

ASYMMETRIC INFORMATION, INTERVIEWER BEHAVIOR, AND UNIT NONRESPONSE¹

Arthur B. Kennickell, Federal Reserve Board

Mail Stop 153, Federal Reserve Board, Washington, DC 20551; email: Arthur.Kennickell@frb.gov

Key words: Interviewer effects, unit nonresponse

Unit nonresponse in surveys is generally taken to be the result of an unsuccessful interaction, or series of interactions, between an interviewer and a respondent. However, in the case of in-person surveys, the subject of this paper, the intensity of interactions and the pursuit of situations that increase the likelihood of interactions are to a large degree under the control of interviewers. Normally, survey managers are able to target field resources only at a relatively gross level, and they observe only a very limited part of the set of characteristics of sample members that are known more fully to interviewers. Often the managers only know not much more than whether or not a case is completed and the total amount of time an interviewer worked on all cases over a given time period.

Because the structure of information usually makes it impossible to reward interviewers for the appropriateness of their efforts across all cases, their performance evaluations very often turn critically on the degree to which they complete the sample cases in their assignments. Thus, all other things being equal, interviewers have an incentive to pursue cases that have the highest likelihood of success. More generally, the absence of full monitoring implies that interviewers will arrange their work to suit their own preferences as far as such behavior is consistent with any overall constraints imposed by the field managers. To the extent that variations in effort are correlated with anything measured in a survey, the result will be a type of implicit field stratification of the original sample observations, where the dimensions of stratification are related to the level of difficulty or other features of the cases perceived by the interviewers.

In the economics literature, a great deal has been written on behavior under uncertainty in situations with asymmetric information, and contracts between various actors that may be written to reduce incentives for partially observed agents to act against the wishes of a principal.² Although the findings are often quite complex and particular, the essence of this work is that the presence of private information makes it difficult to align the incentives of the actors; consequently, attention must be paid to understanding the incentive effects of the contracts between the actors, and the cost and value of obtaining additional information. This paper attempts to bring some of the reasoning in this literature to bear on understanding the behavior of interviewers and the consequences of their behavior for nonresponse and the distributions that are ultimately measured.

The first section of the paper discusses the potential consequences of the structure of interviewers' incentives in the presence of asymmetric information about the sample members. The next section illustrates the issue with data from the 1998 Survey of Consumer Finances (SCF). The final section summarizes the findings in the paper and points toward related areas in need of additional research.

I. A discussion of interviewer behavior

The response rate and the distribution of completed cases in a survey are functions of many factors: the number of interviewers, the amount time they are allowed to work, the degree to which interviewers choose to allocate effort to particular cases, the nature of interviewers' interactions with respondents, and the level and variation of resistance among respondents. Building on their earlier work and that of others, Groves and Couper (1996, 1998) have built a very clear and useful framework for characterizing the nature of interviewer-respondent interactions and respondent resistance. The more elementary question of the allocation of resources, either by management or by the interviewer, has received remarkably little attention.

Ideally, one would like effort to be applied to all cases in a sample such that the distribution of the completed cases (along dimensions relevant for the motivating research) is, at worst, a completely random subset of the original sample.³ If resources were always evenly matched with expected resistance in the sample, it would be only surprises in the required level of effort for a given survey that would keep completion rates below 100 percent. It does appear that there is an important element of surprise in the response rates for most surveys, but there are also deeper complications. First, there are probably always respondents who will not agree to be interviewed no matter how much attention they are given, and respondents who would require so much effort to persuade that they far exceed the practical limits of effort. Second, for many reasons, interviewers are not often distributed in equal proportion to required effort, and it is costly to move interviewers over long distances. Hiring and retention can be unusually difficult in some areas, and it may well be that labor market tightness is also related to the overall level of respondents' degree of resistance. Third, interviewers may adjust the level of effort that they apply to particular cases in response to the incentives they face and given the set of information they have about the sample and individual cases. Because the information about the level of interviewers' efforts and the level of difficulty of cases is not the same for interviewers and

managers, it is not possible for the managers to make clear decisions about whether redeploying resources might actually change outcomes. This paper focuses on the third issue.

