

Sample Bias Related to Household Role

Marcin Hitczenko

Working Paper 2021-9

February 2021

Abstract: This paper develops a two-stage statistical analysis to identify and assess the effect of a sample bias associated with an individual's household role. Survey responses to questions about the respondent's role in household finances and a sampling design in which some households have all members take the survey enable the estimation of distributions for each individual's share of household responsibility. The methodology is applied to the 2017 Survey of Consumer Payment Choice. The distribution of responsibility shares among survey respondents suggests that the sampling procedure favors household members with higher levels of responsibility. A bootstrap analysis reveals that population mean estimates of monthly payment instrument use that do not account for this type of sample misrepresentation are likely biased for instruments often used to make household purchases. For checks and electronic payments, analysis suggests it is likely that unadjusted estimates overstate true values by 10 percent to 20 percent.

JEL classification: C11, C83, D12

Key words: survey error, Bayesian interference, Survey of Consumer Payment Choice, bootstrap, household economics

<https://doi.org/10.29338/wp2021-09>

The views expressed here are those of the authors and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Any remaining errors are the author's responsibility.

Please address questions regarding content to Marcin Hitczenko, Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree Street NE, Atlanta, GA 30309, marcin.hitczenko@atl.frb.org.

Federal Reserve Bank of Atlanta working papers, including revised versions, are available on the Atlanta Fed's website at www.frbatlanta.org. Click "Publications" and then "Working Papers." To receive e-mail notifications about new papers, use frbatlanta.org/forms/subscribe.

1 Introduction

An increasingly important aspect of survey-based inference involves recognizing and adjusting for discrepancies between respondents and the target population. For one, a general trend of decreasing response rates has magnified the influence of differential nonresponse, so that those who participate are inherently different from those who choose not to (Cornesse and Bosnjak 2018; Curtin et al. 2005; Tourangeau and Plewes 2013). In addition, the relative ease and low cost of data collection has led to the rising popularity of nonprobability samples, in which recruitment is based on convenience (Groves 2006). Such samples risk misrepresentation due to selection bias, if those targeted in recruiting are not representative of the intended population. A 2012 Task Force Report by The Journal of Survey Statistics and Methodology stressed the need to develop methodologies to use nonprobability samples for population inference (Baker et al. 2013).

Population mean inference inherently requires assumptions about the nature of sample representation. When sampling weights are available, the Horvitz-Thompson estimator of the population mean is unbiased (Horvitz and Thompson 1952), but only under a perfect response rate. In practice, analysis is often based on an assumption of ignorability, or that certain conditional distributions in the population are preserved in the sample (Gelman et al. 2004; Rubin 1987). The direct implication is that within strata defined by auxiliary information, the observed sample is effectively a random selection from the population and the observation mechanism need not be modeled explicitly. Then, design weights can be generated by calibrating for nonresponse (Deville and Sarndal 1992; Lundstrom and Sarndal 1999) or through response propensity adjustments (Seaman and White 2011; Little 1986; Valliant and Dever 2011). For nonprobability samples especially, poststratification (Little 1993; Gelman and Carlin 2002; Gelman 2007) or raking (Deming and Stephan 1940; Lohr 1999) use population information to generate strata weights, with strata means estimated through the data, often through modeling approaches (Wang et al. 2009; Park et al. 2004). Extensions of these methods that allow for nonignorable nonresponse also rely on the inclusion of auxiliary information (Kott and Chang 2010; Pfeffermann and Sikov 2011; Qin et al. 2002).

Research suggests that auxiliary variables relating to observation propensity and to behavior of

interest work best for reducing estimate bias (Kott and Chang 2010; Gelman 2007; Qin et al. 2002; Little and Vartivarian 2005). Out of practical considerations, sample realignment in general population studies is usually limited to a few demographic variables, such as age, gender, or income. Potential biases across less standard dimensions are overlooked. In this work, we bring attention to sample misrepresentation with respect to individuals' household roles, specifically the degree of responsibility for household finances. To our knowledge, there is no existent literature on this topic. Our analysis is based on a case study of the 2017 Survey of Consumer Payment Choice (SCPC). The SCPC recruits respondents through a nonprobability version of address-based sampling, in which addresses and subsequently individuals living at those addresses are invited to participate (AAPOR 2016). As such, it serves as a natural backdrop for examining observation disparity within households. Understanding the key features and limitations of address-based sampling is particularly important in light of its increased use, fueled by the development of extensive databases and improved coverage relative to random digit dialing (AAPOR 2016). Our analysis relies on a unique aspect of the SCPC sample, namely that some households feature multiple members who completed the survey.

The paper proceeds as follows. Section 2 and 3 introduce the framework of analysis and the SCPC data, respectively. Section 4 uses a Bayesian model to estimate household roles in the SCPC sample and compare their distribution to hypothetical samples that preserve desirable recruitment qualities when sampling at random from within households. Section 5, provides a statistical methodology, based on resampling the observed data, to estimate and adjust for bias introduced by household misrepresentation for poststratified estimates of mean number of payment instrument uses within the general population. A discussion of the findings and future directions for research is provided in Section 6.

2 Framework and Notation

Consider a finite population of individuals clustered within households, which are indexed by the subscript h . Household h has n_h individuals, enumerated by $i = 1, \dots, n_h$, so that the pair (h, i) uniquely identifies each member of the population. Each individual is associated with a

measurement, y_{hi} , which has a population mean μ , the parameter of interest. For the purposes of consumer research, n_h counts only the adults, those who are at least 18 years old.

2.1 Household Dynamics

A key part of our analysis involves quantifying individuals' roles within their households. To do so, we adopt a general construct in which each individual is assigned a share, λ_{hi} , so that $\sum_{i=1}^{n_h} \lambda_{hi} = 1$. As seen in more detail in Section 4, the natural limitation on household sums imposed by the shares serves as a key mechanism in estimating sample and population distributions of λ_{hi} . The particular interest of this work leads to λ_{hi} being defined as an individual's average level of responsibility for four household financial activities featured in a particular SCPC question. However, the adopted framework can be used to define household role with respect to other metrics, such as time use or health outcomes.

2.2 Sampling

Inference about μ is derived by sampling the population. The set of individuals for whom y_{hi} is observed in any particular sampling effort is denoted by $I = \{I_{hi}\}$, where a value of $I_{hi} = 0$ indicates that y_{hi} is unobserved, and $I_{hi} = 1, \dots, n_h$ corresponds to the order within the household with which that individual was recruited, if such a distinction exists. Most recruiting strategies select only one member from any household, reducing I_{hi} to a binary variable, but the SCPC sample requires the extension to sampling of multiple household members. The adopted notation does not distinguish why y_{hi} is unobserved; nonresponse and nonselection are both coded with $I_{hi} = 0$.

The sampling distribution of any estimating statistic depends on the stochastic nature of the particular process used to determine which individuals are observed. Therefore, it can be useful to distinguish samples that arise through different recruitment strategies. In this work, we let I^o denote the sampling strategy used in the SCPC, and \tilde{I} denote a hypothetical strategy in which each member of a randomly selected household has an equal chance of being observed.

2.3 Sample Bias and Population Inference

To facilitate discussion, we adopt a super-population framework and let $P(\cdot)$ denote the distribution of its arguments. We define a sample bias with respect to λ_{hi} as occurring when household responsibility shares for a given household size have a different distribution in the sample than in the population:

$$P(\lambda_{hi} \mid I_{hi} > 0, n_h = n) \neq P(\lambda_{hi} \mid n_h = n). \quad (1)$$

This can occur if $P(I_{hi} \mid \lambda_{hi}, n_h = n)$ varies with λ_{hi} , so that certain types of household members are more likely than others to be sampled or respond.

A sample bias with respect to λ_{hi} can lead to nonignorable sampling (see Gelman et al. (2004); Rubin (1987) for background on ignorable and nonignorable sampling). For the purposes of illustration, we consider the case where no variables other than n_h and y_{hi} are observed. Then, ignorability corresponds to a lack of sample bias with respect to y_{hi} , as defined in (1). Under ignorable sampling, estimates are well-behaved. For example, sample averages for each household size are unbiased, and a weighted average according to population prevalence of each household size is an unbiased estimated for the population mean. The sample distribution can be expressed as

$$\begin{aligned} P(y_{hi} \mid I_{hi} > 0, n_h = n) &= \int P(y_{hi} \mid \lambda_{hi}, I_{hi} > 0, n_h = n)P(\lambda_{hi} \mid I_{hi} > 0, n_h = n)d\lambda_{hi} \\ &= \int P(y_{hi} \mid \lambda_{hi}, n_h = n)P(\lambda_{hi} \mid I_{hi} > 0, n_h = n)d\lambda_{hi}, \end{aligned} \quad (2)$$

where the simplification in the second line assumes only that the relationship between λ_{hi} and y_{hi} is preserved in the sample. The identity in (2) reveals that if a sample bias with respect to λ_{hi} exists, ignorability generally occurs only if $P(y_{hi} \mid \lambda_{hi}, n_h = n) = P(y_{hi} \mid n_h = n)$. Within the schema of missing data, this last condition corresponds to missing at random, since observation mechanisms within each household size do not relate to the behavior of interest (Rubin 1987). If sampling is not ignorable, analysis based on a greater set of demographic variables may compensate for the sample bias, but only if the sample distribution of λ_{hi} conditional on the additional variables well-represents that in the population.

3 Data

The data considered in this paper come from the 2017 Survey of Consumer Payment Choice, a survey fielded annually by researchers in the Federal Reserve. SCPC respondents are individuals aged 18 or older from the Understanding America Survey (UAS Various Years), a panel of individuals recruited by the Center for Economic and Social Research (CESR) to participate in a variety of surveys. The SCPC is a 30-minute survey that collects information on payment instrument adoption and use, with the goal of tracking national trends. Data from every SCPC is publicly available at the SCPC website (SCPC Various Years). In the following sections, we introduce the financial responsibility questions and provide a brief overview of the 2017 SCPC sample.

