 1.07 |
| tot_equity_companies_r | 0.94 | 0.73 | 0.56 | 0.55 | 0.75 | 1.09 | 0.70 | 1.07 |
| f4_eq_amt_companies_allyrs | 1.45 | 1.44 | 0.84 | 0.58 | 0.81 | 1.46 | 1.25 | 2.08 |
| tot_equity_companies_allyrs_r | 0.94 | 1.38 | 0.84 | 0.39 | 0.58 | 1.46 | 1.25 | 2.08 |
| f5_perc_owned_companies | 0.94 | 0.90 | 0.61 | 0.39 | 0.70 | 1.46 | 1.25 | 2.08 |
| f3_eq_invest_govt | 0.72 | 0.45 | 0.68 | 0.48 | 0.30 | 0.54 | 0.53 | 0.89 |
| f4_eq_amt_govt | 0.81 | 0.50 | 0.68 | 0.55 | 0.30 | 0.54 | 0.53 | 0.89 |
| tot_equity_govt_r | 0.77 | 0.45 | 0.68 | 0.48 | 0.30 | 0.54 | 0.53 | 0.89 |
| f4_eq_amt_govt_allyrs | 0.81 | 0.47 | 0.81 | 0.74 | 0.46 | 0.91 | 0.90 | 1.67 |
| tot_equity_govt_allyrs_r | 0.77 | 0.47 | 0.81 | 0.65 | 0.46 | 0.91 | 0.90 | 1.67 |
| f5_perc_owned_govt | 0.72 | 0.47 | 0.81 | 0.74 | 0.46 | 0.91 | 0.90 | 1.67 |
| f3_eq_invest_other | 0.89 | 0.78 | 0.74 | 0.69 | 0.83 | 1.01 | 0.88 | 1.61 |
| f4_eq_amt_other | 1.15 | 0.84 | 0.74 | 0.76 | 0.90 | 1.01 | 0.88 | 1.61 |
| tot_equity_other_r | 0.94 | 0.84 | 0.74 | 0.76 | 0.90 | 1.01 | 0.88 | 1.61 |
| f4_eq_amt_other_allyrs | 1.15 | 0.94 | 0.96 | 1.02 | 1.34 | 1.79 | 1.53 | 3.04 |
| tot_equity_other_allyrs_r | 0.94 | 0.94 | 0.96 | 1.02 | 1.34 | 1.79 | 1.53 | 3.04 |
| f5_perc_owned_other | 0.81 | 0.82 | 0.96 | 1.02 | 1.34 | 1.67 | 1.53 | 3.04 |
| f3_eq_invest_parents | 1.02 | 0.67 | 0.37 | 0.34 | 0.15 | 0.31 | 0.62 | 1.16 |
| f4_eq_amt_parents | 1.70 | 1.12 | 0.56 | 0.41 | 0.15 | 0.31 | 0.62 | 1.16 |
| tot_equity_parents_r | 1.23 | 0.95 | 0.50 | 0.34 | 0.15 | 0.31 | 0.62 | 1.16 |
| f4_eq_amt_parents_allyrs | 1.70 | 1.29 | 0.70 | 0.59 | 0.24 | 0.55 | 1.14 | 2.34 |
| tot_equity_parents_allyrs_r | 1.23 | 1.23 | 0.62 | 0.49 | 0.24 | 0.55 | 1.14 | 2.34 |
| f5_perc_owned_parents | 1.19 | 0.92 | 0.47 | 0.59 | 0.24 | 0.55 | 1.14 | 2.34 |
| f3_eq_invest_spouse | 1.06 | 0.56 | 0.31 | 0.21 | 0.15 | 0.47 | 0.26 | 0.80 |
| f4_eq_amt_spouse | 1.62 | 0.78 | 0.31 | 0.21 | 0.23 | 0.47 | 0.35 | 0.80 |
| tot_equity_spouse_r | 1.32 | 0.67 | 0.31 | 0.21 | 0.23 | 0.47 | 0.35 | 0.80 |
| f4_eq_amt_spouse_allyrs | 1.62 | 0.78 | 0.46 | 0.29 | 0.35 | 0.92 | 0.62 | 1.54 |
| tot_equity_spouse_allyrs_r | 1.32 | 0.72 | 0.38 | 0.29 | 0.35 | 0.79 | 0.47 | 1.54 |
| f5_perc_owned_spouse | 1.19 | 0.72 | 0.46 | 0.29 | 0.23 | 0.79 | 0.47 | 1.54 |

| Follow-Up | _0 | _1 | _2 | _3 | _4 | _5 | _6 | _7 |
|---|---|---|---|---|---|---|---|---|
| f3_eq_invest_vent_cap | 0.94 | 0.56 | 0.56 | 0.48 | 0.08 | 0.39 | 0.26 | 0.89 |
| f4_eq_amt_vent_cap | 1.36 | 0.67 | 0.62 | 0.55 | 0.15 | 0.47 | 0.26 | 0.89 |
| tot_equity_vent_cap_r | 1.19 | 0.67 | 0.62 | 0.48 | 0.08 | 0.39 | 0.26 | 0.89 |
| f4_eq_amt_vent_cap_allyrs | 1.36 | 0.71 | 0.74 | 0.74 | 0.23 | 0.77 | 0.45 | 1.67 |
| tot_equity_vent_cap_allyrs_r | 1.19 | 0.71 | 0.74 | 0.65 | 0.23 | 0.64 | 0.45 | 1.67 |
| f5_perc_owned_vent_cap | 1.15 | 0.65 | 0.74 | 0.74 | 0.11 | 0.64 | 0.45 | 1.67 |
| f5a_seek_equity | | | | | | 0.50 | 0.05 | 0.10 |
| f6a_personal_use | 20.21 | 0.30 | 0.29 | 0.24 | 0.35 | 0.58 | 0.19 | 0.45 |
| f6b_personal_use_amt | 20.21 | 1.85 | 1.03 | 0.93 | 0.81 | 1.12 | 0.56 | 1.15 |
| tot_personal_use_r | 20.21 | 1.85 | 0.62 | 0.34 | 0.46 | 0.75 | 0.24 | 0.75 |
| f6z_family_owned | | | | | 18.04 | | | |
| f7a_bus_credcard | 0.93 | 0.50 | 0.44 | 0.55 | 0.42 | 0.66 | 0.52 | 0.55 |
| f7b_bus_credcard_numused | 2.56 | 3.78 | 4.13 | 4.32 | 8.60 | 7.81 | 7.90 | 7.57 |
| f8a_bus_credcard_line | 2.86 | 2.07 | 1.81 | 1.78 | 1.37 | 1.77 | 1.77 | 1.85 |
| tot_bus_credcard_line_resp_r | 1.50 | 1.03 | 0.90 | 1.13 | 0.71 | 1.24 | 1.31 | 1.11 |
| f8b_bus_credcard_bal | 2.19 | 1.84 | 1.45 | 1.31 | 0.93 | 1.50 | 1.31 | 1.38 |
| tot_bus_credcard_bal_resp_r | 1.58 | 1.03 | 0.80 | 0.89 | 0.66 | 1.11 | 1.15 | 1.11 |
| f7a_pers_credcard | 1.01 | 0.53 | 0.83 | 0.65 | 0.54 | 0.54 | 0.61 | 0.90 |
| f7b_pers_credcard_numused | 2.76 | 2.33 | 2.74 | 2.71 | 4.60 | 5.27 | 5.17 | 4.78 |
| f8a_pers_credcard_line | 4.18 | 2.85 | 2.52 | 2.36 | 1.43 | 1.81 | 1.90 | 2.52 |
| tot_pers_credcard_line_resp_r | 2.25 | 1.41 | 1.62 | 1.48 | 1.17 | 1.21 | 1.71 | 2.09 |
| f8b_pers_credcard_bal | 4.26 | 2.57 | 2.52 | 2.09 | 1.37 | 1.28 | 2.19 | 2.52 |
| tot_pers_credcard_bal_resp_r | 2.19 | 1.28 | 1.75 | 1.15 | 0.98 | 1.21 | 1.62 | 2.09 |
| f7a_pers_loan_bank | 1.10 | 0.48 | 0.56 | 0.51 | 0.61 | 0.62 | 0.47 | 0.65 |
| f7b_pers_loan_bank_numused | 2.05 | 1.25 | 1.56 | 1.54 | 2.03 | 2.08 | 1.60 | 1.94 |
| f8c_pers_loan_bank_amt | 2.62 | 0.90 | 0.91 | 1.17 | 1.23 | 1.37 | 0.56 | 0.90 |
| tot_pers_loan_bank_resp_r | 1.56 | 0.63 | 0.71 | 0.99 | 1.07 | 1.25 | 0.52 | 0.80 |
| f8d_pers_loan_bank_owed | 2.62 | 1.35 | 1.35 | 1.13 | 1.20 | 1.22 | 0.99 | 1.79 |
| tot_pers_loan_bank_owed_resp_r | 1.56 | 1.35 | 1.06 | 0.75 | 1.08 | 1.15 | 0.91 | 1.51 |
| f7a_pers_loan_fam | 0.91 | 0.70 | 0.88 | 0.55 | 0.46 | 1.00 | 0.56 | 0.75 |
| f7b_pers_loan_fam_numused | 1.54 | 1.50 | 1.50 | 1.20 | 1.61 | 2.45 | 1.74 | 1.84 |
| f8c_pers_loan_fam_amt | 1.46 | 0.95 | 0.97 | 0.75 | 0.65 | 1.04 | 0.66 | 0.80 |
| tot_pers_loan_fam_resp_r | 1.14 | 0.78 | 0.91 | 0.62 | 0.58 | 1.04 | 0.61 | 0.75 |
| f8d_pers_loan_fam_owed | 1.46 | 1.05 | 1.16 | 0.99 | 0.61 | 1.47 | 0.98 | 1.27 |
| tot_pers_loan_fam_owed_resp_r | 1.14 | 1.05 | 1.09 | 0.73 | 0.61 | 1.47 | 0.91 | 1.19 |
| f7a_pers_loan_other | 0.91 | 0.63 | 0.71 | 0.51 | 0.50 | 0.83 | 0.38 | 0.85 |
| f7b_pers_loan_other_numused | 1.16 | 0.95 | 0.91 | 0.65 | 0.77 | 1.37 | 0.80 | 1.44 |
| f8c_pers_loan_other_amt | 1.03 | 0.78 | 0.77 | 0.58 | 0.58 | 1.00 | 0.38 | 0.85 |
| tot_pers_loan_other_resp_r | 0.93 | 0.73 | 0.74 | 0.51 | 0.50 | 0.96 | 0.38 | 0.85 |
| f8d_pers_loan_other_owed | 1.03 | 0.77 | 0.82 | 0.67 | 0.65 | 1.06 | 0.49 | 1.25 |
| tot_persloan_other_owed_resp_r | 0.93 | 0.77 | 0.76 | 0.59 | 0.60 | 1.06 | 0.49 | 1.18 |
| f7a_pers_other | 1.75 | 1.63 | 1.24 | 1.48 | 1.27 | 1.74 | 1.13 | 1.30 |
| f7b_pers_other_numused | 1.95 | 1.75 | 1.33 | 1.58 | 1.46 | 2.08 | 1.36 | 1.44 |

