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

Survey Incentives, Survey Effort, and Survey Costs

Jesse Bricker

2014-74

Survey Incentives, Survey Effort, and Survey Costs*

Jesse Bricker
Federal Reserve Board
20th & C Street, NW
Mail Stop 153
Washington, DC 20551
jesse.bricker@frb.gov

Abstract
This paper uses the 2007 and 2010 waves of the Survey of Consumer Finances (SCF) to investigate how monetary incentives affect the time and effort that interviewers expend during the survey field period, and how these incentives affect effort expended by the survey respondent. The results imply that a larger monetary incentive offer helps reduce contact attempts and time in the field while maintaining data quality and effort during the survey by the respondent. Our results are based on a quasi-experiment that varies which families receive an incentive offer letter. Supporting evidence is given through a comparison of field effort outcomes between 2010 and 2007 after the base incentive increased from $20 in 2007 to $50 in 2010.

Keywords: incentives, contact attempts, data quality, record-of-call paradata.

* 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. I am grateful to Arthur Kennickell, John Sabelhaus, and my colleagues at the Board of Governors for providing helpful comments and suggestions, to the SCF respondents for providing their data for research purposes, to Micah Sjoblom and NORC staff for extra work on record of call and monetary incentive data, and to Hannah Bricker for editorial assistance. Any errors and shortcomings are those of the author.

1. Introduction

Participating in a household survey can be its own reward for some respondents. For many other potential respondents, though, a monetary incentive can help induce participation, especially if the survey is particularly burdensome (Singer, 2002; Shettle and Mooney, 1999; Groves, Singer, and Corning, 2000; Singer, 2011; Rogers, 2011). As response rates to surveys have fallen over time, the cost to run a large household survey has increased (Curtin Presser, and Singer, 2005). Incentive payments have been proposed as a way of increasing response rates in a cost-efficient manner. This work estimates the number of contact attempts that can be saved by offering an incentive, which can help survey administrators decide if offering an incentive is cost-efficient.

The Survey of Consumer Finances (SCF) asks families about topics that are especially technical and private and can be a taxing on a family's time because of its length.[1] Accordingly, the SCF offers monetary incentives to many participants. This offer is conveyed to respondents in a mailing and in person by a field interviewer. The SCF includes an area-probability (AP) sample and a list (LS) oversample of expectedly wealthy families. All families in the AP sample are offered an incentive to participate, while only a small fraction of LS families are offered incentives.

This paper uses a quasi-experiment that compares some LS families that received the incentive offer to other observably identical families that did not receive the offer. The families that received the initial incentive agreed to participate quicker than families that did not receive the initial offer, both in terms of the number of contact attempts and in time since first contact. Data quality measures and respondent effort are little affected by the offered incentive.

The quasi-experiment used here derives from the manner in which the expectedly-wealthy LS families are sampled. There is no sampling frame for wealth, so wealth cannot be used in the

[1] The median interview length in the 2010 SCF was about 85 minutes.

sampling process. A sample frame for income does exist, though, and the oversampling mechanism depends on modeling wealth as a function of income.[2] Families are then arranged into one of seven strata of increasing predicted wealth and are oversampled according to this wealth prediction.

While families in strata one and two were offered the incentive, the families in strata three through seven were believed to be too wealthy for an incentive to have an impact on response. These families received the same advance mailing as the strata one and two families, except the incentive offer was not included. Field staff was also not authorized to verbally offer an incentive to these families.

Once wealth is measured in the SCF, though, some of the families in stratum three were actually as wealthy as (and observably equivalent to) families in stratum two that received the incentive offer.[3] Comparing the stratum two families (who received the incentive offer) to the observably equivalent stratum three families (who did not) serves as the basis for the quasi-experiment.

Further evidence comes from a change in the base incentive rate from $20 in 2007 to $50 in 2010. In our preferred specification, on average, the 2010 families treated with a $50 offer needed four fewer attempted contacts before agreeing to participate, relative to the untreated families. In 2007, families treated with a $20 offer also agreed to respond more readily than the untreated families, but the difference was smaller: only two contacts were saved. Increasing from a $20 offer to a $50 offer saved two attempted contacts.

[2] The sample frame for income comes at a two-year lag, though.

[3] There are two sources of imperfection in the oversample. First, the sampling is based on income and income is not a perfect substitute for wealth. Second, in the sample frame income is measured in years t-2, t-3, and t-4 (for example, the 2013 SCF oversample frame is derived from income in 2011, 2010, and 2009). Third, the input data to the model that correlates income and wealth is based on wealth and income data that are three years old; the contemporaneous correlation may be different.

We also compare the 2010 AP sample to the 2007 AP sample. In 2010 AP families were offered $50 while in 2007 the offer was $20. The 2010 AP families agreed to participate much quicker than the 2007 AP families, supporting the idea that a larger incentive is of more assistance than a smaller incentive (Singer, 2002).

Typically, very few wealthy people respond to surveys. But because the quasi-experiment is among an over-sample of expectedly wealth families, we can comment on the impact of incentives at varying degrees of wealth. Our results imply that a $50 incentive offer is most salient to families above the median but below the top decile of wealth.

It is likely that our main dependent variable (the number of contact attempts) is mis-measured as it derives from paradata that are not the result of a random sampling process (Biemer, Chen and Wang, 2011; Wang and Biemer, 2010; Kennickell, 2012). However, the nature of the mis-measurement should lead our results to be a lower bound for the true impact of the incentive offer.

2. Literature review

Monetary incentives are typically employed to help increase survey response rates. As more surveys turn to incentives, interest has turned to theories of survey response and the role of monetary incentives in this process (Groves, Singer, and Corning, 2000; Singer, 2011; Singer, 2002; Church, 1993). Fewer studies, though, have examined how incentives help make the survey experience more efficient by reducing the effort that field staff must exert to complete an interview (and, hence, possibly reduce survey costs).

Incentives and effort

An increase from a $20 to a $50 incentive in the 2000 Health and Retirement Study (HRS) partially offset other field costs by reducing the number of interviewer conversion attempts with HRS respondents (Rodgers, 2002). Members of the 1996 Survey of Income and Program

Participation (SIPP) that randomly received a $20 incentive offer also needed fewer in person visits than did the members that did not receive an incentive (James, 1997).

An advance letter that arrives *with a prepaid incentive* reduced the average number of interviewer contacts needed to complete an interview from 11.5 to 8.75 in the Survey of Consumer Attitudes (SCA), though the advance letter itself (without a prepaid incentive) showed no reduction in contacts (Singer, van Hoewyk, and Maher, 2000).

Incentives also may affect the quality of the collected data by encouraging disinterested families to participate (Hansen, 1980). Yet, most previous work shows either no relationship between measures of data quality and survey incentives (Shettle and Mooney, 1999; Davern, Rockwood, Sherrod, and Campbell, 2003) or that survey incentives increased data quality (Singer et al, 2000).

Theories of survey participation and monetary incentives

Leverage salience theory (Groves, Couper, and Singer, 2000) presents the decision to respond to a survey as a function of individual-specific traits and beliefs coupled with how salient the interview traits are made by the field materials and interviewing staff. In particular, a potential respondent may be influenced (positively or negatively) to participate in a survey by the survey's sponsor or topic, the respondent's civic duty, privacy concerns, an incentive, or many other reasons. Each input in the participation decision has a weight (*leverage*) and, all else equal, the potential respondent participates if the weighted benefits are larger than the weighted costs. In the leverage-salience framework, incentives can help offset some negative influences on participation and a larger incentive has a larger impact than a smaller incentive.

