EFFECTS OF THE LEVEL OF INTERVIEWER EFFORT ON THE CHARACTERISTICS OF COMPLETED RESPONSES: AN EXPERIMENT USING THE 1998 SURVEY OF SMALL BUSINESS FINANCES

Marianne P. Bitler and John D. Wolken
NICHD Post-Doctoral Fellow, RAND, Federal Reserve Board of Governors

Abstract

All surveys contain a mix of interviews that required more or less effort to complete. Difficult cases are often assigned to interviewers specializing in refusal conversion or are given incentives to cooperate. Use of these refusal conversion experts can increase response rates among hard to reach subjects. Analysis not explicitly controlling for interviewer effort assumes that the easy to interview and hard to interview cases are statistically identical except for difficulty in completing the interviews. Violations of the assumption of "no interviewer treatment effect" question the validity of comparisons not controlling for ease of interview completion. The 1998 Survey of Small Business Finances allows us to test this assumption. We present analyses suggesting that interviewer effort effects are not important for this data set. We also present analysis of the effects of incentives on interview difficulty.

Key Words: Business Survey, Refusal Conversion Experts, Incentives

Introduction

This paper examines the effects of interviewer effort on the characteristics of completed interviews in the 1998 Survey of Small Business Finances (SSBF), an enterprise level survey conducted by the National Opinion Research Center (NORC) and sponsored by the Board of Governors of the Federal Reserve System. Most analysts assume that completed interviews drawn from different parts of the distribution of interviewer effort are comparable. However, if difficulty in completing the interview is correlated with other characteristics of the firms, then firms that were easy or hard to interview may not be comparable. Due to ex ante assignment of a subset of refusal conversion experts and other specialists to minority interview strata in the 1998 SSBF, we can test this assumption.

Background on the SSBF and the Minority Sub-Sample

The 1998 SSBF is the third in a series of surveys of small firms sponsored by the Board of Governors of the Federal Reserve System. One use of the SSBF will be to assess the availability of credit for small and minority owned businesses.[1] Along with demographic information about the owners of the firms, data are collected on the firms' financial relationships, credit experiences, lending terms and conditions, income and balance sheet information, and the characteristics of the financial institutions used.

One of the goals of the 1998 SSBF is to collect information about minority-held businesses, which are rare.[2] The 1993 National Survey of Small Business Finances found that only 2.9% of businesses were

---

1. Small businesses are defined as non-financial, non-farm, for-profit, privately-owned businesses with fewer than 500 employees.
2. For the 1998 SSBF, minority firms are defined as those with more than 50 % of their shares held by Hispanics (an ethnic category) or by Asians or African-Americans (racial categories).

owned by African-Americans, 3.4% by Asians, and 4.3% by Hispanics (Cole and Wolken, 1995). The rare incidence of minority-owned businesses led to their being oversampled in the 1998 SSBF. Larger small businesses (whether minority-owned or not) are also rare and were also oversampled.[3] The list frame for drawing the sample was Dun & Bradstreet's "Dun's Market Identifiers®" file, which does not contain a reliable indicator for minority-owned businesses (Haggerty et al., 2001). Thus, a sample of 39,240 firms was screened for eligibility, race, and ethnicity. Expectations were that these 39,240 firms would yield 6,000 completed interviews with eligible firms, with over 400 completed interviews from each minority group and from larger non-minority firms. Early experience in the field signaled a serious decline in response rates relative to previous rounds of the survey. Since credible statistical comparisons among subgroups require sufficient sample size, there was extra concern about the ability to complete sufficiently many interviews with minority-owned and larger firms. The final sample contained 3,561 eligible completed interviews.

Incentive Experiments

NORC tried a variety of experiments to increase response rates. A small number of initial refusals were offered monetary incentives to complete the interviews (post-incentive cases). A larger number of newly released cases were offered an incentive in the form of a check for $20 before the interview started (pre-incentive cases).[4] We will investigate below the effects of the incentives on characteristics of the final eligible completed interviews.