| Follow-Up | _0 | _1 | _2 | _3 | _4 | _5 | _6 | _7 |
|---|---|---|---|---|---|---|---|---|
| f8c_pers_other_amt | 1.81 | 1.63 | 1.30 | 1.51 | 1.30 | 1.74 | 1.13 | 1.30 |
| tot_pers_other_resp_r | 1.66 | 1.63 | 1.27 | 1.48 | 1.27 | 1.74 | 1.13 | 1.30 |
| f8d_pers_other_owed | 1.81 | 1.67 | 1.45 | 1.81 | 1.63 | 2.33 | 1.56 | 1.90 |
| tot_pers_other_owed_resp_r | 1.66 | 1.67 | 1.41 | 1.72 | 1.58 | 2.33 | 1.56 | 1.90 |
| f9a_bus_credcard | 1.75 | 1.65 | 1.33 | 0.21 | 3.33 | 3.06 | 2.98 | 3.96 |
| f9b_bus_credcard_numused | 2.43 | 3.60 | 2.30 | 2.04 | 6.53 | 5.42 | 5.96 | 7.92 |
| f10a_bus_credcard_line | 2.97 | 2.98 | 2.66 | 1.63 | 6.83 | 7.18 | 6.96 | 10.68 |
| tot_bus_credcard_line_others_r | 2.36 | 2.33 | 2.39 | 0.72 | 6.38 | 6.35 | 6.59 | 10.26 |
| f10b_bus_credcard_bal | 2.83 | 2.79 | 2.39 | 1.45 | 6.61 | 7.18 | 6.59 | 10.26 |
| tot_bus_credcard_bal_others_r | 2.43 | 2.33 | 2.26 | 0.36 | 6.38 | 6.35 | 6.59 | 10.26 |
| f9a_pers_credcard | 2.16 | 2.03 | 1.51 | 0.54 | 3.45 | 3.34 | 3.31 | 4.13 |
| f9b_pers_credcard_numused | 3.03 | 2.93 | 2.22 | 1.29 | 5.30 | 5.42 | 4.14 | 5.51 |
| f10a_pers_credcard_line | 6.14 | 4.37 | 3.39 | 3.09 | 7.04 | 8.00 | 8.37 | 11.76 |
| tot_pers_credcard_line_othrs_r | 3.44 | 3.28 | 2.97 | 1.55 | 6.80 | 6.86 | 7.97 | 10.86 |
| f10b_pers_credcard_bal | 4.99 | 3.97 | 3.53 | 1.55 | 7.04 | 7.71 | 7.97 | 11.31 |
| tot_pers_credcard_bal_others_r | 3.64 | 2.98 | 2.97 | 0.97 | 6.80 | 6.86 | 7.97 | 11.31 |
| f9a_pers_loan_bank | 2.36 | 1.65 | 1.24 | 0.21 | 3.69 | 3.20 | 2.98 | 4.48 |
| f9b_pers_loan_bank_numused | 2.83 | 2.18 | 1.51 | 0.75 | 4.43 | 3.76 | 3.31 | 5.34 |
| f10c_pers_loan_bank_amt | 3.03 | 1.88 | 1.42 | 0.43 | 3.82 | 3.20 | 2.98 | 4.99 |
| tot_pers_loan_bank_others_r | 2.76 | 1.73 | 1.33 | 0.21 | 3.82 | 3.20 | 2.98 | 4.82 |
| f10d_pers_loan_bank_owed | 3.03 | 2.39 | 2.26 | 0.90 | 6.74 | 6.30 | 6.27 | 11.37 |
| tot_persloan_bank_owed_othrs_r | 2.76 | 2.39 | 2.00 | 0.36 | 6.52 | 6.30 | 6.27 | 10.98 |
| f9a_pers_loan_fam | 2.16 | 2.03 | 1.86 | 1.18 | 3.69 | 3.48 | 3.15 | 4.48 |
| f9b_pers_loan_fam_numused | 2.43 | 2.18 | 1.95 | 1.29 | 4.19 | 3.62 | 3.15 | 4.65 |
| f10c_pers_loan_fam_amt | 2.16 | 2.18 | 1.86 | 1.18 | 3.82 | 3.48 | 3.15 | 4.82 |
| tot_persloan_fam_othrowners_r | 2.16 | 2.03 | 1.86 | 1.18 | 3.69 | 3.48 | 3.15 | 4.82 |
| f10d_pers_loan_fam_owed | 2.16 | 2.53 | 2.59 | 1.82 | 6.02 | 6.27 | 6.07 | 9.64 |
| tot_persloan_fam_owed_othrs_r | 2.16 | 2.53 | 2.59 | 1.82 | 6.02 | 6.27 | 6.07 | 9.64 |
| f9a_pers_loan_other | 1.82 | 1.80 | 1.51 | 0.54 | 3.69 | 3.20 | 3.15 | 3.96 |
| f9b_pers_loan_other_numused | 1.82 | 1.88 | 1.68 | 0.54 | 3.69 | 3.34 | 3.15 | 4.13 |
| f10c_pers_loan_other_amt | 1.89 | 1.80 | 1.51 | 0.54 | 3.69 | 3.20 | 3.15 | 3.96 |
| tot_pers_loan_other_owners_r | 1.89 | 1.80 | 1.51 | 0.54 | 3.69 | 3.20 | 3.15 | 3.96 |
| f10d_pers_loan_other_owed | 1.89 | 2.14 | 2.05 | 0.82 | 5.84 | 5.56 | 5.86 | 8.10 |
| tot_persloan_othr_owed_othrs_r | 1.89 | 2.14 | 2.05 | 0.82 | 5.84 | 5.56 | 5.86 | 8.10 |
| f9a_pers_other | 2.16 | 2.18 | 1.33 | 1.07 | 3.82 | 3.89 | 3.64 | 4.82 |
| f9b_pers_other_numused | 2.16 | 2.25 | 1.33 | 1.07 | 4.06 | 4.03 | 3.64 | 5.16 |
| f10c_pers_other_amt | 2.16 | 2.18 | 1.42 | 1.07 | 3.82 | 3.89 | 3.64 | 4.82 |
| tot_pers_other_other_owners_r | 2.16 | 2.18 | 1.42 | 1.07 | 3.82 | 3.89 | 3.64 | 4.82 |
| f10d_pers_other_owed | 2.16 | 2.59 | 1.83 | 1.63 | 6.07 | 6.83 | 6.90 | 10.00 |
| tot_pers_other_owed_others_r | 2.16 | 2.59 | 1.83 | 1.63 | 6.07 | 6.83 | 6.90 | 10.00 |
| f11a_bus_cred_line | 0.83 | 0.58 | 0.47 | 0.62 | 0.73 | 0.83 | 0.42 | 0.85 |
| f11b_bus_cred_line_numused | 1.38 | 1.60 | 1.89 | 1.68 | 3.95 | 4.11 | 3.95 | 3.99 |
| f12a_bus_cred_line | 1.28 | 0.89 | 1.03 | 1.18 | 1.44 | 1.40 | 0.89 | 2.03 |