The survey materials and the field interviewers themselves need to make *salient* the aspects of the survey that are attractive to the respondent. The SCF uses a contacting strategy where interviewers first inform and introduce the SCF to the respondent but then tailor their interactions

to persuade and make salient the potential benefits to the respondent (Kennickell, 2005).[4] An advance mailing with the incentive offer may also make conversions easier in the field by allowing the field interviewer to have more confidence in their sales pitch.

Incentives influence participation in other theories, too, including cost-benefit models of survey participation (Singer, 2011), social exchange theory (Dillman, 1978) and the norm of reciprocity (Gouldner, 1960). In each theory, an incentive may induce respondents with low internal motivation to participate in the survey, though at indeterminate cost to data quality. In a social exchange, for example, a higher incentive may lead the respondent to exert more effort during the survey: as the respondent is rewarded with a larger incentive, the respondent is more willing to help. But, these incentivized respondents may have less interest in the survey topic, so their responses may be incomplete.

In general, stylized facts from meta analyses show that incentives are particularly useful when other reasons to participate in the survey are not salient (Singer et al, 1999); that incentives are effective in interviewer-mediated face-to-face surveys and in mailed survey contexts, though they are most effective in mailed surveys (Singer et al, 1999); that respondents prefer prepaid monetary incentives over both conditional monetary incentives and gifts (Church, 1993, Singer et al, 1999); that larger incentives imply larger response rates and the effect of increasing the incentive is linear (Church, 1993);[5] Groves, Singer and Corning, 2001); and that incentives *themselves* appear to influence the response rates of potential respondents, apart from any gain in confidence that interviewers may feel through having the incentive (Singer et al, 2000).

[4] The SCF contacting strategy is discussed in detail later in this paper.

[5] The impact of incentives on participation, on average, is positive the impact is also heterogeneous. Eight of the 74 studies in Church's (1993) review find either a negative impact or no impact on participation. And though incentives are useful in gaining cooperation of low income and minority families (Singer et al, 1999), small monetary incentives may have the opposite effect on high net worth and high income families by trivializing the importance of the survey

This paper builds off of the Rodgers (2002), James (1997), and Singer, et al (2000) literature by analyzing the effect of offering an incentive on the number of contacts needed to compete the survey. The incentive offer is sent in an advance mailing and is conveyed by field staff in person, but is not pre-paid. This paper also builds off of the on past work by Shettle and Mooney (1999) and Davern, et al (2003) by examining the role of incentives on data quality, as measured by the respondent's use of documents during the interview, by interviewer-perceived interest on the part of the respondent, and by the fraction of missing dollar value data in the survey.

3. SCF Data

The primary data are from the 2007 and 2010 SCF surveys. The SCF is typically conducted by NORC on behalf of the Federal Reserve Board (FRB) as a cross-sectional survey every three years and provide the most comprehensive and highest quality microdata available on U.S. household wealth.[6] The survey has a CAPI instrument and is administered by an NORC field interviewer either in person or over the phone.

SCF respondents answer questions about financial and nonfinancial assets, debts, employment, income, and household demographics. In subsequent regression analysis, this information is used to help ensure that we are comparing similar families. The SCF questionnaire is very detailed and can be time-consuming for the respondent. The median length of an SCF interview is nearly 90 minutes, and families with complicated finances can be engaged with the survey for more than two hours.

The SCF combines a geographically stratified and nationally-representative area probability (AP) sample and a list sample (LS) that oversamples households that are likely to be wealthy. The AP

[6] See Bricker, Kennickell, Moore and Sabelhaus (2012) for results from the most recent triennial SCF.

sample is drawn by NORC at the University of Chicago and provides a nationally-representative sample of families.[7]

The LS is drawn by FRB staff using a frame of statistical records derived from tax returns at the Statistics of Income (SOI) Division of the Internal Revenue Service.[8] Using models that correlate income to wealth, FRB staff generates a wealth prediction for each SOI tax filer.[9] Filers are ranked by their predicted wealth and are placed into one of seven wealth strata of increasing expected wealth; the probability of being sampled increases for as the value (one through seven) of the strata increases. The wealth of filers in the lowest strata is often comparable to the AP sample and strata two through seven become increasingly wealthy. The AP and LS samples are combined and weighted to represent the population of households.

The ranking process generally does a very good job but data limitations make a perfect rank ordering nearly impossible (Kennickell, 1998). For instance, the most recent SOI data available during the sampling process is two years old, yet we need to sample contemporaneous wealth. Further, the model that generates the correlation between income and wealth uses SCF data from three years ago and SOI data from five to seven years ago; these correlations are applied to the contemporaneous sample frame data to generate a wealth rank ordering. The relationship between wealth and income would need to be completely stable in the contemporaneous sampling period for the ranking to work perfectly.

---

[7] See Tourangeau, et al (1993) for more information about the NORC national sample.

[8] Prior to use, the data are edited by SOI to support research at the Office of Tax Analysis and the Joint Economic Committee of the Congress (Kennickell, 1998). A great degree of security is involved with this sampling procedure and formal contract govern the agreement between the FRB, NORC and SOI. The FRB selects the sample from a data file with no names but with a unique identifier; the FRB sends the list of those selected to SOI along with a different temporary unique identifier, SOI then sends the temporary unique identifier to NORC along with contacting information; NORC collects the survey information and sends to FRB. Thus, the FRB never knows any contacting information, SOI never knows any survey responses, and NORC never knows anything more than survey responses and location information. See page 2 of Kennickell (1998) for more information.

[9] These models are described in more detail in Kennickell and McManus (1993) and Kennickell (1998, 2001).

If the rank ordering of wealth was perfect then there would be no overlap between the wealth distributions across neighboring strata. In practice, though, there is a small amount of overlap across the actual wealth distributions in the 2010 SCF (figure 1) and in past SCFs (Kennickell, 1998).

4. Background on Interview and Case Working Process

Incentives

An incentive offer is sent to all sampled AP addresses and to sampled families in strata one and two of the LS. In early April these families receive a mailing with a letter describing the project, an incentive offer, a brochure about the survey, and a letter from the current Federal Reserve chairman asking for the family's cooperation. In 2010 the incentive offer was $50; in 2007, the offer was $20. Each respondent will receive at least one in-person visit from a trained field interviewer, often to explain the purpose of the SCF and to gain cooperation.[10] Field interviewers are free to discuss the incentive offer when contacting the AP families and the LS families in strata one and two.

The LS families in strata three and above receive an identical package without the monetary incentive offer. As part of the case management strategy (described in detail below), the field staff were allowed to escalate the incentive for any family when a larger incentive was deemed necessary for participation.

Record of calls

Field interviewers also record the details of their case work in the "record of calls" (ROC) database; from these data we can measure the number of times that an interviewer attempted to contact a respondent.[11] Each ROC entry is dated, so we can also measure the number of days between first contact and the interview date.

[10] However, about half of SCF interviews are completed over the telephone.
[11] These contacts run the gamut from an in-person chat with the respondent, to talking to the respondent's neighbor, to leaving a voicemail with the respondent, to a refusal.

Each interviewer is trained on how to use the ROC software and about the protocol for entering each contact in the ROC. However, despite these best efforts, it is the case that ROC data are sometimes entered selectively (Biemer, Chen and Wang, 2011; Kennickell, 2012).[12]

Case work

The SCF also has a three phase case-working strategy to help ensure that each case is worked, regardless of the perceived difficulty (Kennickell, 2005).[13] In the first phase, interviewers inform and introduce the SCF to the respondent, and some limits are placed on the amount of effort that field staff can put into any one case. Field managers work closely with field interviewers as the survey is being introduced and will oversee effort. Casework often ends in phase one with a successful interview.

Absent a completed interview, phase two begins with an express mailing of survey materials (identical to that received prior to the field period) and continues with limited follow-up to persuade the respondent to participate. Cases with continued refusals (or that cannot be contacted) reach phase three: the stage at which case efforts are reassessed, often through ROC data. Phase three allows unlimited effort to get resolution to the case. Field interviewers work with the guidance of field managers to craft a strategy to finally gain cooperation.

