Research Data Reports

FEDERAL RESERVE BANK OF BOSTON™

No. 15-7

Identifying and Evaluating Sample Selection Bias in Consumer Payment Surveys

Marcin Hitczenko

Abstract:

This paper develops a two-stage statistical analysis to identify and assess the effect of a sample selection bias associated with an individual's household role. The methodology is applied to the 2012 Survey of Consumer Payment Choice. Survey responses to a series of questions about the respondent's role in household finances are combined and adjusted for response error to estimate a latent variable that represents each individual's share of household responsibility. The distribution of this variable among survey respondents suggests that the sampling procedure favors household members with higher levels of financial responsibility. We also find that the estimated household role relates to the number of noncash payments made. As a result, estimators that do not account for the selection bias overestimate population means for these payment instruments by around 10 percent. We discuss ways to account for this selection bias in future surveys.

JEL Classifications: C11, C42, D12, D13

Keywords: survey error, household economics, Dirichlet regression, Survey of Consumer Payment Choice

---

Marcin Hitczenko is a survey methodologist and a member of the Consumer Payments Research Center in the research department of the Federal Reserve Bank of Boston. His e-mail address is marcin.hitczenko@bos.frb.org.

This paper, which may be revised, is available on the web site of the Federal Reserve Bank of Boston at http://www.bostonfed.org/economic/rdr/index.htm.

This version: November 23, 2015

CONSUMER PAYMENTS Research Center

1 Introduction

The quality of inference based on survey data depends on the ability to recognize and adjust for discrepancies between respondents and the target population, an issue that is relevant to both probability and nonprobability samples. Probability samples are increasingly subject to low response rates, so such samples often represent a random selection of individuals willing to participate in the survey rather than a random selection of the population (Curtin, Presser, and Singer 2005; Tourangeau and Plewes 2013). As a result, researchers often rely on techniques such as propensity modeling (Valliant and Dever 2011) and calibration (Deville and Sarndal 1992) to adjust for systematic nonresponse.

At the same time, nonprobability sampling, which bases selection primarily on availability, has been gaining popularity due to the relative ease and low cost of data collection (Groves 2006). In a notable example of its potential effectiveness, Wang et al. (2009) used multilevel regression and poststratification (Park, Gelman, and Bafumi 2004) to make election forecasts for the 2012 U.S. presidential election based on the responses of X-box users, a subset of the population that is highly unrepresentative of the voting public. More recently, the *Journal of Survey Statistics and Methodology* devoted a Task Force Report and discussion to the use of nonprobability samples in surveys, stressing the need to develop methodologies for using such data to make population inferences (Baker et al. 2013). A major component to accomplishing this goal is understanding how samples differ from the population with respect to heterogeneous clusters of the population.

In the context of surveys, ex post adjustments are usually based on easily measured variables with known distributions in the target population, such as age, gender, or income. In this work, we focus on identifying and assessing the effect of an overlooked source of heterogeneity related to household structure and dynamics. In accordance with the theories on household economics, most notably those outlined by Becker (1991), efficiency dictates that the performance of household tasks is often concentrated among a subset of household members. For example, having one person responsible for purchasing all weekly groceries limits trips to the store and reduces the need for coordination. As a result, the heterogeneity of peoples' behavior is at least partially related to having different roles within a household.

The sampling of individuals inherently involves selection from within households, so a tendency to select certain types of members results in unrepresentative samples. Failing to account for this tendency in estimation is likely to produce biased estimates for any variable related to the respondent's household role. This paper takes advantage of a unique aspect of the 2012 Survey of Consumer Payment Choice (SCPC), in which certain respondents come from the same household, to study the presence and ramifications of this type of selection bias within the context of consumer payments studies.

The paper proceeds as follows. In Section 2, we describe the 2012 SCPC sample attributes and survey variables that are used in our analysis. Section 3 introduces a model that maps these variables to a measure of a respondent's true level of household responsibility and discusses a procedure to estimate all model parameters. In Section 4, we use our model estimates to argue that the 2012 SCPC selection process favors household members with higher shares of financial responsibility. Section 5 studies how the estimated household contributions of the respondents relate to the share of household payments for four types of payment instruments, while Section 6 offers insight into the magnitude of the bias in population mean estimates due to the apparent sample selection bias. Finally, a discussion of the results, including implications for the effective design and analysis of consumer surveys, is featured in Section 7.

2 Data

The data considered in this paper come from the 2012 Survey of Consumer Payment Choice (Schuh and Stavins 2014), a survey fielded annually by the Consumer Payments Research Center (CPRC) at the Federal Reserve Bank of Boston. Respondents to the SCPC are adults (18 years of age or older) selected from the RAND Corporation's American Life Panel (ALP). The SCPC asks respondents to provide information about adoption and use, both current and past, of different payment instruments, along with other information about their consumer decisions and preferences. A fundamental goal of the SCPC is to estimate parameters, such as the average number of monthly payments made, for the entire U.S. adult population.

2.1 Sample Selection

This paper's analysis is strictly limited to the 2012 SCPC sample. However, the mechanisms of how respondents are recruited determine to what extent our findings can be generalized to other samples or sampling methodologies. In 2012, an invitation to participate in the SCPC was sent to 2,473 individuals who had taken the survey in previous years and to 1,197 newly recruited individuals selected randomly from within certain demographic groups. Of those who had taken the SCPC before, around 900 were part of the 1,100 or so original SCPC panelists in 2008, with the rest, again, chosen at random from within certain demographic subsets of the ALP at some point from 2009 to 2011. All the respondents were offered a financial incentive to complete the survey. More details can be found in Schuh and Stavins (2014).

Because around 85 percent of the invitees agreed to participate and over 50 percent of the 2012 ALP panelists took the 2012 SCPC, the results likely reflect how well the ALP represents the U.S. adult population. The ALP is a well-established source of respondents, used by academics and government agencies in over 400 surveys spanning behavioral, political, and economic studies. Perhaps the most recognizable is the 2014 Midterm Election Opinion Study, a collaboration between the RAND Corporation and the New York Times's website, The Upshot (Carman and Pollard 2014). The ALP's popularity makes insight into the panel's potential deviations from the U.S. population of sufficient interest.

The ALP is composed of two fundamentally different types of panelists. Around 80 percent were recruited through well-recognized procedures, albeit from a series of as many as eight different methodologies, rather than one concerted effort. Initially, RAND contacted individuals from other already-existing cohorts, most notably respondents of the University of Michigan's Survey of Consumers, who were originally chosen through stratified random digit dialing. From 2011 to 2012, the ALP relied on address-based sampling in ZIP codes with high percentages of Hispanics and low-income households to recruit around 2,000 individuals. Other sources of panelists include experiments in mail-based and phone recruiting. The set of the above panelists most closely resemble those found in other panels and probability samples.

In addition, RAND invited individuals already in the ALP to volunteer fellow household members as panelists. Around 20 percent of the ALP was recruited in this manner. The availability of multiple respondents from the same household allows for interesting household-based studies, and this feature is crucial to this paper's analysis. Nevertheless, this multiplicity is a rare feature for online, opt-in panels and probability samples. Thus, to make conclusions as generalizable as possible, we study selection bias among those SCPC respondents who represent the first point of contact for a household. While recruitment into the 2012 SCPC is a complicated, multi-stage process, there is no obvious reason why the results for this subset could not apply to other respondent sources.

2.2 2012 Survey of Consumer Payment Choice

The SCPC is a consumer-focused survey, as respondents provide answers as individuals and not as household representatives. The 2012 SCPC cohort features 3,176 respondents from 2,822 households. Table 1 shows the breakdown of these households by a cross-section of the total number of adults in the household and the number that are featured in the sample. While most households are represented by one individual, 303 multi-person households are represented by at least two individuals, and 243 multi-person households were fully sampled, meaning all adult members in those households participated in the survey.

	Household Size (# Adults)							
# Adults in Sample	1	2	3	4	5	6	7	8
1	754	1,269	303	135	48	6	2	2
2		207	28	19	2	4	1	0
3			27	5	1	0	0	0
4				9	0	0	0	0

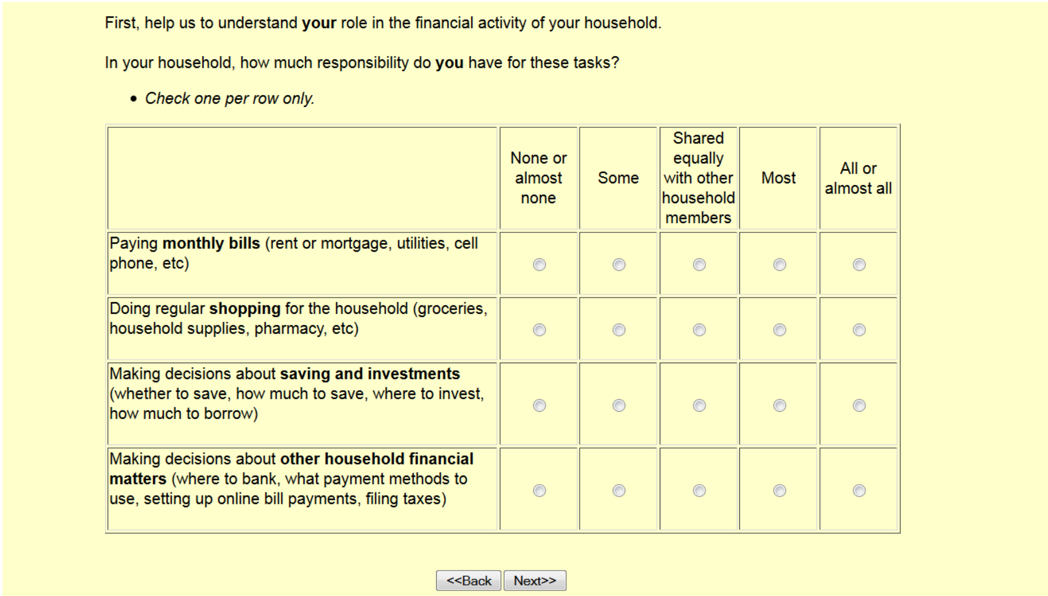
Table 1: Sample Counts of Households by the Number of Adults in the Household and the Number of Adults in the SCPC

Each ALP member completes the "My Household Questionnaire" (MHQ) every three months to keep individual and household demographic information up to date. The data from the MHQ can be used to determine the ages of every household member. Occasional discrepancies in the information provided by members of the same household arise, though this is usually limited to the ages and number of children. If inconsistencies cannot be clarified via other means, the number of adults is estimated by rounding the average number of adults reported by each member.

Insight into an individual's role within the household is gained from a set of four "financial responsibility" questions, a screenshot of which is shown in Figure 1. Specifically, the individual is asked to rank his or her responsibility for "paying bills," "household shopping," "making decisions about saving and investments," and "making decisions about other household financial matters" on a scale of 1 to 5, with 1 representing "no or almost no responsibility" and 5 representing "all or almost all responsibility." A summary of the response distributions among 2012 SCPC respondents is provided in Table 2. For each of the four questions, the responses are notably skewed toward higher degrees of responsibility. The MHQ also provides information for the relative rank of the respondent's income within the household, but this data is not used in the following analysis.

