

**Finance and Economics Discussion Series  
Divisions of Research & Statistics and Monetary Affairs  
Federal Reserve Board, Washington, D.C.**

**Alternative Methods of Unit Nonresponse  
Weighting Adjustments:  
An Application from the 2003 Survey of Small Business  
Finances**

**Lieu N. Hazelwood, Traci L. Mach, and John D. Wolken**

**2007-10**

NOTE: Staff working papers in the Finance and Economics Discussion Series (FEDS) are preliminary materials circulated to stimulate discussion and critical comment. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors. References in publications to the Finance and Economics Discussion Series (other than acknowledgement) should be cleared with the author(s) to protect the tentative character of these papers.

**Alternative Methods of Unit Nonresponse  
Weighting Adjustments: An Application from  
the 2003 Survey of Small Business Finances**

Lieu N. Hazelwood, Traci L. Mach, and John D. Wolken\*

February 22, 2007

\*The views expressed herein are those of the authors. They do not necessarily reflect the opinions of the Federal Reserve Board or its staff.

## **1 Introduction**

In business surveys, as in all surveys, firms selected to participate often fail to respond to the survey. When left unaccounted for, this unit nonresponse can lead to biased estimates (both relative and absolute), inappropriate variances in the weights, as well as invalid confidence statements. However, there is no single accepted method to account for unit nonresponse. Whenever nonresponse adjustments are made, there is a risk of introducing further or different problems to the data.

The 2003 Survey of Small Business Finances (SSBF) screening interview had significant unit nonresponse and therefore some type of nonresponse adjustment was deemed necessary. The approach used in the 2003 survey differed from that used in previous surveys. The current paper uses the data from the 2003 SSBF to investigate the impact of various nonresponse adjustment methods on the weights, point estimates and variance of the estimates by comparing the approach ultimately implemented for the 2003 survey to alternative approaches, including one similar to what was done in the 1998 SSBF.

The remainder of this paper is organized as follows. We begin in Section 2 with a brief description of the SSBF, followed in Section 3 by a discussion of different nonresponse adjustment methods, highlighting the general advantages and the disadvantages of each method. We provide a detailed description of the four methods examined in this study in Section 4, explaining how each was applied to the 2003 SSBF data. We also describe the evaluation methodology. Section 5 provides the results of the four methods using the 2003 SSBF screener data and Section 6 concludes.

## 2 2003 SSBF

The 2003 SSBF was conducted to collect information from the owners of a nationally representative sample of up to 5,000 business enterprises. Small businesses were asked about financial relationships, credit experiences, lending terms and conditions, income and balance sheet information, the location and types of financial institutions used, and other firm characteristics.<sup>1</sup>

The target population was defined as for-profit, nongovernmental, nonfinancial, and nonagricultural enterprises with fewer than 500 employees that were either single establishments or the headquarters of a multiple establishment company. Firms also had to be in business during December 2003 and at the time of the interview, which for most firms occurred between June and December 2004.

Data collection had two phases: a screening phase to determine eligibility and a main interviewing phase. A stratified systematic sample of 37,600 businesses was selected from the Dun and Bradstreet market identifier file. The sample was stratified according to employment size, census division, and urban/rural status. In addition, the Standard Industrial Classification code (SIC) was used to sort the frame before systematic selection within each stratum. This implicit stratification helped improve the representativeness of the sample with respect to industry.

The screening was designed to verify the firm's eligibility to participate in the main study. During the main screening effort, screening interviews were attempted with 23,798 firms. 14,061 firms completed the screening interview, yielding an unweighted response rate of 59%.

---

<sup>1</sup> For detailed information about the 2003 SSBF, see 2003 Methodology Report (National Opinion Research Center, 2005). Selected survey results are summarized in Mach and Wolken, 2006.

This paper focuses solely on the unit nonresponse in the screener interview. A future study may be conducted which will deal with results from the main interview.

### **3 Nonresponse Adjustment Methods**

In this section, we describe four nonresponse weighting adjustment techniques for unit nonresponse: 1) traditional weighting class, 2) direct propensity weighting, 3) propensity stratification (Beck, and Bienias., 2000), and 4) 2003 SSBF hybrid propensity stratification method.