The underlying information model may be quite complicated. For example, in a world where the difficulty of cases is uncertain and it is costly for interviewers to determine the true level, they may extrapolate from their experience to create a working model of the expected payoff for cases with common characteristics and adjust their effort accordingly. Sometimes managers may know things about cases that are not—or cannot be—told to interviewers, and this information might contradict the interviewer's mental model if it were known. The critical factor in generating distortions in effort relative to the ideal is that interviewers and managers conditioning on their available information have different expectations about the difficulty of cases.

Incentives for interviewers typically mix an hourly rate of compensation that may vary with experience, with a required number of completed cases to remain employed. There is almost surely great variation in the level of difficulty across cases, and some cases may so intimidate interviewers that the benefit in hourly pay is less than the psychic cost. Interviewers' preferences and abilities also vary, but because such variation is not directly observable to managers, there is uncertainty in comparing behavior across interviewers. Because the "types" of cases may not be completely obvious to interviewers, they may start cases that they expect will be easier than they turn out to be, and the opposite situation may also occur. Some information about the difficulty of cases is revealed when interviewers classify a case as a nonrespondent, but it is virtually impossible to gauge whether this subjective classification indicates that an interview is impossible or that it is simply some order of magnitude more difficult than other cases. Competing opportunities for interviewers may arise during the field period, and if only more difficult cases remain, the minimum additional benefits needed to get the interviewer to switch jobs will be smaller. Numerous sources of uncertainty may raise costs and budgets are always ultimately limited: these facts are sufficient to make clear to everyone that the field period will almost surely not continue until every case is completed or even until every possibility of refusal conversion is exhausted. Finally, it is important to note that some interviewers clearly display an unusually high level of altruism in working on cases that are not necessarily in a narrow view of their interest.

An essential thread in many of the innumerable models that could be built to describe behavior of interviewers in the field is the informational asymmetry between interviewers and survey managers. As long as managers and interviewers have different information, their conditional estimates of the level of difficulty involved in completing various cases will differ, and

consequently interviewers' desired levels of effort on individual cases may differ from the effort managers would desire in the presence of full information. Where interviewers have information not known to the managers, the managers are unable to estimate sufficiently accurately the true level of difficulty of cases, and consequently they cannot make direct decisions about allocating time to individual cases. If the difficulty of obtaining cases is correlated with analytically interesting variables, this implicit field stratification will almost certainly lead to bias in the distribution of the final data.

There are several possibilities for dealing with the problems induced by the asymmetry of information. First, in some cases it may be that factors sufficiently correlated with difficulty can be observed for each observation during the field period, and that these factors can be used *ex post* to make nonresponse corrections through post-stratification of the analysis weights. Second, if such factors are observable sufficiently promptly during the field period, they might be used to develop explicit quotas. A third possibility is to design contracts with interviewers either to raise the level of their incentives and manipulate their constraints so that they are economically motivated to complete all their cases, or to divide the interviewing staff into units with different operational responsibilities. Each of these possibilities is addressed in turn below.

Whether explicitly or not, post-stratification is very often used with the aim of reducing nonresponse bias, including any component attributable to differential application of effort. Research more focused on this dimension of nonresponse could lead to the discovery of better *ex-post* adjustments.

Although there are many references to the use of quotas in market research, the subject has rarely been touched in a positive way in more formal survey research. Sudman (1966) develops the idea of using quotas within the overall constraints of a scientifically drawn sample. In essence, such an approach is analogous to post-stratification. Research could be devoted to identifying reliably measurable indicators, and computer-assisted data collection in the field could make implementation practical. There may be a particular advantage to this approach over post-stratification if the strata are not homogeneous, as one would ideally like them to be. With either the *ex-ante* or *ex-post* approach, there may still be variations in the level of difficulty within strata. Pressing for additional cases in the *ex-ante* approach would have the advantage of obtaining more relatively difficult cases. However, to the degree that either approach failed fully to condition outcomes in terms of difficulty, bias would remain.