	None or Almost None	Some	Shared Equally	Most	All or Almost all
Paying Bills	13.0	9.5	13.4	7.3	56.8
Shopping	6.5	13.0	20.1	15.0	45.4
Saving and Investments	9.5	8.1	29.1	12.1	41.2
Other Decisions	8.1	7.9	28.4	12.7	42.9

Table 2: Percent of Respondents Providing Each Rating


First, help us to understand **your** role in the financial activity of your household.

In your household, how much responsibility do **you** have for these tasks?

- *Check one per row only.*

	None or almost none	Some	Shared equally with other household members	Most	All or almost all
Paying **monthly bills** (rent or mortgage, utilities, cell phone, etc)	○	○	○	○	○
Doing regular **shopping** for the household (groceries, household supplies, pharmacy, etc)	○	○	○	○	○
Making decisions about **saving and investments** (whether to save, how much to save, where to invest, how much to borrow)	○	○	○	○	○
Making decisions about **other household financial matters** (where to bank, what payment methods to use, setting up online bill payments, filing taxes)	○	○	○	○	○

<<Back Next>>

Figure 1: Screenshot of the Financial Responsibility Questions in the 2012 SCPC

3 Modeling Household Financial Responsibility

A household member's share of responsibility for any particular activity can be defined through a variety of metrics. For the purposes of this work, however, it is only important that the reported categorizations of share responsibility provided by each household in the 2012 SCPC offer insight into the true responsibility of shares for some well-defined metric. We define the "financial responsibility quotient" (FRQ) as the average of the shares of responsibility held by an individual for each of the four activities featured in the SCPC (shown in Figure 1). The FRQ is just one possible measure of an individual's role in the household. Other measures, based on different sets of activities, different weights placed on those activities, or predetermined metrics used to define the responsibility share, can also be generated and are briefly discussed in later sections.

This section is devoted to developing a statistical model that permits posterior distributions to be generated for the FRQs of all respondents in the 2012 SCPC, as well as a posterior distribution for FRQs in the entire population. The presence of households with several sampled members is fundamental to this process because the extent to which reported shares from such households are consistent with each other provides valuable insight into the reliability of the responses. Information about the response accuracy can then be used to define posterior distributions for the FRQs of respondents who are the sole representatives of a given household.

3.1 Financial Responsibility Quotients

As the SCPC focuses on adult consumers, households, indexed by subscript h, are classified by the number of adults in the household, n_h. The FRQ of individual i in household h is depicted as λ_{hi}. Because the FRQs represent proportions, there are the added conditions that $\lambda_{hi} \in [0, 1]$ and

$$\sum_{i=1}^{n_h} \lambda_{hi} = 1,$$

for all h. We let $\lambda_h = \{\lambda_{h1}, \ldots, \lambda_{hn_h}\}$ represent a random permutation of FRQs for the entire household. The permutation means that λ_h defines the absolute distribution of the FRQs without considering to which household members these likely correspond.

As a flexible model for finite-dimensional, discrete probability distributions, we assume a Dirichlet distribution for the vector of FRQs of a household, λ_h:

$$\lambda_h \sim \text{Dirichlet}(n_h, \tau_h). \tag{1}$$

The parameterization in (1) corresponds to a Dirichlet distribution for a vector of length n_h with the concentration parameter for each component being τ_h. This model can be derived as an approximation to a model in which the vector of FRQs, ordered according to known and expected sizes, is randomly permuted to create λ_h. Details of this derivation are found in Appendix A.

The model in (1) implies that the expectation $\text{E}[\lambda_{hi}] = \frac{1}{n_h}$ for all i, and the concentration parameter τ_h defines the degree of dispersion of the FRQs within a household. Larger values of τ_h condense the probability mass of λ_{hi} around $\frac{1}{n_h}$ for each i. More specifically, the λ_{hi} are exchangeable and follow a $\text{Beta}\left(\tau_h, \tau_h(n_h - 1)\right)$ distribution. In this formulation, the concentration parameter differs across households, allowing for heterogeneity across different types of households.

For most households in the 2012 SCPC, λ_h is not observed in full, just a subset of size $s_h \leq n_h$. Since which household members are selected into the sample and the order in which they enter might relate to their role within the household, the assumption of exchangeability no longer applies. Thus, unlike for λ_h, the order of FRQs for sampled members, $\lambda_h^o \subseteq \lambda_h$, does matter: λ_{hi}^o, the i^{th} element of λ_h^o, represents the FRQ of the i^{th} household member selected into the ALP. In particular, the FRQs of the first recruits, those members who were the original point of contact for the household, are represented by λ_{h1}^o. Throughout the remainder of the paper, the superscript "o" serves to signify observed data.

3.2 Financial Responsibility Scores

We now propose a model that relates the responses to the four financial responsibility questions shown in Figure 1 to the FRQs. We let F_{hij} represent the response given to the j^{th} financial responsibility question by individual i in household h. The statistic that naturally corresponds to the FRQ is the financial responsibility score, defined as $F_{hi} = F_{hi1} + \ldots + F_{hi4} - 4$. These scores take integer values on $[0, 16]$, and higher scores represent higher reported levels of responsibility within the household.

Ideally, the financial responsibility score directly reflects the true FRQ through the relationship $F_{hi} = [16\lambda_{hi}]$, where the brackets denote rounding to the nearest integer. Of course, a certain amount of divergence is expected between λ_{hi} and the survey-implied quotient, $\frac{F_{hi}}{16}$. There are many sources of discrepancy, which we collectively dub "reporting error," despite the fact that not all discrepancies reflect a mistake made by the respondent. For example, household members younger than 18 years who have some household responsibility will introduce a disparity between the survey results and λ_{hi} as we have defined it. Reasonable differences in interpretations of the four areas of financial responsibility, the use of different metrics to define the responsibility share, and disagreements on the difference between "almost none" and "some" among members of the same household may skew household scores as well. Finally, more traditional sources of response error such as straightlining and misunderstanding the question also affect the distribution of responses (Schaeffer and Presser 2003). Nevertheless, we expect that F_{hi} and λ_{hi} will be correlated, with the former serving as a noisy estimate of λ_{hi}.

A sense of this noise can be gathered from Figure 2, which shows the reported values of F_{hi} for individuals from the same household. As expected due to the fact that more responsibility for one household member generally implies less for the others, the plot shows a negative correlation between the reported scores of one household member with the sum of scores from all other members. At the same time, there are instances in which there is an apparent inconsistency. For example, there are two households of size $n_h = 2$, in which both adults are sampled and both adults provide scores of $F_{hi} = 16$. Other insight into response error comes from a study of individuals representing one-adult households. Roughly 80 percent of these respondents claimed full financial responsibility ($F_{hi} = 16$), and 90 percent claimed a large majority of such responsibility($F_{hi} \geq 12$). Overall, despite evidence of some response errors, we believe that survey responses offer useful insight into financial responsibility levels within the household.

The F_{hi} follow some probability distribution on the integers from 0 to 16, defined at least partly by the

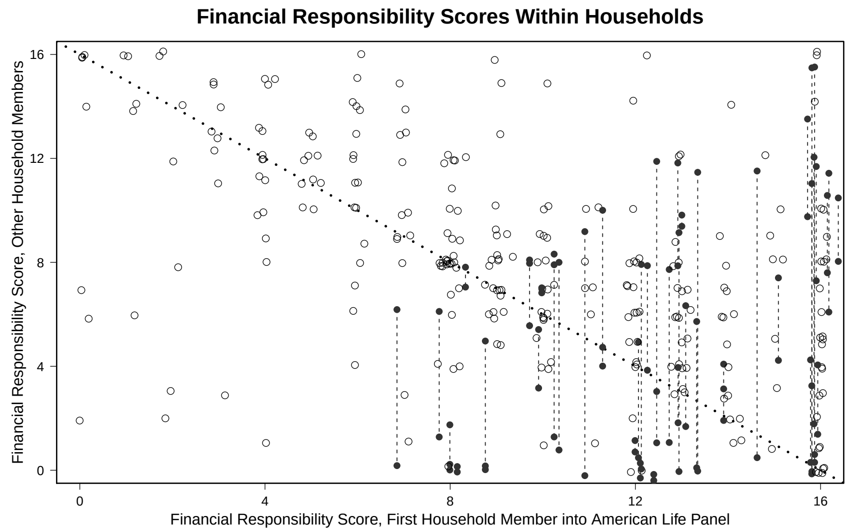


Figure 2: Financial Responsibility Scores in the 2012 SCPC for Members of the Same Household
Note: Filled circles and dotted lines show scores for households with more than two members in the sample.

true λ_{hi}. In this work, we consider the following model:

$$\mathrm{P}(F_{hi} = f \mid \lambda_{hi}) \propto \begin{cases} \exp\left(-\phi_\ell[h]\left|\lambda_{hi} - \frac{f}{16}\right|\right), & \text{if } \lambda_{hi} \geq \frac{f}{16} \\ \exp\left(-\phi_u[h]\left|\lambda_{hi} - \frac{f}{16}\right|\right), & \text{if } \lambda_{hi} < \frac{f}{16}. \end{cases} \tag{2}$$

The assumed model in (2) has several attractive features. First, under the assumptions that $\phi_u[h], \phi_\ell[h] \geq 0$, it assures that $F_{hi} = [16\lambda_{hi}]$ is at least as likely as all responses. In fact, if both parameters are greater than 0, $[16\lambda_{hi}]$ will be the mode. Second, by virtue of using two parameters, $\phi_\ell[h]$ and $\phi_u[h]$, the model allows for different likelihoods of over- and under-estimation of one's household responsibility. This is important, as the data shown in Figure 2 suggest that respondents are more likely to overestimate their share of financial responsibility. As with the distribution of λ_h, the degree and direction of the reporting error associated with $F_{hi} \mid \lambda_{hi}$ is allowed to depend on the household to which the individual belongs. Again, this flexibility recognizes that certain types of households might have more ambiguous household dynamics, resulting in greater discord between true and reported responsibility.