3.1 Financial Responsibility Questions

Included in the SCPC is a set of questions that characterize each respondent's level of responsibility for four activities related to household finances. Insight into an individual's role within the household is provided by a set of four financial responsibility questions, asking each respondent 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 5-point Likert scale ranging from "no or almost no responsibility" to "all or almost all responsibility." A replica of the survey questions is shown in Figure 1.

3.2 SCPC Sampling

3.2.1 SCPC Recruitment

Although the annual SCPC samples have a longitudinal component, recruitment is done anew every year. In 2017, an invitation for the SCPC survey, along with an offer of a \$20 incentive, was emailed to every member of the "Nationally Representative" UAS panel. This panel was generated through nine waves of address-based sampling, starting in 2014, with the goal of recruiting a set of individuals that well-represent those aged 18 and older living in the United States. The address-based sampling protocol employed by CESR follows best practices

UnderstandingAmericaStudy

Financial Responsibility

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.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Doing regular shopping for the household (groceries, household supplies, pharmacy, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Making decisions about saving and investments (whether to save, how much to save, where to invest, how much to borrow)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Making decisions about other household financial matters (where to bank, what payment methods to use, setting up online bill payments, filing taxes)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figure 1: A replica of the financial responsibility questions from the 2017 SCPC.

established by Dillman et al. (2014) and is detailed on the UAS webpage (UAS Various Years). Briefly, addresses are drawn at random from purchased post office delivery sequence files. If a name is not linked with the address, mail is sent to “Current Resident.” An introductory postcard is followed by a package that includes a cover letter, a \$5 cash reward, and a promise of \$15 for the completion of an “intake” survey. Those who complete this survey and agree to more surveys are asked to provide demographic information about their household, which comes with a \$20 incentive. CESR offers to provide internet-connected tablets for those without internet access. Everyone who completes the household survey is an official panelist, until the individual either opts out or has not responded to any contact for 10 months. If an individual declines an invitation, an offer for other household members to join the panel is extended. Across all recruitment waves, around 11 percent of mailings yield a primary contact in the panel. Once a primary contact is established, CESR makes an effort to recruit additional members from each household. In general, about 20 percent of households featured in the UAS include more than one household member.

3.2.2 2017 SCPC Sample

At the time of recruitment for the 2017 SCPC, the UAS panel consisted of 4,759 individuals. The 2017 SCPC was completed by 3,099 individuals, corresponding to a participation rate of about 65 percent. The 2017 SCPC cohort represents 2,688 unique households. Table 1 shows the breakdown of these households according to household size, n_h , and number of members featured in the 2017 SCPC, s_h^o . While most households are represented by one individual, 349 multi-person households are represented by at least two individuals, and 267 multi-person households had all adult members participate in the survey.

The subset of primary respondents, identified by $I_{hi}^o = 1$, most closely resembles typical samples, in which only one household member is selected. Thus, to better generalize, analysis is often restricted to the primary respondents.

# Adults in Sample	Household Size (# Adults)							
	1	2	3	4	5	6	7	8
1	457	1,385	313	132	34	11	2	1
2		229	38	29	3	4	0	0
3			36	6	2	0	0	0
4				2	0	0	0	0

Table 1: Number of households in the 2017 SCPC by the total number of adults (n_h) and number of sampled adults (s_h^o).

3.3 Raw Data Estimates of λ_{hi}

The financial responsibility questions in the survey suggest a particular measure of household responsibility share, λ_{hi} . We define the sample statistic F_{hi} as the sum of all four responses coded from 0 (“None or almost none”) to 4 (“All or almost all”). This statistic takes on integer values in $[0, 16]$, and $\frac{F_{hi}}{16} \in [0, 1]$ serves as an estimate of λ_{hi} , defined as an individual’s average share of household responsibility across the four featured activities. Although somewhat arbitrary, this definition of λ_{hi} undoubtedly gives some insight into household financial dynamics.

Ideally, $F_{hi} = [16\lambda_{hi}]$, where the brackets denote rounding to the nearest integer. Subject to this rounding, the sum of reported household scores should be 16 if responses represent true

dynamics. Figure 2 shows the reported values of F_{hi} for individuals from the same household. Although there are clear intra-household inconsistencies, the general trend suggests that the reported scores relate to true household dynamics. Among two-adult households, the correlation of reported scores is -0.40 , and the mode total score is 16, with about two-thirds of households yielding a total between 13 and 19.

Discrepancies of the kind observed in Figure 2 are to be expected and have many potential sources. Reasonable differences in interpretations of the four activities, the use of different metrics to define responsibility share, and disagreements on the difference between “almost none” and “some” among members of the same household may skew household scores as well. More traditional sources of response error such as straightlining and misunderstanding the question completely also contribute to response variation (Schaeffer and Presser 2003). Finally, the fact that 64 percent of two-adult households and 88 percent of larger households have total scores greater than 16 suggests the presence of the well-studied “overconfidence effect,” in which respondents tend to overstate their contribution (Svenson 1981).

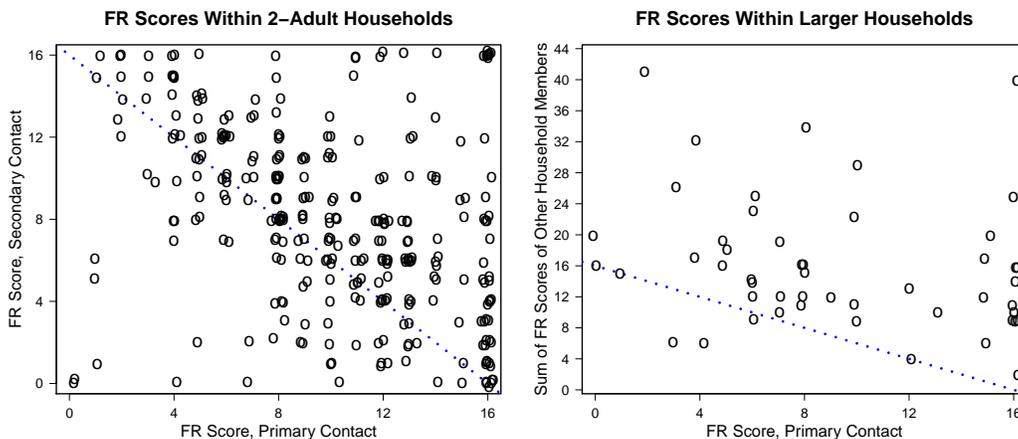


Figure 2: Financial responsibility scores, F_{hi} , in the 2017 SCPC for primary contacts and all other sampled household members.

4 Estimating Sample Bias

In the following section, we detail a Bayesian methodology to estimate the distributions of responsibility shares, as measured by λ_{hi} , in the SCPC sample as well as the general population.

For single-member households, prior assumptions require that $\lambda_{hi} = 1$ with probability 1. However, for multi-member households, a model relating F_{hi} to λ_{hi} , as well as priors for λ_{hi} and any hyper-parameters, are required.

Sections 4.1-4.5 define the model components, Section 4.6 describes estimation, and Section 4.7 discusses posterior results. Finally, in Section 4.8, we use posterior estimates to compare λ_{hi} in the observed sample to what would be expected in hypothetical samples that draw uniformly from within households.

4.1 Modeling λ_{hi}

As a set of shares, a flexible class of models for $\lambda_h = \{\lambda_{h1}, \dots, \lambda_{hn_h}\}$ is,

$$\lambda_h \mid \tau_h \sim \text{Dirichlet}(\{\tau_h, \dots, \tau_h\}), \quad (3)$$

a Dirichlet distribution for a vector of length n_h with the concentration parameter for each component τ_h . Therefore, λ_{hi} are exchangeable and follow a Beta($\tau_h, \tau_h(n_h - 1)$) distribution. Exchangeability follows from the fact that intra-household indices, i , are assigned arbitrarily. The parameter τ_h defines the degree of dispersion of the λ_{hi} within a household. Larger values of τ_h condense the probability mass of λ_{hi} around $E[\lambda_{hi}] = \frac{1}{n_h}$ for each i .

4.2 Modeling Responses

The reported scores, F_{hi} , are modeled as independent, both across households and within households, with

$$P(F_{hi} = f \mid \lambda_{hi}, \phi_{hl}, \phi_{hu}) \propto \begin{cases} \exp(-\phi_{hl} \mid [\lambda_{hi}] - \frac{f}{16} \mid), & \text{if } \lambda_{hi} \geq \frac{f}{16} \\ \exp(-\phi_{hu} \mid [\lambda_{hi}] - \frac{f}{16} \mid), & \text{if } \lambda_{hi} < \frac{f}{16}, \end{cases} \quad (4)$$

where $[\lambda]$ is the value of $\frac{0}{16}, \frac{1}{16}, \dots, \frac{16}{16}$ that minimizes distance to λ_{hi} .

The probability distribution on integers from 0 to 16 in (4) has several attractive features. First, restricting $\phi_{hu}, \phi_{hl} \geq 0$, $F_{hi} = [16\lambda_{hi}]$ is at least as likely as any other response. In fact, if $\phi_{hu}, \phi_{hl} > 0$, $[16\lambda_{hi}]$ will be the mode. Second, by virtue of using two parameters, 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 respondents are more likely to overestimate their share of financial responsibility.