| Follow-Up | _0 | _1 | _2 | _3 | _4 | _5 | _6 | _7 |
|---|---|---|---|---|---|---|---|---|
| tot_cred_line_bus_line_r | 0.99 | 0.71 | 0.63 | 0.80 | 1.10 | 1.22 | 0.76 | 1.64 |
| f12b_bus_cred_line_bal | 1.48 | 0.94 | 1.06 | 1.44 | 1.24 | 1.45 | 0.82 | 1.95 |
| tot_cred_line_bus_bal_r | 0.99 | 0.76 | 0.73 | 0.80 | 1.05 | 1.28 | 0.69 | 1.64 |
| f11a_bus_credcard | 0.63 | 0.38 | 0.65 | 0.62 | 0.96 | 1.16 | 0.33 | 0.60 |
| f11b_bus_credcard_numused | 1.60 | 2.58 | 2.95 | 2.54 | 6.14 | 6.10 | 4.84 | 4.68 |
| f12a_bus_credcard_line | 2.44 | 1.72 | 2.58 | 2.51 | 2.37 | 2.42 | 1.55 | 2.17 |
| tot_credcard_line_bus_r | 1.36 | 0.90 | 1.31 | 1.20 | 1.99 | 2.12 | 1.01 | 1.45 |
| f12b_bus_credcard_bal | 1.89 | 1.50 | 1.97 | 2.04 | 2.44 | 2.71 | 1.01 | 1.66 |
| tot_credcard_bal_bus_r | 1.20 | 0.85 | 1.00 | 1.05 | 1.92 | 2.20 | 0.82 | 1.24 |
| f11a_bus_loans_bank | 0.79 | 0.75 | 0.65 | 0.69 | 0.58 | 1.12 | 0.80 | 1.05 |
| f11b_bus_loans_bank_numused | 1.08 | 1.25 | 1.15 | 1.44 | 1.30 | 2.20 | 1.83 | 2.04 |
| f12c_bus_loans_bank_amt | 1.54 | 0.93 | 1.18 | 1.17 | 0.73 | 1.20 | 0.94 | 1.20 |
| tot_loan_bank_bus_r | 1.08 | 0.80 | 0.88 | 0.86 | 0.58 | 1.12 | 0.89 | 1.10 |
| f12d_bus_loans_bank_owed | 1.54 | 1.05 | 1.16 | 1.17 | 0.91 | 1.68 | 1.44 | 1.86 |
| tot_bus_loans_bank_owed_r | 1.08 | 1.05 | 0.90 | 0.88 | 0.76 | 1.57 | 1.37 | 1.70 |
| f11a_bus_loans_nonbank | 0.71 | 0.58 | 0.50 | 0.65 | 0.77 | 0.79 | 0.61 | 0.65 |
| f11b_bus_loans_nonbank_numused | 0.77 | 0.73 | 0.59 | 0.79 | 1.11 | 1.12 | 0.89 | 0.80 |
| f12c_bus_loans_nonbank_amt | 0.85 | 0.65 | 0.59 | 0.79 | 0.77 | 0.87 | 0.66 | 0.75 |
| tot_loan_nonbank_bus_r | 0.77 | 0.60 | 0.53 | 0.69 | 0.77 | 0.83 | 0.66 | 0.65 |
| f12d_bus_loans_nonbank_owed | 0.85 | 0.72 | 0.60 | 0.83 | 0.94 | 1.13 | 1.00 | 0.98 |
| tot_bus_loans_nonbank_owed_r | 0.77 | 0.72 | 0.54 | 0.75 | 0.94 | 1.07 | 0.88 | 0.91 |
| f11a_bus_loans_emp | 5.07 | 6.89 | 5.84 | 7.86 | 1.81 | 1.68 | 0.92 | 1.25 |
| f11b_bus_loans_emp_numused | 5.07 | 6.93 | 5.98 | 7.92 | 1.95 | 1.84 | 1.02 | 1.25 |
| f12c_bus_loans_emp_amt | 5.07 | 6.89 | 5.84 | 7.92 | 1.81 | 1.68 | 1.02 | 1.25 |
| tot_loan_emp_bus_r | 5.07 | 6.89 | 5.84 | 7.86 | 1.81 | 1.68 | 0.92 | 1.25 |
| f12d_bus_loans_emp_owed | 5.07 | 11.68 | 12.16 | 16.93 | 5.05 | 4.98 | 3.24 | 4.48 |
| tot_bus_loans_emp_owed_r | 5.07 | 11.68 | 12.16 | 16.93 | 5.05 | 4.98 | 2.94 | 4.48 |
| f11a_bus_loans_fam | 0.75 | 0.73 | 0.50 | 0.82 | 0.61 | 0.87 | 0.47 | 0.75 |
| f11b_bus_loans_fam_numused | 1.10 | 1.13 | 0.77 | 0.99 | 1.07 | 1.50 | 0.85 | 1.30 |
| f12c_bus_loans_fam_amt | 0.93 | 0.78 | 0.53 | 0.99 | 0.73 | 0.87 | 0.52 | 0.75 |
| tot_loan_fam_bus_r | 0.83 | 0.78 | 0.50 | 0.89 | 0.69 | 0.87 | 0.47 | 0.75 |
| f12d_bus_loans_fam_owed | 0.93 | 0.80 | 0.61 | 1.05 | 0.86 | 1.15 | 0.64 | 1.07 |
| tot_bus_loans_fam_owed_r | 0.83 | 0.80 | 0.55 | 0.97 | 0.81 | 1.15 | 0.64 | 1.07 |
| f11a_bus_loans_govt | 0.89 | 0.98 | 0.74 | 0.86 | 0.81 | 0.96 | 0.61 | 0.70 |
| f11b_bus_loans_govt_numused | 0.97 | 1.05 | 0.80 | 0.86 | 0.88 | 1.04 | 0.66 | 0.85 |
| f12c_bus_loans_govt_amt | 1.01 | 0.98 | 0.80 | 0.89 | 0.81 | 1.00 | 0.71 | 0.70 |
| tot_loan_govt_bus_r | 0.91 | 0.98 | 0.80 | 0.86 | 0.81 | 0.96 | 0.66 | 0.70 |
| f12d_bus_loans_govt_owed | 1.01 | 1.07 | 0.79 | 1.02 | 0.97 | 1.27 | 0.86 | 0.96 |
| tot_bus_loans_govt_owed_r | 0.91 | 1.07 | 0.79 | 0.98 | 0.97 | 1.21 | 0.86 | 0.96 |
| f11a_bus_loans_other_ind | 0.89 | 0.83 | 0.68 | 0.82 | 0.81 | 0.83 | 0.56 | 1.10 |
| f11b_busloans_otherind_numused | 0.99 | 0.93 | 0.68 | 0.86 | 0.84 | 1.08 | 0.71 | 1.10 |
| f12c_bus_loans_other_ind_amt | 0.93 | 0.90 | 0.68 | 0.89 | 0.81 | 0.83 | 0.56 | 1.10 |
| tot_loan_other_ind_r | 0.91 | 0.85 | 0.68 | 0.82 | 0.81 | 0.83 | 0.56 | 1.10 |