A full model for the observed household scores must establish the dependence of reporting error among members of the same household. Assuming that each individual's survey responses are not influenced by those of other household members implies that

$$\mathrm{P}(F_h^o \mid \lambda_h^o) = \prod_{i=1}^{s_h} \mathrm{P}(F_{hi}^o \mid \lambda_{hi}^o). \tag{3}$$

However, a study of the data reveals an interesting pattern relating to the reported scores of individuals who are the sole representatives of their household. Figure 3 depicts the distribution of financial responsibility scores for the primary individuals sampled from two-adult households and households of three or more adults. A distinction is made between households in which multiple members responded to the survey and those in which only one member responds. Although the general shape of the distribution is similar in both cases, it is clear that for households with only one sampled member ($s_h = 1$), the proportion of cases for which $F^o_{h1} = 16$ is significantly higher. Unfortunately, it is impossible to determine if this

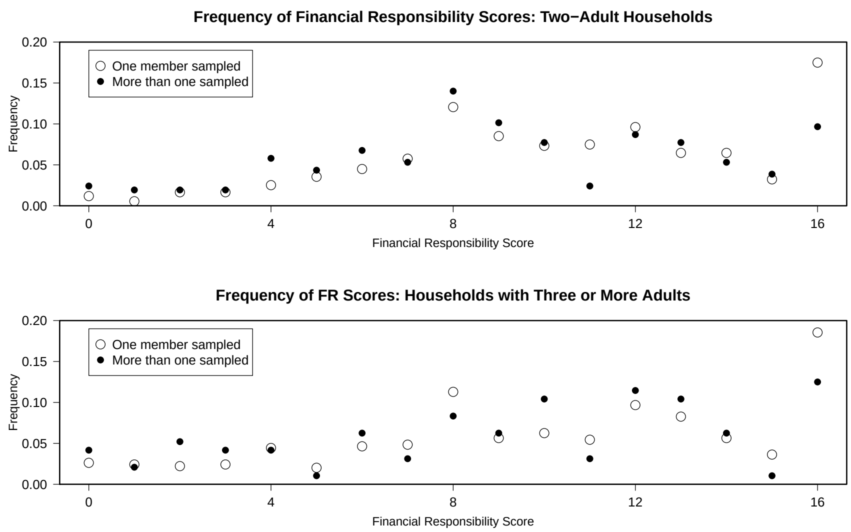


Figure 3: Distribution of Financial Responsibility Scores for the First Household Member Sampled
Note: Results are shown for those who are the sole household representative and for those who have at least one more household member in the 2012 SCPC.

phenomenon is due to differences in selection or in reporting. Selection differences might arise from the fact that households with a dominant decisionmaker are less likely to volunteer other household members into the panel, while reporting differences may stem from the fact that respondents who know that no other household member is taking the survey are more likely to exaggerate their financial role or take the question less seriously in responding. Nevertheless, we adopt the independence assumption given in (3), with the hope that it is largely true. Independence across households is also assumed.

3.3 Estimating Financial Responsibility Quotients

This section presents the methodology for estimating the FRQs of the 2012 SCPC respondents. We let $F^o = \{F^o_h\}$ and $\lambda^o = \{\lambda^o_h\}$ define the set of scores and quotients for all observed members of households featured in the 2012 SCPC. Adopting a Bayesian framework, we would like to sample from

$$\mathrm{P}(\lambda^o \mid F^o) \quad \propto \quad \mathrm{P}(F^o \mid \lambda^o_h)\mathrm{P}(\lambda^o). \tag{4}$$

For one-adult households, determining λ_{hi} is trivial, as our prior assumptions force $\lambda_{hi} = 1$ with a probability of 1. However, for multi-adult households, the posterior distribution in (4) is defined by many components. The three major factors relating to inference of λ^o_{hi} are:

1. The distribution (d) of λ_h within households: $\mathrm{P}(\lambda_h \mid \nu_d)$. In our notation, $\nu_d = \{\tau_h\}$, the Dirichlet concentration parameter in (1).
2. The reporting error (e) in mapping λ_{hi} to F_{hi}: $\mathrm{P}(F^o_h \mid \lambda^o_h, \nu_e)$. In our notation, $\nu_e = \{\phi_\ell[h], \phi_u[h]\}$.
3. The selection process (s), which might favor certain types of individuals: $\mathrm{P}(\lambda^o_h \mid \lambda_h, \nu_s)$.

Without making strong assumptions, the model's multiple parts present identifiability problems for our data, most notably due to confounding between reporting error and the model's other features. Perhaps the most obvious form of unidentifiability relates to cases in which only a subset of household members are featured in the sample, leading to incomplete household financial responsibility scores, thus that $F^o_h \subseteq F_h$. In this case, it is impossible to distinguish between the effects of the selection process and reporting error. For example, in a three-adult household, if two members are sampled and provide scores of $F^o_h = \{2, 2\}$, it is unknown whether the observed scores reflect a process that tends to select individuals who do not shoulder much financial responsibility within the household, or if the observed data is a result of respondents underestimating share of responsibility.

Fully-sampled households can also present identification issues, in which dispersion of the λ_{hi} and error cannot be distinguished. If the observed F_{hi} are distributed uniformly on $0 - 16$, the likelihood $\mathrm{P}(\lambda^o, \tau_h \mid F^o)$ can be made arbitrarily high by claiming independence between F_{hi} and λ_{hi}. Letting $\phi_\ell = \phi_u = 0$ will match the observed distribution of F^o and the whole likelihood can be driven up by increasing τ_h and concentrating all λ_{hi} around $\frac{1}{n_h}$. The resulting inference — that every individual has an equal role within each household — shows that posterior maximization may not correspond to reality.

As a solution, we first estimate the reporting error, ν_e, and then make all inferences about the remaining parameters conditionally on the estimated ν_e. We estimate ν_e so that the reporting error component contributes as little to the variation in the F_{hi} as possible. Specifically, we are interested in the ν_e that maximizes the likelihood,

$$\mathcal{L}(\nu_e, \lambda^o) = \prod_{\{h | n_h = s_h\}} \mathrm{P}(F^o_h \mid \nu_e, \lambda^o_h). \tag{5}$$

Considering only fully-sampled households removes confounding between the reporting error and selection biases. There are no constraints on λ^o_h other than $\sum_{i=1}^{n_h} \lambda^o_{hi} = 1$, so that as much of the variation in the F^o_{hi} as possible is explained by variation in λ^o_{hi} rather than reporting error. We define $\hat{\nu}_e$ as the set of parameters that optimizes the likelihood in (5), which will be highest when some measure of distance from λ^o_{hi} to $\frac{F^o_{hi}}{16}$ is minimized. Details on how this likelihood is optimized are provided below in Section 3.4.

Fixing the parameters relating to the reporting error is akin to making a very strong prior assumption about the reporting error's role in determining the observed financial responsibility scores, as we are forcing as little distinction between truth and reporting as possible in the data. Once $\hat{\nu}_e$ is estimated, we consider the probability distribution:

$$\mathrm{P}(\lambda^o, \nu_d, \nu_s \mid F^o, \hat{\nu}_e) \propto \mathrm{P}(F^o \mid \lambda^o, \hat{\nu}_e)\mathrm{P}(\lambda^o \mid \nu_s, \nu_d)\mathrm{P}(\nu_s, \nu_d).$$

A priori, we assume that the number of individuals and their respective FRQs are chosen at random from the household, independently from the quotient values themselves. Considering the probability function describing the selection process conditionally on the full set of household quotients, λ_h, this means:

$$\begin{aligned} \mathrm{P}(\lambda^o_h \mid \lambda_h, \nu_d, \nu_s) &= \sum_{s_h=1}^{n_h} \mathrm{P}(\lambda^o_h \mid \lambda_h, s_h, \nu_d, \nu_s)\mathrm{P}(s_h \mid \lambda_h, \nu_d, \nu_s) \\ &\propto \sum_{s_h=1}^{n_h} \frac{1}{\binom{n_h}{s_h}} \frac{1}{n_h}. \end{aligned} \tag{6}$$

As stated, the second line in (6) makes two key assumptions: first, that conditional on the number of sampled household members, s_h, every set of s_h members is equally likely to be selected and second, that the number of household members selected into the SCPC, s_h, is uniform on integers from 1 to n_h. Making these two assumptions effectively defines the prior for ν_s and indicates a belief that there is no selection bias. Because (6) implies that λ^o_h is independent of λ_h, integrating with respect to λ_h yields

$$\mathrm{P}(\lambda^o_h \mid \nu_d, \nu_s) \propto \mathrm{P}(\lambda^o_h \mid \nu_d),$$

with the right-hand side defining the distribution of a randomly chosen subset of size s_h from a vector of size n_h itself drawn from the $\mathrm{Dirichlet}(n_h, \tau_h)$ distribution. The aggregation property of the Dirichlet distribution makes this a tractable distribution. Because we are not making predictions for unobserved respondents, there is no need to specify and estimate the selection process. If there is a selection bias in the SCPC, it will be reflected in the posterior estimates of λ^o. The posterior distribution is thus given by

$$\mathrm{P}(\lambda^o, \nu_d \mid F^o, \hat{\nu}_e) \quad \propto \quad \mathrm{P}(F^o \mid \lambda^o, \hat{\nu}_e)\mathrm{P}(\lambda^o \mid \nu_d)\mathrm{P}(\nu_s, \nu_d). \tag{7}$$

A discussion of how ν_d and ν_e are defined, along with details of the parameter estimation, are provided in Section 3.4.

3.4 Model Specification

For the purposes of this work, households are classified into two groups, defined by the number of adults in the household. Specifically, we differentiate between households with two adults and those with three or more adults. To this end, we define

$$c_h = \begin{cases} 1, & n_h = 2 \\ 2, & n_h \geq 3, \end{cases}$$

to identify the classification of household h. This delineation is primarily motivated by economic intuition. The set of households with two adult members is primarily composed of life partners for whom cooperation and the sharing of household decisions is a major component of the relationship. Households with three or more adults will feature a wider array of family arrangements and, as a result, might show a higher level of variation in financial responsibility levels within the household. A better sense of the total number of financial decisions being made within a household might also enable members of smaller households to better assess their own contribution. While it would be interesting to make finer distinctions by household size for the larger households, the practical limitation of having only 27 fully-sampled three-adult households and nine fully-sampled four-adult households makes this difficult.

Based on this reasoning, parameters relating to the distribution of quotients, ν_d, and to the reporting error, ν_e, are defined separately for each class of households. Thus, we define $\nu_d = \{\tau_1, \tau_2\}$, with τ_h in (1) equal to τ_{c_h}. Similarly, we define the error parameters as $\nu_e = \{\phi_{\ell 1}, \phi_{u1}, \phi_{\ell 2}, \phi_{u2}\}$, with the first two elements defining the error function in (2) for households with $n_h = 2$ and the remaining two doing the same for households for which $n_h \geq 3$.

We estimate ν_e for the two types of households by sampling $1{,}000$ draws from the $\text{Dirichlet}(1, n_h)$ for each household. For $c = 1, 2$, we conduct a grid search over $\phi_c = \{\phi_{\ell c}, \phi_{uc}\} \in [0, 2] \times [0, 2]$ to identify the pair that maximizes

$$\sum_{\{h|n_h = s_h\}} \max_{k=1}^{1{,}000} \log \mathrm{P}(F_h^o \mid \lambda_h^o[k], \phi_c), \tag{8}$$

with $\lambda_h^o[k]$ representing the k^{th} draw of λ_h^o and each evaluation corresponding to the log of (5).