There are some similarities in the various approaches to calculating the nonresponse adjustment. Each method assumes that nonresponse occurs randomly at some level of aggregation. The nonresponse adjustment inflates the sampling weight initially assigned to the respondents in order to compensate for those observations lost through nonresponse. Consequently, both the respondents and the nonrespondents are represented through weight-adjusted respondents. The general assumption for all four nonresponse adjustment methods is that if it were possible to obtain the responses of the nonrespondents within each cell, their response would mimic exactly the responses of the respondents within the same cell.

#### *Traditional Weighting Class Adjustments*

The most widely used adjustment is the traditional weighting class adjustment. This method was employed in the 1998 SSBF. Traditional weighting class adjustments are based on the assumption that sample members can be partitioned into homogeneous cells, or weighting classes. These cells are usually formed using observable characteristics of respondents and nonrespondents that are thought to be correlated with survey responses. The weights for respondents are inflated by the inverse of the response rate within each cell. A weight of zero is

given to each nonrespondent. Given this weight adjustment, the summation of the nonresponse-adjusted weights is the same as the summation of original sampling weights. This method preserves cell counts as well as population counts (Hansen, Hurwitz, and Madow, 1953).

### *Propensity Scoring Methods*

More recently, nonresponse adjustments have been considered in the context of multiple-regression analysis, where the zero-one response indicator is regressed on a set of independent variables. The independent variables are generally similar to the design variables used to define the weighting class cells in the traditional weighting class method. The predicted value obtained from the regression equation, referred to as the propensity score, is the estimated response probability. All cases (population members) with the same observed characteristics, that is within the same cell, are assigned the same propensity score (Iannacchione, and Folsom, 1991).

Direct propensity weighting uses the propensity score to construct the nonresponse adjustments applied to the sampling weights for respondents. A spectrum of nonresponse adjustments can be constructed using the propensity scores. The simplest application assigns all respondents a nonresponse adjustment factor equal to the inverse of the mean propensity score; that is, there is a single adjustment factor for all respondents (one cell). The other extreme is to assign each individual case an adjustment factor equal to the inverse of its propensity score, allowing every respondent in the sample to have a different nonresponse adjustment factor ( $n$  cells, where  $n$  is the sample size). Another alternative is to partition the sample into “ $m$ ” cells and assign each respondent within each cell a nonresponse adjustment factor equal to the inverse of the mean propensity score of the cell; that is,  $m$  nonresponse adjustment factors are applied to the data. Weights of zero are assigned to all nonrespondents. Note that direct propensity weighting does not necessarily preserve cell or population counts.

An alternative to direct propensity weighting is propensity stratification. In this method, the propensity scores are used only to stratify the sample into  $m$  propensity classes (cells). The adjustment factor for the propensity stratification method can then be calculated as the inverse of the fraction of respondents in a cell (response rate). Although the adjustment cells are formed using a regression model to determine their boundaries, this adjustment closely mimics the traditional weighting class method. The nonresponse adjustment is applied to the sampling weights for the respondents and the nonrespondents are given a weight of zero. This method preserves cell counts within the propensity strata. However, the cells have no analytical meaning because they were formed solely by the propensity scores and do not correspond to the original design stratification variables. The sample population count is also maintained in this method.

Both weighting class and propensity stratification methods preserve population counts for the entire population and within each nonresponse-adjustment cell or propensity strata. In the weighting class method, the nonresponse-adjustment cells are often defined in terms of strata variables. If strata variables are used to define the nonresponse-adjustment cells and raking is implemented, then the sub-strata counts will be preserved. In propensity methods, the nonresponse-adjustment cells are not defined in terms of strata variables. Therefore, the nonresponse-adjustment cells most likely cross sampling-strata boundaries and as a result, the sub-strata counts may not be preserved after nonresponse adjustments. In cases where it is important to preserve sub-strata counts, a hybrid propensity-scoring weighting class procedure may be used. A logistic model is run on the entire data set and propensity scores are calculated. The sample is then divided into “super” strata based on one or more of the sampling stratification variables and then the cells are defined according to the propensity score within each “super”

stratum (propensity strata). This procedure preserves population counts in each propensity strata, sample partition (sample sub-strata), and the overall sample count.