| Follow-Up | _0 | _1 | _2 | _3 | _4 | _5 | _6 | _7 |
|---|---|---|---|---|---|---|---|---|
| f12d_bus_loans_other_ind_owed | 0.93 | 0.97 | 0.75 | 1.01 | 0.96 | 1.05 | 0.73 | 1.49 |
| tot_bus_loans_otherind_owed_r | 0.91 | 0.97 | 0.75 | 0.93 | 0.96 | 1.05 | 0.73 | 1.49 |
| f11a_bus_loans_owner | 0.88 | 0.88 | 15.37 | 18.08 | 18.25 | 27.81 | 27.84 | 25.49 |
| f11b_bus_loans_owner_numused | 1.29 | 1.76 | 15.76 | 18.35 | 18.86 | 28.59 | 28.36 | 26.19 |
| f12c_bus_loans_owner_amt | 1.60 | 1.07 | 15.37 | 18.08 | 18.25 | 27.92 | 27.84 | 25.49 |
| tot_loan_owner_bus_r | 1.34 | 0.94 | 15.37 | 18.08 | 18.25 | 27.92 | 27.84 | 25.49 |
| f12d_bus_loans_owner_owed | 1.60 | 1.44 | 18.83 | 25.32 | 28.06 | 42.74 | 44.92 | 45.16 |
| tot_bus_loans_owner_owed_r | 1.34 | 1.36 | 18.83 | 25.32 | 28.06 | 42.74 | 44.92 | 45.16 |
| f11a_bus_loans_other_bus | 0.81 | 0.73 | 0.53 | 0.79 | 0.73 | 0.91 | 0.52 | 0.75 |
| f11a_busloans_otherbus_numused | 0.87 | 0.78 | 0.59 | 0.79 | 0.77 | 0.96 | 0.66 | 0.80 |
| f12c_bus_loans_bus_amt | 0.85 | 0.75 | 0.53 | 0.79 | 0.73 | 0.96 | 0.52 | 0.75 |
| tot_loan_other_bus_r | 0.83 | 0.75 | 0.53 | 0.79 | 0.73 | 0.96 | 0.52 | 0.75 |
| f12d_bus_loans_bus_owed | 0.85 | 0.79 | 0.56 | 0.89 | 0.87 | 1.15 | 0.67 | 1.02 |
| tot_bus_loans_otherbus_owed_r | 0.83 | 0.79 | 0.56 | 0.89 | 0.87 | 1.15 | 0.67 | 1.02 |
| f11a_bus_other | 1.01 | 1.10 | 0.80 | 1.20 | 1.38 | 1.58 | 0.85 | 1.49 |
| f11b_bus_other_numused | 1.08 | 1.25 | 0.88 | 1.23 | 1.50 | 1.70 | 0.85 | 1.54 |
| f12c_bus_other_amt | 1.06 | 1.13 | 0.91 | 1.27 | 1.38 | 1.58 | 0.85 | 1.49 |
| tot_bus_debt_other_r | 1.03 | 1.13 | 0.88 | 1.23 | 1.38 | 1.58 | 0.85 | 1.49 |
| f12d_bus_other_owed | 1.06 | 1.12 | 0.85 | 1.41 | 1.67 | 2.03 | 1.12 | 2.08 |
| tot_bus_loans_other_owed_r | 1.03 | 1.12 | 0.85 | 1.37 | 1.67 | 2.03 | 1.12 | 2.08 |
| f12e_collateral | | | | | | 1.03 | 0.27 | 0.45 |
| f12f_bus_real_estate | | | | | | 6.91 | 3.36 | 2.96 |
| f12f_business_equip_veh | | | | | | 6.25 | 2.94 | 3.45 |
| f12f_business_sec_dep | | | | | | 6.25 | 2.52 | 3.45 |
| f12f_intellectual_prop | | | | | | 6.58 | 2.52 | 3.45 |
| f12f_inventory_acctrec | | | | | | 6.25 | 3.36 | 3.45 |
| f12f_other | | | | | | 6.25 | 2.52 | 2.96 |
| f12f_other_pers_assets | | | | | | 5.92 | 2.10 | 3.94 |
| f12f_pers_real_estate | | | | | | 6.58 | 2.94 | 3.94 |
| f13_trade_fin | 0.43 | 0.45 | 0.59 | 0.27 | 0.27 | 0.37 | 0.09 | 0.65 |
| f14a_trade_fin_amt | 3.35 | 2.03 | 1.47 | 1.06 | 0.92 | 1.00 | 0.61 | 1.10 |
| tot_trade_finan_r | 0.85 | 0.70 | 0.62 | 0.31 | 0.27 | 0.50 | 0.24 | 0.80 |
| f14d_new_loans | | | | 0.27 | 0.27 | 0.37 | 0.09 | 0.35 |
| f14g_didnotapply | | | | 0.14 | 0.23 | 0.42 | 0.19 | 0.40 |
| f14h_loan_guarantees | | | | | 0.27 | 0.33 | 0.09 | 0.25 |
| f14e_approved_denied | | | | 0.28 | 0.29 | 0.00 | 0.00 | 0.00 |
| f14f_bus_credit_hist | | | | 11.01 | 6.84 | 12.07 | 4.71 | 17.14 |
| f14f_inadeq_doc | | | | 11.01 | 8.55 | 11.21 | 7.06 | 18.57 |
| f14f_insuff_coll | | | | 10.09 | 8.55 | 12.07 | 3.53 | 20.00 |
| f14f_loan_toolarge | | | | 11.01 | 10.26 | 11.21 | 3.53 | 18.57 |
| f14f_new_bus | | | | 11.93 | 9.40 | 12.07 | 7.06 | 18.57 |
| f14f_other | | | | 11.01 | 8.55 | 12.93 | 5.88 | 20.00 |
| f14f_pers_credit_hist | | | | 11.01 | 8.55 | 12.07 | 3.53 | 18.57 |