A significant amount of interviewer effort is typically needed to guarantee adequate response rates in the SCF. In the 2010 SCF less than 10 percent of cases were completed on the first contact

[12] These ROC data will serve as the dependent variable in subsequent regression analysis. In an OLS regression a classically-mismeasured dependent variable will imply that model coefficients are estimated inefficiently but are still consistent (Bound, Brown, Mathiowetz, 2001). If the mis-measurement in the dependent variable is correlated with the dependent variable then the bias will be the correlation between the two. If the mis-measurement in the dependent variable is correlated with the independent variable of interest (in our case a dummy for stratum two group membership) then the estimated coefficient is inconsistent with possibly indeterminate bias. We will consider the potential correlation in detail in the results section, but ultimately will argue that this mis-measurement in the dependent variable may lead to the reported effects being a lower bound.

[13] If field interviewers are paid by the completed case then they will have incentive to target the most promising cases first and not work the more difficult cases (Kennickell, 2004). On the SCF, though, field interviewers are also paid for the hours that they put into a case, even if they do not complete the interview.

and the majority of cases required at least 10 attempted contacts. And among cases that entered phase two, the majority needed at least 17 attempted contacts. Even more interviewer effort was needed in the 2007 SCF, as the majority of cases needed at least 16 attempted contacts.

Wealthier and higher-income households are often harder to persuade to take the SCF. The opportunity cost of time for higher income families is high, and wealthier households can often have gatekeepers (literal and figurative) that restrict access. The average number of contacts until a successful interview is about 7 for AP cases, 10 for stratum one cases, 11 for stratum two cases, and 12 to 13 for stratum three and above.

5. A thought experiment to identify the usefulness of an incentive offer

The majority of results in this paper are derived from comparing the differences between LS families in stratum 2 and stratum 3 in terms of contact attempts and data quality. Stratum 2 families receive an incentive offer while stratum 3 families do not; if incentives impact these outcomes then we expect to find a difference between the two groups.

In principle, an unbiased estimate of the group difference is possible by randomizing group membership and giving one group a treatment (e.g. an incentive offer) and leaving the other group untreated to serve as the control (e.g. let S2=Treatment and S3=Control). The control group allows us to consider the counterfactual: the expected outcome of the treatment group had they not received the treatment. But randomization is not possible in this setting and, in fact, observable characteristics form the basis for assignment into stratum 2 and stratum 3.[14] The design of the LS, however, allows us to estimate the impact of the incentive offer in a quasi-experimental setting.

[14] In the SCF, it is not possible to randomize a larger incentive payment for some respondents. Interviewers are at least partially remunerated by the number of interviews that they complete and issues of fairness arise when some interviewers would have randomly large incentive cases to work while others do not. Much of the fieldwork on the SCF is collaborative so the perceived lack of fairness would not be concealed.

In general, the raw gap in contacts between stratum two (S2) families and stratum three (S3) families is expressed as:

(1) $\Pi(S2) = E(\text{contacts}_{S2} \mid S2) - E(\text{contacts}_{S3} \mid S3)$.

In our case, the difference in the mean of these two groups will be a biased measure of the true treatment effect because membership in the groups S2 and S3 is a function of traits that are observable and unobservable to the statistician. That is, differences in observable traits (X) of stratum two and three can explain some of the gap in contacts:

(2) $\Delta(S2 \mid X = x) = E(\text{contacts}_{S2} \mid X = x, S2) - E(\text{contacts}_{S3} \mid X = x, S3)$.

Because SCF wealth strata are based on *predictions* of wealth by using SOI income data, we will not observe a perfect correlation with our pre-survey wealth rank-ordering (which is used to place families into either stratum 2 or 3) and the rank-ordering of family wealth realized in the SCF. The distributions of wealth observed in the SCF for stratum two and three families are plotted in figure 1. In general most strata two families are less wealthy than strata three families but there is noticeable overlap between the distributions; that is, some strata two families are actually as wealthy as strata three families.

Using this insight we can compare families in strata two and three that are observably equivalent in every way, including wealth and income in the SCF, except that the stratum two families receive an incentive offer while the stratum three families do not. In the experimental nomenclature from above, the stratum two families are the treated group while stratum three families are the control group. After conditioning on observables, our insight allows us to re-write equation (2) from

$E(\text{contacts}_{S2} \mid X = x, S2) - E(\text{contacts}_{S3} \mid X = x, S3)$

to

(3) $\Delta(S2 \mid X = x) = E(\text{contacts}_{S2} \mid X = x, S2) - E(\text{contacts}_{S3} \mid X = x, S2)$.

That is, the number of contacts for stratum three families can stand as the missing counterfactual for the number of contacts that stratum two families would have needed had they not received the incentive offer. We use this quasi-experiment to find the impact of receiving the initial incentive offer.

In a fully randomized trial, an OLS regression of the number of contacts on treatment status (i.e. being in strata two) and other covariates will recover an unbiased estimate of $\Delta(S2 \mid X = x)$ in equation (3).

However, the wealth distributions of stratum two and three families do not always overlap. In region A of figure 1 there are many stratum two families but there are few stratum three families, and in region B there are many stratum three families but few comparable stratum two families. In region C, though, there are many comparable stratum 2 and 3 families. The linear parametric assumption made in a typical OLS regression won't distinguish between regions A, B, and C and may lead to poor comparisons between the treatment and control groups in regions A and B where there is little (or no) common support (DiNardo, Fortin and Lemieux, 1996; Barsky, Bound, Charles and Lupton, 2002).

As such, we turn our attention to the nonparametric and semi-parametric cell matching estimator of Black, Haviland, Sanders, and Taylor (BHST 2006, 2008), which will force comparisons be on common support, will give low weight to comparisons made with thin data (regions A and B), and higher weight to comparisons made in the overlap area with strong common support (region C). In effect, the weighting of the BHST estimator alters the distribution of observables so that the distribution of stratum three families' traits look like the distribution of stratum two families' traits.

For example, in the simple univariate case of figure 1 the reweighting will make the stratum 3 wealth distribution look like the stratum 2 wealth distribution.[15]

The estimator works as a weighted least squares estimator. Stratum two and three observations are pooled in the weighted regression model:

(4) $\text{contacts} = \alpha + \beta \text{Stratum2} + \varepsilon$.

Weights are generated by finding a propensity score P(X): the probability of being treated (i.e. in stratum two) conditional on observable X. Data with no match (that is, off common support) get zero weight; stratum two families that can be matched with an observably equivalent stratum three family get weight of one; and stratum three families that can be matched with an observably equivalent stratum two family get weight of P(X)/[1-P(X)].

The estimator can be either fully nonparametric or semiparametric. In the nonparametric estimator, the propensity score is generated by a comparison of cell means of the Xs between stratum two and three families. The cost of the nonparametric estimator is the "curse of dimensionality" – the small number of observable traits that one can use to create the propensity score (Rosenbaum and Rubin, 1983). The semiparametric estimator uses a logit model to find a propensity score and allows the comparison to include more observable traits (X). In the results that follow we use nonparametric and semi-parametric regressions (along with OLS) to find the effect of an incentive offer on interviewer and respondent effort.

6. **Results**

In the 2010 SCF, families in stratum two needed to be contacted 2.2 fewer times than stratum three families before agreeing to the SCF interview (table 1). This difference may be due to

[15] The same weights are applied to the outcome so the reweighting will change the distribution of contacts for stratum 3 families, too.

differences in observable characteristics between families in these strata (appendix table), and may be due to the difference in the incentive offer. Results described in this section control for differences in characteristics and use the quasi-experiment described above to find the impact of the incentive offer.