4.3 Model Specification

In this work, we distinguish households with two adults and those with three or more adults. Specifically, we define

$$\tau_h = \begin{cases} \tau_1, & n_h = 2 \\ \tau_2, & n_h \geq 3, \end{cases} \quad \phi_{hl} = \begin{cases} \phi_{1l}, & n_h = 2 \\ \phi_{2l}, & n_h \geq 3, \end{cases} \quad \phi_{hu} = \begin{cases} \phi_{1u}, & n_h = 2 \\ \phi_{2u}, & n_h \geq 3. \end{cases}$$

Such a delineation is partly motivated by economic intuitions. The set of households with two adults 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 responsibility levels within the household. Individuals in smaller households may also be able to better assess their share of contribution. While it would be interesting to make finer distinctions along household size, we are limited by having only 36 fully-sampled three-adult households and 2 fully-sampled four-adult households.

4.4 Model Identifiability

As a matter of notation, we let $\lambda_h^o = \{\lambda_{hi} \mid I_{hi}^o > 0\}$ and $F_h^o = \{F_{hi} \mid I_{hi}^o > 0\}$ represent the household responsibility shares and reported statistics of the observed sample from household h , respectively. The subscript h is dropped to indicate observed data for all households, so the posterior distribution of interest is

$$P(\lambda^o, \tau, \phi \mid F^o) \propto P(F^o \mid \lambda, \phi) P(\lambda^o \mid \tau) P(\tau, \phi). \quad (5)$$

However, certain identifiability issues arise in (5) when not all household members are observed. Then, it is impossible to determine to what extent observed patterns in F_{hi} reflect

patterns in the sampled λ_{hi} , defined by τ and any selection effects, or response error, defined by ϕ . For example, with only one observation per household, the observed tendency for higher values of F_{hi} in the 2017 sample could result from a recruitment methodology that favors those with greater responsibility or systematic over-estimation of contribution by respondents. With fully-sampled households, there is no need to model selection, implicit in $P(\lambda^o | \tau)$ of (5). In addition, the additive constraint on λ_{hi} limits how λ_{hi} relates to F_{hi} , significantly restricting the parameter space.

In this work, we address identifiability issues in two ways. The first is through data augmentation, discussed in van Dyk and Meng (2012), in which a model for F_{hi} of unobserved household members is supplemented into analysis. Specifically, if $F_h^\phi = \{F_{hi} | I_{hi} = 0\}$ represents data from unobserved household members, we model $F_h^\phi | F_h^o$ and make inferences about τ and ϕ conditional on data from the full household. A second option is to estimate τ and ϕ using only the fully-sampled households. Such a restriction is suitable if the set of partially-observed households is a random selection of all households. Data augmentation does not require such an assumption and will improve inference to the extent that differences between fully- and partially-sampled households are reflected in differences in the distributions of the observed F_{hi} for both sets of households. However, if the data augmentation model is misspecified, inference quality can suffer, especially with such a large share of households being partially-observed. Posterior estimates for both approaches are compared in Section 4.6.

4.4.1 Modeling F_h^ϕ

Letting I_h^ϕ represent the ordered set of household indices for those not sampled, the general data augmentation model relies on decomposing

$$P(F_h^\phi | F_h^o) = \prod_{i \in I_h^\phi} P(F_{hi} | F_h^o \cup \{F_{hj} | j < i\}). \quad (6)$$

For each item on the right-hand side of (6), we use ordered probit regression (Agresti 2002; Hoff 2009).

In the case of two-adult households, we let

$$\begin{aligned} P(F_{hi} \leq f_o \mid F_{hj} = f_o) &= \text{Prob}(Z_h \leq g_{k[f_o], f_o}), \\ Z_h &\sim \text{Normal}(\beta_{k[f_o]} \times f_o). \end{aligned}$$

By fitting separate models for groupings of observed F_{hi} values, we have more flexibility to adequately capture dynamics between the primary and secondary observations. The groupings, $k[f]$, are chosen so that observed and simulated distributions of household totals, $F_{h1} + F_{h2}$ matched in key ways, most notably by a high and roughly uniform concentration between 14 and 20. Appendix A provides specifics of $k[f]$.

For larger households, we assume that the distribution of F_{hi} depends on the number of observed household members, n_{hi}^o , the sum of their scores, T_{hi} , and the household size, n_h . The full model is

$$\begin{aligned} P(F_{hi} = f \mid F_h^o \cup \{F_{hj} \mid j < i\}) &= P(Z_{hi} \leq g_f), \\ Z_{hi} &\sim \text{Normal}(\beta_1 n_h + \beta_2 n_{hi} + \beta_3 t_{hi} + \beta_4 t_{hi}^2, 1). \end{aligned}$$

Again, the exact model specification is chosen so that observed and simulated distributions of household totals shared key traits.

4.5 Priors

For τ and ϕ , our primary choice for priors guides the posterior distribution by reflecting a general intuition for plausible values. The assumed priors are

$$\text{Prior A: } \phi_{1\ell}, \phi_{2\ell}, \phi_{1u}, \phi_{2u} \stackrel{\text{iid}}{\sim} \text{Gamma}(24, 9.8) \quad \text{and} \quad \tau_1, \tau_2 \stackrel{\text{iid}}{\sim} \text{Gamma}(2, 1.4), \quad (7)$$

with the Gamma distribution parameterized by its mean and standard deviation, respectively. By having the prior likelihood of ϕ_ℓ and ϕ_u monotonically increasing away from 0, the prior favors models in which there is greater correspondence between λ_{hi} and F_{hi} . The prior mode of 20 in (7), though somewhat arbitrary, represents a value safely greater than one consistent with

the observed data.

The parameter τ defines the heterogeneity in how households share financial responsibility among adult members. The chosen prior effectively precludes extreme cases, in which responsibility is allotted similarly across households, but is relatively unassuming among more realistic choices. Within two-adult households, the central 95 percent of prior mass falls roughly on the range $\tau_1 \in [0.25, 5.5]$. A value of $\tau_1 = 0.25$ assumes that about 25 percent of two-adult households entrust one person with almost all financial responsibility ($\lambda_{hi} > 0.95$) and about 8 percent of households split responsibility more or less evenly ($\lambda_{hi} \in [0.4, 0.6]$). If $\tau_1 = 5.5$, these quantities are less than 1 percent and almost 50 percent respectively. Our intuition, as reflected by the prior, is that τ likely falls between these value, corresponding to greater variability in household patterns.

To test the influence of prior assumptions, we also generate posterior estimates based on

$$\begin{aligned} \text{Prior B: } & \phi_{1\ell}, \phi_{2\ell}, \phi_{1u}, \phi_{2u} \stackrel{\text{iid}}{\sim} \text{Gamma}(24, 9.8) \quad \text{and} \quad \tau_1, \tau_2 \stackrel{\text{iid}}{\sim} \text{Gamma}(6, 2.4) \\ \text{Prior C: } & \phi_{1\ell}, \phi_{2\ell}, \phi_{1u}, \phi_{2u} \stackrel{\text{iid}}{\sim} \text{Gamma}(1, 1) \quad \text{and} \quad \tau_1, \tau_2 \stackrel{\text{iid}}{\sim} \text{Gamma}(2, 1.4). \end{aligned} \quad (8)$$

Prior B in (8) puts greater mass on larger values of τ , which translates to more equal sharing of responsibility within households. Prior C is characterized by favoring values of ϕ closer to 0, thus minimizing the correspondence between the reported F_{hi} and the true responsibility shares, λ_{hi} . The impact of the priors on posterior estimates is discussed in Section 4.6.

Prior distributions for the data augmentation variables, β and g , are uninformative, consistent with defaults used by the R package `bayespl`. Specifically, β are assumed to arise from a generalized t-distribution with 1 degree of freedom and a scale parameter of 4. Priors for the cutpoints, g , are characterized by prior counts at each response value, which we set to $\frac{1}{17}$.

4.6 Model Estimation

For the data augmentation model, we run an MCMC based on the prior distributions in (7).

The MCMC algorithm iterates as follows:

1. Sample $\beta, g \mid F^o \sim P(F^o \mid \beta, g)P(\beta, g)$.
2. For each partially-sample household, sample $F_h^o \mid \beta, g, F_h^o$.

3. Sample $\phi \mid \lambda^o, F^o \sim P(F^o \mid \lambda^o, \phi)P(\phi)$.
4. Sample $\tau \mid \lambda \sim P(\lambda \mid \tau)P(\tau)$.
5. For each h , sample $\lambda_h \mid F_h, \tau, \phi \sim P(F_h \mid \lambda_h, \phi)P(\lambda \mid \tau)$.

Sampling is done using a combination of Gibbs and Metropolis-Hasting sampling. Four distinct chains are run for 16,000 iterations in R, with every 10th draw stored. Figure 3 shows the running averages of τ and ϕ at each stored iteration for each chain.

Using the last 500 draws from each chain, we calculate \hat{R} , a measure of convergence based on the ratio of average within-chain variance to the inter-chain variance of means (Gelman et al. 2004). Values of \hat{R} are shown in Figure 3. For two-adult households, \hat{R} is essentially 1, suggesting almost perfect convergence. For bigger households, \hat{R} is larger, though still suitably close to 1 to accept the sample. We thus combine the last 500 iterations from each of the four runs as our posterior sample.