| Follow-Up | _0 | _1 | _2 | _3 | _4 | _5 | _6 | _7 |
|---|---|---|---|---|---|---|---|---|
| f14f_restr_on_lending | | | | | 90.60 | 7.76 | 2.35 | 14.29 |
| f14i_economy_effect | | | | | 0.23 | | | |
| f14j_most_challenging | | | | | 1.53 | 1.62 | 1.51 | 1.79 |
| f15_revenue | 0.73 | 0.78 | 0.65 | 0.27 | 0.19 | 0.46 | 0.24 | 0.15 |
| f16a_rev_amt | 12.62 | 7.20 | 3.66 | 3.26 | 9.02 | 3.65 | 2.40 | 2.94 |
| tot_revenue_r | 3.79 | 2.00 | 1.42 | 0.93 | 6.83 | 1.45 | 0.66 | 0.90 |
| f17a_total_exp_amt | 16.54 | 9.88 | 4.90 | 4.08 | 3.53 | 3.28 | 2.59 | 2.89 |
| tot_expenses_r | 4.10 | 2.00 | 1.12 | 0.65 | 0.58 | 1.00 | 0.52 | 1.10 |
| f18a_wage_exp_amt | 8.28 | 5.88 | 4.22 | 3.16 | 1.92 | 2.28 | 1.88 | 2.09 |
| tot_wages_r | 4.36 | 2.65 | 2.12 | 1.48 | 0.46 | 1.00 | 0.61 | 1.20 |
| f19_res_dev | 0.39 | 0.38 | 0.29 | 0.31 | 0.15 | 0.17 | 0.05 | 0.30 |
| f19a_res_dev_amt | | | | 0.86 | 0.81 | 0.87 | 0.52 | 0.60 |
| tot_res_dev_r | | | | 0.38 | 0.27 | 0.37 | 0.14 | 0.40 |
| f19b_a_design | | | | | 0.44 | 0.99 | 0.96 | 1.60 |
| f19b_b_investments | | | | | 0.65 | 1.72 | 2.23 | 1.60 |
| f19b_c_brand_dev | | | | | 1.09 | 1.72 | 0.96 | 1.60 |
| f19b_d_org_dev | | | | | 0.65 | 1.72 | 1.27 | 1.60 |
| f19b_e_worker_training | | | | | 1.09 | 1.72 | 0.64 | 1.92 |
| f19b_f_other | | | | | 2.18 | 2.46 | 3.50 | 4.79 |
| f19c_a_design_amt | | | | | | 6.77 | 4.84 | 4.76 |
| tot_intangassets_design_r | | | | | | 2.90 | 1.61 | 2.78 |
| f19c_b_investments_amt | | | | | | 10.84 | 8.81 | 4.67 |
| tot_intangassets_invest_r | | | | | | 5.91 | 4.40 | 3.33 |
| f19c_c_brand_dev_amt | | | | | | 9.17 | 6.57 | 5.49 |
| tot_intangassets_branddev_r | | | | | | 5.83 | 2.53 | 3.85 |
| f19c_d_org_dev_amt | | | | | | 13.75 | 9.09 | 9.21 |
| tot_intangassets_orgdev_r | | | | | | 11.25 | 6.49 | 6.58 |
| f19c_e_worker_training_amt | | | | | | 11.18 | 4.84 | 7.96 |
| tot_intangassets_wkrtrng_r | | | | | | 6.58 | 2.42 | 6.19 |
| f19c_f_other_amt | | | | | | 75.00 | 68.75 | 75.00 |
| tot_intangassets_other_r | | | | | | 68.75 | 68.75 | 75.00 |
| f19c_intangassets_amt | | | | | 14.81 | | | |
| tot_intang_assets_r | | | | | 1.09 | | | |
| f20_mach | 0.47 | 0.73 | 0.62 | 0.48 | 0.27 | 0.54 | 0.61 | 0.25 |
| f21_land_rent | 0.45 | 0.58 | 0.59 | 0.41 | 0.35 | 0.42 | 0.47 | 0.55 |
| f22_mach_rent | 0.45 | 0.68 | 0.59 | 0.55 | 0.19 | 0.33 | 0.24 | 0.40 |
| f23_profit_or_loss | 2.70 | 1.93 | 1.24 | 0.51 | 0.38 | 1.04 | 0.89 | 1.10 |
| f24_profit_amt | 16.80 | 9.17 | 6.13 | 4.97 | 3.42 | 5.15 | 4.01 | 4.96 |
| tot_profit_r | 9.75 | 4.75 | 2.89 | 1.62 | 1.27 | 2.77 | 2.08 | 2.88 |
| f26_loss_amt | 19.39 | 13.60 | 8.71 | 7.37 | 6.17 | 6.54 | 6.15 | 5.95 |
| tot_loss_r | 9.05 | 6.65 | 4.91 | 2.19 | 2.12 | 3.27 | 3.21 | 4.69 |
| f24_profitloss_amt | 15.95 | 9.20 | 5.87 | 5.42 | 4.14 | 4.69 | 3.90 | 4.24 |
| f28a_asset_cash | 1.28 | 0.68 | 0.53 | 0.21 | 0.88 | 0.42 | 0.66 | 0.65 |

| Follow-Up | _0 | _1 | _2 | _3 | _4 | _5 | _6 | _7 |
|---|---|---|---|---|---|---|---|---|
| f29_assetval_cash | 11.10 | 6.65 | 3.60 | 2.92 | 2.84 | 2.95 | 2.45 | 2.99 |
| tot_asset_cash_r | 5.03 | 2.48 | 1.47 | 0.82 | 1.46 | 1.25 | 1.27 | 1.84 |
| f28b_asset_acct_rec | 2.03 | 1.23 | 1.12 | 0.72 | 1.73 | 1.62 | 1.08 | 1.69 |
| f29_assetval_acctrec | 8.64 | 5.35 | 3.69 | 2.85 | 3.42 | 3.45 | 2.73 | 3.44 |
| tot_asset_acct_rec_r | 4.71 | 2.60 | 1.95 | 1.30 | 2.07 | 1.95 | 1.55 | 2.54 |
| f28c_asset_inv | 1.10 | 1.45 | 1.47 | 1.27 | 1.88 | 1.58 | 1.51 | 1.59 |
| f29_assetval_inv | 5.19 | 4.48 | 2.89 | 2.81 | 3.11 | 3.24 | 2.73 | 2.79 |
| tot_asset_inv_r | 2.70 | 2.30 | 1.80 | 1.58 | 2.07 | 1.99 | 1.65 | 2.09 |
| f28d_asset_equip | 0.83 | 0.95 | 0.74 | 0.55 | 1.27 | 1.29 | 0.89 | 1.10 |
| f29_assetval_equip | 6.21 | 4.80 | 2.80 | 2.57 | 2.80 | 3.16 | 2.30 | 2.79 |
| tot_asset_equip_r | 2.80 | 2.00 | 1.33 | 0.93 | 1.65 | 1.62 | 1.18 | 1.69 |
| f28e_asset_landbuild | 1.36 | 2.28 | 1.92 | 1.34 | 2.30 | 2.24 | 2.12 | 1.99 |
| f29_assetval_landbuild | 2.31 | 2.93 | 2.51 | 1.85 | 2.65 | 2.74 | 2.35 | 2.69 |
| tot_asset_landbuild_r | 1.77 | 2.40 | 2.04 | 1.44 | 2.34 | 2.37 | 2.12 | 2.34 |
| f28f_asset_veh | 1.18 | 1.85 | 1.59 | 1.27 | 1.88 | 2.12 | 1.65 | 1.44 |
| f29_assetval_veh | 2.56 | 3.23 | 2.57 | 1.89 | 2.69 | 2.95 | 2.40 | 2.09 |
| tot_asset_veh_r | 1.68 | 2.25 | 1.95 | 1.30 | 2.15 | 2.28 | 1.79 | 1.59 |
| f28g_other_bus_prop | 4.30 | 3.18 | 2.80 | 2.54 | 3.30 | 3.20 | 3.10 | 3.19 |
| f29_assetval_othbusprop | 4.44 | 3.25 | 2.86 | 2.71 | 3.30 | 3.24 | 3.20 | 3.39 |
| tot_asset_other_bus_prop_r | 1.77 | 3.18 | 2.80 | 2.57 | 3.30 | 3.20 | 3.15 | 3.34 |
| f28h_other_assets | 3.49 | 3.23 | 3.10 | 2.54 | 3.30 | 3.61 | 3.29 | 3.49 |
| f29_assetval_other | 3.90 | 3.65 | 3.45 | 2.85 | 3.53 | 3.90 | 3.39 | 3.69 |
| tot_asset_other_r | 1.93 | 3.30 | 3.22 | 2.64 | 3.49 | 3.61 | 3.34 | 3.54 |
| f30a_liab_acctpay | 1.46 | 0.60 | 0.65 | 0.48 | 0.84 | 0.75 | 0.42 | 0.50 |
| f31_value_acctpay | 6.03 | 3.65 | 2.39 | 1.92 | 1.84 | 2.24 | 1.46 | 1.54 |
| tot_liab_acct_pay_r | 3.23 | 1.38 | 1.09 | 0.69 | 1.19 | 1.04 | 0.61 | 0.95 |
| f30b_liab_pension | 1.14 | 1.55 | 1.24 | 0.99 | 1.69 | 1.29 | 1.51 | 1.10 |
| f31_value_pension | 1.44 | 1.80 | 1.47 | 1.20 | 1.77 | 1.54 | 1.88 | 1.20 |
| tot_liab_pension_r | 1.28 | 1.68 | 1.33 | 0.99 | 1.69 | 1.41 | 1.60 | 1.10 |
| f30c_liab_other | 1.38 | 1.93 | 1.74 | 1.61 | 1.96 | 1.87 | 1.83 | 1.79 |
| f31_value_other | 1.66 | 2.05 | 1.80 | 1.78 | 1.96 | 1.91 | 1.88 | 1.94 |
| tot_liab_other_r | 1.54 | 1.98 | 1.77 | 1.68 | 1.96 | 1.87 | 1.83 | 1.84 |
| f32_chap11_bankruptcy | | | | | 0.27 | 1.08 | 0.24 | 0.25 |
| f33_expected_growth | | | | | 0.19 | | | |
| f34_future_revenue | | | | | 0.77 | | | |
| respondent | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| website | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| email | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| credrisk | 26.83 | 2.28 | 13.04 | 14.48 | 8.10 | 16.45 | 16.56 | 20.63 |
| fssp | 26.87 | 5.18 | 5.93 | 7.41 | 8.21 | 16.53 | 16.60 | 20.63 |
| paysc | 95.70 | 84.49 | 76.28 | 59.18 | 51.92 | 49.50 | 46.71 | 41.85 |