Each of the nonresponse adjustment methods has advantages and disadvantages. The main advantage of implementing the traditional weighting class method for the 2003 survey is that there is prior knowledge gained from the 1998 survey (National Opinion Research Center, 2001). Even though our choice of cell collapsing rules and variables used in the adjustment might differ from previous surveys, the basic cell collapsing rules and the best possible variables used to form nonresponse adjustment cells have already been defined. The disadvantages of the traditional weighting class method compared to the propensity weighting or stratification methods are that (i) continuous variables must be collapsed into categorical variables, (ii) practically only a limited number of variables may be used, and (iii) cell collapsing may be necessary if the cell size is small. In contrast, the major advantages of propensity weighting or stratification methods compared to the traditional weighting class method are that (i) continuous variables can be used to define cells, (ii) a large number of variables may be used in the model and (iii) the technique is simple to apply. The 2003 SSBF hybrid stratification method shares the same advantages as the propensity weighting and stratification method, as well as maintains size class counts - -one of the major stratification variables. This method was deemed appropriate for the 2003 SSBF since there was a great deal of information available for both respondents and nonrespondents, including both categorical and continuous variables, and it preserved size class totals for future analysis.<sup>2</sup>

---

<sup>2</sup> There are several other types of weighting adjustments that are not discussed in detail in this paper. They include but are not limited to: weighting classes formed by prior survey knowledge, post stratification, raking procedures, and population weighting (Hansen, Hurwitz, and Madow, 1953).

The four methods can be described as either using stratification variables or propensity scores to classify nonresponse adjustment cells and using either the inverse of the response rate or the inverse of the propensity score as the nonresponse adjustment factor applied to the eligibility screened weight. The traditional cell weighting method uses stratification variables to formulate nonresponse adjustment cells. In contrast the direct propensity, propensity stratification, and the 2003 SSBF hybrid method uses the propensity scores to formulate nonresponse adjustment cells. The direct propensity score uses the inverse of the average propensity score as the nonresponse adjustment factor and does not preserve the sample count or the analysis cell counts. In contrast the traditional cell weighting, propensity stratification, and the 2003 SSBF hybrid method use the inverse of the response rate as the nonresponse adjustment factor and preserve the sample count and nonresponse adjustment cell counts.<sup>3</sup>

#### **4 Weighting Adjustments Used in This Study**

This section describes the implementation of four unit nonresponse adjustment techniques described in Section 3. Using the 2003 screener data, the screener nonresponse adjusted weights are calculated using four methods: (1) traditional weighting class adjustments; (2) direct propensity weighting; (3) propensity stratification; and (4) 2003 SSBF hybrid propensity stratification method.

##### *Weighting Class Method*

The first method we implemented is the traditional weighting class method. The adjustment cells could be defined by the original stratification variables (as received from D&B) and the credit score percentiles because they are available for all cases. Defining cells by

---

<sup>3</sup> Although the propensity stratification method preserves nonresponse adjustment cell counts, these cells are not useful for analysis purposes and should not be confused with the analysis cell counts based on stratification variables.

sampling strata or other criteria related to response patterns relaxes the random nonresponse assumption somewhat, allowing response rates to vary by the characteristics of the adjustment cells.

Experience in the 1998 SSBF indicated that response rates may vary by state or by smaller size classes than those employed in the 2003 SSBF sampling stratification. In the 1998 SSBF screener, the nonresponse adjustment was based on 7 size classes (0-2, 3-4, 5-9, 10-19, 20-49, 50-99, 100-499, and 500 or more employees), state and urban/rural status. Similar to 1998, we define the nonresponse adjustment cells using the four size classes (0-19, 20-49, 50-99, 100-499 employees), census division, credit score percentiles and urban/rural status. Other variables, such as minority indicator, are not available for all respondents at the screener stage.

Correlations between the response rate and the stratification variables (size class, census division, and urban/rural status) were compared. Adjustment cells were formed by examining the correlation with response rates and stratification variables and collapsing groups, when necessary, into strata with similar response rates.

Since the response rate among very small groups can vary dramatically and cause an increase in the variability of the weights, cell collapsing was required. Like the 1998 survey, the 2003 adjustment cells were collapsed if either the following conditions were met: (1) a cell had fewer than 20 cases or (2) a cell had a collapsing factor greater than 2.0, where the collapsing factor for a cell is defined by the following two rules:

$$\text{collapsing factor}_{\text{respondents}} = \frac{\sum_c w_i}{\sum_c w_i \delta_i},$$

$w$  = screener eligibility adjusted weights,

$i = 1, 2, \dots, n_c$ ,

$\delta = 1$  for respondents,

and  $\delta = 0$  for nonrespondents

$$\text{collapsing factor}_{\text{nonrespondents}} = 0$$

The response rate was calculated for each nonresponse adjustment cell,  $c$ , and the inverse of the response rate was applied to the screener eligibility-adjusted weight.<sup>4</sup> Cell collapsing was implemented as necessary. A raking procedure was applied to preserve the original cell totals within the size strata.<sup>5</sup>

---

<sup>4</sup> Thirty-one percent of firms who completed the screener were ineligible to participate in the 2003 SSBF. Some portion of the nonrespondents were expected to have been ineligible had they completed the screener. The weights associated with the businesses at this point are referred to as the eligibility-adjusted weights. For more information on how the initial sampling weights were adjusted to reflect this ineligibility refer to the 2003 Methodology Report (National Opinion Research Center, 2005).