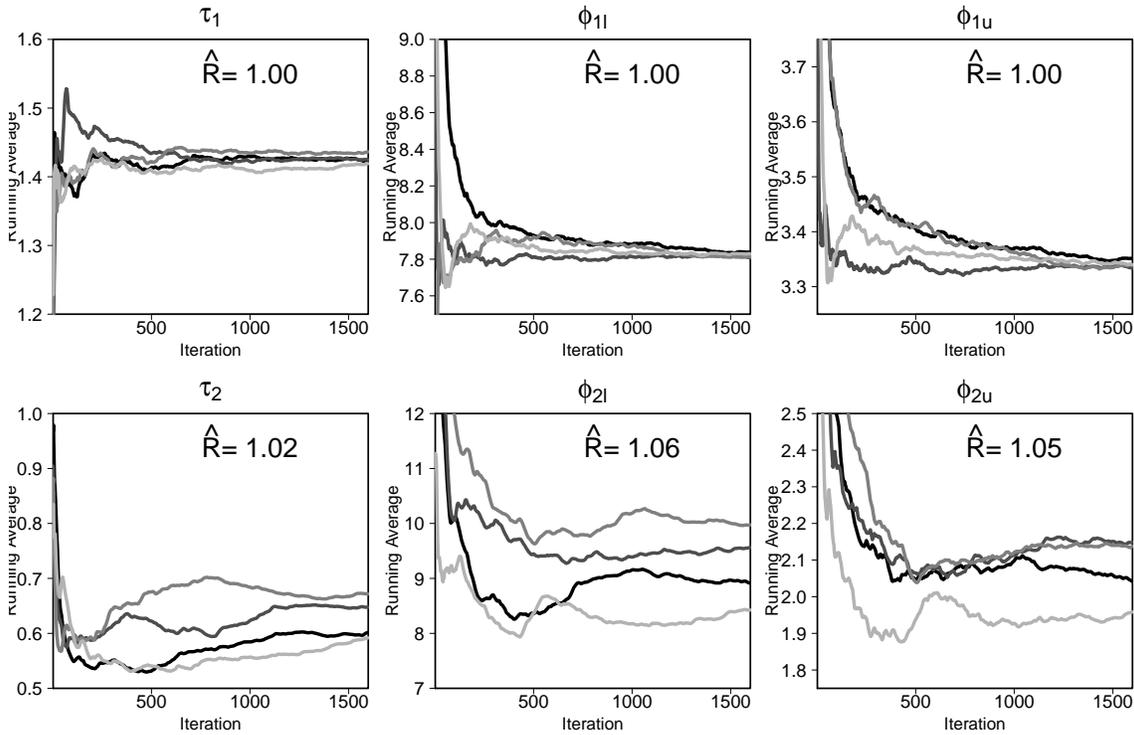


Figure 3: Running average of posterior draws of τ and ϕ from the MCMC algorithm for all four chains. Values of \hat{R} are calculated using the last 500 draws of each chain.

We also estimate posterior distributions of τ and ϕ using only the fully-sampled households. To ease computation, we evaluate probabilities on a grid, taking advantage of the fact that

$\text{Prob}(F_{hi} | \lambda_{hi}, \phi)$ is a step function. Along with the priors in (7), posterior distributions are estimated with the prior combinations given in (8).

Figure 4 shows prior and posterior credible intervals for τ and ϕ based on all four estimation methods. In general, point estimates for τ are similar, though the interval is much shorter when using data augmentation, presumably due to additional information provided by the partially-sampled households. In the case of ϕ , there is greater variability in point estimates. Interestingly, the inclusion of partially-sampled households via data augmentation does not significantly reduce uncertainty about ϕ , as measured by credible interval lengths. The most noticeable difference in estimates of ϕ occurs when a $\text{Gamma}(1, 1)$ distribution is assumed. However, a prior that so strongly favors less correspondence between F_{hi} and λ_{hi} is fairly nonsensical. Perhaps most importantly, the posterior distributions, and especially posterior means, of the observed λ_{hi} are not fundamentally different across models. Overall, the posterior results suggest a fair degree of robustness to reasonable prior choices and estimation methods. All following results are based on the data augmentation approach.

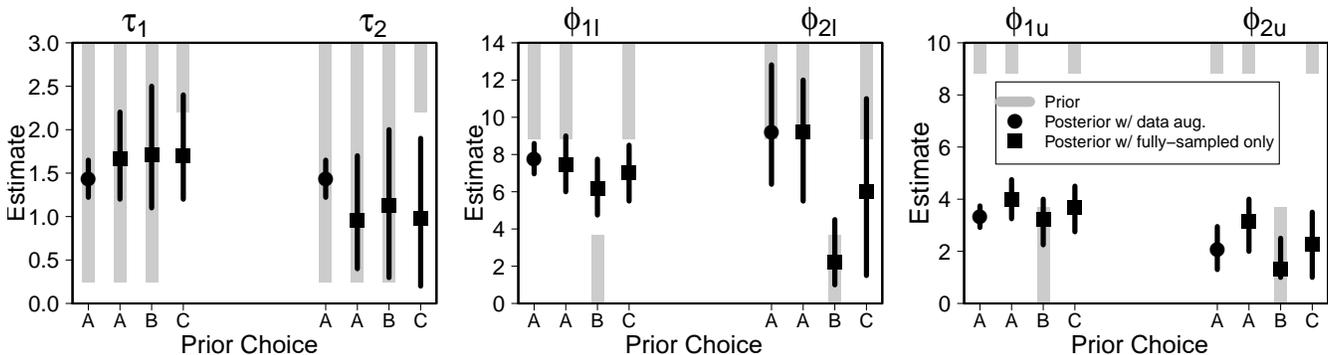


Figure 4: 95% credible intervals for τ , ϕ based on priors and various posterior estimates. Prior ‘A’ corresponds to (7), and Priors ‘B’ and ‘C’ are given in (8), respectively.

4.7 Posterior Distributions

Below, we discuss posterior estimates of τ , ϕ , and λ_{hi} based on the data augmentation model. Posterior distributions for the parameters in the data augmentation model itself, β and g , are summarized in Appendix A.

Posterior intervals of τ , included in Figure 4, show τ_1 around 1.5 and τ_2 around 0.7. Therefore,

two-adult households tend to share financial responsibility more evenly (higher concentration around $\frac{1}{n_h}$) than larger households. This is perhaps due to the fact that larger households include life partners housing adult children or elderly relatives, who might have less say in financial decisions of the household.

The fact that $\phi_\ell > \phi_u$, as Figure 4 makes clear, confirms that individuals are much more likely to overstate their share of household responsibility than understate it. The finding that $\phi_1 > \phi_2$ suggests that the F_{hi} more closely correspond to λ_{hi} in two-adult households than in larger households.

Figure 5, depicting posterior draws of λ_{hi} for five households, illustrates how posterior distributions of λ_{hi} relate to the observed data, $\frac{F_{hi}}{16}$. In cases where all household members are observed, the posterior mean is similar to $\frac{F_{hi}}{\sum_{i=1}^{n_h} F_{hi}}$, a natural estimate, and posterior uncertainty relates to the degree of intra-household consistency (determined by how close the sum of the scores is to 16). The first three panels illustrate this by showing two fully-sampled households in which members' financial responsibility scores show a fair amount of internal consistency and one fully-sampled two-adult household with clearly inconsistent scores, $F_h = \{15, 12\}$. The lack of internal consistency in the third household results in much more uncertainty about the true values of λ_{hi} than for the first two households. Nevertheless, individuals with higher values of F_{hi} have higher posterior means. Finally, the last two panels feature partially-sampled households, in which only one representative completed the survey. In both cases, $F_{hi} = 16$, but evidence of greater reporting error among larger households results in more uncertainty and a lower posterior mean for the individual from the three-person household (panel 5) than for the individual from the two-person household (panel 4).

4.8 Sample Bias in the SCPC

To assess sample bias among the primary respondents of the SCPC, we compare posterior statistics of the observed $\lambda_{hi} \mid I_{hi}^o = 1$ to those for hypothetical samples in which household members are selected without bias. For any given sample I , with $m_n(I)$ primary contacts

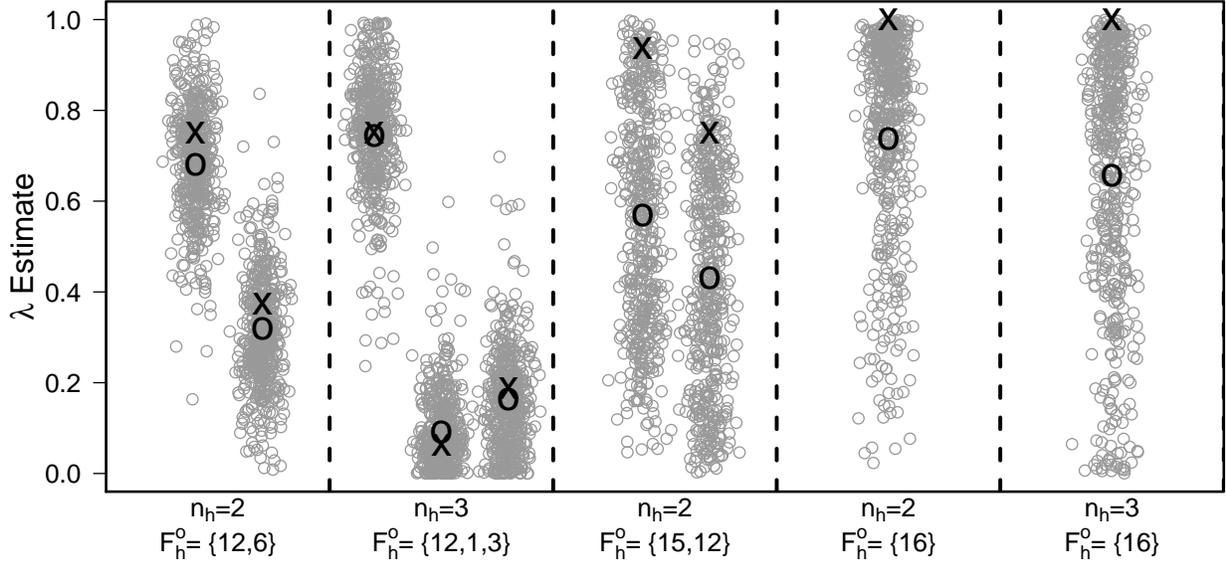


Figure 5: Posterior estimates of λ_h^o for five different households. 'X' represents the value $\frac{F_{hi}^o}{16}$, and 'O' represents the posterior mean under the data augmentation model.

among households of size n ,

$$\bar{\lambda}_n(I) = \frac{1}{m_n(I)} \sum_{\{h,i\}} \lambda_{hi} 1 [I_{hi} = 1 \text{ and } n_h = n].$$

is the average financial responsibility share among observed primary respondents.