**Appendix B**

| The question appear in follow-up | | | | | | | | Question | Variable Name |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |
| **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | In what calendar year did [NAME BUSINESS] close? | a11_year_closed |
| **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | Did [NAME BUSINESS] file for bankruptcy? | a11a_bankruptcy |
| **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | **No** | How much did the nation's recent financial problems, which became highly visible in YYYY, affect [NAME BUSINESS] during calendar year YYYY? Would you say . . . | a11b_economy_effect |
| **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | Our records show that [NAME BUSINESS] had a legal status of [LEGAL STATUS]. As of December 31, YYYY, is that still the legal status of [NAME BUSINESS]? | c1z_confirm_legal_status |
| **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | I'm going to read you a list of some different forms of legal status a business can have. As of December 31, YYYY, which form of legal status did [NAME BUSINESS] have? Was it a . . . Something else? (SPECIFY) | c1z2otherspecify |
| **No** | **Yes** | **Yes** | **Yes** | **No** | **No** | **No** | **No** | Was this change an increase, a decrease, or no change in the number of people who worked for [NAME BUSINESS] on December 31, YYYY compared to December 31, YYYY? | c5b_num_employees_change |
| **No** | **Yes** | **Yes** | **Yes** | **No** | **No** | **No** | **No** | And what was the (increase/decrease) in the number of people who worked for [NAME BUSINESS] on December 31, YYYY compared to December 31, YYYY? Your best estimate is fine | c5c_num_employees_change_amt |
| **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | Our records show that the primary location where [NAME BUSINESS] operates is [PRIMARY LOCATION]. Is that correct? | c8z_primary_loc_confirm |
| **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | What was the main reason for the change of location? | c9_loc_change_reason |

**Appendix B – Continued**

| The question appear in follow-up | | | | | | | | Question | Variable Name |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |
| **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | What was the main reason for the change of location? OTHER (SPECIFY) | c9otherspecify |
| **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | As of December 31, YYYY, did [NAME BUSINESS] operate in more than one location? | c10_morelocations |
| **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | And as of December 31, YYYY, how many locations did [NAME BUSINESS] operate in? | c11_num_locations |
| **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | In what month and year did you open your second location?, Month | c11a_2ndopening_month |
| **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | In what month and year did you open your second location?, Year | c11a_2ndopening_year |
| **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | **No** | There are many programs available to help new businesses. I am going to read some possible sources of training and assistance that may have been used to help [NAME BUSINESS]. Have you (or any of the other owners) ever received any business training, mentoring, or technical assistance sponsored by (READ ITEM) to help [NAME BUSINESS]? - The Small Business Administration or SBA | c12a_sba |
| **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | **No** | There are many programs available to help new businesses. I am going to read some possible sources of training and assistance that may have been used to help [NAME BUSINESS]. Have you (or any of the other owners) ever received any business training, mentoring, or technical assistance sponsored by (READ ITEM) to help [NAME BUSINESS]? - A Federal government agency other than SBA | c12b_fed_gov |

**Appendix B – Continued**

| The question appear in follow-up | | | | | | | | Question | Variable Name |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |
| **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | **No** | There are many programs available to help new businesses. I am going to read some possible sources of training and assistance that may have been used to help [NAME BUSINESS]. Have you (or any of the other owners) ever received any business training, mentoring, or technical assistance sponsored by (READ ITEM) to help [NAME BUSINESS]? - A state or local government | c12c_statelocal_gov |
| **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | **No** | There are many programs available to help new businesses. I am going to read some possible sources of training and assistance that may have been used to help [NAME BUSINESS]. Have you (or any of the other owners) ever received any business training, mentoring, or technical assistance sponsored by (READ ITEM) to help [NAME BUSINESS]? - A non-profit association for small businesses such as SCORE | c12d_non_profit |
| **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | **No** | There are many programs available to help new businesses. I am going to read some possible sources of training and assistance that may have been used to help [NAME BUSINESS]. Have you (or any of the other owners) ever received any business training, mentoring, or technical assistance sponsored by (READ ITEM) to help [NAME BUSINESS]? - A community college or university | c12e_college_univ |
| **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | **No** | There are many programs available to help new businesses. I am going to read some possible sources of training and assistance that may have been used to help [NAME BUSINESS]. Have you (or any of the other owners) ever received any business training, mentoring, or technical assistance sponsored by (READ ITEM) to help [NAME BUSINESS]? - A chamber of commerce | c12f_chamber_of_comm |

**Appendix B – Continued**

| The question appear in follow-up | | | | | | | | Question | Variable Name |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |
| **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | **No** | There are many programs available to help new businesses. I am going to read some possible sources of training and assistance that may have been used to help [NAME BUSINESS]. Have you (or any of the other owners) ever received any business training, mentoring, or technical assistance sponsored by (READ ITEM) to help [NAME BUSINESS]? - A for-profit organization such as an accounting firm | c12g_for_profit_org |
| **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | **No** | There are many programs available to help new businesses. I am going to read some possible sources of training and assistance that may have been used to help [NAME BUSINESS]. Have you (or any of the other owners) ever received any business training, mentoring, or technical assistance sponsored by (READ ITEM) to help [NAME BUSINESS]? - Another Source | c12h_other |
| **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | **No** | There are many programs available to help new businesses. I am going to read some possible sources of training and assistance that may have been used to help [NAME BUSINESS]. Have you (or any of the other owners) ever received any business training, mentoring, or technical assistance sponsored by (READ ITEM) to help [NAME BUSINESS]? - Other source specify | c12other_specify |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | As of December 31, YYYY, what state is [NAME BUSINESS] chartered in? | c1z3_state_chartered |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | First, during calendar year YYYY, did (BUSINESS NAME) introduce any products or services that were new or significantly improved? | d1_a_new_product |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | During calendar year YYYY, were any of the products or services new to any market or markets [NAME BUSINESS] competes in? | d1_b_new_to_market |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | Were any of the new or significantly improved products or services introduced in YYYY new to [ITEM]? a) A regional market such as nearby cities or counties | d1c_a_regional |

**Appendix B – Continued**

| The question appear in follow-up | | | | | | | | Question | Variable Name |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | Were any of the new or significantly improved products or services introduced in YYYY new to [ITEM]? b) A national-wide market | d1c_b_national |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | Were any of the new or significantly improved products or services introduced in YYYY new to [ITEM]? c) An international market | d1c_c_international |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | During calendar year YYYY, did [BUSINESS NAME] introduce any new or significantly improved processes in the production of goods or providing services? Please include any new or improved processes, even if [NAME BUSINESS] was not the first to introduce it. | d1d_new_processes |
| **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | Was the competitive advantage [NAME BUSINESS] had in calendar year YYYY related in any way to teaming up with another company? | d2a_compadv_comp_reason |
| **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | Was the competitive advantage [NAME BUSINESS] had in calendar year YYYY related in any way to teaming up with a government lab or research center? | d2a_compadv_govlab_reason |
| **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | Was the competitive advantage [NAME BUSINESS] had in calendar year YYYY related in any way to patents that [NAME BUSINESS] owns, has applied for, or licensed? | d2a_compadv_patents_reason |
| **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | Was the competitive advantage [NAME BUSINESS] had in calendar year YYYY related in any way to teaming up with a college or university? | d2a_compadv_univ_reason |
| **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | Do you consider this to have given [NAME BUSINESS] a major or a minor competitive advantage in calendar year YYYY? | d2b_compadv_comp_strength |