<sup>5</sup> An evaluation of the pre- and post-raked weights was performed. There were minimal differences between the overall screener nonresponse adjusted weight before and after raking. There was a slight increase in the variance of the post-raked screener nonresponse adjusted weight (181,810) compared to the pre-raked screener nonresponse adjusted weight (181,173). Also the pre- and post-raked weights showed small differences at the 2003 SSBF strata level, which was defined by the cross classification of total employee size class, census division, and urban/rural status.

### *Propensity Scoring Methods*

Each of the three propensity methods began by estimating the propensity to respond using a logistic model. A logistic regression model was fit to predict the response propensity and the propensity score was used to form nonresponse adjustment cells. The adjustment was carried out within each adjustment cell. The adjustment factor for the respondents was the inverse of the average propensity score or inverse of the response rate for each nonresponse adjustment cell. All nonrespondents were given an adjustment factor of zero. Variables were limited to information available from the frame and credit score percentiles. The model used the urban/rural indicator, state, organizational form, status indicator (single location, head quarters, or branch), size of firm, credit score percentiles and industry classification.

For each case, an estimated response propensity was obtained from the model. The sample was then sorted in ascending sequence by propensity score. The sample was further divided into five equal sized strata based on propensity scores. These strata were the nonresponse adjustment cells<sup>6</sup>. The direct propensity method used the inverse of the average propensity score in each nonresponse adjustment cell as the weight adjustment for the respondents. An adjustment factor of zero was applied to the nonrespondents.

The propensity stratification method used the inverse of the weighted screener response rate per adjustment cell as the weight adjustment for each respondent. An adjustment factor of zero was applied to the nonrespondents.

For the 2003 SSBF hybrid propensity method, firm size was deemed an important design and analysis variable and it was decided that nonresponse adjustments should preserve size class

---

<sup>6</sup> Comparisons of point estimates indicate the extent to which the adjustment estimates are sensitive to the number of adjustment cells chosen. For five or more adjustment cells, the reported point estimates are relatively similar. This is consistent with the suggestion that five cells may provide the most effective bias reduction to be obtained from a given cell-formation method (Rosenbaum and Rubin, 1983).

counts. The nonresponse adjustment cells (propensity strata) were formed within four total firm employment size class partitions (“super” strata) based on propensity scores. A total of 40 nonresponse adjustment cells were formed. Size class 0-19 was divided into 25 cells and each of the remaining size classes were divided into 5 cells. The weight adjustment was the inverse of the response rate in each nonresponse adjustment cell for the respondents and the nonrespondents received an adjustment factor equal to zero.

## **5 Results**

In this section, we present the analysis of key point estimates, the variance of the weights and the variance of the point estimates for unit nonresponse weighting adjustments.

Comparisons for the total number of firms, the variance of the screener nonresponse adjusted weights for the entire universe and certain subcategories, the average number of employees and the variance of the number of employees are presented for each of the four methods. It is important to note that this study only deals with the screener interview and only a few analysis variables were available at that time.

Table 1 illustrates the comparison of total firms and the variance of the screener nonresponse adjusted weight. The traditional cell weighting, the propensity stratification, and the 2003 SSBF hybrid methods all preserve the same total number of firms (6,647,602) with similar variances. The direct propensity weighting method increased the overall sample count (7,148,235) and also displayed the largest variance.

Table 2 shows how the total number of firms and the percent of the total firms by business type (sole proprietorship, partnership, and corporation) varies by adjustment method. Our study shows small differences among the four methods. There is an increase in the number of firms and the percentage of sole proprietorships for the 2003 SSBF hybrid method (3,539,360

firms and 53.24%) relative to the traditional cell weighting method (3,389,465 firms and 50.99%).