Fundamentally, we want to compare $\bar{\lambda}_n^o = \bar{\lambda}_n(I^o) | F^o, I^o$ to the distribution of $\tilde{\lambda}_n = \bar{\lambda}_n(\tilde{I}) | F^o, I^o$, in which \tilde{I} is defined by selecting one member with uniform probability from each of $m_n^o = m_n(I^o)$ households of size n .

Based on the assumed population model introduced in Section 4.1, the sample average of randomly selected household members follows the same distribution as the average of m_n^o independent draws from $\text{Beta}(\tau, (n-1)\tau)$. Thus, a posterior estimates of $\tilde{\lambda}_n$ are easily generated for each posterior draw of τ . Similarly, for $\bar{\lambda}_n^o$, the estimate is easily calculated through posterior draws of $\lambda_{hi} | I_{hi}^o = 1$.

Figure 6 compares 2,000 posterior draws for $\bar{\lambda}_n^o$ and $\tilde{\lambda}_n$ for households of size $n = 2, 3, 4$. As expected, samples based on unbiased draws have expectation $\frac{1}{n}$. For all three household sizes, the sample average λ_{hi} in the SCPC falls above the expected range, even when uncertainty

about the observed λ_{hi} is accounted for. Despite the fact that our model did not assume intra-household selection bias a priori, posterior results suggest that the SCPC is oversampling household members with higher financial responsibility. Similar results are found when theoretical samples are derived by randomly sampling one individual, and thus a λ_{hi} , from each of the households observed in the 2017 SCPC.

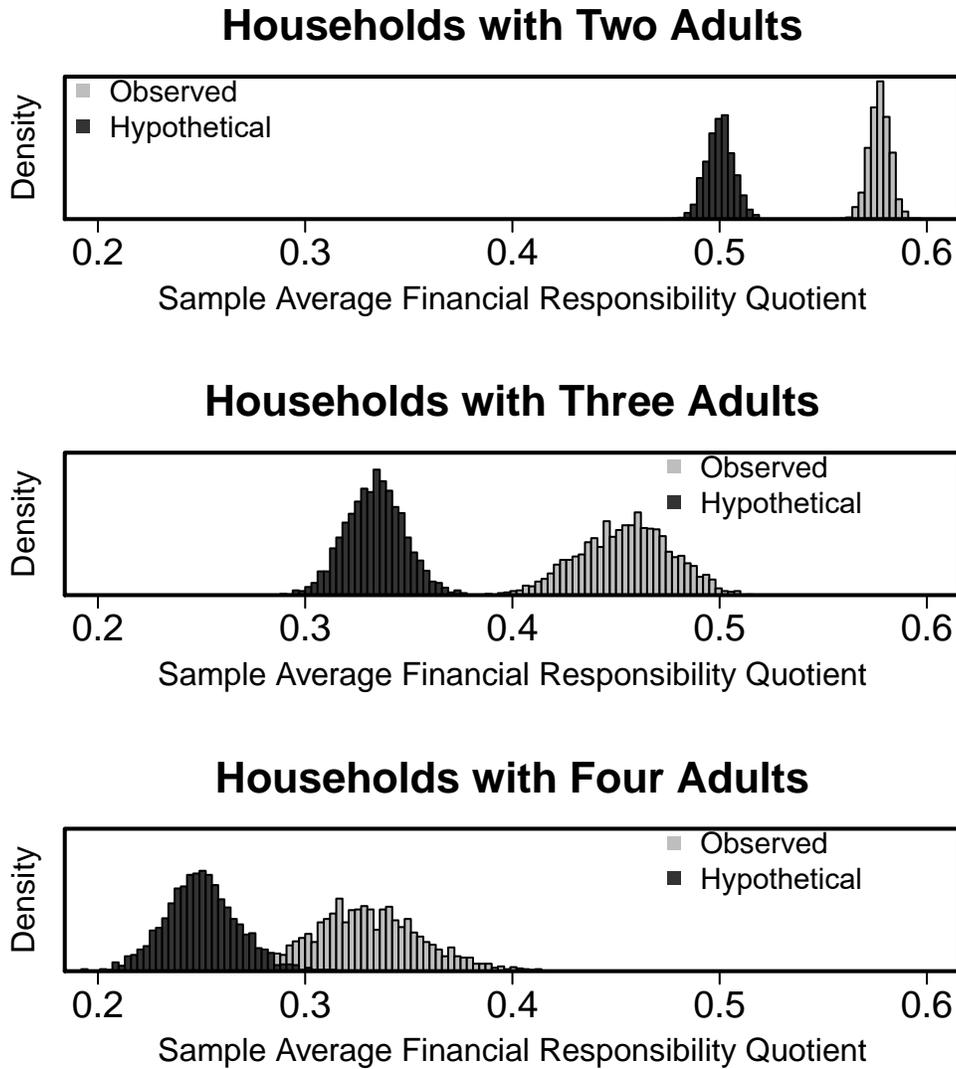


Figure 6: Histograms of 2,000 posterior draws of $\bar{\lambda}_n^o$ (“Observed”), the observed average of λ_{hi} among primary respondents in the 2017 SCPC sample, and $\bar{\lambda}_n^h$ (“Hypothetical”), the expected average based on uniform sampling within households, for household sizes $n = 2, 3,$ and 4 . Exact definitions of these statistics are given in Section 4.8.

While it is impossible to know the extent that this result generalizes to other samples, it does

serve as some evidence that address-based sampling can yield a skewed sample with respect to household role. Perhaps, this is not surprising. Beyond the relatively low response rate (less than 10 percent), there is a legitimate concern that names linked with addresses or those replying to “Current Resident” are not a random sample of household members. Although a probability-based version of address-based sampling, in which, for example, the individual with the next birthday is selected, might fare better, Battaglia et al. (2008) show that surveys are often completed by the wrong person.

5 Effect of Sample Bias on Population Estimates

Below, we estimate the effect of the observed sample bias on population estimates for the mean number of monthly payments. We consider four types 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.

We let y_{hi} represent the number of monthly payments that individual i in household h makes using a generic payment instrument, as reported in the SCPC. We opt against a model-based approach, which requires estimating the joint distribution of demographics, y_{hi} , and λ_{hi} . Instead, we compare a poststratified estimate based on the primary responders in the SCPC data to hypothetical samples that preserve certain distributional aspects of population λ_{hi} . While the focus on primary responders is intended to highlight the impact of the household bias in a sample more representative of typical recruiting efforts, inherently housed in the analysis is a methodology for generating estimates that account for sample bias with respect to λ_{hi} .

5.1 Generating Samples

We again take \tilde{I} to be a recruitment strategy that selects individuals from each household with equal probability. It is assumed that \tilde{I} preserves the number of respondents from households of size n , m_n^o . By conditioning in this way, differences in calculated statistics better reflect discrepancies in who is selected from within households rather than in the types of households.

The desired sampling distribution can be expressed as

$$P(y_{hi} | \tilde{I}_{hi} = 1) \propto \int P(y_{hi} | \lambda_{hi}, \tilde{I}_{hi} = 1)P(\lambda_{hi} | \tilde{I}_{hi} = 1)d\lambda_{hi}, \quad (9)$$

which suggests samples can be generated by first drawing λ_{hi} from its target distribution and then conditionally drawing y_{hi} .

Assuming an infinite population, one way to generate the desired sample for households of size n is to determine the number of sampled individuals with $\lambda_{hi} \in b_{nj}$, m_{nj} , for some disjoint partition of $[0, 1]$, $\{b_{nj}\}_j$. These counts have distribution

$$\{m_{nj}\} \sim \text{Multinomial}(m_n^o, \{p_{nj}\}), \quad (10)$$

where $p_{nj} = \text{Prob}(\lambda_{hi} \in b_{nj} | n_h = n, \tau)$, a function of τ . Given m_{nj} , one can then draw independent samples from $P(y_{hi}, \lambda_{hi} | \lambda_{hi} \in b_{nj}, n_h = n)$. We approximate this step via a bootstrap, sampling with replacement from the observed individuals in the 2017 SCPC sample. For the bootstrap to represent the desired distribution of household roles, it is necessary that

$$P(\lambda_{hi} | \lambda_{hi} \in b_{nj}, n_h = n, I_{hi}^o = 1) = P(\lambda_{hi} | \lambda_{hi} \in b_{nj}, n_h = n, \tilde{I}_{hi} = 1),$$

for all n and b_{nj} (Efron and Tibshirani 1993).

We use subscript k to denote \tilde{y}^k , a hypothetical sample from $y(\tilde{I}) | F^o, I^o, y^o$, based on parameters from the k^{th} posterior draw. Sample k is generated as follows:

Step 1: Sample for $n = 1$. Bootstrap m_1^o pairs of (λ_{hi}, y_{hi}) from all observed individuals for whom $n_h = 1$.

Step 2: Sample for $n = 2, 3, 4$. We define b_{nj} , shown in Table 2, such that the number of individuals in the SCPC sample with estimated posterior means of $\lambda_{hi} | I_{hi}^o > 0$ in each interval is at least 75. For the k^{th} sample:

- I. For each j , determine $p_{nj}^k = \text{Prob}(\lambda_{hi} \in b_{nj} | n_h = n, \tau^k)$. Values of p_{nj} based on posterior means of τ are also shown in Table 2.