**Appendix B – Continued**

| The question appear in follow-up | | | | | | | | Question | Variable Name |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |
| **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | Do you consider this to have given [NAME BUSINESS] a major or a minor competitive advantage in calendar year YYYY? | d2b_compadv_govlab_strength |
| **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | Do you consider this to have given [NAME BUSINESS] a major or a minor competitive advantage in calendar year YYYY? | d2b_compadv_patents_strength |
| **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | Do you consider this to have given [NAME BUSINESS] a major or a minor competitive advantage in calendar year YYYY? | d2b_compadv_univ_strength |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | Was the competitive advantage [NAME BUSINESS] had in calendar year YYYY related in any way to [ITEM]? e) Cost advantages | d2c_compadv_cost_reason |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | Was the competitive advantage [NAME BUSINESS] had in calendar year YYYY related in any way to [ITEM]? f) Product or service design or quality | d2c_compadv_design_reason |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | Was the competitive advantage [NAME BUSINESS] had in calendar year YYYY related in any way to [ITEM]? g) Specialized or range of expertise, products or service | d2c_compadv_expertise_reason |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | Was the competitive advantage [NAME BUSINESS] had in calendar year YYYY related in any way to [ITEM]? b) Marketing and promotion | d2c_compadv_marketing_reason |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | Was the competitive advantage [NAME BUSINESS] had in calendar year YYYY related in any way to [ITEM]? a) Price | d2c_compadv_price_reason |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | Was the competitive advantage [NAME BUSINESS] had in calendar year YYYY related in any way to [ITEM]? d) Established reputation | d2c_compadv_reputation_reason |

**Appendix B – Continued**

| The question appear in follow-up | | | | | | | | Question | Variable Name |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | Was the competitive advantage [NAME BUSINESS] had in calendar year YYYY related in any way to [ITEM]? c) Speed of service | d2c_compadv_speed_reason |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | Do you consider this to have given [NAME BUSINESS] a major or a minor competitive advantage in calendar year YYYY? e) Cost advantages | d2d_compadv_cost_strength |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | Do you consider this to have given [NAME BUSINESS] a major or a minor competitive advantage in calendar year YYYY? f) Product or service design or quality | d2d_compadv_design_strength |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | Do you consider this to have given [NAME BUSINESS] a major or a minor competitive advantage in calendar year YYYY? g) Specialized or range of expertise, products or service | d2d_compadv_expertise_strength |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | Do you consider this to have given [NAME BUSINESS] a major or a minor competitive advantage in calendar year YYYY? b) Marketing and promotion | d2d_compadv_marketing_strength |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | Do you consider this to have given [NAME BUSINESS] a major or a minor competitive advantage in calendar year YYYY? a) Price | d2d_compadv_price_strength |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | Do you consider this to have given [NAME BUSINESS] a major or a minor competitive advantage in calendar year YYYY? d) Established reputation | d2d_compadv_reput_strength |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | Do you consider this to have given [NAME BUSINESS] a major or a minor competitive advantage in calendar year YYYY? c) Speed service | d2d_compadv_speed_strength |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | Was [NAME BUSINESS] founded around a new or customized product or service that was created by you or one of the founders of the business? | d5a_founded_newprod |

**Appendix B – Continued**

| The question appear in follow-up | | | | | | | | Question | Variable Name |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | Thinking about the new or customized product or service, around which [NAME BUSINESS] was founded, why was it originally developed? Was it because… a) You or one of the founders needed it for personal use? | d5b_a_personaluse |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | Thinking about the new or customized product or service, around which [NAME BUSINESS] was founded, why was it originally developed? Was it because… b) You or one of the founders needed it for use at a previous job or business? | d5b_b_previousjob |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | Thinking about the new or customized product or service, around which [NAME BUSINESS] was founded, why was it originally developed? Was it because… c) You or one of the founders thought about starting a business based on it or to sell it to someone else? | d5b_c_startingbus |
| **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | Now, I will read you a list of customer locations. When I am done reading, please select only one answer choice for your response. During calendar year YYYY, where were most of [NAME BUSINESS]'s customers located? Would you say . . . | d8_customer_locations |
| **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | During calendar year YYYY, were any of [NAME BUSINESS]'s sales made to individuals, businesses, or governments outside the United States? | d8a_international_sales |
| **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | What percent of [NAME BUSINESS]'s total sales were to individuals, businesses, or governments outside of the United States? Would you say . . . | d8b_perc_international_sales |

**Appendix B – Continued**

| The question appear in follow-up | | | | | | | | Question | Variable Name |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |
| No | No | No | Yes | Yes | Yes | Yes | Yes | During calendar year YYYY, were any of [NAME BUSINESS]'s sales made to customers through the internet, such as through the business' website or an online retailer site? | d9_internet_sales |
| No | No | No | Yes | Yes | Yes | Yes | Yes | What percent of [NAME BUSINESS]'s total sales were sales made to customers through the internet? Would you say . . . | d9a_perc_internet_sales |
| No | No | No | No | No | Yes | Yes | Yes | During calendar year YYYY, did [BUSINESS NAME] actively seek but not obtain equity from companies, government agencies, venture capitalists, angel investors, or any other individuals who are not spouses, life partners, parents, in-laws, or children of the owners? | f5a_seek_equity |
| No | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Have you (or other owners) withdrawn money from the business for personal use in calendar year YYYY? | f6a_personal_use |
| No | No | No | No | Yes | No | No | No | #N/A | f6z_family_owned |
| No | No | No | No | No | Yes | Yes | Yes | Was collateral required to obtain any of the debt financing options that were used in calendar year YYYY? That is, were you or was [NAME BUSINESS] required to pledge as security any personal or business assets that can be taken should the business fail to repay the debt? | f12e_collateral |
| No | No | No | No | No | Yes | Yes | Yes | What collateral was required? Was it... e) Business real estate? | f12f_bus_real_estate |
| No | No | No | No | No | Yes | Yes | Yes | What collateral was required? Was it... b) Business equipment or vehicles? | f12f_business_equip_veh |
| No | No | No | No | No | Yes | Yes | Yes | What collateral was required? Was it... c) Business securities or deposits? | f12f_business_sec_dep |
| No | No | No | No | No | Yes | Yes | Yes | What collateral was required? Was it... d) Patents, copyrights or trademarks? | f12f_intellectual_prop |

**Appendix B – Continued**

| The question appear in follow-up | | | | | | | | Question | Variable Name |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | What collateral was required? Was it... a) Inventory or Accounts receivable? | f12f_inventory_acctrec |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | What collateral was required? Was it... h) Some other type of collateral? | f12f_other |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | What collateral was required? Was it... g) Other personal assets? | f12f_other_pers_assets |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | What collateral was required? Was it... h) Some other type of collateral? (SPECIFY) | f12f_otherspecify |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | What collateral was required? Was it... f) Personal real estate? | f12f_pers_real_estate |
| **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | Did [NAME BUSINESS] make any applications for new or renewed loans or lines of credit in calendar year YYYY? | f14d_new_loans |
| **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | Were these applications always approved, sometimes approved and sometimes denied, or always denied? | f14e_approved_denied |
| **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | Consider the most recent time [NAME BUSINESS]'s credit application was denied. Officially, was the application denied because of business credit history? | f14f_bus_credit_hist |
| **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | Consider the most recent time [NAME BUSINESS]'s credit application was denied. Officially, was the application denied because of inadequate documentation provided? | f14f_inadeq_doc |
| **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | Consider the most recent time [NAME BUSINESS]'s credit application was denied. Officially, was the application denied because of insufficient collateral? | f14f_insuff_coll |
| **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | Consider the most recent time [NAME BUSINESS]'s credit application was denied. Officially, was the application denied because of the loan requested was too large? | f14f_loan_toolarge |

**Appendix B – Continued**

| The question appear in follow-up | | | | | | | | Question | Variable Name |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |
| **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | Consider the most recent time [NAME BUSINESS]'s credit application was denied. Officially, was the application denied because of not being in business long enough? | f14f_new_bus |
| **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | Consider the most recent time [NAME BUSINESS]'s credit application was denied. Officially, was the application denied because of other reason? | f14f_other |
| **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | Consider the most recent time [NAME BUSINESS]'s credit application was denied. Officially, was the application denied because of other reason (specify)? | f14f_otherspecify |
| **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | Consider the most recent time [NAME BUSINESS]'s credit application was denied. Officially, was the application denied because of personal credit history? | f14f_pers_credit_hist |
| **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | Consider the most recent time [NAME BUSINESS]'s credit application was denied. Officially, was the application denied because of banks putting stricter restrictions on lending? | f14g_didnotapply |
| **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | During calendar year YYYY, was there any time when [NAME BUSINESS] needed credit, but did not apply because you or others associated with [NAME BUSINESS] thought the application would be denied? | f14g_didnotapply |
| **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | In calendar year YYYY, did [NAME BUSINESS] have any loan guarantees from a federal government agency, such as the Small Business Administration, or any state or local government agencies? | f14h_loan_guarantees |
| **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | What was the most challenging problem your business faced in calendar year YYYY? Was it . . . | f14j_most_challenging |