All other parameters are estimated conditionally on the estimated $\hat{\phi}_c$ through a Markov Chain Monte Carlo (MCMC) algorithm. In each iteration of the algorithm we do the following:

1. For $c = 1, 2$ draw τ_c from

$$\mathrm{P}(\tau_c \mid \lambda_h^o) \quad \propto \quad \left[\prod_{\{h|c_h = c\}} \mathrm{P}(\lambda_h^o \mid \tau_c) \right] \mathrm{P}(\tau_c).$$

   We assume a noninformative prior $\mathrm{P}(\tau_c) \propto 1$.

2. For each household, h, draw λ_h^o from

$$\mathrm{P}(\lambda_h^o \mid F_h^o, \tau_{c_h}, \hat{\phi}_{c_h}) \quad \propto \quad \mathrm{P}(\lambda_h^o \mid \tau_c) \mathrm{P}(F_h^o \mid \lambda_h^o, \hat{\phi}_{c_h}).$$

3.5 Results

The first step of the estimation process is to determine ϕ_c for $c = 1, 2$ by maximizing the log-likelihood in (8). The resulting contour plots from the grid search, shown in Figure 4, indicate that the log-likelihood is maximized at $\phi_1 = \{1.3, .4\}$ and $\phi_2 = \{0.8, 0.2\}$. Larger values for ϕ_1 than for ϕ_2 suggest that the distribution of survey-implied quotients is more closely concentrated around the true quotients in two-adult households than for larger households. The fact that $\phi_{\ell c} < \phi_{uc}$ in both cases confirms that individuals are much more likely to overstate their household share of financial responsibility than to understate it. As an illustration of both phenomena, Figure 5 shows the distribution of $F_{hi} \mid \lambda_{hi}$ for $\lambda_{hi} = 0.2$ and $\lambda_{hi} = 0.8$ as estimated for an individual from each type of household. Despite nontrivial error and a tendency to overstate responsibility, the results suggest that F_{hi} is, generally speaking, a good proxy for λ_{hi}.

We run the MCMC algorithm with a burn-in period of 1,000 iterations, sampling 200 posterior draws for each parameter of interest by recording every 15^{th} iteration of the algorithm. The posterior estimates of τ are shown in Figure 6. Again, the results suggest that the dynamic in two-adult households is different

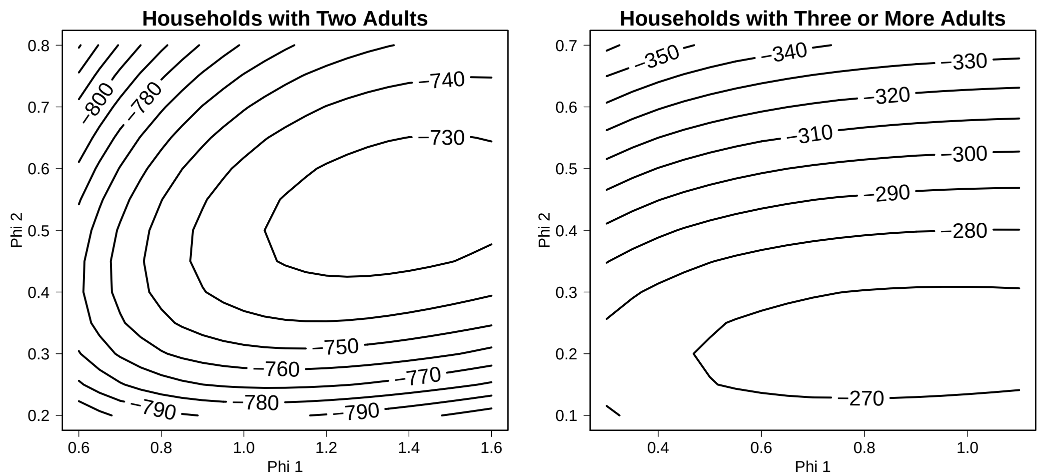


Figure 4: Contours of the Log-Likelihood of Observed Financial Responsibility Scores based on Different Error Parameters, ν_e, by Household Type

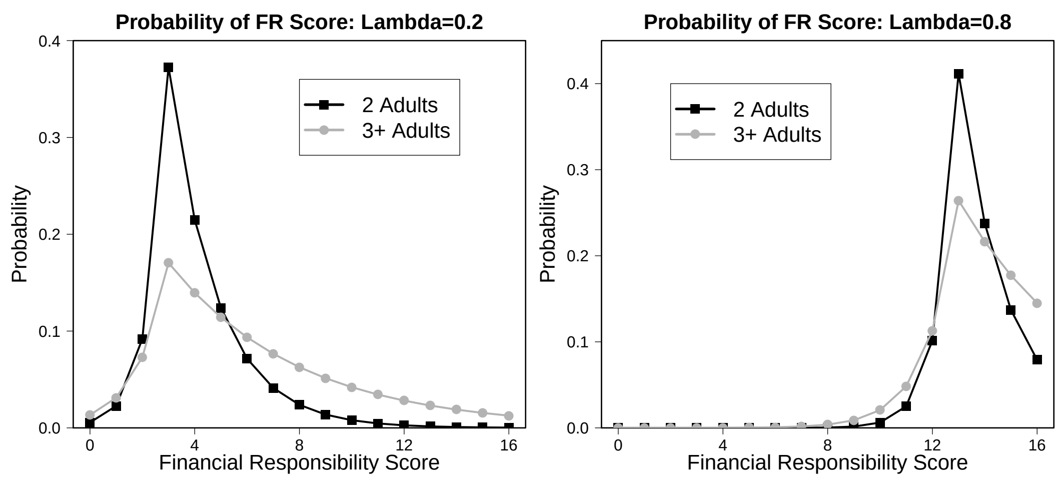


Figure 5: Distribution of Financial Responsibility Scores for λ_{hi} of 0.2 and 0.8 for Individuals from Two-Adult and Multi-Adult Households

from those in households with more adults. In the former, financial responsibility tends to be shared more evenly (meaning a higher concentration around $\frac{1}{n_h}$), perhaps because larger households include life partners housing adult children or elderly relatives, who might have less say in the household's financial decisions.

Figure 7, depicting the posterior draws of λ^o_{hi} for five households, illustrates how the estimation procedure links likely FRQs to the survey-implied quotients, $\frac{F_{hi}}{16}$, in different situations. The first two panels show two fully-sampled households, one consisting of two adult members and one with three adults, in which the members' financial responsibility scores show a fair amount of internal consistency (defined by how close the sum of scores is to 16). The result is that the posterior estimates of λ^o_{hi} are dispersed

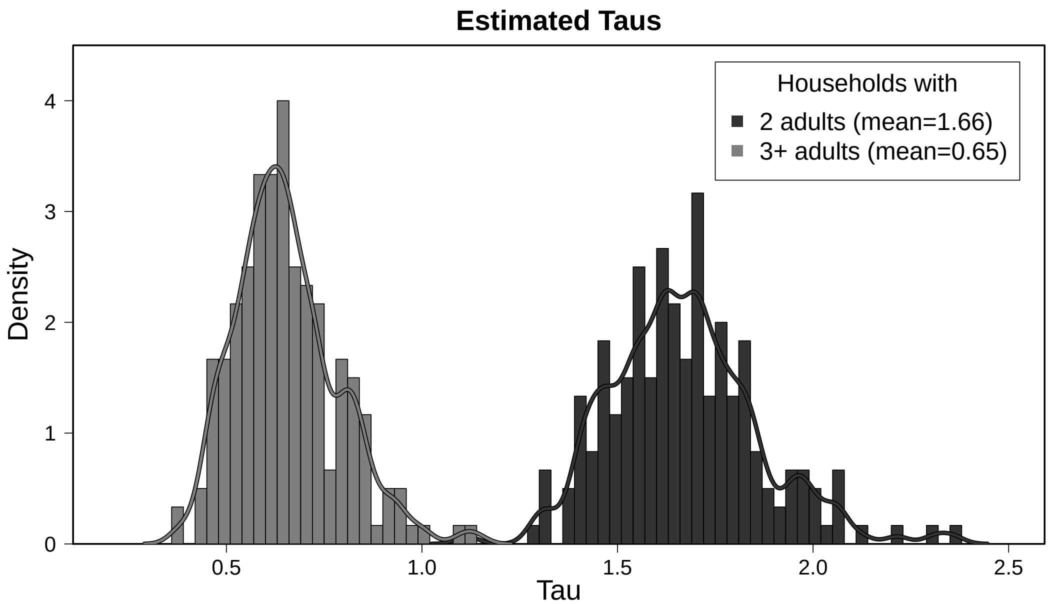


Figure 6: Posterior Estimates of τ_0 and τ_1

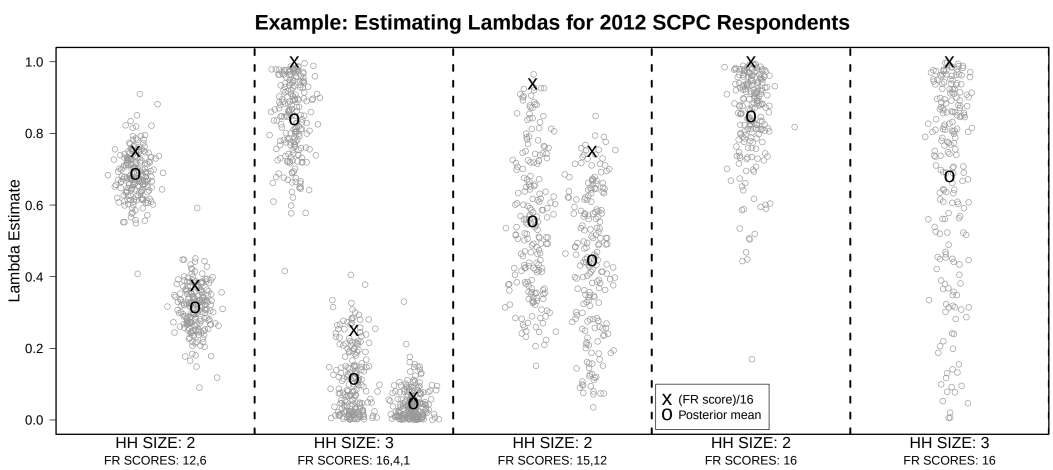


Figure 7: Posterior Estimates of λ_h^o for Five Households in the 2012 SCPC

relatively near the implied scores, $\frac{F_{hi}^o}{16}$. The third panel shows a fully-sampled two-adult household with clearly inconsistent scores, in this case $F_h^o = \{15, 12\}$. The lack of internal consistency results in much more uncertainty about the true values of the FRQs than for the first two-adult household. Nevertheless, the relationship between the members is preserved, as the individual with a higher score also has a higher posterior mean. Finally, the last two panels feature partially-sampled households, in which only one representative took the survey. In both cases, the individual responded with a score of $F_{hi}^o = 16$. However, in the case of the three-person household depicted in panel 5, the overall evidence of greater reporting error results in more variation in the final value and a lower posterior mean than in the two-adult household shown in panel 4. As might be desired, in cases where all household members were featured in the survey, the posterior mean is essentially proportional to the individual's score over the total household score, $\frac{F^o_{hi}}{\sum_i F^o_{hi}}$.