II. Sample $\{m_{nj}^k\}$ according to (10).

III. For each j , bootstrap m_{nj}^k pairs (λ_{hi}, y_{hi}) from the relevant subset of the k^{th} posterior draw of λ_{hi}^k in the 2017 SCPC sample: $\{\lambda_{hi}^k \in b_{nj} \mid n_h = n, 1_{hi}^o > 0\}$.

Step 3: Sample for $n \geq 5$. Households of size $n \geq 5$ are bootstrapped without consideration of sampled λ_{hi} . However, because only 3 percent of consumers belong to such households, doing so will have virtually no effect on population estimates.

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

Table 2: Partitions of $[0, 1]$, b_{nj} , and the population proportion of λ_{hi} in each, p_{nj} , based on posterior means of τ for households of size $n = 2, 3, 4$.

5.2 Generating Estimates

For each dataset, y , we estimate the mean number of monthly payments among U.S. consumers, μ , via poststratification (Little 1993; Gelman and Carlin 2002; Gelman 2007). Most generally, an estimate of μ based on a partition of the population into D disjoint strata with known population proportions, w_d , takes the form

$$\hat{\mu}(y) = \sum_{d=1}^D w_d \hat{\mu}_d(y), \quad (11)$$

where $\hat{\mu}_d(y)$ is an estimate of μ_d , the stratum-specific population mean. Using household size, household income, gender, and age to stratify, we generate two estimates of μ , described below.

Estimate A: Based on a 31-strata, we estimate μ_d with the observed stratum average, \bar{y}_d . The population mean estimates are designated by $\hat{\mu}_A(y)$ and represent simple poststratification estimates.

Estimate B: Based on a 75-strata, we use a hierarchical model of the observed stratum means, \bar{y}_d and standard errors, $\sqrt{v_d}$:

$$\bar{y}_d \sim \text{Normal}(\mu_d, \sqrt{v_d}) \quad \text{and} \quad \mu_d \sim \text{Normal}(\mu, \sigma).$$

Posterior means of μ_d are generated through RSTAN. Assumed priors are $\mu \sim \text{Normal}(20, 7)$, a diffuse distribution over the general range of population means estimated from previous years' data, and $\sigma \sim \text{half-Cauchy}(0, 2.5)$, a weakly-informative prior for between group variance (Gelman 2006; McElreath 2016). The hierarchical approach benefits from “partial pooling,” so that estimates for strata with few or no observations are stabilized by incorporating information from other strata (Elliott and Little 2000). The resulting population mean estimate is given by $\hat{\mu}_B(y)$.

The choice of demographic variables on which to stratify is important. It is possible that the observed sample bias with respect to λ_{hi} only relates to the distribution of sampled demographics, so that sample bias exists only because $P(d_{hi} \mid I_{hi}^o = 1) \neq P(d_{hi})$. In this case, poststratification on the appropriate set of demographics, d_{hi} , should adjust for bias.

Particularly important candidates are those demographic variables that vary sufficiently within household and have a non-trivial disparity between the sample and population distribution.

An interesting aspect of intra-household sampling is that in households featuring life partners many of the traditional demographic variables used in poststratification, such as age, education, and household income, tend to be similar, and thus are unlikely to adequately distinguish household members. The obvious exception is gender, and since 58 percent of 2017 SCPC respondents are female, it is a potentially vital stratification variable.

5.3 Results

For the k^{th} bootstrapped sample, a useful way of comparing the SCPC-based estimate, calculated using the first responders, $y^{op} = \{y_{hi} \mid I_{hi}^o = 1\}$, to bootstrapped estimates is through the percent deviation:

$$\Psi_A^k = 100 \times \frac{\hat{\mu}_A(\tilde{y}^k) - \hat{\mu}_A(y^{op})}{\hat{\mu}_A(\tilde{y}^k)} \quad \text{and} \quad \Psi_B^k = 100 \times \frac{\hat{\mu}_B(\tilde{y}^k) - \hat{\mu}_B(y^{op})}{\hat{\mu}_B(\tilde{y}^k)}$$

Histograms of the Ψ^k are shown in Figure 7 for both stratifications. Overall, the simulated estimates, which do not oversample high values of λ_{hi} , are generally lower than those based on the observed sample for all instruments except cash. Electronic payments shows the most evidence of overestimation, with estimated deviations centered around -14 percent. Checks and card payments show respective percent errors of about -8.5 and -5 percent respectively. For cash, the estimated percent deviations average to about 3 percent. Unlike for the other three payment instruments, there is little evidence that the sample bias observed in the 2017 SCPC sample significantly affects estimates of cash use. Results are robust to prior choices for μ and σ in Estimate B.

A simple explanation for these findings is that electronic payments, checks, and, to a lesser extent, cards are commonly used for bills and large-ticket items, which are more likely to be household purchases, such as cars, appliances, or services. In accordance with the theories of household economics, most notably those outlined by Becker (1991), efficiency dictates that household tasks often be concentrated among a subset of household members. As such, we hypothesize that it is natural for households to designate particular members, likely those with higher levels of responsibility and who are more involved in decision-making, to make these types of payments. Cash, on the other hand, is associated with daily, small-value purchases, which are less related to household needs and more to individual ones. As a result, greater responsibility within a household does not necessarily translate to more cash purchases.

5.4 Supplementary Analysis

Below, we briefly describe a supplementary analysis, detailed in Appendix B, that links posterior estimates of λ_{hi}^o to the share of household payments reported by each household member. We assume the reported number of payments $\{y_{hi}\}$ in each household follow a Multinomial distribution conditional on their sum with probabilities, $\{\gamma_{hi}\}$, representing the average share of household payments made by individual i . The share of payments made are linked to household roles through

$$\gamma_h \sim \text{Dirichlet}(\{a\lambda_h + b\}), \tag{12}$$

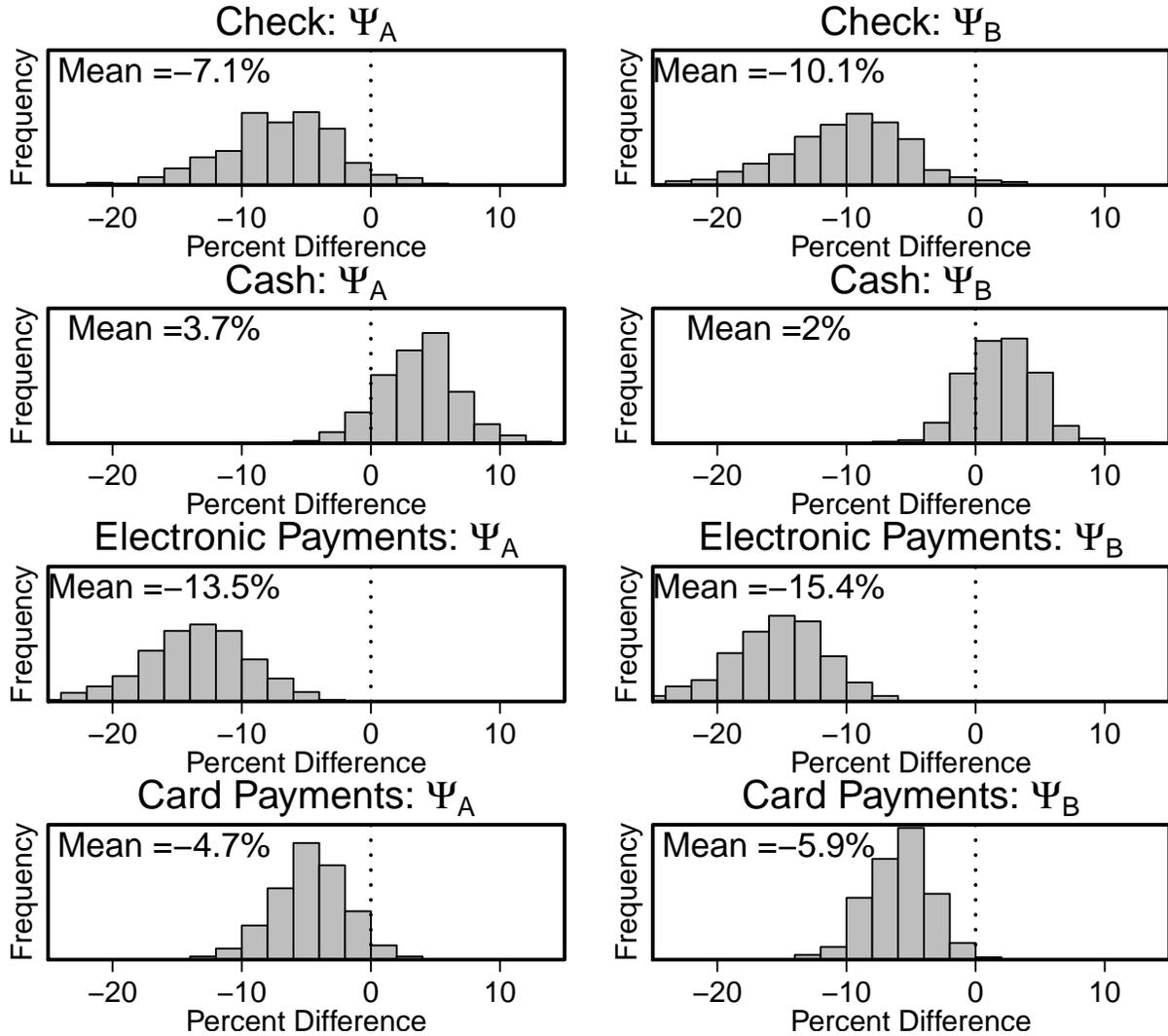


Figure 7: Estimated distributions of Ψ , the percent difference between mean estimates based on the primary respondents in the 2017 SCPC sample and those based on a hypothetical sample that draws individuals at random from within households, for each of the four payment instruments and based on two different postratification estimates (A and B).

where $\{a\lambda_h + b\}$ is the vector of Dirichlet concentration parameters that define the strength of the relationship between λ_{hi} and γ_{hi} . Large values of a , relative to b , suggest a strong correspondence between λ_{hi} and γ_{hi} . Alternatively if $a = 0$, λ_{hi} does not relate to the share of payments at all.