In an OLS regression that controls for differences in net worth percentile, income percentile, age class dummies, education dummies, race dummies, marital status, and urban status, on average the stratum two families needed 2.45 fewer contact attempts before completing the survey (table 2, panel a).[16] And after interviewer fixed effects are included (second column), the estimated effect of the offer is reduced to about one fewer attempted visit. Thus, the effectiveness of the incentive offer appears to be partly mediated through the interviewer herself.

The nonparametric weighting estimator, though, shows stratum two families needed 4.65 fewer attempted contacts, on average, than stratum three families. This estimator forces comparison across like families, but only a small subset of covariates can be used; net worth category dummies, education dummies, and urban status indicator are used in these regressions. An OLS regression with the same set of covariates as in the nonparametric regression indicates that stratum two families needed about 2.5 fewer attempted contacts before a completed interview (not shown), similar to column 1 of table 2. That is, the assumptions made in an OLS model—and lack of assumptions in the nonparametric model—drives the difference in nonparametric and OLS results in table 2, not the smaller set of covariates needed for the nonparametric model.[17]

The differences between the nonparametric and OLS results show the importance of making comparison on common support. At higher levels of net worth (region B in figure 1), the difference

[16] The note of the appendix table contains a detailed description of each regressor.
[17] A regression discontinuity quasi-experimental design (where the cut-off between stratum 2 and 3 is the discontinuity) yields a similar result.

in total attempted contacts between stratum two and three families is less than one; the difference in total attempted contacts in region C, though, is close to six (not shown). The nonparametric estimator places a higher weight where there are many comparable families (region C), and down weights differences where there are very few comparable families (regions A and B). An OLS estimate does not discern between the three regions.

The nonparametric model allows only a parsimonious set of covariates, but the semi-parametric correlate to the nonparametric model allows us to use all of the variables from the OLS model in column one. In this case, treated families agreed to participate with 3.1 fewer contact attempts. The semi-parametric estimates that incorporate an interviewer fixed effect (not reported) are comparable to the OLS results with an interviewer fixed effect. In the semi-parametric model many observations are dropped due to common support issues.

Results using a smaller incentive (2007 SCF)

We carry out a parallel exercise using the 2007 SCF in panel B of table 2, when the incentive offer was only $20. In 2007, it is still the case that field staff needed to attempt contact with stratum two families fewer times than they needed with stratum three families. The OLS regression implies 1.66 fewer attempts were needed, and the nonparametric and semi-parametric results imply 2 to 2.34 fewer attempts.

The larger magnitudes of the 2010 estimates in the top panel indicate that the larger $50 incentive offer in 2010 was more effective than the smaller $20 incentive offer was in 2007.[18]

Contacting strategy

[18] The magnitudes of the estimates in panel A is considerably larger than those in panel B. However, the point estimates lie within the 95% confidence interval of the respective estimates in panel B.

The multi-phase SCF contacting strategy has its clearest break point between phase one and phase two when the contacted family receives an express-mailed package of survey materials.[19] After the limited phase two follow up attempts, phase three can begin with unlimited contact attempts. When we split the total number of contacts into phase one contacts and post-phase one contacts (table 3), the effect of the incentive offer does not appear in the phase one "introduce and inform" stage, but is apparent after phase one.

Very few stratum two and three interviews are agreed-to during phase one. These are typically busy families with relatively complex finances and the field staff must build a strong rapport in order to convince them to participate. But, the incentive offer appears to give the field staff a foot in the door that pays off in the long run.

Elapsed time-to-interview

Another way of looking at the effectiveness of the incentive offer is the amount of time it takes to complete the interview: the time from first contact to interview date. Arguably, this measure is less likely to be measured with error as there is little incentive to under-report the first contact and interview date is known.

The families that were offered the incentive in 2010 agreed to the interview in 9 to 23 fewer days than the families that were not initially offered the incentive (table 4). And, similar to total contact attempts, the $50 offer in 2010 also appears to be more effective than the $20 offer from 2007.

Data quality

[19] The express mailing is seen as a helpful conversion tool by the interviewers.

A worry is that incentives bring in respondents that are less motivated and less likely to provide high-quality data. But measures of data quality and respondent motivation appear to be little affected (negatively or positively) by the offered incentive (table 5).

The outcomes in table 5 are a mixture of objective outcomes (whether the respondent used documents to support responses during the interview, and the fraction of dollar-value questions that are answered "don't know" or "refused" during the interview) and interviewer observation (the respondent's interest in the survey and the respondent's suspicion prior to the interview). The private financial nature questions in the SCF require an interested and non-suspicious respondent.

The data show that stratum two families, on average, are neither more nor less likely to use documents, express interest prior to the survey, or be suspicious prior to the survey. Though the estimates are positive, the standard errors are relatively large are the point estimates are small. The changes implied by the point estimates in columns 1 through 3 of table 5 are between 1 and 3 percent of the respective means.

However, column 4 shows that treated families have higher rates of missing dollar-value data in the survey: about 3 percentage points. The mean of missing dollar values across stratum 2 and 3 is about 10 percent, meaning that the incentive may move families from reporting on 90 percent of dollar values to reporting on 87 percent on dollar values.[20] Dollar values are a sensitive topic so both levels of item response (87 and 90) can be considered relatively high.

7. **Supporting evidence: a comparison of AP cases between 2007 and 2010.**

All families in the AP sample of the SCF are offered a participation incentive; the incentive offer was $20 in 2007 and increased to $50 in 2010. Coincident with this incentive increase was a

[20] The algorithm that determines the missing dollar value rate penalizes refuse/don't know responses to "do you have…" questions. For example, if respondent did not know whether the family had directly-held mutual funds our algorithm would say that she did not know to all seven follow up questions about the dollar value of each of seven types of mutual funds in the questionnaire.

decrease in the number of attempts needed to gain final cooperation (table 1, top row). As the SCF field period progresses, the base incentive can increase if field staff determine that a larger incentive may be effective. In 2007, then, most AP families received $20 for participation but some received a $50 participation incentive.[21]

The next set of results compare the number of contact attempts in 2010 for the group that received the base $50 incentive in 2010 (the "2010-50" group) to the 2007 respondents who completed the interview for either the base $20 incentive or the escalated $50 incentive (the "2007-20-50" group). The 2010-50 group serves as the counterfactual for the 2007-20-50 group in the thought experiment of "what would the change in field effort have been in 2007 if the 2007 respondents were offered a $50 base incentive instead of a $20 base incentive?"

How effective will the 2010-50 group be as a comparison for the 2007-20-50 group? Observably these two groups appear very similar (table 6). The SCF sampled the same areas in 2007 and 2010, so any underlying difficulties in gaining cooperation should be controlled-for by geographic variables and local macroeconomic variables can help control for economic changes over time. An interviewer ID, unique across waves, allows us to control for interviewer efficacy across years.

However, unobserved propensities to participate in the SCF may have changed between 2007 and 2010 and a comparison of these two groups may pick up this unobserved propensity. However, proxies for willingness-to-participate are available in the interviewer debriefing instrument, which is completed by the interviewer after the survey. Interviewers assess the respondent's interest, use of documents (cooperativeness), and pre-interview suspicion level during a post-interview debriefing.

[21] About 19 percent of 2007 SCF AP families received $50 and 80 percent received $50 or less. About 81 percent of AP families in 2010 received $50.

Figure 2 shows a stylized description of the distribution of unobserved willingness to participate; the respondents who complete the survey for $20 proxy for the "high" willingness group while those that need $50 are a less willing group.[22] In 2007 sixty-four percent of the base incentive ($20) group was judged to be "not suspicious" while the $50 group in 2007 was generally more suspicious (table 7, top panel). Similarly, the base incentive ($20) group in 2007 was more interested and more cooperative during the survey than was the $50 group in 2007.