4 Selection Bias in the SCPC

The analysis in Section 3 yields estimates for the distribution of household FRQs in the population, λ_h, as well as estimates for the particular FRQs of respondents in the 2012 SCPC, λ^o_{hi}. Unbiased selection from within households should yield sampled FRQs that are consistent with the population distribution. Thus, comparing the average FRQs in the 2012 SCPC sample to the expected distribution of such a mean under a random sampling of household members is akin to testing for the presence of a selection bias by using the sample average as the test statistic. We do so for households of size $n = 2, 3, 4$, which for each there are $m_n = \sum_h 1\left[n_h = n\right]$ individuals in the 2012 SCPC. The values of m_n are shown in Table 1. Focusing on the first recruits, we define

$$\bar{\lambda}^o(n) \quad = \quad \frac{1}{m_n} \sum_h \lambda^o_{hi} 1\left[n_h = n\right]$$

as the average FRQ of all first recruits from households of size n. The 200 posterior draws of each λ^o_{h1} can be used to generate a posterior distribution for $\bar{\lambda}^o(n)$.

The expected distribution under random sampling from within households can be calculated based on our modeling assumptions and parameter estimates. Specifically, if the unordered FRQs of a household of size n follow a $\mathrm{Dirichlet}(\tau_n)$ distribution, then the FRQ of a randomly selected household member follows a $\mathrm{Beta}(\tau_n, (n-1)\tau_n)$. Within this context, $\bar{\lambda}^o(n)$ can be be viewed as one outcome from the random variable

$$\bar{\lambda}^*(n) \quad = \quad \frac{1}{m_n} \sum_{k=1}^{m_n} Z_k(n),$$

where $Z_k(n)$ are independent draws from $\mathrm{Beta}(\tau_n, (n-1)\tau_n)$. Figure 8 shows histograms for 200 posterior draws of $\bar{\lambda}^o(n)$ and $\bar{\lambda}^*(n)$, the latter being simulated using posterior mean estimates, $\hat{\tau}_n$. For all three household sizes the sample average FRQ in the SCPC is significantly greater than the expected value of $\frac{1}{n}$, even when sampling variation is accounted for.

Despite the fact that our model does not assume intra-household selection bias a priori, the posterior results suggest that the SCPC is oversampling household members with higher financial responsibility.

In other words, household members who have more financial responsibility are more likely to appear in the panel. Such a phenomenon could be a result of panelist characteristics, if those more involved in household matters are also more likely to agree to being surveyed, or be unique to the recruitment methods used, say if mailed panel invitations are more likely to be filtered through a head of household. In Section 6 we consider whether over-representation of high FRQs in the 2012 SCPC relates to an over-representation of certain demographic variables, but otherwise leave the study of what type of household members report higher shares of financial responsibility for future work.

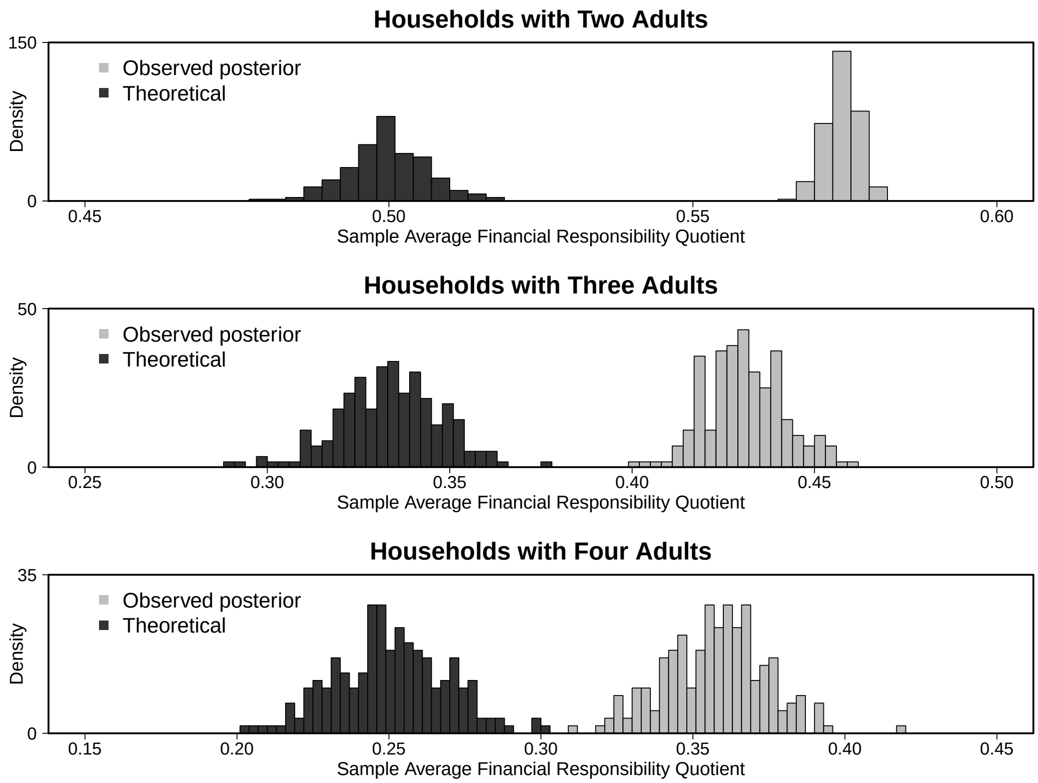


Figure 8: Histogram of 200 Posterior Draws of $\bar{\lambda}^o(n)$ and $\bar{\lambda}^*(n)$

5 Relating Financial Responsibility Quotients to Household Payments

The presence of a selection bias is interesting in its own right, but it becomes even more so if the bias affects a variable of interest. In this section, we motivate and fit a model to determine the extent to which the estimated FRQs relate to the reported share of household payments. It is possible that they do not, either because the variable λ_{hi} does not correlate with payments behavior or because our methodology provides poor estimates of λ_{hi}. Again, the analysis relies on having more than one respondent from a subset of households.

Consider Y_{hi} to be the number of payments made in a "typical" month by individual i of household h, as measured in the SCPC. Below, the analysis is done for four types of payment instruments, but for the purposes of presenting the model, we let Y_{hi} correspond to an unspecified payment instrument. We let $T_h = \sum_{i=1}^{n_h} Y_{hi}$ represent the total number of household payments made, or at least the number that would be reported if all members participated in the survey. Additionally, $\gamma_h = \{\gamma_{hi}\}$ represents the share of household payments made by each adult member. Insight into γ_{hi} is revealed in the data through the relationship between Y_{hi} and T_h.

One simple model expresses this relationship by

$$Y_h \mid T_h, \gamma_h \quad \sim \quad \text{Multinomial}(T_h, \gamma_h). \tag{9}$$

If the FRQ of an individual relates to the share of household payments, then λ_{hi} informs γ_{hi} and vice versa. To test this, we use Dirichlet regression (Campbell and Mosimann 1987; Hijazi and Jernigan 2009), which defines the distribution of one vector of shares with respect to a second vector of shares:

$$\gamma_h \quad \sim \quad \text{Dirichlet}(\{a\lambda_h + b\}), \tag{10}$$

where $\{a\lambda_h + b\}$ is the vector of the Dirichlet concentration parameters. Because the Dirichlet concentration parameters must be positive and $\lambda_{hi} \in [0, 1]$, we impose the restrictions that $b > 0$ and $a + b > 0$. As parameterized in (10), a and b define the strength of the relationship between financial responsibility and the share of payments made. Large values of a, relative to b, suggest a strong correspondance between λ_{hi} and γ_{hi}, and increasing a implies greater predictive power between the two variables. Alternatively if $a = 0$, the FRQ does not relate to the share of payments at all.

In the case where data for only a subset of household members is observed, we define $T_h^o = \sum_{i=1}^{s_h} Y_{hi}^o$ as the total number of payments reported by the sampled members of household h. Then, the assumed multinomial model in (9) further implies that

$$Y_h^o \mid T_h^o, \gamma_h^o \quad \sim \quad \text{Multinomial}(T_h^o, {\gamma'}_h^o), \tag{11}$$

where ${\gamma'}_{hi}^o = \frac{\gamma_{hi}^o}{\sum_{i=1}^{s_h} \gamma_{hi}^o}$. By the aggregation properties of the Dirichlet distribution,

$${\gamma'}_h^o \quad \sim \quad \text{Dirichlet}(\{a\lambda_{hi}^o + b\}). \tag{12}$$

Combining (11) and (12), the likelihood function for (a, b) for household h can be expressed as

$$\mathrm{L}_h(a,b) \quad = \quad \int \mathrm{P}(Y_h^o \mid T_h^o, {\gamma'}_h^o)\mathrm{P}({\gamma'}_h^o \mid a, b, \lambda_h^o, Y_h^o) d{\gamma'}_h^o, \tag{13}$$

and the integration of (13) leads to the closed-form likelihood,

$$\mathrm{L}_h(a,b) \quad = \quad \frac{\Gamma(a\sum_{i=1}^{s_h}\lambda_{hi}^o + s_h b)\prod_{i=1}^{s_h}\Gamma(Y_{hi} + a\lambda_{hi}^o + b)}{\Gamma(a\sum_{i=1}^{s_h}\lambda_{hi}^o + s_h b + T_h^o)\prod_{i=1}^{s_h}\Gamma(a\lambda_{hi}^o + b)}. \tag{14}$$

If only one household member is sampled ($s_h = 1$), the likelihood in (14) reduces to 1 for all a and b. Yet these cases can be used to estimate the parameters if prior information about T_h is incorporated. Knowledge about T_h combined with Y_h^o provides information about the number of payments made by unsampled household members. Similarly, λ_{hi}^o and τ_{c_h} define the likelihood of λ_{hi} for the unobserved household members. The conditional likelihood in (14) can be integrated with respect to likely outcomes for the unobserved household members to estimate a and b. However, other than a general sense of an upper bound, we know little about the distribution of T_h. A relatively uninformative prior seems unlikely to aid much in our estimation. This fact, combined with the relative ease of estimation if prior information on T_h is ignored, justifies our decision to use only data from cases in which at least two household members were sampled.

As before, we fit the model separately for households with two adults and for households with three or more adults. Assuming independence across households, the likelihood function of a and b in each case take the respective forms:

$$\mathrm{L}_2(a,b) = \prod_{\{h \mid n_h = 2, s_h > 1\}} \mathrm{L}_h(a,b) \quad \text{and} \quad \mathrm{L}_{3+}(a,b) = \prod_{\{h \mid n_h > 2, s_h > 1\}} \mathrm{L}_h(a,b). \tag{15}$$

Therefore, for a given set of values of λ_h, it is fairly straightforward to use numerical techniques to perform the linearly constrained optimization necessary to generate estimates of the parameters (a, b).