Refusal Conversion Experts and Assignment of Minority Pool Interviewers to Minority Interviews

Like any large survey organization, NORC has refusal conversion experts among its interviewers. These are individuals deemed particularly skilled at convincing reluctant respondents to complete interviews. In the hopes of boosting minority response rates which had historically been low, NORC assigned some refusal conversion experts (RCEs) skilled at interviewing minorities to focus on interviewing minority-owned firms (minority pool interviewers). Some regular interviewers with language skills were also assigned to the minority interview pool. As with any survey, one hopes use of RCEs makes it possible to collect information from respondents from the entire distribution of willingness to participate. Without representation from the harder to convince participants, the information collected would not be representative of the population. One also hopes that designating some interviewers (those in the minority pool) to concentrate primarily on minority interviews did not disproportionately induce more reluctant minority strata respondents to complete the interview (relative to non-minority strata respondents).

Classification of Interviewers and Interviews

An interviewer can be either a response converter (RCE) or an ordinary interviewer (ORD). Additionally, the interviewer could have been assigned to the specific minority strata (IWPOOL). The

---

3. The major strata for the main 1998 SSBF interview (derived from screener interview values) were 3 minority strata (Hispanic, Asian non-Hispanic, and African-American non-Asian non-Hispanic firms), and 4 non-minority strata (firms with 1-19 employees, 20-49 employees, 50-99 employees, and 100-499 employees. See Haggerty et al. (2001b), especially chapter VIII.

4. See Groves (1989) on incentives and interviewer effects or Knopf et al. (2000) for a discussion of the use of incentives in telephone surveys.

categories of RCE and IWPOOL are not mutually exclusive. Some minority pool interviewers are not otherwise RCEs, but were assigned as minority pool specialists because of other characteristics, such as language expertise. Furthermore, the IWPOOL interviewers can complete interviews for other strata than the specific minority one they were assigned. Finally, there is a chance that interviews from majority strata could differ from those in minority strata if only because of the different routing mechanism for minority interviews (namely that many of them were steered to the IWPOOL interviewers). We could expect as many as four main effects on the outcome of the interview and subsequently possibly on the characteristics of completed interviews. These are: 1) the effect of the interview being a majority strata interview (MAJ); 2) the effect of the interview having been completed by a response converter (RCE); 3) the effect of an interview having been completed by a IWPOOL interviewer when the interview was not in that interviewer's assigned minority strata (IWPNOTSPEC); and 4) the effect of the interview having been completed by an IWPOOL interviewer in that interviewer's assigned minority strata (IWPSPEC).

A priori, we expect RCE interviewers to be more efficient than other interviewers at completing interviews. Clearly, the fact that the IWPOOL interviewers were specially assigned to certain minority strata implies that the IWPOOL interviewers will most likely be more efficient than other interviewers at completing interviews in their assigned minority strata (for IWPSPEC interviews). We do not have such expectations about the effect of being an IWPNOTSPEC interview. Finally, the minority pools were created in anticipation of the minority interviews being difficult to complete. In order to analyze the possible effects of assigning more persuasive interviewers to minority strata, the contractors monitored the workload of these minority IWPOOL interviewers to make sure that they also completed some non-minority interviews. As an artifact of the interview room technology, other interviewers (both RCE and non-RCE as well as IWPOOL not in their designated special strata) completed some minority interviews.

Distribution of Interviewers and Interviews

Of the 60 regular interviewers who completed interviews in the final data, 20 were RCEs and 12 were assigned to the minority pool (9 of these 12 were RCEs).[5] Out of the 746 minority-strata interviews that were delivered as final eligible interviews, 450 cases were completed by IWPOOL interviewers in their specialized minority strata (IWPSPEC) and another 70 cases were completed by IWPOOL interviewers not in their assigned minority strata (IWPNONSPEC). Of the 746 minority interviews, 478 cases were completed with interviewers that were response converters (RCEs), many of which were also IWPOOL interviewers. More precisely, 390 cases were completed by interviewers who were RCEs and in the IWPOOL (of these 390, 329 were in the assigned strata of the IWPOOL interviewer or were IWPSPEC cases). There were 130 cases that used IWPOOL interviewers who were not RCEs, 88 cases that used interviewers that were non-IWPOOL RCEs, and 138 cases that used ordinary interviewers. For the non-minority strata, referred to as the majority (MAJ) strata, 1,621 interviews were completed with RCEs and 1,194 with non-RCEs.   Although there was no analog to the special IWPOOL for the majority strata, there were 908 interviews completed by the IWPOOL interviewers in the non-minority strata. All of these figures are calculated for the last interviewer to work a case and are calculated for eligible complete cases. More than one interviewer may have interacted with a particular firm. Future work will examine the possible effects of an RCE