The 2007 $20 and $50 groups are combined into the 2007-20-50 group and compared to the 2010-50 group (table 7, columns c and d). Overall, among those with either a $20 or $50 incentive in 2007, 59 percent were not suspicious, 33 percent were "somewhat suspicious" and nine percent were "very suspicious" (column c). Importantly, the 2010 base incentive ($50) group had the same distribution of suspicion (column d), which indicates that the unobservable of the $20-to-$50 group from 2007 appear to have the same distribution as their comparison group (the $50 group in 2010). Comparisons of columns c and d in the second panel (interest in the survey) and the third panel (use of supporting documents during the interview) also indicate that the 2007-20-50 group and the 2010-50 group have similar distributions of unobservable propensities to respond.

Similar to the quasi-experiment results, the $50 incentive group (here, the 2010-50 group) needed fewer contact attempts before agreeing to the survey (table 8); here, these families needed 5 to 6 fewer contact attempts.[23]

8. Discussion

The incentive offer in the SCF helps interviewers gain cooperation, as in Rodgers (2002) and James (1997). And the incentive offer appears to have little cost to data quality, as in Shettle and

[22] The 2010-50 group will have both families that would have participated for $20 and those that did participate for $50 in 2010. We cannot separate these families in 2010 but we can in 2007.
[23] The 2010-50 group needed 2 to 3 fewer contact attempts than the 2007-20 group.

Mooney (1999) and Davern, et al (2003). The SCF families that received an initial incentive offer agreed to participate quicker than families that did not receive the initial offer. The incentivized and non-incentivized families expended the same amount of effort during the interview, though the incentivized had higher rates of missing dollar value data.

The $50 incentive offered in 2010 was more effective than the smaller $20 offer in 2007, similar to results found by Church (1993) and Singer (2002). The incentive, offered in person and through an advance mailing, had a small impact in phase one of the SCF contacting strategy but had it biggest impact post-phase one, similar to Singer, et al (2000).

Our preferred estimates imply that 3 to 6 contact attempts can be saved by offering the $50 incentive. As the costs to run a large household survey has increased (Curtin Presser, and Singer, 2005) incentive payments have been proposed as a way of increasing response rates in a cost-efficient manner. Survey administrators can use this reduction in contact attempts to help decide if offering an incentive is cost-efficient.

A synthesis of results in table 2 and table 8 also implies that the incentive offer is most salient for families that are not in the top decile of wealth. In table 2, the impact of the incentive in the nonparametric model is larger than in the OLS model. The nonparametric model gives low weight to comparisons made with little common support (like region B of figure 1) and gives more weight to observations on common support (like region C). Region B roughly corresponds to the top decile of the wealth distribution and region C roughly corresponds to the 50^{th}-90^{th} percentiles. Stratum 2 families in region C complete the SCF with about 6 fewer attempted contacts than stratum 3 families in region C; in region B the difference is less than one attempted contact. The incentive offer appears to be less salient to wealthier families. Further, the results comparing AP families in table 8

(very few of whom are in the top decile) shows that about 6 fewer contact attempts were needed when the incentive was increased in 2010.

Errors in record-of-call (ROC) data should be considered carefully (Biemer, et al, 2011; Wang and Biemer, 2010; Kennickell, 2012). Our main dependent variable (the number of attempted contacts) is likely to be mis-measured because a ROC is only entered once the interviewer *chooses* to enter it, not as part of a random process. If the mis-measurement is random and not correlated with observable or unobservable traits of families then the resulting OLS regressions should be unbiased, but standard errors may be too large (Bound, et al, 2001).

However, a survey of interviewers indicates that ROC reporting error is not random and that interviewers typically choose to under-report the true number of contacts with a family (Wang and Biemer, 2010).[24] Further, the interviewer survey implies that there may be more underreporting among the stratum 3 cases: interviewers may choose to under-report because they fear that field managers will close out a case that the interviewer believes can be converted. If stratum 3 cases are more difficult to work (which is a reasonable assumption considering the higher income and wealth of these families) then the extent of under-reporting will be larger for stratum 3 cases.

The general form for bias in OLS estimates when measurement error in X (denoted as u) and measurement error in Y (denoted by v) are both potentially correlated with X and Y is characterized by:

$$\hat{\beta}^{OLS} = \beta + (\tilde{X}^T\tilde{X})^{-1}\tilde{X}^T(-u\beta + v) \text{ where } \tilde{X} = X + u \text{ and } \tilde{Y} = Y + v\,.$$

[24] The survey of interviewers by Wang and Biemer (2010) shows that interviewers are more likely to under-report rather than over-report because it is known that field managers will examine time sheets as verification of effort. Interviewers may choose to under-report because they fear that field managers will close out a case that the interviewer believes can be converted, and because they do not want to appear to be unproductive or unable to complete their tasks. The survey results also show that unfruitful attempted contacts are most likely to go unreported.

That is, the bias in an OLS estimate is determined by the correlation between mis-measured X ($\tilde{X}$) and its mis-measurement (u), and by the correlation between mis-measured X ($\tilde{X}$) and the mis-measurement in Y (v). Considering the case of equation (4) and assuming no mis-measurement in the Stratum2 dummy variable (the X variable), the correlation between X and v should be positive.[25] The OLS estimate of β from equation (4) is negative in table 2 (theory also suggests it will be negative) so if mis-measurement is accurately described in this example then the estimates reported in table 2 are likely to be a lower bound on the true point estimate.

[25] In this example the dependent variable is positive and under-reported, so $v<0$. If under-reporting is less common in ROCs of stratum 2 families, then v should be less negative when X (the stratum 2 indicator variable) takes the value 1 (indicating a stratum two family) and should be more negative when X takes the value 0 (indicating a stratum three family).

References

Barsky, Robert, John Bound, Kerwin Kofi Charles, and Joseph P. Lupton (2002) "Accounting for the Black-White Wealth Gap: A Nonparametric Approach," *The Journal of the American Statistical Association*, Vol. 97, No. 459, pp. 663-673.

Black, Dan, Amelia Haviland, Seth Sanders, and Lowell Taylor (2008) "Gender Wage Disparities among the Highly Educated" *The Journal of Human Resources*, Vol. 43, pp. 630-658.

Black, Dan, Amelia Haviland, Seth Sanders, and Lowell Taylor (2006) "Why do Minority Men Earn Less? A Study of Wage Differentials of the Highly Educated" *Review of Economics and Statistics*, Vol. 88, No. 2, pp. 300-313.

Biemer, Paul, Patrick Chen, and Kevin Wang, 2011. "Errors in Recorded Number of Call Attempts and Their Effect on Nonresponse Adjustments Using Callback Models" In *Proceedings of 58th World Statistical Congress*, Dublin – Session IPS033.

Bound, John, Charles Brown, and Nancy Mathiowetz, 2001 "Measurement Error in Survey Data" *Handbook of Econometrics Vol. 5*, J. Heckman and E. Leamer, Eds. pp. 3706-3843.

Bricker, Jesse, Arthur Kennickell, Kevin Moore, and John Sabelhaus (2012) "Changes in U.S. Family Finances from 2007 to 2010: Evidence from the Survey of Consumer Finances," *Federal Reserve Bulletin*, Vol. 98, No. 2.

Church, Allan (1993) "Estimating the Effect of Incentives on Mail Survey Response Rates: A Meta-Analysis" *Public Opinion Quarterly*, Vol. 57, No. 1, pp. 62-79.

Curtin, Richard, Stanley Presser, and Eleanor Singer (2005) "Changes in Telephone Survey Nonresponse" *Public Opinion Quarterly*, Vol. 69, No. 1, pp. 87-98.