5.1 Results

We fit the model described above to the SCPC data for four groups of payment instruments: checks, cash, electronic payments, and card payments. The electronic payments category includes online banking and bank account number payments, while card payments combines credit, debit, and prepaid card payments. Rather than treating the draws of λ_{hi}^o from each iteration of the MCMC together, we order the posterior draws for each household according to the distances from the posterior means, thus arranging the posterior draws according to their likelihood. For individual i, we let $\bar{\lambda}^o_{hi}$ be the estimated posterior mean determined by averaging the 200 posterior draws. Then, for each household, we order the posterior draws, $\tilde{\lambda}^o_h[1], \ldots, \tilde{\lambda}^o_h[200]$ such that $k \leq k'$ implies that $\sum_{i=1}^{s_h} \left| \tilde{\lambda}^o_{hi}[k] - \bar{\lambda}^o_{hi} \right| \leq \sum_{i=1}^{s_h} \left| \tilde{\lambda}^o_{hi}[k'] - \bar{\lambda}^o_{hi} \right|$. We group household draws according to ordering rank, so that $\tilde{\lambda}^o[k] = \{\tilde{\lambda}^o_{hi}[k]\}$ is the set of posterior draws that are k^{th} closest to the posterior means. This provides a meaningful differentiation between the 200 constructed samples. For both classes of households, those with two adult members and those with three or more, we estimate the parameters (a, b) in (15) based on each of the 200 ordered samples as well as on the set of posterior means $\bar{\lambda}^o$.

One measure of the relationship between the FRQ and the share of payments made is the expected proportion of variance in γ_{hi} explained by λ_{hi}:

$$\mathrm{VE} \quad = \quad 1 - \frac{\mathrm{E}[\mathrm{Var}[\gamma_{hi} \mid \lambda_{hi}]]}{\mathrm{Var}[\gamma_{hi}]}. \tag{16}$$

Akin in spirit to R-squared estimates, $\mathrm{VE} \in [0, 1]$, with higher values indicating a stronger relationship between λ_{hi} and γ_{hi}. On one end of the extreme, if knowledge of λ_{hi} predicts γ_{hi} with absolute certainty, then $\mathrm{Var}[\gamma_{hi} \mid \lambda_{hi}] = 0$, resulting in $\mathrm{VE} = 1$. On the other end, if λ_{hi} does not relate to γ_{hi} at all, then $\mathrm{Var}[\gamma_{hi} \mid \lambda_{hi}] = \mathrm{Var}[\gamma_{hi}]$ and $\mathrm{VE} = 0$.

The value of VE depends on a, b as well as the distribution of λ_{hi}, which itself depends on τ_{c_h} and n_h. The exact representation of VE is provided in Appendix B. Figure 9 shows the explained variance for households with two, three, and four adult members using the posterior means of τ_{c_h} and the 200 ordered, posterior draws of $\lambda^o[k]$. Figure 9 reveals that the link between the share of payments and the FRQ gets stronger as the posterior estimates get closer to the posterior means. This fact can be interpreted as validating that this is a sensible methodology for estimating λ^o_{hi} from F^o_{hi}.

There is a marked difference in the strength of the relationship between λ_{hi} and γ_{hi} in two-adult households and this relationship in households with three or four adults. Two-adult households show a weaker relationship for all four payment instruments. This finding may be partly due to the higher concentration of λ_{hi} around $\frac{1}{n_h}$ ($\hat{\tau}_1 = 1.66$ vs. $\hat{\tau}_2 = 0.66$) and the fact that a more equal distribution of responsibility makes it more difficult to distinguish between differences in responsibility and noise. This phenomenon could be attributed to the fact that in larger households, a clear delegation of tasks is more important to efficient household management. In two-adult partnerships, easier intra-household communication and more fluid dynamics may make it easier to share tasks.

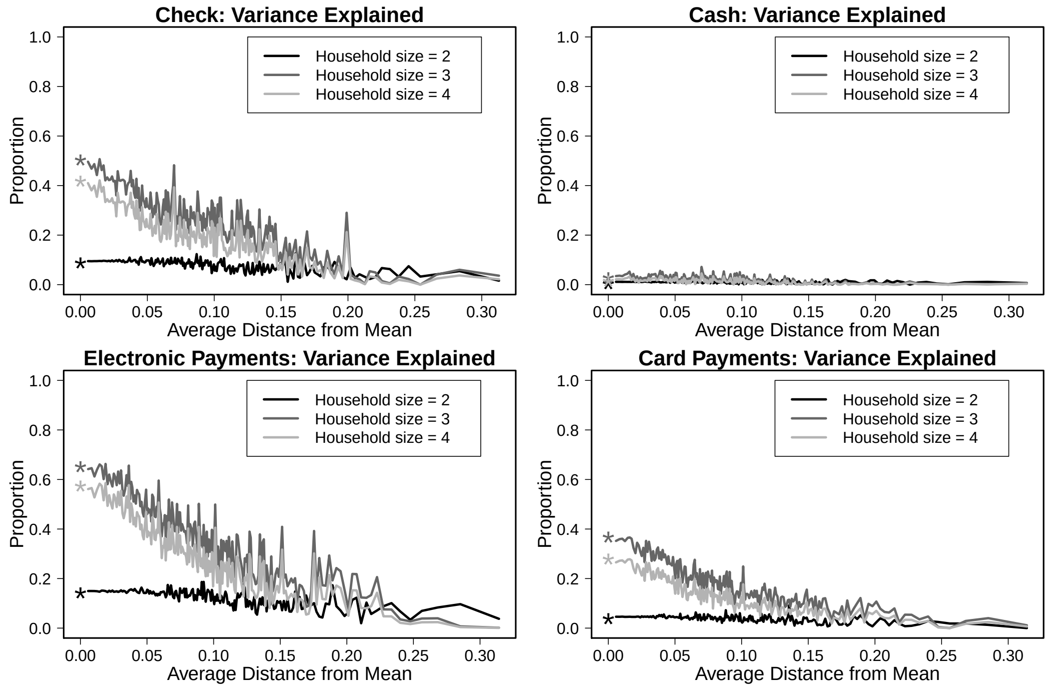


Figure 9: Estimated Proportion of Variance in F_{hi} Explained by Knowing λ_{hi} for Each of the Four Payment Instrument Groups

Note: The $*$ represent calculations using posterior means of the λ^o_{hi}, $\bar{\lambda}_{hi}$.

In general, the results show that the household FRQ is most closely related to the share of payments in the case of checks and electronic payments. The model in (11) and (12) can also be fit with the parameter $a = 0$, corresponding to the case where λ_{hi} does not help predict γ_{hi}. As the models are nested, a likelihood ratio test with one degree of freedom measures the improved quality of fit when considering the FRQ. For households of size $n = 2$ and $n = 3$, all p-values comparing the full and restricted model are found to be less than 0.01, except for cash payments, for which p-values are 0.21 and 0.11 respectively. Indeed, for cash, the proportion of variance explained does not rise above 0.05. For checks and electronic payments, the proportion of variance explained for the best fits range from 0.12 to 0.18 for two-adult households to 0.40 to 0.65 for larger households. Even for credit cards, the proportion of variance explained gets as high as around 0.35 for larger households.

A simple explanation for these findings is that check and electronic payments are most commonly used for bills and large-ticket items, which are more likely to be household purchases like appliances, cars, or household services. It seems natural for these types of payments to be made by household members with higher levels of responsibility and who are more involved in household decisionmaking. Cash, on the other hand, is associated with daily, small-value purchases, which are less related to household needs and more to the immediate needs of individual household members. As a result, having greater financial responsibility within a household does not necessarily translate to the same individual making more cash purchases. Somewhat intuitively, the frequency of using credit and debit cards falls somewhere between the use of cash, check and electronic payments.

6 Effect of Selection Bias on Population Estimates

If not appropriately accounted for, a sampling methodology that favors respondents with certain behavioral tendencies is likely to produce biased population estimates. In this section, we assess how population mean estimates for the number of monthly payments made by U.S. consumers is affected by likely selection bias. This is done by comparing an estimate based on the observed SCPC data to estimates based on simulated samples that preserve, in expectation, certain distributional aspects of the population's FRQs.

The estimate of the mean number of monthly payments among U.S. consumers is generated via poststratification, with the strata means estimated by sample averages for respondents in each stratum. Therefore, if the population is divided into $s = 1, \ldots, S$ disjoint strata, with w_s representing the proportion of U.S. adult consumers belonging to stratum s, then the estimate based on a sample, $\{Y_{hi}\}$, takes the form

$$\hat{\mu} \quad = \quad \sum_{s=1}^{S} w_s \frac{\sum_{h,i} Y_{hi} 1\,[\text{ individual i in household h is in stratum s}]}{\sum_{h,i} 1\,[\text{ individual i in household h is in stratum s}]}. \tag{17}$$

The choice of demographic variables on which to stratify is important. A skew in sample FRQs could be a result of a skew in a demographic variable that can be easily adjusted for via poststratification, thus negating the effect of the selection bias on appropriately weighted estimates. To do so, candidate variables must vary sufficiently within households and have a nontrivial disparity between the sample and population distribution. In the case of the 2012 SCPC, in which 58 percent of respondents are female, gender stands out as an important stratification variable. In addition, we consider household size, household income, and age to generate two possible stratifications: one with 75 strata and one with 31 strata. The strata are constructed to ensure that the minimum number of SCPC respondents in each stratum is 10.

Again, to maximize comparability to common sampling procedures, the estimator based on the SCPC, $\hat{\mu}_0$, is limited to the set of first-recruits, $\{Y^o_{h1}\}$. Simulated samples are conditioned on m_n, the number of respondents from households of size n, so that the observed differences reflect differences in who is selected from within households as opposed to which types of households are selected. To generate these samples, we proceed as follows. For each household size n, we partition the interval $[0,1]$ into disjoint intervals $\{b_{nj}\}_j$ so that the number of individuals in the SCPC sample with estimated posterior mean FRQ in each interval is at least 75. Households of size $n = 1$ represent a trivial case as $\lambda_{hi} = 1$ for all individuals. Table 3 shows the chosen partitions for households of size $n = 2, 3$, and 4.