A uniform rise in the wage structure for interviewers to ensure a high level of effort on all cases, though arguably highly desirable for other reasons, would be difficult to implement in the competitive world of

survey research operations.⁴ Potentially, a nonlinear compensation scheme with progressively larger reward for additional completed cases might be even more effective. However, both arrangements could be seen as handicapping subsequent surveys, particularly if there is substantial variation in difficulty across surveys, and developing a general formula that might apply across many surveys would probably be very complicated. Still, there appears to have been very little systematic exploration of the range of feasible compensation plans for interviewers, and it may well be that the large economic literature on contracting developed in other areas could have significant payoffs. Another possibility would be to offer interviewers a menu of compensation plans. For example, interviewers might have an option of receiving a higher wage by agreeing to deal only with cases that other interviewers have classified as refusals, or receiving the base wage by agreeing to the normal conditions of completing a certain percentage of cases etc. Effectively, this approach would codify the use of “refusal converters,” a common arrangement on many surveys. One negative effect of such a formal arrangement might be that the “regular” interviewers would tend to give up on moderately difficult cases that they would otherwise have completed. Other literature on asymmetric information suggests that sometimes incentives can be structured to make it in agents’ own interest to reveal at least some part of their “private” information, as is commonly the case in the design of insurance contracts. A better understanding of the structure of respondents’ propensity to be “difficult” would probably be needed to make significant progress in this direction.

II. Illustrative data from the SCF

A. Background on the survey

The data used here derive from the 1998 Survey of Consumer Finances (SCF), which was sponsored by the Federal Reserve Board (FRB) in cooperation with the Statistics of Income Division of the Internal Revenue Service (SOI). The SCF is designed to collect detailed information on households’ assets, liabilities, pensions, work history and use of financial services. The survey also collects extensive demographic and other data useful in understanding the more narrowly economic data.⁵ The sample is a dual-frame design. One part of the sample is selected using a standard multi-stage area-probability (AP) technique (see Tourangeau *et al.*, 1993). The other part is a list sample developed from statistical records created by SOI based on individual tax returns (see Kennickell, 1998a). This list sample oversamples wealthy families. Unlike the AP sample, respondents in the list sample were given an opportunity to refuse participation in the survey by returning a postcard; respondents who returned the postcard were not subjected to any type of refusal conversion. In 1998, about 21 percent of the list sample cases opted out of the survey by returning the

postcard. Over both samples, 4,309 interviews were completed, with about a third of those observations coming from the list sample. In the AP sample, the response rate among eligible cases was about 66 percent. In the list sample overall, the rate was about 29 percent, but this figure ranges from 41 percent in the least wealthy stratum to about 8 percent in the wealthiest stratum. Data for the survey were collected between June and December of 1998 by the National Opinion Research Center (NORC) at the University of Chicago using CAPI.

B. Illustrations

Because the signals generated by many types and levels of behavior by respondent, interviewers, and management personnel are superimposed in the observable data, it is very difficult to produce statistics that give an unambiguous picture of how limited information in particular may have real effects on the distribution of outcomes in the SCF. Nonetheless, there are some examples that indicate a role for the effects of informational limitations. Three such examples are given here.

Illustration 1

Although interviewers could easily distinguish list sample cases (for which they located a specific person or couple) from AP cases (for which they had only an address), they did not have any information that allowed them to discriminate directly between the strata of the list sample. Such information was available only to the management and supervisory staff and to the FRB project staff. Although there were specific minimum completion targets set for each of the list sample strata, these targets in each survey were projected based on what was actually achievable in earlier surveys, and there was a very clear impression at the time of the survey that it was a struggle to reach these goals.