Table 3 reports the variance of the screener nonresponse adjusted weights by organizational form for each of the four adjustment methods. The variance of the screener nonresponse adjusted weights by business type appears to be quite similar for the four methods. The 2003 SSBF hybrid method shows an increase in variance for the 2003 sole proprietorships (284,922 versus 264,989) but a decrease in variance for the partnerships (146,757 versus 176,135) and corporations (114,799 versus 125,665) compared to the traditional cell weighting method.

Table 4 illustrates the comparison of the total number of firms and the percent of the total firms by size classes (0-19, 20-49, 50-99, and 100-499 employees). The distribution of firms by size class remains similar for the four methods. The number of firms and distribution of firms by size class remains the same for the 2003 SSBF hybrid method and the traditional cell weighting. This is because the traditional cell weighting is raked at the firm size class level and the sample is stratified before forming the propensity strata. The propensity stratification method shows a decrease in the number and the percentage of firms for the 0-19 employee size class (6,183,346 versus 6,211,287) and the 100-499 employee size class (49,434 versus 50,522) but an increase for the 20-49 employee size class (327,514 versus 300,705) and the 50-99 employee size class (87,308 versus 85,088) compared to the 2003 SSBF hybrid method and the traditional cell weighting method. These differences are minimal. The differences between all the size class totals for the direct propensity weighting method compared to the other three methods can be explained by the inflation of the overall sample count calculated from the direct propensity weighting method.

The variance of the screener nonresponse adjusted weights by size class across adjustment methods are presented in Table 5. Except for the direct propensity weighting method, the differences in the variance of the screener nonresponse adjusted weight by size class are minimal. The 2003 SSBF hybrid method shows an increase in variance for the 100-499 employee size class (290 versus 285) and the 0-19 size class (278,114 versus 278,087) compared to the traditional cell weighting method. The 2003 SSBF hybrid method shows a decrease in the variance for the 20-49 employee size class (13,978 versus 14,044) and the 50-99 employee size class (794 versus 802) compared to the traditional cell weighting method. The direct propensity weighting method results in greater weight variance for all four size classes than any of the other three methods.

Table 6 presents results for the average number of employees and the variance of the number of employees reported in the screener interview. The mean number of employees is virtually unchanged for all four methods (approximately 9) and the variance remains similar for the four methods. There is a reduction in variance for the 2003 SSBF hybrid method (398,019) compared to the traditional cell weighting method (424,719). Point estimates for the average credit score and variance of credit score are listed in Table 7. The average credit score (51-53) and variance of the credit score between all four methods is very similar. There was a slight decrease in the average credit score for the 2003 SSBF hybrid method (52.20) compared to the traditional cell weighting method (53.18). There was a slight increase in the variance of the credit score for the 2003 SSBF hybrid method (309,117) compared to the traditional cell weighting method (299,986).

## 6 Conclusions

Because the 2003 SSBF oversampled firms with more than 50 employees, it was deemed important to preserve within-size class population totals. While traditional weighting class adjustments are capable of doing so, this adjustment technique limits the number of variables that can be practicably be used to form the adjustment classes. Propensity scoring models, on the other hand, allow a large number of variables to be used to form the adjustment cells, thereby improving the likelihood that the responses of the respondents mimic those of the nonrespondents they represent in the final sample. However, traditional propensity scoring models do not preserve the within-size class population totals. For the 2003 SSBF, a hybrid method was used to make nonresponse weighting adjustments.

The current paper examines the impact of this technique on weights, point estimates and variance of the estimates by comparing the approach ultimately implemented for the 2003 survey to alternative approaches. The results using the 2003 SSBF hybrid method are very similar to the traditional weighting class method and the propensity stratification methods. Even though the hybrid technique did in some instances increase the variance of the weights over the traditional weighting class adjustment method, the differences were quite small. In addition the hybrid method decreased the variance for the weights as well as the some of the point estimates. The variance reduction by the hybrid method relative to the direct propensity method can most likely be attributed to the stratification. From an implementation standpoint, the 2003 hybrid method was only slightly more difficult to implement than the direct propensity method, but easier and less time-consuming than the traditional weighting class method. In sum, the 2003 SSBF hybrid method performed similarly to the traditional weighting class method, controlled for population and size class totals, easily accommodated a large number of variables to

determine the response adjustment cells and was only slightly more difficult to implement than the direct propensity method while decreasing the variances in the weights.