---

5. A supervisor, operator, or editor (rather than a regular interviewer) completed 148 of the 3,561 interviews. All the regressions include a dummy variable for the interview being one of these 148 cases.

or IWPOOL interviewer having started the interview. It will also expand the current analysis to consider whether there were effects of interviewers on response and time to completion for all cases in the main interview rather than the subset of eligible complete cases considered here.

In addition to the four main effects discussed above (RCE interviewer, MAJ interview, IWPSPEC interview, IWPNONSPEC interview), there are four two-way interactions possible (RCE interviewer and IWPSPEC interview; RCE interviewer and IWPNONSPEC interview, RCE interviewer in non-minority strata, and IWPOOL interviewer not in assigned minority strata and non-minority strata interview), as well as one three-way interaction (RCE interviewer, IWPOOL interviewer not in assigned minority strata, non-minority strata interview).[6]

In a regression framework, the main effects can be identified by including four dummy variables: one for the interview having been completed by an RCE; one for the interview having been completed by an IWPOOL interviewer; one for the interview having been a minority or non-minority strata firm (MAJ); and one for the interview having been completed by an IWPOOL interviewer in their assigned minority pool (IWPSPEC interview). The interactions will be interactions between these variables. Including a dummy variable for the interview being done by an IWPOOL interviewer and one for the interview being an IWPSPEC interview is computationally equivalent to including dummies for the IWPSPEC and IWPNONSPEC interviews. In this framework, the coefficient on the IWPOOL dummy represents the effect of having an IWPOOL interviewer complete an interview from a stratum other than their specially assigned stratum (IWPNONSPEC) while the sum of the coefficients on the IWPOOL dummy and the IWSPEC dummy represents the effect of having and IWPOOL interviewer complete an interview from their assigned minority stratum.

If the characteristics of completed interviews do not vary between any of these groups in a systematic fashion, it is evidence that there may not be an interviewer effort effect. A significant effect of being a non-minority interview will be evidence that non-minority strata firms have different characteristics than minority strata firms. Evidence of effects on interview length of having an RCE interviewer or IWPOOL interviewer in their assigned minority strata may show that using RCE or IWPOOL interviewers has cut non-response or increased the speed of completion. Furthermore, evidence that these main effects or interactions are significantly associated with completed interview characteristics may raise concern that their use has affected response.

Analytic Framework

For this analysis, we compare the distribution of various firm characteristics as a function of controls and the four main effects as well as interactions in some specifications. Within the major strata, we calculated medians for the number of employees, sales last year, number of owners, Dun & Bradstreet credit rating, firm age, number of calls to complete the main interview, and number of calendar days to complete the interview. Within these strata, we also constructed indicators of whether the value for an individual firm was greater than the median for that firm's sampling stratum. These indicators are the firm characteristics of interest. The regressions predicting firm characteristics allow us to see if the use of RCEs and IWPOOL interviewers influenced the final distribution of eligible completed interviews. The regressions predicting interview length (measured by the number of calls and

---

6. The rest of the usual three-way interactions and the four way interactions possible with four main effects are identical to some of the above two-way and three-way interactions because the groups IWPNONSPEC and IWPSPEC are mutually exclusive.

calendar time) allow us to see if the use of RCEs and IWPOOL interviewers shortened the interview length. The regressions are run for two samples; all eligible completed interviews and minority-strata eligible completed interviews. For comparison purposes, Table 1 contains means for firm characteristics as well as the results of T-tests for the equality of means between RCE completed interviews and non-RCE completed interviews.