A useful measure of the relationship between λ_{hi} and the average share of payments made is

the proportion of variation in γ_{hi} explained by λ_{hi} :

$$\text{VarExp} = 1 - \frac{\text{E}[\text{Var}[\gamma_{hi} | \lambda_{hi}]]}{\text{Var}[\gamma_{hi}]} \quad (13)$$

Akin in spirit to R-squared, $\text{VarExp} \in [0, 1]$, 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 $\text{Var}[\gamma_{hi} | \lambda_{hi}] = 0$ and $\text{VarExp} = 1$. On the other end, if λ_{hi} does not relate to γ_{hi} at all then $\text{Var}[\gamma_{hi} | \lambda_{hi}] = \text{Var}[\gamma_{hi}]$ and $\text{VarExp} = 0$.

For a given set of estimates of λ_{hi} in the sample and τ , we generate maximum likelihood estimates of a, b , and, in turn VarExp . This is done separately for each payment instrument and for household sizes of $n = 2, 3, 4$. In doing so, posterior draws of λ_h^o are ranked by the total absolute deviation of the λ_{hi} from their respective posterior means, thus arranging posterior draws roughly according to perceived likelihood. The k^{th} estimation is based on the set of household draws of λ_h^o whose total absolute deviation from posterior means is k^{th} largest. Figure 8 shows the estimates of VarExp as a function of distance to posterior mean as well as an estimate based on posterior means of λ_h^o . Most generally, these results lend validity to our above results. First, the relationships between household role and use of a payment instrument are confirmed. In addition, the fact that these relationships strengthen as one assumes values of λ_{hi} closer to their posterior mean supports our methodology for estimating household responsibility shares. Interestingly, λ_{hi} and γ_{hi} have a weaker relationship in two-adult households than in larger ones. This may be partly due to more equal distribution of responsibility in two-adult households or the fact that easier intra-household communication and more fluid dynamics make sharing of tasks easier. By contrast, in larger households, clear and efficient delegation of tasks is more important to efficient household management.

6 Discussion

In our view, the most meaningful contribution of this work is the recognition of a potential source of sample bias that, as far as we are aware, is generally ignored. Researchers who rely on surveys should be generally aware of intra-household variability in behavior and the

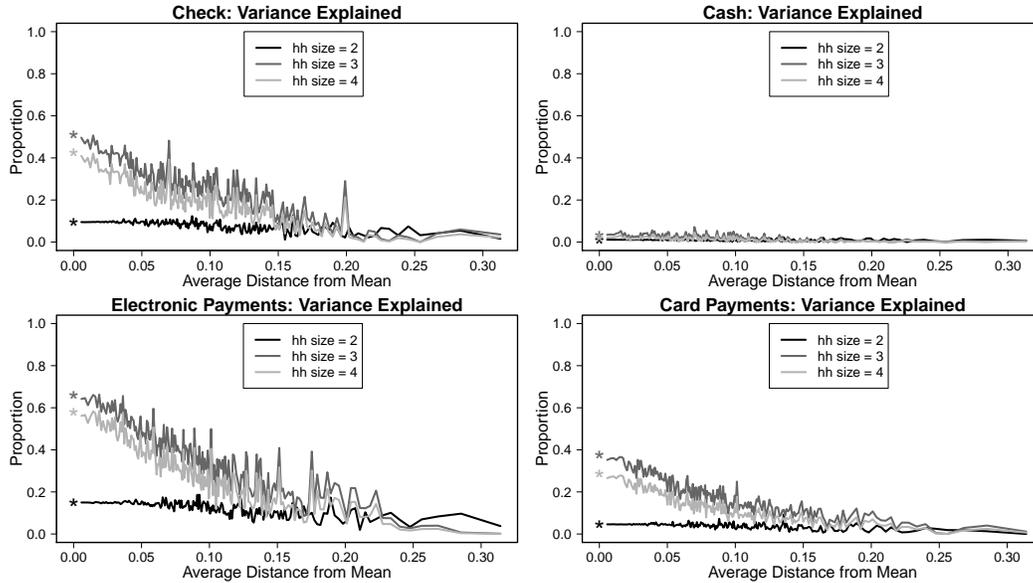


Figure 8: Estimated proportion of variation in F_{hi} explained by λ_{hi} for each of the four payment instrument groups. The * represent calculations using posterior means of the observed λ_{hi} , $\bar{\lambda}_{hi}$.

possibility of introducing bias by over- or under-sampling individuals with certain types of roles. It is possible that many sampling schemes, not just those based on address-based sampling, are affected by such a bias. In fact, all patterns observed in the 2017 SCPC sample are also found in the 2016 SCPC sample and the 2012 SCPC sample, the latter of which was fielded to a set of respondents recruited via an entirely different methodology than CESR uses (ALP Various Years).

The need to address the issue depends on how closely the variable of interest relates to household role, but it seems likely that many social and economic variables will be affected. A simple consideration for researchers studying variables that are potentially related to household role is to consider the relative values of individual-based questions and 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 (CES Various Years) or the Survey of Consumer Finance (SCF Various Years). Of course, the quality of such estimates depends on the information available to the respondent. 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 everyone’s cash spending.

Alternatively, this paper identifies a methodology to identify and adjust for the effects of

sample bias related to household role. This process not only requires a particular question format, but also relies on a unique data structure in which a non-trivial fraction of households feature several respondents. As such, it is impractical to implement in most surveys. Improvement in the process depends on simultaneously devising survey variables that are less prone to measurement error and developing a better statistical framework for incorporating household role information into inference. In particular, adjustments based on choices of λ_{hi} that more closely relate to the outcome of interest will yield less uncertainty about the true value.

References

- AAPOR (2016), "Address-Based Sampling," <http://www.aapor.org/Education-Resources/Reports/Address-based-Sampling.aspx>.
- Agresti, A. (2002), *Categorical Data Analysis*, New York, NY: Wiley, 2nd ed.
- ALP (Various Years), "American Life Panel," <https://mmicdata.rand.org/alp/>.
- Baker, R., Brick, J. M., Bates, N. A., Battaglia, M., Couper, M. P., Dever, J. A., Gile, K. J., and Tourangeau, R. (2013), "Summary Report of the AAPOR Task Force on Non-Probability Sampling," *Journal of Survey Statistics and Methodology*, 1(2), 90–105.
- Battaglia, M. P., Link, M. W., Frankel, M. R., Osborn, L., and Mokdad, A. H. (2008), "An Evaluation of Respondent Selection Methods for Household Mail Surveys," *Public Opinion Quarterly*, 72, 459–469.
- Becker, G. (1991), *A Treatise on the Family*, Cambridge, Massachusetts: Harvard University Press.
- Campbell, G. and Mosimann, J. (1987), "Multivariate Methods for Proportional Shape," *ASA Proceedings of the Section on Statistical Graphics*, 10–17.
- CES (Various Years), "Consumer Expenditure Survey," <http://www.bls.gov/cex/home.htm>.
- Cornesse, C. and Bosnjak, M. (2018), "Is There an Association Between Survey Characteristics and Representativeness? A Meta-Analysis," *Survey Research Methods*, 12(1), 1–13.

- Curtin, R., Presser, S., and Singer, E. (2005), "Changes in Telephone Survey Nonresponse Over the Past Quarter Century," *Public Opinion Quarterly*, 69(1), 87–98.
- Deming, W. E. and Stephan, F. F. (1940), "On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Tables are Known," *The Annals of Mathematical Statistics*, 11, 427–444.
- Deville, J. C. and Sarndal, C. (1992), "Calibration Estimators in Survey Sampling," *Journal of the American Statistical Association*, 87, 376–382.
- Dillman, D. A., Smyth, J. D., and Christian, L. M. (2014), *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*, New York, New York: Wiley & Sons.
- Efron, B. and Tibshirani, R. (1993), *An Introduction to the Bootstrap*, Boca Raton, FL: Chapman and Hall/CRC.
- Elliott, M. R. and Little, R. J. A. (2000), "Model-Based Alternatives to Trimming Survey Weights," *Journal of Official Statistics*, 16, 191–209.
- Gelman, A. (2006), "Prior Distributions for Variance Parameters in Hierarchical Models," *Bayesian Analysis*, 1(3), 515–533.
- (2007), "Struggles with Survey Weighting and Regression Modeling," *Statistical Science*, 22(2), 153–164.
- Gelman, A. and Carlin, J. B. (2002), "Poststratification and Weighting Adjustments," in *Survey Nonresponse*, eds. Groves, R. M., Dillman, D. A., Eltinge, J. L., and Little, R. J. A., Wiley.
- Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004), *Bayesian Data Analysis*, Boca Raton, FL: Chapman and Hall, 2nd ed.
- Groves, R. M. (2006), "Nonresponse Rates and Noresponse bias in Household Surveys," *Public Opinion Quarterly*, 70, 646–675.
- Hijazi, R. and Jernigan, R. (2009), "Modelling Compositional Data Using Dirichlet Regression Models," *Journal of Applied Probability and Statistics*, 4, 77–91.