HH Size	b_{nj} / p_{nj}	$j = 1$	$j = 2$	$j = 3$	$j = 4$
2 Adults		[0,0.25) 0.18	[0.25,0.5) 0.32	[0.5,0.75) 0.32	[0.75,1] 0.18
3 Adults		[0,0.2) 0.41	[0.2,0.5) 0.31	[0.5,1] 0.28	
4 Adults		[0,0.2) 0.53	[0.2,0.4) 0.23	[0.4,1] 0.24	

Table 3: Population Proportions of FRQs in Different Intervals by Household Size

We use posterior estimates of τ_c to determine $p_{nj} = \text{Prob}(\lambda_{hi} \in b_{nj} \mid n_h = n)$, the probability of a randomly selected individual from households of size n having a FRQ in the interval b_{nj}. These probabilities are also shown in Table 3. Given the vector of probabilities, $p_n = [p_{nj}]$, a simple random sample of m_n individuals from households of size n will result in e_{nj} respondents with a FRQ in b_{nj}. Then, $e_n = [e_{nj}]$ has a distribution of

$$e_n \sim \text{Multinomial}(m_n, p_n).$$

For each sample, we first draw e_n for each n and then, for each j, bootstrap e_{nj} individuals from the relevant subset of the full 2012 SCPC sample. Households of size $n \geq 5$ are bootstrapped without consideration of the FRQ. However, because only 3 percent of consumers belong to such households, this strategy will have virtually no effect on the population estimates. For the k^{th} bootstrapped sample, we use (17) to generate the population estimate, $\hat{\mu}_k^*$. A useful way of comparing the SCPC-based estimate, $\hat{\mu}_0$, to the bootstrapped estimates is through the percent deviation:

$$\Psi_k = 100 \times \frac{\hat{\mu}_k^* - \hat{\mu}_0}{\hat{\mu}_k^*}.$$

Histograms of the Ψ_k are shown in Figure 10 for both stratifications. Overall, the direction and magnitude of the relative errors are consistent with the findings presented in Section 5. As expected, the simulated estimates, which do not oversample high FRQs, are generally lower than those based on the observed sample for all payment instruments except cash. Electronic payment shows the most evidence of overestimation, with estimated deviations close to 12 percent. Checks and card payments show errors of about 8 and 7 percent respectively, though the range of estimated errors extends to larger values for checks. At least in the case of checks, the selection bias might partly explain a notable discrepancy between the SCPC's estimate of 6.0 and the Federal Reserve Payment Study estimate of 4.2. For cash, the estimated percent deviations average to about 1 percent, suggesting that the selection bias observed in the 2012 SCPC sample does not affect estimates. Naturally, the bias's effect on narrow categories of payments more closely related to household payments, such as bills or grocery shopping, is likely to be more dramatic.

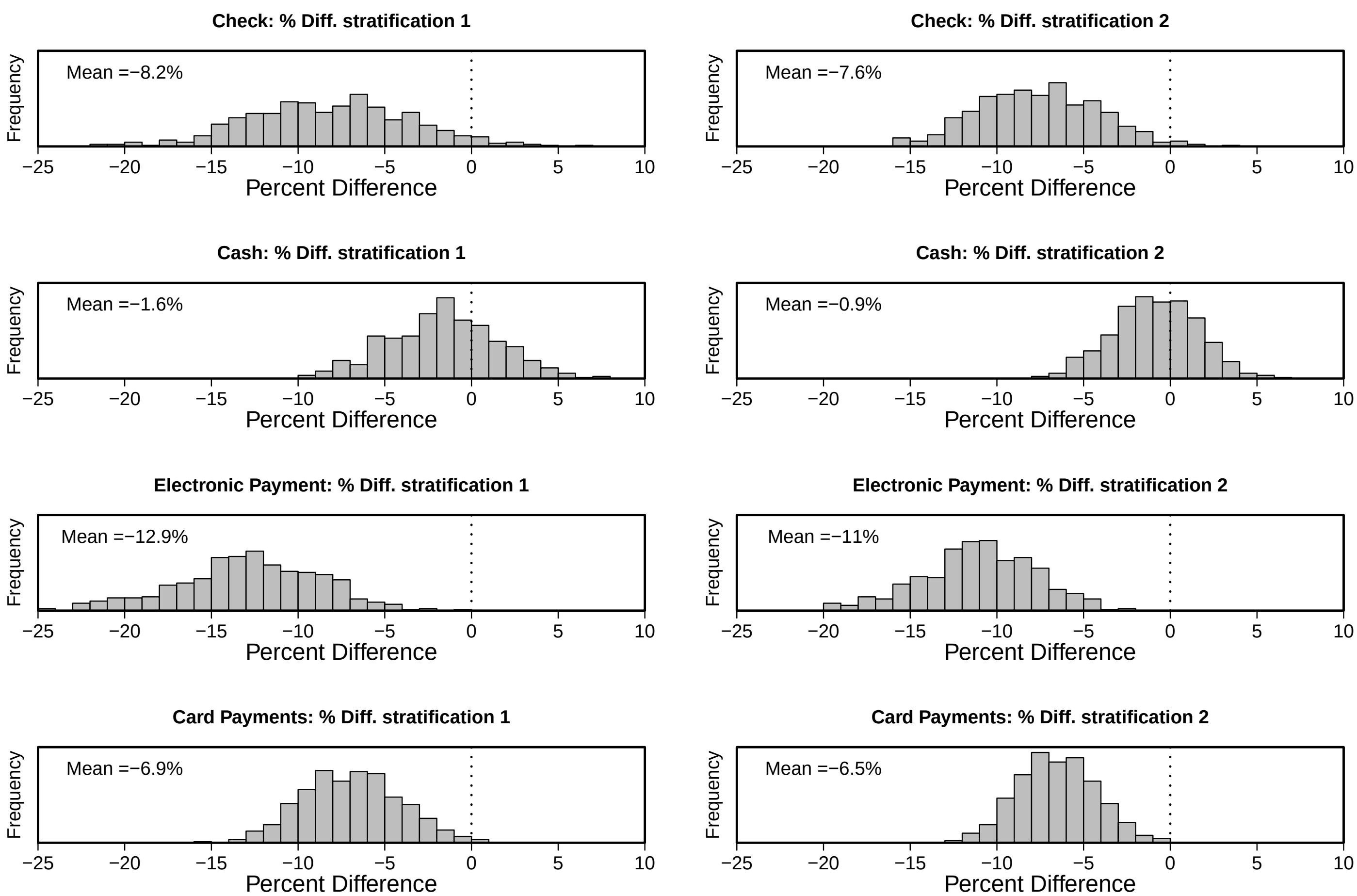


Figure 10: Estimated Percent Deviation for Each of the Four Payment Instrument Groups under two Different Post-stratifications

The use of the bootstrap means we avoid the complicated task of modeling conditional distributions of Y_{hi}, λ_{hi} alongside demographic variables, though with a potential loss of accuracy. A more precise estimate for the size of the selection effect could also be gained by resampling with respect to a household responsibility measure that more closely correlates to the reported number of payments. In the context of our data, one such option is a measure that assigns nonuniform weights to the four financial responsibility questions. However, as detailed in Appendix C, we find little gain in the ability to predict the share of household payments made when using nonuniform weights.

The fact that poststratification of the original SCPC sample on gender does not eliminate the discrepancy due to selection deficiencies broadly suggests that gender is not a strong determinant of an individual's FRQ. However, this does not mean that gender, especially when considered alongside other demographic variables, is completely unrelated to household financial responsibility. In particular, response distributions could differ across the four financial responsibility questions.

7 Discussion

While this paper focuses on the survey methodological implications of any selection bias, there are several interesting findings that relate to the cognitive process involved in survey responses and to household dynamics. First, our data suggest that individuals in our sample consistently overestimate their contributions to household financial decisions, as might be predicted by the well-studied "overconfidence effect" (Svenson 1981). The modeling also found that two-adult households tend to divide responsibilities more evenly than larger households and to provide responses about household responsibility that are more internally consistent. As expected, an adult's level of financial responsibility in a household generally relates to the share of payments he or she makes, although with varying degrees for different payment instruments. Checks, electronic payments, and card payments show a stronger relationship with FRQ than cash, a result that is consistent with the notion that the first three payment types are more likely to be used for less frequently made purchases made by the heads of household.

In our view, the most meaningful contribution of this work is the recognition of a potential selection bias that is generally ignored, at least in the field of consumer surveys. As discussed in Section 2, the degree to which the observed selection bias extends to other samples is difficult to determine, and the need to address the issue depends on how closely the variable of interest relates to the survey respondent's household role. Overall, we believe researchers who rely on surveys should be generally aware of intra-household variability in behavior and the possibility of introducing bias by over- or under-sampling individuals with certain types of roles.

The methodology used in this paper to identify and estimate the effect of the selection bias involves two stages of modeling and relies on a unique data structure in which a nontrivial fraction of households feature several respondents. As such, it is impractical to implement in most surveys. Improving the process depends on simultaneously devising survey variables that provide better insight into household roles and better understanding of how those variables relate to behavior. Different measures of interest are likely to require different survey questions.

A second consideration for researchers studying variables that are known to relate to a respondent's household role is to move away from individual-based questions and toward household-based ones. Interviewing the head of the household to determine economic variables for the entire household is an approach already used by the Consumer Expenditure Survey or the Survey of Consumer Finance. Estimates based on household data can be easily converted to those on a per-consumer basis as long as the average number of consumers per household is known. Of course, the quality of such estimates depends on the accuracy of the reported household data. While it seems reasonable for the head of a household to be able to generate estimates for the number of bills paid, it seems harder to tally every member's cash expenditures. Further research might reveal that question scope should depend on the particular variable and how closely it relates to household role.

References

Baker, Reg, J. Michael Brick, Nancy A. Bates, Mike Battaglia, Mick P. Couper, Jill A. Dever, Krista J. Gile, and Roger Tourangeau. 2013. "Summary Report of the AAPOR Task Force on Non-Probability Sampling." *Journal of Survey Statistics and Methodology* 1(2): 90–105.

Becker, Gary. 1991. *A Treatise on the Family*. Cambridge, MA: Harvard University Press.

Campbell, G., and J. Mosimann. 1987. "Multivariate Methods for Proportional Shape." *ASA Proceedings of the Section on Statistical Graphics* 10–17.

Carman, Katherine Grace, and Michael Pollard. 2014. "Methodology of the RAND Midterm 2014 Election Panel." Available at `http://www.rand.org/pubs/research_reports/RR854.html`.

"Consumer Expenditure Survey". Various Years. Washington, DC: U.S. Bureau of Labor Statistics. Available at `http://www.bls.gov/cex/home.htm`.

"Current Population Survey". 2012 Washington, DC: United States Census Bureau. Available at `http://www.census.gov/cps/`.

Curtin, Richard, Stanley Presser, and Eleanor Singer. 2005. "Changes in Telephone Survey Nonresponse Over the Past Quarter Century." *Public Opinion Quarterly* 69(1): 87–98.

Deville, Jean-Claude, and Carl-Erik Sarndal. 1992. "Calibration Estimators in Survey Sampling." *Journal of the American Statistical Association* 87 (418): 376–382.

Groves, Robert M. 2006. "Nonresponse Rates and Nonresponse Bias in Household Surveys." *Public Opinion Quarterly* 70 (5): 646–675.

Hijazi, Rafiq H., and Robert W. Jernigan. 2009. "Modeling Compositional Data Using Dirichlet Regression Models." *Journal of Applied Probability and Statistics* 4 (1): 77–91.

Park, David K., Andrew Gelman, and Joseph Bafumi. 2004. "Bayesian Multilevel Estimation with Post-stratification: State-Level Estimates from National Polls." *Political Analysis* 12(4): 375–385.

"RAND American Life Panel". Various Years. Santa Monica, CA: RAND Corporation. Available at `https://alpdata.rand.org/`.

Schaeffer, Nora Cate, and Stanley Presser. 2003. "The Science of Asking Questions." *Annual Review of Sociology* 29: 65–88.