We ran logistic regressions of the form

$$Y = \alpha \,'I + \beta \,'X + \chi \, MAJ + \delta \, RCE + \eta \, IWPSPEC + \phi \, IWPOOL + \gamma \,'Z + \varepsilon.$$

Y is some characteristic of the firm. MAJ indicates the interview was from the non-minority strata, RCE that the interview was completed by an RCE, IWPSPEC that it was completed by an IWPOOL interviewer in her assigned minority strata, IWPOOL that it was completed by an IWPOOL interviewer in any strata; all of these are dummy variables. I is a vector of interviewer dummy variables for interviewers who completed more than 3% of the interviews. X is a vector of other controls. Z includes interactions among the main effects and is not always present in the regressions. When the sample is restricted to the minority strata interviews, the MAJ dummy and all interactions including it are inapplicable to the model.

The control variables (X) include dummy variables for the number of employees of the firm in categories (except when the outcome is having larger than median employment); for whether the firm was offered an incentive of either kind to participate; for the Census region in which the firm is located; for whether the firm was located in an MSA; for the race, ethnicity, and sex of the firms' owners holding more than 50 percent of the shares (African-American or Asian, Hispanic, male owner); for whether the screener interview took more than the median number of calls to complete (the screener was difficult to complete); and for when the case was released. The regressions are adjusted for the complex nature of the sample and use the final weights (which adjust for non-response and ineligibility).[7] The seven strata were derived from the screener interview with the firm. There were almost no non-white Hispanic firms.

---

7. We chose to use the final sampling weights when analyzing these data because the data represent only the eligible completed interviews; thus unweighted sample means unadjusted for the complex nature of the sample would not be representative of the full population of small firms.

Means– Table 1

The means are presented in Table 1. T-tests (reported in column 3) show that the firms whose interviews were completed by RCEs and those whose interviews were completed by non-RCEs do not differ along most dimensions in a statistically significant fashion. RCE completed interviews were significantly more likely (at 1% level) to have been with minority pool assigned interviewers whether in their assigned strata or in any strata (not surprising since 9 of the 12 IWPOOL interviewers were RCEs). The only other significant differences between RCE completed interviews and non-RCE completed interviews are that the RCE completed interviews were more likely to have received a pre-incentive (5% level) and were more likely to have larger than median sales (10% level). The RCE completed interview means did not differ from the non-RCE completed interview means for any of the other firm or interview characteristics. The T-tests in Table 1 suggest that RCE interviews do not differ from non-RCE interviews along most observable dimensions.

Results from Logistic Regressions

This section discusses the results of the logistic regressions predicting firm characteristics or interview characteristics as a function of the IWPOOL and RCE status of the interviewers. The four main effects of RCE, MAJ, IWPOOL and IWPSPEC were all insignificant (as were the pre-incentive indicator and the indicator for the screener interview having been difficult) in the regressions for having larger than median employment, larger than median sales, larger than the median number of owners, being older than the median firm, and having a higher than median credit score. With a few exceptions listed below, this held in both the full sample and the minority strata sub-sample with or without the interaction terms.

One exception was that in the full sample, having had an RCE interview complete the interview was positively associated (at the 10% level) with having had larger than median sales last year; this was no longer true when the interactions were included. Another exception was having had a pre-incentive which was positively associated with being an older firm at the 10% level; this also held when the interactions were included. In the minority sub-sample, having had a pre-incentive was negatively associated with the firm having higher than median employment both with and without interactions. In the minority sub-sample, having had a hard screener was negatively associated with having larger than median sales (but only at the 10% level). Again, in the minority sub-sample, having had an RCE complete the interview was associated (again at the 10% level) with the firm having larger than the median number of owners when interactions were included, but not when they were absent. Finally, in the minority sub-sample having had a hard screener was negatively associated with the interview having taken a long time in calendar days to complete (again only at the 10% level).

The fact that the main effects (and interactions) are almost never significant in these regressions suggests that having had a IWPOOL or an RCE interviewer complete the interview had no effect on the median firm characteristics of eligible completed interviews. The fact that there are a few instances when the coefficients are significant does not significantly weaken this conclusion since there is no systematic pattern to the exceptions.