The completion rates for the bottom two strata of the list sample are particularly interesting. These two strata overlap in terms of wealth with the great majority of the interviewed part of the AP sample (see Kennickell 1998b). The cases interviewed in stratum 1 have a median that lies between the median and 25th percentile of the cases interviewed for the AP sample (table 1). By construction, the list sample excludes people who did not file a tax return—about 10 to 15 percent of all households—and almost all of that group could be expected to have wealth sufficiently low that they would have been included in stratum 1 if they had filed a tax return. Stratum 2 appears to be roughly equivalent to the wealthier end of the AP sample. As noted earlier, a key difference between the AP and list sample cases is the fact that before the interviewers were sent into the field, the list cases were given an opportunity to refuse participation by returning a postcard. Because the postcard refusals should have removed cases most determined not to be interviewed, one might expect the response rates in stratum 1 to be substantially above the rate for at least the

Table 1: Unweighted quantiles of net worth and response rate (percent) for AP sample and strata 1 and 2 of the list sample; 1998 SCF.

Full AP sample	
Net worth	
25 th percentile	5,800
Median	58,800
75 th percentile	202,000
Response rate	65.9
List sample	
Stratum 1	
Net worth	
25 th percentile	2,700
Median	21,600
75 th percentile	105,000
Response rate	
Overall	41.3
Given PC	40.9
Stratum 2	
Net worth	
25 th percentile	117,000
Median	284,000
75 th percentile	530,000
Response rate	
Overall	39.2
Given PC	34.3
All strata	
Net worth	
25 th percentile	456,000
Median	2,182,000
75 th percentile	10,430,000
Response rate	
Overall	28.6
Given PC	33.1

tax filers in the AP sample. However, even if one assumes that all of the cases in the full AP sample who did not file tax return were actually interviewed, the implied response rate for the remainder of the AP sample falls to only about 60 percent—still far above the rates in strata 1 and 2 adjusted for the postcard refusals. It is possible that there were systematically greater problems in locating the list sample cases, but there is no sign of a sufficient level of such problems in the case records. It is also possible that the fact that list sample respondents were requested by name made some people more suspicious than would be the cases for AP cases, whose names were unknown a priori since they were selected only on the basis of their address.

Another interpretation is that the data show the effects of limited information. Because the samples for both stratum 1 and stratum 2 are relatively small—each comprises well under 10 percent of the list sample—most of the list sample cases interviewers see are much

wealthier than the great majority of AP sample cases. If, as appears strongly to be the case, wealth is correlated with difficulty, interviewers may have attributed to stratum 1 and 2 cases a subjective estimate of the level of difficulty that was higher than the correct conditional level.

Illustration 2

The case records allow one to construct a measure of the application of effort across different parts of the sample. These data show that the distribution of the number of attempts needed to obtain an interview in the AP sample and the bottom two strata of the list sample are similar (table 2).⁶ For completed cases in list sample strata 3 and above, the data indicate that the distribution of the number of attempts needed to complete the case is shifted up by one relative to AP cases across most of the distribution. One would expect that even with the removal of many of the potentially most hostile cases via the postcard refusal, the wealthy families in these strata would be relatively more difficult to persuade to complete an interview.

Although the evidence for the completed cases could be taken to suggest that an overall higher level of effort was devoted to the list sample cases, a similar analysis for the final refusals and the “censored” cases (eligible cases that were classified as neither completed nor refused at the end of the field period) suggests the contrary (table 2).⁷ For both of the subgroups of the list sample, the distribution of the number of attempts lies distinctly below that for the AP sample alone. Even so, it could be that refused and censored list sample cases were so much more difficult than the completed cases that the expected payoff to addition effort was negligible. Perhaps surprisingly, the available data do not support this hypothesis (table 3). For the list sample refusals and for the censored cases, there are far fewer references to a hard refusal.⁸ Moreover, there were relatively fewer attempts with the list sample refusals to convert the refusal than was the case for the AP sample.⁹ Although the case records were used for some types of interviewer supervision during the field period, there is no sign that these data were used in a systematic way to regulate the efforts on individual cases. Thus, there was an effective asymmetry of information between interviewers and managers, and one interpretation of the data is that interviewers may have shaded their efforts away from the list sample.