Table 1. Total Firms and Variance of the Screener Non-Response Adjusted (NRADJ) Weight

Non-Response Adjustment	Total Number of Firms	Variance of Screener NRADJ Weight
Traditional Cell Weighting	6,647,602	181,955
Direct Propensity	7,148,235	235,005
Propensity Stratification	6,647,602	180,193
2003 SSBF Hybrid Method	6,647,602	181,810

Table 2. Total Number of Firms and Percent of Total Firms by Business Type

Non-Response Adjustment	Sole Proprietorships	% of Total	Partnerships	% of Total	Corporations	% of Total
Traditional Cell Weighting	3,389,465	50.99	401,375	6.04	2,856,761	42.97
Direct Propensity	3,851,927	53.89	386,725	5.41	2,909,584	40.70
Propensity Stratification	3,513,706	52.86	373,851	5.62	2,760,045	41.52
2003 SSBF Hybrid Method	3,539,360	53.24	369,679	5.56	2,738,563	41.20

Table 3. Variance of the NRADJ Screener Weight by Business Type

Non-Response Adjustment	Sole Proprietorships	Partnerships	Corporations
Traditional Cell Weighting	264,989	176,135	125,665
Direct Propensity	364,123	176,115	147,474
Propensity Stratification	279,904	149,351	115,165
2003 SSBF Hybrid Method	284,992	146,757	114,799

Table 4. Total Number of Firms and Percent of Total Firms by Size Class

Non-Response Adjustment	0-19 Employees	% of Total	20-49 Employees	% of Total	50-99 Employees	% of Total	100-499 Employees	% of Total
Traditional Cell Weighting	6,211,287	93.44	300,705	4.52	85,088	1.28	50,522	0.76
Direct Propensity	6,661,437	93.19	345,043	4.83	90,784	1.27	50,973	0.71
Propensity Stratification	6,183,346	93.02	327,514	4.93	87,308	1.31	49,434	0.74
2003 SSBF hybrid method	6,211,287	93.44	300,705	4.52	85,088	1.28	50,522	0.76

Table 5. Variance of the NRADJ Screener Weight by Size Class

Non-Response Adjustment	0-19 Employees	20-49 Employees	50-99 Employees	100-499 Employees
Traditional Cell Weighting	278,087	14,044	802	285
Direct Propensity	347,753	20,795	1,041	353
Propensity Stratification	275,311	16,943	850	281
2003 SSBF hybrid method	278,114	13,978	794	290

Table 6. Average Number of Employees and the Variance of the Number of Employees

Non-Response Adjustment	Mean	Variance
Traditional Cell Weighting	8.94	424,719
Direct Propensity	8.75	419,941
Propensity Stratification	8.86	395,131
2003 SSBF hybrid method	8.78	398,019

Table 7: Average Credit Score and the Variance of the Credit Score

Non-Response Adjustment	Mean	Variance
Traditional Cell Weighting	53.18	299,986
Direct Propensity	51.12	334,372
Propensity Stratification	52.37	309,130
2003 SSBF hybrid method	52.20	309,117

## References

- Beck, Todd L. and Bienias, Julia L. (2000). Comparison of Nonresponse Adjustment Methods in Health Survey. *Proceedings of the Annual Meeting of the American Statistical Association, Section on Survey Research Methods*, pp.630-635.
- Hansen, M.H., Hurwitz, W.N., and Madow, W.G. (1953), Sample Survey Methods and Theory, Vol II. John Wiley & Sons, New York.
- Iannacchione, V.G., Milne, J.G., and Folsom, R.E. (1991). Response Probability Weight Adjustments using Logistic Regression, *Proceedings of the Annual Meeting of the American Statistical Association, Section on Survey Research Methods*, pp.637-642.
- Mach, Traci L, and Wolken, John D. (2006). Financial Services Used by Small Businesses: Evidence from the 2003 Survey of Small Business Finances, *Federal Reserve Bulletin*, pp. 167-195.
- National Opinion Research Center (April 2001). "The 1998 Survey of Small Business Finances Methodology Report," mimeo,  
<http://www.federalreserve.gov/pubs/oss/oss3/ssbf98/ssbf98home.html#ssbf98results>
- National Opinion Research Center (April 2005). "The 2003 Survey of Small Business Finances Methodology Report," mimeo,  
<http://www.federalreserve.gov/pubs/oss/oss3/ssbf03/ssbf03home.html#ssbf03results>
- Rosenbaum, P.R. and Rubin, D.B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*. 70, 41-55.