The only regressions where any of these interviewer effects are significant consistently across both samples and both specifications are those predicting whether the number of calls to complete the interview was larger than the median. Table 2 contains the coefficients and standard errors for

coefficients of interest from logistic regressions of whether the interview took more than the median number of calls for both samples (all firms and minority strata firms) and both specifications (main effects only and interaction effects included). Table 2 also contains F-tests (and degrees of freedom) for whether the interviewer dummies were jointly significantly different from zero. It also contains F-tests for whether the sum of the coefficient on being an IWPOOL interview and the coefficient on being an IWPSPEC interview is zero, which is a test of whether being an IWPOOL interview in the assigned minority strata is significant.

We do not expect the RCE completed interview coefficients, or the sum of the IWPSPEC and IWPOOL coefficients (the effect of the IWPOOL interviewer in their assigned minority strata completing the interview) to be significant if the "no interviewer effect" hypothesis is correct. Similarly, we do not expect the coefficient on having been a hard screener to complete to be significantly different from zero except in the regressions where the dependent variable measures how hard the interview was to complete (e.g., how long it took in days or how many calls it took); for this regression we anticipate a positive and significant coefficient as firms that are hard to screen will likely be hard to interview. We expect that incentives will increase cooperation. Since only 16 firms received the post-incentive, we do not expect its coefficient to be significant. If the dependent variable is whether it was a difficult interview to complete (measured by the number of calls exceeding the median calls in a stratum), we expect the coefficient on the pre-incentive payment dummy to be negative.

Table 2 shows that having received a pre-incentive made the interview easier to complete when difficulty of completion is measured by number of calls. This held for both the full sample and the minority strata sub-sample with and without the interaction terms. Since the releases of batches of new cases were not uniform during the field period, the queue was more or less crowded at different times, perhaps explaining why this effect was not present when ease of completion was measured by how many days the case took to complete. The interaction terms are not jointly significant in these regressions in either sample, suggesting that they can be safely excluded from the regressions. The individual interview effects are jointly significant in all four regressions. Finally, in all four regressions, the coefficient on being a minority pool interviewer (IWPOOL) is positive and significant. Since this coefficient is computationally equivalent to one for having been an IWPOOL interviewer completing an interview in a strata other than the specially assigned one (coefficient on IWPNONSPEC), this suggests that having the IWPOOL interviewers complete interviews in strata other than their assigned ones was associated with the interview taking longer. The effect of the interview being completed by an IWPOOL interviewer in their assigned minority strata is the sum of the coefficients on IWPOOL and IWPSPEC; this is never significant according to the F-tests in the final row of table 2.

Conclusion

The results presented here are encouraging. The fact that the interviewer effects are jointly significant suggests that some interviewers differentially affect the number of calls that an interview takes. However, the logistic analysis of firm characteristics as a function of RCE or IWPOOL status of the interviewer suggest that there was no significant effect of the RCE and IWPOOL interviewers on the distribution of characteristics of firms who were eligible and completed the interviews. This analysis also suggests that incentives reduce the number of calls necessary to complete the interview as has been found in other settings.

Acknowledgements and Disclaimer

References

Cole, R. A. and J. D. Wolken (1995), "Financial Services Used by Small Businesses: Evidence from the 1993 National Survey of Small Business Finances," Federal Reserve Bulletin, 81, pp. 629-67.

Elliehausen, G. E. and J. D. Wolken (1990), "Banking Markets and the Use of Financial Services by Small Businesses and Medium-Sized Businesses," Federal Reserve Bulletin, 76, pp. 801-17.

Groves, R. M. (1989), Survey Errors and Survey Costs, New York: John Wiley & Sons, Inc.

Haggerty, C., K. Grigorian, R. Harter, and J. Wolken (2001a), "The 1998 Survey of Small Business Finances: Sampling and the Level of Effort Associated with Gaining Cooperation from Minority-Owned Businesses," Conference Proceedings, International Conference on Establishment Surveys II.

Haggerty, C., K. Grigorian, R. Harter, and A. Stewart (2001b), "The 1998 Survey of Small Business Finance Methodology Report," Unpublished mimeo, http://www.federalreserve.gov/pubs/oss/oss3/ssbf98/ssbf98home.html#ssbf98results.