Illustration 3

Further evidence of possible aversion of interviewers to the list sample cases may also be gleaned from looking at the application of effort at the beginning of the field period. During the first three weeks of the field period, all cases were supposed to be “touched” in

Table 2: Quantiles of distribution of number of attempts to interview cases, AP sample, list sample, strata 1 and 2 of list sample, and strata 3-7 of list sample, by final case disposition, 1998 SCF.

	Distribution of no. of attempts		
	25 th %ile	Median	75 th %ile
AP cases			
Completes	3	5	9
Refusals	6	9	13
Censored	6	9	14
List sample cases			
Completes	4	6	10
Refusals *	5	8	11
Censored	4	7	10
List strata 1 and 2			
Completes	3	5	8
Refusals *	5	7	10
Censored	4	6	9
List strata >2			
Completes	4	6	10
Refusals *	5	8	11
Censored	4	7	10

* Excludes postcard refusals.

Table 3: Percent of cases classified as giving "hard" refusals, percent where refusal conversion was attempted, and percent in group, by final refusal status and final censored status, and by AP sample, and refusals other than postcard refusals in the list sample, strata 1 and 2 of list sample, and strata 3-7 of list sample, 1998 SCF.

	Sample group			
	AP	All list	Strata 1+2	Strata 3-7
<i>Refusals *</i>				
Hard refusal	55.0	42.6	39.8	42.9
Ref. conversion attempted	80.1	63.3	62.7	63.3
Ref. as % elig. cases in samp. grp.	24.9	24.8	16.9	27.9
<i>Censored</i>				
Hard refusal	39.1	28.4	27.4	28.5
Ref. conversion attempted	36.5	19.8	18.8	20.0
Censored as % elig. cases in samp. grp.	2.3	27.1	49.2	29.1

* Excludes postcard refusals

some way. The list cases were released a few weeks after the AP cases to ensure that interviewers had sufficient experience to be able to deal with the list sample cases, a practice recommended by earlier experience. Of the AP cases, there is a record of some attempt to reach about 3/4 of the cases during the first 28 days of the field period, and about half had at least some sort of contact. In contrast, 28 days after work had begun on the list sample cases, there had been attempts on only about a fifth, and only 10 percent had any sort of contact. The interviewers' direct supervisors were supposed to have enforced work on all cases, but it is very clear from the data that effort was more focused on the AP cases. Like the interviewers, the supervisors were under pressure to produce completed cases.

III. Summary and future research

This paper focuses on the role of asymmetric information between interviewers and survey managers in determining the distribution of completed cases in a survey. When interviewers and survey managers do not have the same information about the sample cases, and there is a fixed incentive scheme for interviewers, there will be a tendency for the interviewers to arrange their work to suit their own preferences. Those arrangements may well differ from the behavior desired by managers with complete information. As noted in the paper, such behavior does not deny the clear fact that many interviewers exhibit dedicated behavior that goes well beyond what would normally be seen as being in their narrow self interest. Indeed, I believe it is very likely that the managers involved in data collection do not normally give sufficiently clear information about the sampling objectives of a project for the interviewers even to develop a notion of how suiting themselves might lead to suboptimal outcomes in the distribution of completed interviews.

Future research should particularly target improving knowledge of the extent and the structure of interviewers' decisions to allocate effort, developing new means to monitor and control progress on cases in the field, and investigating incentive schemes for interviewers that might be more likely to align interviewers' decisions with the interests of survey managers. Undoubtedly, improvements in monitoring and more complex incentive plans will be inadequate alone. Deeper structural understanding of the problem may lead to improved nonresponse adjustments, most likely through post-stratification.