Schuh, Scott, and Joanna Stavins. 2014. "The 2011 and 2012 Surveys of Consumer Payment Choice." Research Data Report 14-1. Boston: Federal Reserve Bank of Boston. Available at `http://www.bostonfed.org/economic/rdr/2014/rdr1401.pdf`.

"Survey of Consumer Finances". Various Years. Washington, DC: Board of Governors of the Federal Reserve System. Available at `http://www.federalreserve.gov/econresdata/scf/scfindex.htm`.

"Survey of Consumers". Various Years. Survey Research Center. Ann Arbor, MI: University of Michigan. Available at `https://data.sca.isr.umich.edu/reports.php`.

Svenson, Ola. 1981. "Are We All Less Risky and More Skillful Than Our Fellow Drivers?" *Acta Psychologica* 47(2): 143–148.

"The 2013 Federal Reserve Payments Study: Recent and Long-Term Trends in the United States: 2000-2012". 2014 Washington, DC: Federal Reserve System. Available at `https://www.frbservices.org/files/communications/pdf/general/2013_fed_res_paymt_study_detailed_rpt.pdf`.

Tourangeau, Roger, and Thomas J. Plewes, eds. 2013. *Nonresponse in Social Science Surveys: A Research Agenda.* Washington, DC: The National Academies Press.

Valliant, Richard, and Jill A. Dever. 2011. "Estimating Propensity Adjustments for Volunteer Web Surveys." *Sociological Methods and Research* 40(1): 105–137.

Wang, Wei, David Rothschild, Sharad Goel, and Andrew Gelman. 2009. "Forecasting Elections with Non-Representative Polls." *Public Opinion Quarterly* 73(5): 895–916.

A Appendix A

The model in (1) can be derived as an approximation to a model in which some demographic-based information can be used to order household members according to the expected financial responsibility factor. In other words, given information, such as age of the members or incomes, the distributions are no longer exchangeable. For example, middle-aged individuals with higher personal income levels might be expected to hold higher values of λ_{hi} than other individuals in the household.

Specifically, consider households defined by $n_h = n$. One can assume that, based on some household member information, a rank of the members for each of these households can be established according to the expected level of financial responsibility. Thus, we let $\lambda_{r[h]} = \{\lambda_{r[h][1]}, \ldots, \lambda_{r[h][n]}\}$ represent a permutation of λ_h, where $\mathrm{E}[\lambda_{r[h][i]}] > \mathrm{E}[\lambda_{r[h][i']}]$ for all $i \leq i'$. The ordering of household members within each household does not necessarily correspond to ranking of the true values, but rather a ranking of the expected values based on the provided information. We consider

$$\lambda_{r[h]} \sim \mathrm{Dirichlet}(\alpha_1, \alpha_2, \ldots, \alpha_n),$$

so that $\alpha_i = \alpha\, p_i$ and

$$p_1 \geq p_2 \geq \ldots \geq p_n \quad \text{and} \quad \sum_{i=1}^{n} p_i = 1.$$

Therefore, $\mathrm{E}[\lambda_{r[h][i]}] = p_i$, and the variance of the FRQs is determined by the scalar α. Now, we consider λ_{hi} to be randomly selected from $\lambda_{r[h]}$. The moment-generating function for λ_{hi} then takes the form

$$\mathrm{E}\left[\exp t\lambda_{hi}\right] = \frac{1}{n}\sum_{j=1}^{n}\mathrm{E}\left[\exp t\lambda_{r[h][j]}\right],$$

since each member is likely to be chosen with probability $\frac{1}{n}$. Now, given the assumption of the Dirichlet distribution, it is known that $\lambda_{r[h][j]}$ is itself a Beta distribution with parameters α_j and $\sum_{k=1}^{n}\alpha_k - \alpha_j = \alpha - \alpha_j$. The moment-generating function for $\lambda_{r[h][j]}$ is known as

$$\mathrm{E}\left[\exp t\lambda_{r[h][j]}\right] = 1 + \sum_{k=1}^{\infty}\left(\prod_{r=0}^{k-1}\frac{\alpha_j + r}{\alpha + r}\right)\frac{t^k}{k!}.$$

Therefore,

$$\begin{aligned}\mathrm{E}\left[\exp t\lambda_{hi}\right] &= 1 + \sum_{k=1}^{\infty}\sum_{j=1}^{n}\frac{1}{n}\left(\prod_{r=0}^{k-1}\frac{\alpha_j + r}{\alpha + r}\right)\frac{t^k}{k!} \\ &= 1 + \sum_{k=1}^{\infty}\frac{t^k}{k!}\frac{\sum_{j=1}^{n}\frac{1}{n}\prod_{r=0}^{k-1}(\alpha_j + r)}{\prod_{r=0}^{k-1}(\alpha + r)}.\end{aligned}$$

This last form closely resembles the known form of the moment-generating function of a Beta distribution, and in fact the first moment, corresponding to $k = 1$, will be identical $\left(\frac{1}{n}\right)$. As α_j converges to $\frac{\alpha}{n}$, the form of the moment-generating function and thus the distribution of λ_{hi} converges to that of a Beta distribution with parameters $\frac{\alpha}{n}$ and $\frac{(n-1)\alpha}{n}$. If λ_{hi} comes from a Beta distribution, this implies that the random permutation of $\lambda_{r[h]}$ is a Dirichlet distribution with concentration parameter $\frac{\alpha}{n}$ for all elements. The degree to which the two distributions resemble each other depends, as stated, on the relationship of α_j. So, the more p_j are similar to one another (the closer α_j are to one another), the more the two distributions will resemble one another. Similarity of p_j depends on the extent to which known demographic variables predict expected FRQs.

B Appendix B

The form for VE in (16) depends on the first two moments of λ_{hi} through the form:

$$\mathrm{VE}(a, b, \tau_{c_h}) = \frac{\kappa_1\mathrm{Var}[\lambda_{hi}]}{\kappa_1\mathrm{Var}[\lambda_{hi}] + \kappa_2(\mathrm{E}[\lambda_{hi}]^2 + \mathrm{Var}[\lambda_{hi}]) + \kappa_3\mathrm{E}[\lambda_{hi}] + \kappa_4},$$

where

$$\kappa_1 = \frac{a^2}{(a+bn_h)^2} \quad \text{and} \quad \kappa_2 = \frac{-a^2}{(a+bn_h)^2(a+bn_h+1)}$$
$$\kappa_3 = \frac{a^2+abn_h-2ab}{(a+bn_h)^2(a+bn_h+1)} \quad \text{and} \quad \kappa_4 = \frac{ab+(n_h-1)b^2}{(a+bn_h)^2(a+bn_h+1)}.$$

The assumption that $\lambda_h \sim \text{Dirichlet}(n_h, \tau_{c_h})$, implies

$$\text{E}[\lambda_{hi}] = \frac{1}{n_h} \quad \text{and} \quad \text{Var}[\lambda_{hi}] = \frac{n_h - 1}{n_h^2(\tau_{c_h} n_h + 1)}.$$

C Appendix C

Our paper estimates an individual's FRQ by studying unweighted sums of the scores for the financial responsibility questions, F_{hi}, with the assumption that this observable statistic is correlated with the FRQ. This assertion was strong enough to argue that the data sample favored household members with higher shares of responsibility. However, when trying to adjust sample-based estimates, perhaps this is suboptimal; in other words, it is entirely possible that a weighted sum of the scores more closely corresponds to the actual share (and thus number) of monthly payments. Forming a better understanding of this relationship will naturally lead to better adjustments.

To this end, we compared the predictive ability of the estimated λ_{hi} through the model in (10) to one in which the prediction is based on linear combinations of the raw financial responsibility scores, $\alpha^T V_{hi} = \alpha_1 F_{hi1} + \alpha_2 F_{hi2} + \alpha_3 F_{hi3} + \alpha_3 F_{hi4} + \alpha_4$. We define

$$\begin{aligned} V_{hi}^T &= [F_{hi1} \;\; F_{hi2} \;\; F_{hi3} \;\; F_{hi4} \;\; 1] \\ \alpha^T &= [\alpha_1 \;\; \alpha_2 \;\; \alpha_3 \;\; \alpha_3 \;\; \alpha_4]\,. \end{aligned}$$

Then, for a given set of parameters, α, the likelihood of observed shares for a given payment instrument is:

$$\tilde{\text{L}}_h(\alpha) = \frac{\Gamma(\sum_{i=1}^{s_h} \alpha^T V_{hi}) \prod_{i=1}^{s_h} \Gamma(Y_{hi} + \alpha^T V_{hi})}{\Gamma(\sum_{i=1}^{s_h} \alpha^T V_{hi} + T_h^o) \prod_{i=1}^{s_h} \Gamma(\alpha^T V_{hi})}. \tag{18}$$

Again, we separate the households with two adults and those with three or more adults and define:

$$\tilde{\text{L}}_2(\alpha) = \prod_{\{h|n_h=2, s_h>1\}} \tilde{\text{L}}_h(\alpha) \quad \text{and} \quad \tilde{\text{L}}_{3+}(\alpha) = \prod_{\{h|n_h>2, s_h>1\}} \tilde{\text{L}}_h(\alpha). \tag{19}$$

We estimate maximum likelihood estimates for α, $\hat{\alpha}$, and compare maximum log-likelihoods with those in the models based on the FRQs, given in (15):

$$\begin{aligned}\Delta_2 &= \log \tilde{\mathrm{L}}_2(\hat{\alpha}) - \log \mathrm{L}_2(\hat{a}, \hat{b}) \\ \Delta_{3+} &= \log \tilde{\mathrm{L}}_{3+}(\hat{\alpha}) - \log \mathrm{L}_{3+}(\hat{a}, \hat{b}).\end{aligned}$$

The estimated values of α and likelihood differences between the two models are shown in Table 4.

	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\alpha}_4$	Δ_2
Check	0.14	0.07	0.00	0.96	-2.2
Cash	0.00	0.00	0.00	1.04	0.8
Electronic Payments	0.12	0.01	0.03	0.55	-1.5
Card Payments	0.03	0.05	0.03	0.84	-0.3

	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\alpha}_4$	Δ_{3+}
Check	0.04	0.13	0.07	0.14	-1.5
Cash	0.05	0.04	0.00	1.04	0.8
Electronic Payments	0.31	0.08	0.11	0.30	0.6
Card Payments	0.08	0.17	0.09	0.40	-6.0

Table 4: Optimal Weights and Log-Likelihood Differences Between Models Based on Estimated FRQs and Models Based on Financial Responsibility Questions for Two-Adult Households and Larger Households

The numbers in Table 4 suggest that the difference in the fit, as measured by the log-likelihoods, is minimal across the two models, with the lone exception being credit card payments among larger households. In this particular case, it seems that placing more weight on reported household shopping scores improves the model's predictive ability. Otherwise, the optimal weightings have little pattern and are difficult to explain, though this is likely due to the strong correlation between responses to the financial responsibility questions and relatively low sample sizes. Overall, it seems that the FRQ, as calculated, does an adequate job of estimating the effect of the selection bias.