**Appendix B – Continued**

| The question appear in follow-up | | | | | | | | Question | Variable Name |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |
| **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | What was the most challenging problem your business faced in calendar year YYYY? Other specify ... | f14j_other_specify |
| **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | **No** | #N/A | f14i_economy_effect |
| **No** | **Yes** | **Yes** | **Yes** | **No** | **No** | **No** | **No** | Was this an increase, a decrease, or no change in the amount of revenue for [NAME BUSINESS] in YYYY compared to YYYY? | f16b_rev_YYYY_change |
| **No** | **Yes** | **Yes** | **Yes** | **No** | **No** | **No** | **No** | And what was the percentage change in revenue in YYYY compared YYYY? Your best estimate is fine. | f16c_perc_change |
| **No** | **Yes** | **Yes** | **Yes** | **No** | **No** | **No** | **No** | Was this an increase, a decrease, or no change in total expenses for [NAME BUSINESS] in YYYY compared to YYYY? | f17b_total_exp_YYYY_change |
| **No** | **Yes** | **Yes** | **Yes** | **No** | **No** | **No** | **No** | And what was the percentage change in total expenses in YYYY compared to YYYY? Your best estimate is fine. | f17c_perc_change |
| **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | **Yes** | Please estimate [NAME BUSINESS]'s total research and development expenses for calendar year YYYY, including materials, equipment, space, salaries, wages, benefits, and consulting fees? | f19a_res_dev_amt_YYYY |
| **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | Investments in intangible assets are expenditures expected to produce long-term benefits for businesses. I’m going to read you some types of intangible assets. When thinking about each category, please consider the cost of in-house activities in these areas including the time of the business owner(s), as well as services or license fees from outside providers. Did [NAME BUSINESS] have expenditures in [ITEM/The design of new and improved products and services] in calendar year YYYY? | f19b_a_design |

**Appendix B – Continued**

| The question appear in follow-up | | | | | | | | Question | Variable Name |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |
| **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | Investments in intangible assets are expenditures expected to produce long-term benefits for businesses. I'm going to read you some types of intangible assets. When thinking about each category, please consider the cost of in-house activities in these areas including the time of the business owner(s), as well as services or license fees from outside providers. Did [NAME BUSINESS] have expenditures in [ITEM/Investments in software or databases] in calendar year YYYY? | f19b_b_investments |
| **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | Investments in intangible assets are expenditures expected to produce long-term benefits for businesses. I'm going to read you some types of intangible assets. When thinking about each category, please consider the cost of in-house activities in these areas including the time of the business owner(s), as well as services or license fees from outside providers. Did [NAME BUSINESS] have expenditures in [ITEM/Brand development such as advertising or marketing] in calendar year YYYY? | f19b_c_brand_dev |
| **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | Investments in intangible assets are expenditures expected to produce long-term benefits for businesses. I'm going to read you some types of intangible assets. When thinking about each category, please consider the cost of in-house activities in these areas including the time of the business owner(s), as well as services or license fees from outside providers. Did [NAME BUSINESS] have expenditures in [ITEM/Organizational development such as company formation expenses or management consulting] in calendar year YYYY? | f19b_d_org_dev |

**Appendix B – Continued**

| The question appear in follow-up | | | | | | | | Question | Variable Name |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |
| **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | Investments in intangible assets are expenditures expected to produce long-term benefits for businesses. I'm going to read you some types of intangible assets. When thinking about each category, please consider the cost of in-house activities in these areas including the time of the business owner(s), as well as services or license fees from outside providers. Did [NAME BUSINESS] have expenditures in [ITEM/Worker training] in calendar year YYYY? | f19b_e_worker_training |
| **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | Investments in intangible assets are expenditures expected to produce long-term benefits for businesses. I'm going to read you some types of intangible assets. When thinking about each category, please consider the cost of in-house activities in these areas including the time of the business owner(s), as well as services or license fees from outside providers. Did [NAME BUSINESS] have expenditures in [ITEM/Any other intangible asset investments] in calendar year YYYY? | f19b_f_other |
| **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | Investments in intangible assets are expenditures expected to produce long-term benefits for businesses. I'm going to read you some types of intangible assets. When thinking about each category, please consider the cost of in-house activities in these areas including the time of the business owner(s), as well as services or license fees from outside providers. Did [NAME BUSINESS] have expenditures in [ITEM/Any other intangible asset investments - specify] in calendar year YYYY? | f19b_f_other_specify |
| **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | **No** | Thinking about all the intangible asset expenditures [LIST IF NECESSARY] you just told me about, please estimate [NAME BUSINESS]'s total expenses on intangible assets for calendar year YYYY. | f19c_intangassets_amt |

**Appendix B – Continued**

| The question appear in follow-up | | | | | | | | Question | Variable Name |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | During calendar year YYYY, how much money did (BUSINESS NAME) spend on [INTANGIBLE ASSETS ITEM]/The design of new and improved products and services? | f19c_a_design_amt |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | During calendar year YYYY, how much money did (BUSINESS NAME) spend on [INTANGIBLE ASSETS ITEM]/Investments in software or databases? | f19c_b_investments_amt |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | During calendar year YYYY, how much money did (BUSINESS NAME) spend on [INTANGIBLE ASSETS ITEM]/ Brand development such as advertising or marketing? | f19c_c_brand_dev_amt |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | During calendar year YYYY, how much money did (BUSINESS NAME) spend on [INTANGIBLE ASSETS ITEM]/Organizational development such as company formation expenses or management consulting? | f19c_d_org_dev_amt |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | During calendar year YYYY, how much money did (BUSINESS NAME) spend on [INTANGIBLE ASSETS ITEM]/Worker training? | f19c_e_worker_training_amt |
| **No** | **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | During calendar year YYYY, how much money did (BUSINESS NAME) spend on [INTANGIBLE ASSETS ITEM]/Any other intangible asset investments? | f19c_f_other_amt |
| **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | Did [NAME BUSINESS] file for Chapter 11 bankruptcy protection at any time during calendar year YYYY? | f32_chap11_bankruptcy |
| **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | **No** | Now I'd like you to think about how much you expected [NAME BUSINESS] to grow since the business was started. How much do you think [NAME BUSINESS] met your expectations for growth between when the business was started and December 31, YYYY? Would you say [NAME BUSINESS]'s growth . . . | f33_expected_growth |

**Appendix B – Continued**

| The question appear in follow-up | | | | | | | | Question | Variable Name |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |
| **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | **No** | Compared to [NAME BUSINESS]'s revenues for calendar year YYYY, what do you expect [NAME BUSINESS]'s revenues will be in calendar year 2011? Do you think revenues will . . . | f34_future_revenue |
| **No** | **No** | **No** | **Yes** | **No** | **No** | **No** | **No** | What was the primary field of study for this degree? | g9_fieldofstudy_resp |
| **No** | **No** | **No** | **Yes** | **No** | **No** | **No** | **No** | What was the primary field of study for this degree? | g9_fieldofstudy_resp_2nd |
| **No** | **No** | **No** | **Yes** | **No** | **No** | **No** | **No** | | g9_fieldofstudy_cip_desc |
| **No** | **No** | **No** | **Yes** | **No** | **No** | **No** | **No** | | g9_fieldofstudy2nd_cip_desc |
| **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | What is your marital status? | g10b_marital_status |
| **No** | **No** | **No** | **No** | **Yes** | **Yes** | **Yes** | **Yes** | Including the equity in your home and business, what is your approximate total net worth, which are all your assets minus all debts? | g10c_net_worth |
| **No** | **No** | **No** | **No** | **Yes** | **No** | **No** | **No** | How much do you agree with the following statement? In uncertain times, I usually expect the best. Would you say you . . . | g10d_personal_outlook |