Bibliography

- Grossman, Sanford J. and Oliver D. Hart [1983] "An Analysis of the Principal-Agent Problem," *Econometrica*, v. 51, No. 1, pp. 7-46.
- Groves, Robert M. and Mick P. and Couper [1996] "Contact-Level Influences in Face-to-Face Surveys,"

- Journal of Official Statistics*, Vol. 12, No. 1, pp. 63-83.
- Groves, Robert M. and Mick P. and Couper [1998] *Nonresponse in Household Interview Surveys*, John Wiley and Sons, New York.
- Harris, Milton and Artur Raviv [1979] "Optimal Incentive Contracts with Imperfect Information," *Journal of Economic Theory*, v 20, pp. 231-259.
- Holmström, Bengt [1979] "Moral Hazard and Observability," *Bell Journal of Economics*, pp. 74-91.
- Kennickell, Arthur B. [1998a] "List Sample Design for the 1998 Survey of Consumer Finances," working paper, Federal Reserve Board, <http://www.federalreserve.gov/pubs/oss/oss2/method.html>.
- Kennickell, Arthur B. [1998b] "Using Income Data to Predict Wealth," paper presented at the Annual Meetings of the Allied Social Science Associations, New York, 1999a.
- Kennickell, Arthur B, Martha Starr-McCluer and Brian J. Surette [2000] "Recent Changes in Family Finances: Results from the 1998 Survey of Consumer Finances," *Federal Reserve Bulletin*, (January) pp. 1-29.
- Ross, Stephen A. [1973] "The Economic Theory of Agency: The Principal's Problem," *American Economic Review, Papers and Proceedings of the Eighty-fifty Annual Meeting of the American Economic Association*, (May) pp. 134-139.
- Sudman, Seymour [1966] "Probability Sampling With Quotas," *Journal of the American Statistical Association*, v.61 (315), Sept., pp. 749-771.
- Tourangeau, Roger, Robert A. Johnson, Jiahe Qian, Hee-Choon Shin, and Martin R. Frankel [1993] "Selection of NORC's 1990 National Sample," working paper, NORC at the University of Chicago, Chicago, IL.
- partially offset through reduced attrition and consequently lower costs for training, supervision, etc. A more permanent professional staff would presumably also improve data quality by developing and maintaining high quality data collection skills among interviewers.
5. For a more detailed description of the survey and the data, see Kennickell, Starr-McCluer, and Surette (2000).
6. An "attempt" is defined broadly to include any effort devoted to securing a complete interview with a case. In principle, duplicate entries for the same action, and simple comment entries in the case records are excluded from the calculation.
7. Across the censored cases, only about 13 percent of the working disposition codes for the case record entry before the one indicating the suspension of work indicated a "temporary hard refusal." Other codes at this point suggest that additional work might well have been productive.
8. In cases where an initial informant agreed to complete a screener interview designed to identify the appropriate respondent, but the person refused to complete the interview, the interviewer was asked to record whether the refusal was "hard," that is, one that appeared to be relatively poor prospect for refusal conversion. In the case records, interviewers also were able to record such information at any stage of negotiation through the working case disposition codes. If the interviewer recorded a hard refusal either in the screening or at any point in working the case, it is counted here as a hard refusal.
9. Active refusal conversion was indicated by working case disposition codes in the case records.

Endnotes

1. The views presented in this paper are those of the author alone, and they do not necessarily reflect the views of the Board of Governors of the Federal Reserve System or its staff. The author wishes to thank the central office staff—particularly Lisa Thalji and Mary Hess—the field managers and the interviewers at the National Opinion Research Center at the University of Chicago for collecting the data, and the respondents who gave their time for the interview. The author is grateful to Amber Lynn Lytle for research assistance. An unabridged version of this paper is available at <http://www.federalreserve.gov/pubs/oss/oss2/method.html>.
2. For example, see Ross (1973), Harris and Raviv (1979), Holmström (1979), and Grossman and Hart (1983).
3. This argument assumes that any treatment effects resulting from differential applications of effort are negligible or otherwise not important.
4. Approaches that lead to higher pay might have the additional benefit of leading to a more "professionalized" interviewer corps, the cost of which could be at least