Knopf, M. E., J. Schieb, and J. Blair (2000), "The Effects of Alternative Incentives on Cooperation and Refusal Conversion in a Telephone Survey," Unpublished mimeo, University of Maryland Survey Research Center.

Table 1: Population Means of Firm and Interview Characteristics

Characteristic	Overall Population Mean	Standard Error of Mean	T-statistic Difference of Means between Response Converter Interviews and Ordinary Interviews

Used Response Converter	0.580	0.010	N.A.
Minority Pool Interviewer-Any Strata	0.340	0.009	-23.57***
Minority Pool Interviewer- Assigned Minority Pool	0.082	.003	-5.10***
Non-minority Strata	0.862	.002	1.42
Number of calls, main interview	12.655	0.216	1.40
Number of calls, screener interview	3.029	0.045	-0.08
Offered pre-incentive	0.360	0.010	-1.96**
Offered post-incentive	0.005	0.001	0.63
Larger age than median	0.464	0.010	-0.29
Larger than median credit score	0.425	0.010	-0.32
Larger than median number of owners	0.285	0.009	-0.55
Larger than median number of calls (main)	0.463	0.010	0.70
Larger than median number of calls (screener)	0.440	0.010	-0.32
Larger than median sales	0.475	0.010	-1.71*
Larger than median number of employees	0.430	0.010	-1.34
Larger than median number of days between start and end of interview	0.311	0.009	1.45

Notes: The population mean estimates and standard errors presented in this table are constructed accounting for sample weighting and stratification. All variables labeled "Larger than…" are calculated by determining if the case value is larger than the median value of the case's major strata. The last column contains the T-statistic for a test of differences of means between interviews that used a response converter specialist (RCE) and interviews that used ordinary interviewers, assuming equal variances. The difference is calculated as mean of non-response converter interview minus mean of response converter interview. A negative sign on the T-statistic implies that the RCE mean is greater than the non-RCE mean. *, **, and *** indicate the means are significantly different at the 10, 5, and 1 percent level of significance.

Table 2: Logistic Regressions of Hard Main Interview – Selected Coefficients

Variable	All Observations- Full Model	All Observations- No Interactions	Minority Strata – Full Model	Minority Strata – No Interactions
Pre-Incentive	-0.67***	-0.68***	-1.18***	-1.18***
	(0.21)	(0.21)	(0.31)	(0.31)
Hard Screener	0.07	0.07	-0.28	-0.28
	(0.08)	(0.08)	(0.17)	(.017)
Non-Minority Strata	0.11	0.09	N/A	N/A
	(0.35)	(0.32)		
Response Converter	0.03	-0.09	-0.06	-0.10
	(0.30)	(0.12)	(0.37)	(0.29)
Minority Pool-Minority Strata	-1.46*	-.37**	-1.80**	-0.85**
	(.85)	(0.19)	(0.87)	(0.34)
Minority Pool-Any Strata	1.56*	0.45***	1.78**	0.92**
	(0.82)	(0.17)	(0.84)	(0.42)
Number of Observations	3561	3561	740	740
F-test for Joint Significance of Interactions	0.67	N/A	0.76	N/A
	(5,3485)		(2,733)	
F-test for Joint Significance of Individual Interviewer Effects	2.77***	2.84***	1.70*	1.97**
	(9,3481)	(9,3481)	(9,726)	(9,726)
F-test: Sum of Minority Pool- Minority Group & Minority Pool–Any Group=0.0	0.10	0.15	0.00	0.08
	(0.75)	(0.70)	(.96)	(0.78)

Notes: Coefficient (standard error) for variables; F-value (degrees of freedom) for F-tests. *, **, *** indicate significantly different from zero at the 10, 5, and 1 percent level of significance. In addition to the variables shown, the models included controls for race, Hispanic ethnicity, sex, employment categories, individual interviewer dummies for interviewers completing more than three percent of the cases, whether the interviewer was a supervisor or not, case release date category, region dummies, and whether the firm was located in a Metropolitan Statistical Area.