Davern, Michael, Todd Rockwood, Randy Sherrod, and Stephen Campbell (2003) "Prepaid Monetary Incentives and Data Quality in Face-to-Face Interviews: Data from the 1996 Survey of Income and Program Participation Incentive Experiment" *Public Opinion Quarterly*, Vol. 67, No. 1, 139-147.

Dillman, Don (1978) *Mail and Telephone Surveys: The Total Design Method.* New York: Wiley

DiNardo, John, Nicole Fortin, and Thomas Lemieux (1996) "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semi-Parametric Analysis," *Econometrica*, Vol. 64, pp.1001-1044.

Gouldner, Alvin (1960) "The Norm of Reciprocity: A Preliminary Statement" *American Sociological Review*, Vol. 25, No. 2, pp. 161-178.

Groves, Robert, Eleanor Singer, and Amy Corning (2000) "Leverage-Saliency Theory of Survey Participation" *Public Opinion Quarterly*, Vol. 64, No. 3, pp. 299-308

Hansen, Robert (1980) "A Self-Perception Interpretation of the Effect of Monetary and Non-monetary Incentives on Mail Survey Response Behavior" *Journal of Marketing Research*, Vol. 17, 77-83.

James, Tracey (1997) "Results of the Wave 1 Incentive Experiment in the 1996 Survey of Income and Program Participation Incentive Experiment." In *Proceedings of the Survey Research Section of the American Statistical Association*, pp. 834-39

Kennickell, Arthur (2012) "What's the Chance? Interviewers' Expectations of Response in the 2010 SCF" In *Proceedings of the Survey Research Section of the American Statistical Association.*

Kennickell, Arthur (2005) "Darkness Made Visible: Field Management and Non response in the 2004 SCF" In *Proceedings of the Survey Research Section of the American Statistical Association.*

Kennickell, Arthur (2004) "Action at a Distance: Interviewer Effort and Nonresponse in the SCF" Federal Reserve Board mimeo.

Kennickell, Arthur (2001) "Modeling Wealth with Multiple Observations of Income: Redesign of the Sample for the 2001 Survey of Consumer Finances" In *Proceedings of the Survey Research Section of the American Statistical Association.*

Kennickell, Arthur (1998) "List Sample Design for the 1998 Survey of Consumer Finances" Federal Reserve Board mimeo.

Kennickell, Arthur and Douglas McManus (1993) "Sampling for Household Financial Characteristics Using Frame Information on Past Income" Federal Reserve Board mimeo.

Rodgers, Willard (2011) "Effects of Increasing the Incentive Size in a Longitudinal Survey" *Journal of Official Statistics*, Vol. 27, No. 2, pp. 279-299

Rodgers, Willard (2002) "Size of Incentive Effects in a Longitudinal Study" In *Proceedings of American Association for Public Research – Section on Survey Methods.*

Rosenbaum, Paul and Donald Rubin (1983) "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, Vol. 70, pp. 41-55.

Shettle, Carolyn and Geraldine Mooney (1999) "Monetary Incentives in US Government Surveys" *Journal of Official Statistics*, Vol. 15, No. 2, pp. 231-50.

Singer, Eleanor (2011) "Toward a Benefit-Cost Theory of Survey Participation: Evidence, Further Tests, and Implications" *Journal of Official Statistics*, Vol. 27, No. 2, pp. 379-392.

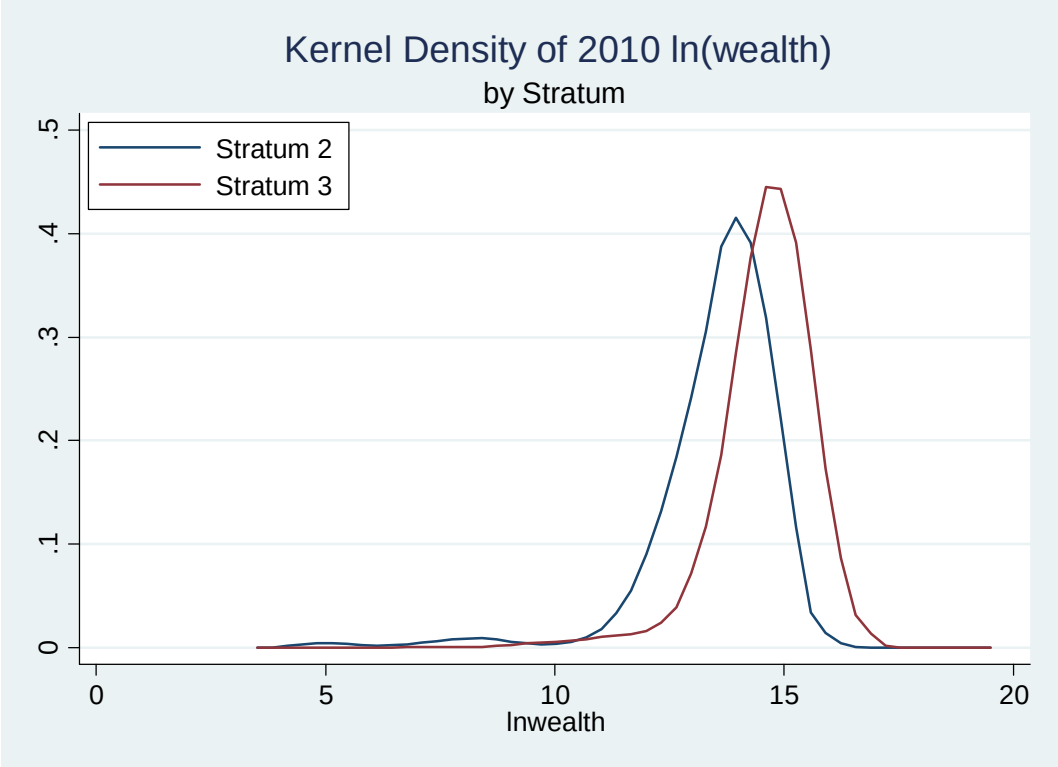
Singer, Eleanor (2002) "The Use of Incentives to Reduce Nonresponse in Household Surveys" Working Paper No. 051, The University of Michigan Institute for Social Research Survey Research Center.

Singer, Eleanor, John van Hoewyk, and Mary Maher, (2000) "Experiments with Incentives in Telephone Surveys" *Public Opinion Quarterly*, Vol. 64, No. 2, pp. 171-188.

Tourangeau, Roger, Robert Johnson, Jiahe Qian, Hee-Choon Shin, and Martin Frenkel (1993) "Selection of NORC's 1990 National Sample" National Opinion Research Center at the University of Chicago working paper.

Wang, Kevin and Paul Biemer (2010) "The Accuracy of Interview Paradata: Results from a Field Investigation" In *Proceedings of American Association for Public Research – Section on Survey Methods.*

Figure 1. Overlap in wealth distribution in stratum two and three



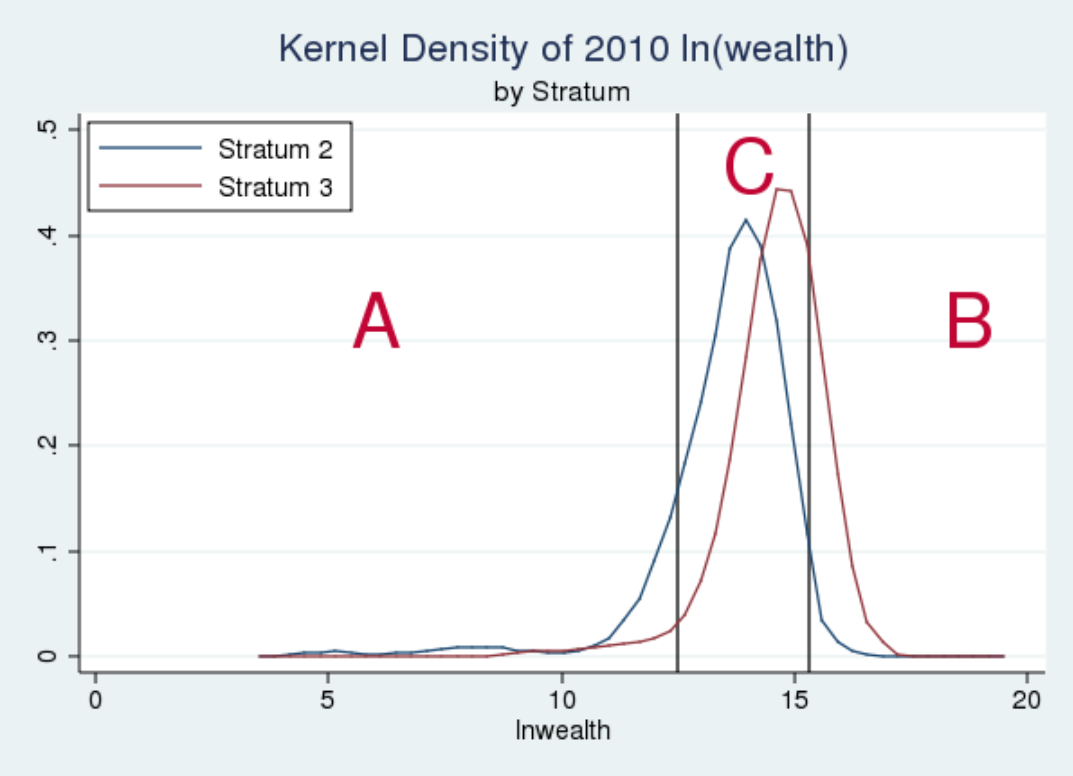
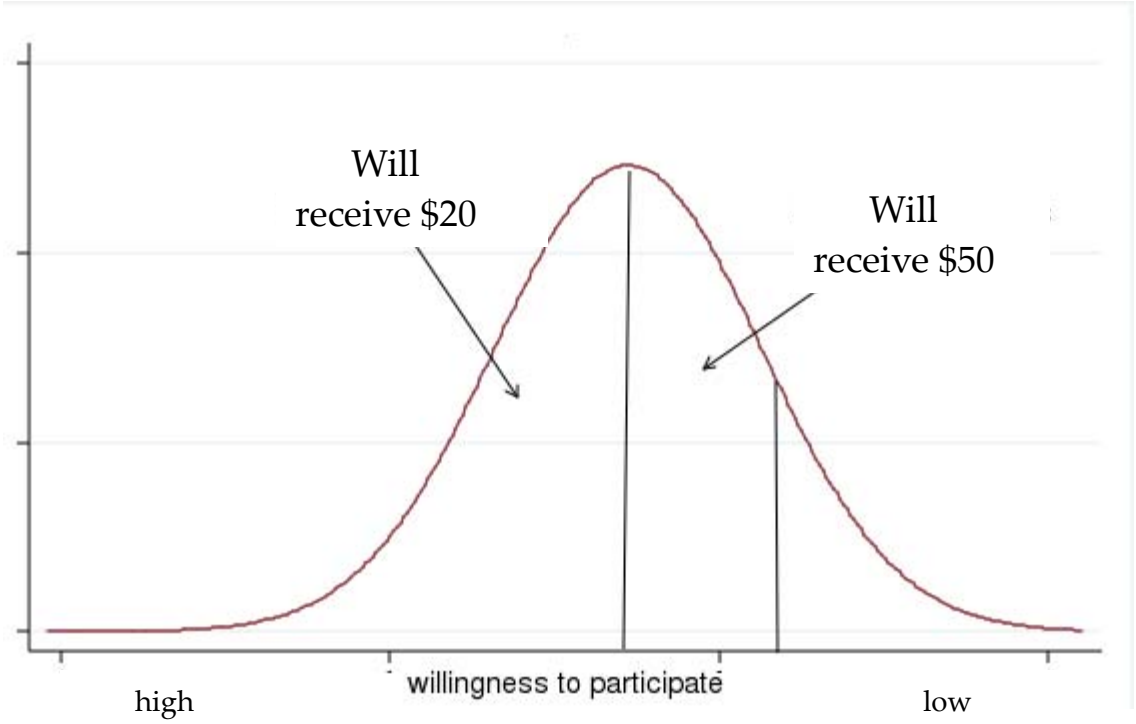


Figure 2. Stylized willingness to participate



Note: In this stylized example, 2007 SCF AP families with high willingness to participate in the SCF will be observed receiving the base incentive ($20) while families with lower willingness will need a higher incentive before participating.

Table 1. Average number of contact attempts in 2010 SCF

Area Probability Sample	7.4
List Sample	
Stratum 1	10.2
Stratum 2	11.3
Stratum 3	13.5
Stratum 4	13.5
Stratum 5	12.9
Stratum 6	13.5
Stratum 7	12.5

Source: 2010 SCF records of call data

Table 2. Difference in contacts between stratum two and three

	2010			
	OLS	OLS	Nonparametric	Semiparametric
Treated (stratum 2)	-2.45	-1.14	-4.65	-3.10
	[-4.77, -0.00]	[-4.68, 1.37]	[-6.44,-3.03]	[-4.56, -1.72]
Demographics	Yes	Yes	Some	Yes
Interviewer FE	No	Yes	No	No
R^2	0.11	0.61	0.04	0.02
N	368	358	366	338
	2007			
	OLS	OLS	Nonparametric	Semiparametric
Treated (stratum 2)	-1.66	1.54	-2.34	-2.06
	[-4.29, 0.72]	[-1.39, 3.89]	[-3.66, -1.28]	[-4.11, -0.51]
Demographics	Yes	Yes	Some	Yes
Interviewer FE	No	Yes	No	No
R^2	0.08	0.55	0.03	0.01
N	381	381	380	376

Bootstrapped 2-sided 95% confidence interval in []; 299 bootstrap replications used. A $50 base incentive was offered in 2010 and a $20 base incentive was offered in 2007. Estimates are based on pooled regressions of stratum two and stratum 3 group members. Full set of demographics includes net worth and income percentile groups, age class dummies, education dummies, race dummies, marital status, and urban status. In models with interviewer fixed effects we also include a variable that reports the fraction of field contacts that the interviewer was responsible for. The nonparametric regressions use a limited set of demographics: net worth categories, education class dummies, and urban status.

Table 3. Difference in contacts between stratum two and three: Compare phase 1 of field period to phase two and above

	2010 Phase one			
	OLS	OLS	Nonparametric	Semiparametric
Treated (stratum 2)	-0.16	0.90	-0.85	-0.35
	[-1.11, 0.92]	[-0.89, 2.71]	[-1.56, -0.25]	[-0.77, 0.09]
Demographics	Yes	Yes	Some	Yes
Interviewer FE	No	Yes	No	No
R^2	0.10	0.42	0.04	0.00
N	368	358	366	338
	2010 Phase two+			
	OLS	OLS	Nonparametric	Semiparametric
Treated (stratum 2)	-2.61	-2.08	-3.53	-2.77
	[-3.96, -1.08]	[-4.45, -0.20]	[-4.54, -2.44]	[-3.63, -1.87]
Demographics	Yes	Yes	Some	Yes
Interviewer FE	No	Yes	No	No
R^2	0.11	0.54	0.06	0.04
N	368	358	366	338

Bootstrapped 2-sided 95% confidence interval in []; 299 bootstrap replications used. Phase one contacts refer to field period contacts prior to express mailing. Estimates are based on pooled regressions of stratum two and stratum 3 group members. Full set of demographics includes net worth and income percentile groups, age class dummies, education dummies, race dummies, marital status, and urban status (see appendix table). In models with interviewer fixed effects we also include a variable that reports the fraction of field contacts that the interviewer was responsible for. The nonparametric regressions use a limited set of demographics: net worth categories, education class dummies, and urban status.

Table 4. Difference in elapsed time between stratum two and three

	2010			
	OLS	OLS	Nonparametric	Semiparametric
Treated (stratum 2)	-11.00	-9.09	-23.19	-14.04
	[-26.35, 11.63]	[-29.07, 3.71]	[-30.93, -14.63]	[-23.37, -4.62]
Demographics	Yes	Yes	Some	Yes
Interviewer FE	No	Yes	No	No
R^2	0.09	0.71	0.06	0.02
N	358	358	356	338
	2007			
	OLS	OLS	Nonparametric	Semiparametric
Treated (stratum 2)	-15.23	-1.83	-16.35	-19.83
	[-28.06, -2.50]	[-14.74, 11.16]	[-22.21, -10.79]	[-30.39, -9.97]
Demographics	Yes	Yes	Some	Yes
Interviewer FE	No	Yes	No	No
R^2	0.06	0.6	0.03	0.03
N	381	381	380	376

Bootstrapped 2-sided 95% confidence interval in []; 299 bootstrap replications used. A $50 base incentive was offered in 2010 and a $20 base incentive was offered in 2007. Estimates are based on pooled regressions of stratum two and stratum 3 group members. Full set of demographics includes net worth and income percentile groups, age class dummies, education dummies, race dummies, marital status, and urban status (see appendix table). In models with interviewer fixed effects we also include a variable that reports the fraction of field contacts that the interviewer was responsible for. The nonparametric regressions use a limited set of demographics: net worth categories, education class dummies, and urban status.

Table 5. Difference in data quality between stratum two and three

	2010			
	Use of documents	Interest in interview	Not suspicious before IW	Pct. dollar values missing
Treated (stratum 2)	0.008	0.06	-0.04	0.03
	[-0.048, 0.061]	[-0.03, 0.16]	[-0.09, 0.02]	[0.02, 0.05]
Demographics	Some	Some	Some	Some
Iwer FE	No	No	No	No
R^2	0	0	0	0.01
N	366	366	366	366

	2007			
	Use of documents	Interest in interview	Not suspicious before IW	Pct. dollar values missing
Treated (stratum 2)	-0.07	0.14	-0.01	0.03
	[-0.13, -0.02]	[0.05, 0.23]	[-0.07, 0.04]	[0.02, 0.05]
Demographics	Some	Some	Some	Some
Iwer FE	No	No	No	No
R^2	0.01	0.01	0	…
N	381	381	381	381

Bootstrapped 2-sided 95% confidence interval in []; 299 bootstrap replications used. A $50 base incentive was offered in 2010 and a $20 base incentive was offered in 2007. Estimates are based on pooled regressions of stratum two and stratum 3 group members. Full set of demographics includes net worth and income percentile groups, age class dummies, education dummies, race dummies, marital status, and urban status (see appendix table). In models with interviewer fixed effects we also include a variable that reports the fraction of field contacts that the interviewer was responsible for. The nonparametric regressions use just some demographics: net worth categories, education class dummies, and urban status.

Table 6. Observable characteristics of 2007 and 2010 AP samples

	2007	2010
Personal Characteristics		
Average age	49.8	50.5
Percent high school graduate	35.6	34.1
Percent some college	18.7	19.6
Percent college graduate	31.0	33.7
Percent white	69.4	66.4
Average family income (thousands $2010)	57.8	56.3
Average family net worth (thousands $2010)	261.6	224.2
Location Characteristics		
Median Census tract income (thousands $2000)	43.3	44.0
Median Census tract house value (thousands $2000)	125.5	125.8

A comparison of 2007 and 2010 AP cases is shown (statistics are weighted). In general, the two samples are observably similar, though the impact of the 2007-2009 recession is evident in family net worth and income statistics. Both the 2007 and 2010 SCF used 2000 Census tract definitions and the same areas were sampled in both years, so differences in median Census tract values between 2007 and 2010 should only be due to differences in sample composition between years.

Table 7. Proxies for unobservable cooperation traits in 2007 and 2010 AP sample

	(a)	(b)	(c)	(d)
	2007			2010
Suspicion before survey	$20	$50	$20-to-$50	$50
Not suspicious	64	45	59	60
Somewhat suspicious	28	43	33	30
Very Suspicious	8	12	9	9
Interest in survey	$20	$50	$20-to-$50	$50
Very high	24	13	21	21
Above average	37	26	34	35
Average	33	52	39	39
Below average	4	9	5	5
Very low	1	1	1	1
Use documents	$20	$50	$20-to-$50	$50
Frequently	9	8	9	9
Sometimes	18	13	16	16
Rarely	14	13	14	12
Never	59	67	62	62

AP cases only. Suspicion, interests, and effort during the survey can proxy for unobserved cooperation traits of SCF families. Similar to stylized description in figure 2, the families that completed the survey for the base fee ($20) in 2007 (column (a)) were less suspicious of the survey before the interview, were more interested in the survey, and were more likely to use documents during the survey than were the families that completed the survey for $50 (column (b)). The families that completed the survey in 2007 for between $20 and $50 (column (c)) had similar cooperation rates as those families that completed the survey for the base incentive ($50) in 2010 (column (d)). Thus, unobserved differences between the comparison groups in table 8 (column (c) and (d)) should be minimal.

Table 8. Change in number of contacts, comparison of 2010 AP families to 2007 AP families

	Base incentive ($50) in 2010, base incentive ($20) and $50 incentive in 2007		
	(1)	(2)	(3)
Treatment (2010)	-6.61	-6.62	-5.68
	[-6.85, -6.54]	[-6.84, -6.54]	[-6.01, -5.45]
Demographics	No	Yes	Yes
IWer FEs	No	No	Yes
R^2	0.23	0.27	0.42
N	6,484	6,484	6,484

Bootstrapped 2-sided 95% confidence interval in []; 299 bootstrap replications used. AP cases only. A comparison of the number of contact attempts needed before a family agrees to participate in the SCF. The comparison is between the 2007 $20-to-$50 group (the comparison group) and the 2010 $50 group (the treated group). Observable traits of these two groups of families are similar (table 6) and unobservable traits of these two groups of families are similar (table 7).

Appendix Table. Characteristics of 2010 stratum 2 and 3 families

	Stratum 2	Stratum 3
Age groups (percent in group)		
% less than age 45	22	14
% between 45 and 65	51	60
% above 65	27	26
% college graduate	65	82
% married	72	85
% white	77	88
Income (thousands $2010)	141	295
Net worth (thousands $2010)	1,093	3,186

Note: Weighted means (stratum 2 and 3 families only). The statistics presented here are an abbreviated set of characteristics used throughout this paper because of disclosure issues. The age dummies used in the regressions are (1) less than 35, (2) between 35 and 45, (3) between 45 and 55, (4) between 55 and 65, (5) between 65 and 75, and (6) above 75. The education dummies are (1) no high school degree, (2) a high school degree, (3) some college, (4) college degree. A series of dummies that describe the family's place in the SCF income distribution (lower than 20th percentile, between 20th and 40th percentile, between 40th and 60th percentile, between 60th and 80th percentile, between 80th and 90th percentile, and above 90th percentile), and the SCF net worth distribution (lower than 25th percentile, between 25th and 50th percentile, between 50th and 75th percentile, between 75th and 90th percentile, between 90th and 95th percentile, and above 95th percentile) are also included. And a race dummy (white/non-white), marital status dummy, and urban status dummy are also included.