- Hoff, P. D. (2009), *A First Course in Bayesian Statistical Methods*, New York, New York: Springer.
- Horvitz, D. and Thompson, D. (1952), "A Generalization of Sampling Without Replacement from a Finite Universe," *Journal of the American Statistical Association*, 47, 663–685.
- Kott, P. S. and Chang, T. (2010), "Using Calibration Weighting to Adjust for Nonignorable Unit Nonresponse," *Journal of American Statistical Association*, 105, 1265–1275.
- Little, R. J. (1986), "Survey Nonresponse Adjustments for Estimates of Means," *International Statistical Review*, 54 (2), 139–157.
- Little, R. J. A. (1993), "Post-stratification: A Modeler's Perspective," *Journal of the American Statistical Association*, 88, 1001–1012.
- Little, R. J. A. and Vartivarian, S. (2005), "Does Weighting for Nonresponse Increase the Variance of Survey Means?" *Survey Methodology*, 31, 161–168.
- Lohr, S. L. (1999), *Sampling: Design and Analysis*, Pacific Grove, CA: Brooks/Cole Publishing Company.
- Lundstrom, S. and Sarndal, C.-E. (1999), "Calibration as a Standard Method for Treatment of Nonresponse," *Journal of Official Statistics*, 15 (2), 305–327.
- McElreath, R. (2016), *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*, Boca Raton, FL: CRC Press.
- Park, D. K., Gelman, A., and Bafumi, J. (2004), "Bayesian Multilevel Estimation with Poststratification: State-Level Estimates from National Polls," *Political Analysis*, 12(4), 375–385.
- Pfeffermann, D. and Sikov, A. (2011), "Imputation and Estimation under Nonignorable Nonresponse in Household Surveys with Missing Covariate Information," *Journal of Official Statistics*, 27 (2), 181–209.
- Qin, J., Leung, D., and Shao, J. (2002), "Estimation With Survey Data Under Nonignorable Nonresponse or Informative Sampling," *Journal of the American Statistical Association*, 97, 193–200.

- Rubin, D. B. (1987), *Multiple Imputation for Nonresponse in Surveys*, New York, NY: Wiley.
- SCF (Various Years), "Survey of Consumer Finances,"
<http://www.federalreserve.gov/econresdata/scf/scfindex.htm>.
- Schaeffer, N. C. and Presser, S. (2003), "The Science of Asking Questions," *Annual Review of Sociology*, 29, 65–88.
- SCPC (Various Years), "2017 Survey of Consumer Payment Choice,"
<https://www.frbatlanta.org/banking-and-payments/consumer-payments/survey-of-consumer-payment-choice>.
- Seaman, S. R. and White, I. R. (2011), "Review of Inverse Probability Weighting for Dealing with Missing Data," *Statistical Methods in Medical Research*, 22(3), 278–295.
- Svenson, O. (1981), "Are We All Less Risky and More Skillful Than Our Fellow Drivers?" *Acta Psychologica*, 47(2), 143–148.
- Tourangeau, R. and Plewes, T. (2013), "Nonresponse in Social Science Surveys: A Research Agenda," http://www.nap.edu/openbook.php?record_id=18293.
- UAS (Various Years), "Understanding America Study," <https://cesr.usc.edu/>.
- Valliant, R. and Dever, J. A. (2011), "Estimating Propensity Adjustments for Volunteer Web Surveys," *Sociological Methods and Research*, 40(1), 105–137.
- van Dyk, D. A. and Meng, X.-L. (2012), "The Art of Data Augmentation," *Journal of Computational and Graphical Statistics*, 10(1), 1–50.
- Wang, W., Rothschild, D., Goel, S., and Gelman, A. (2009), "Forecasting Elections with Non-Representative Polls," *Public Opinion Quarterly*, 73(5), 895–916.

A Appendix A:

Table 3 shows posterior means and standard errors for all parameters in the data augmentation models. The partitions $k[f]$ for the two-adult household model are

1. $k[f] = 1$ for $f \in \{0 - 3\}$
2. $k[f] = 2$ for $f \in \{4 - 7\}$
3. $k[f] = 3$ for $f \in \{8 - 10\}$
4. $k[f] = 4$ for $f \in \{11 - 13\}$
5. $k[f] = 5$ for $f \in \{14 - 16\}$.

$n_h = 2$						$n_h > 2$	
	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	β_1	0.07 (0.2)
β_k	0.74 (0.3)	-0.23 (0.1)	-0.37 (0.1)	-0.26 (0.2)	-0.14 (0.2)	β_2	0.23 (0.2)
g_{k1}	-0.09 (0.6)	-3.00 (0.8)	-5.58 (1.3)	-6.17 (2.7)	-3.23 (3.1)	β_3	-0.14 (0.04)
g_{k2}	0.40 (0.7)	-2.99 (0.8)	-5.06 (1.3)	-4.70 (2.4)	-2.88 (3.1)	β_4	0.003 (0.0)
g_{k3}	0.42 (0.7)	-2.79 (0.8)	-4.78 (1.3)	-4.24 (2.4)	-2.61 (3.1)	g_1	-2.34 (0.7)
g_{k4}	0.45 (0.7)	-2.78 (0.8)	-4.67 (1.3)	-4.16 (2.4)	-2.54 (3.1)	g_2	-1.98 (0.7)
g_{k5}	0.47 (0.7)	-2.77 (0.8)	-4.42 (1.3)	-3.64 (2.4)	-2.30 (3.1)	g_3	-1.57 (0.7)
g_{k6}	0.49 (0.7)	-2.76 (0.8)	-4.23 (1.3)	-3.53 (2.4)	-2.08 (3.1)	g_4	-1.24 (0.7)
g_{k7}	0.52 (0.7)	-2.75 (0.8)	-3.93 (1.3)	-2.83 (2.4)	-1.96 (3.1)	g_5	-0.99 (0.7)
g_{k8}	0.54 (0.7)	-2.40 (0.8)	-3.59 (1.3)	-2.57 (2.4)	-1.95 (3.1)	g_6	-0.85 (0.7)
g_{k9}	0.56 (0.7)	-2.22 (0.8)	-3.08 (1.3)	-2.20 (2.4)	-1.73 (3.1)	g_7	-0.69 (0.7)
$g_{k,10}$	0.58 (0.7)	-2.01 (0.8)	-2.80 (1.3)	-2.05 (2.4)	-1.54 (3.1)	g_8	-0.54 (0.7)
$g_{k,11}$	0.88 (0.7)	-1.54 (0.8)	-2.55 (1.3)	-1.87 (2.4)	-1.26 (3.1)	g_9	-0.16 (0.7)
$g_{k,12}$	0.89 (0.7)	-1.27 (0.8)	-2.11 (1.3)	-1.86 (2.4)	-1.25 (3.1)	g_{10}	-0.00 (0.7)
$g_{k,13}$	1.12 (0.7)	-0.58 (0.8)	-1.77 (1.3)	-1.76 (2.4)	-1.08 (3.1)	g_{11}	0.25 (0.7)
$g_{k,14}$	1.33 (0.7)	-0.27 (0.8)	-1.50 (1.3)	-1.75 (2.4)	-0.99 (3.1)	g_{12}	0.34 (0.7)
$g_{k,15}$	1.54 (0.7)	0.05 (0.8)	-1.30 (1.3)	-1.74 (2.4)	-0.98 (3.1)	g_{13}	0.55 (0.7)
$g_{k,16}$	2.19 (0.8)	0.35 (0.8)	-0.98 (1.3)	-1.61 (2.4)	-0.97 (3.1)	g_{14}	0.67 (0.7)
						g_{15}	0.90 (0.7)
						g_{16}	1.09 (0.7)

Table 3: Estimated means and standard errors of β, g from the data augmentation models.

B Appendix B: Linking λ_{hi} to Average Share of Household Payments

Let $t_h = \sum_{i=1}^{n_h} y_{hi}$ represent the total number of household payments 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 individual. The parameters and data can be linked by the model

$$y_h \mid t_h, \gamma_h \sim \text{Multinomial}(t_h, \gamma_h). \quad (14)$$

By adopting the model in (12), we assume an approach based on Dirichlet regression (Campbell and Mosimann 1987; Hijazi and Jernigan 2009). We impose the restrictions that $b > 0$ and $a + b > 0$ to ensure well-defined Dirichlet distributions.

In the case where data for only a subset of household members are observed, we define $t_h^o = \sum_{\{i|I_{hi}^o > 0\}} y_{hi}^o$ as the total number of payments reported by sampled members of household h . Then, the assumed multinomial model in (14) further implies that

$$y_h^o \mid t_h^o, \gamma_h^o \sim \text{Multinomial}(t_h^o, \gamma_h^o), \quad (15)$$

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

$$\gamma_h^o \sim \text{Dirichlet}(\{a\lambda_h^o + b\}). \quad (16)$$

Combining (15) and (16), the likelihood function for (a, b) for household h can be expressed as

$$L_h(a, b) = \int P(y_h^o \mid t_h^o, \gamma_h^o) P(\gamma_h^o \mid a, b, \lambda_h^o, y_h^o) d\gamma_h^o, \quad (17)$$

and integration of (17) leads to the closed-form likelihood

$$L_h(a, b) = \frac{\Gamma(a \sum_{i|I_{hi}^o > 0} \lambda_{hi} + s_h^o b) \prod_{i|I_{hi}^o > 0} \Gamma(y_{hi} + a\lambda_{hi} + b)}{\Gamma(a \sum_{i|I_{hi}^o > 0} \lambda_{hi} + s_h^o b + t_h^o) \prod_{i|I_{hi}^o > 0} \Gamma(a\lambda_{hi} + b)}. \quad (18)$$

If only one household member is sampled ($s_h^o = 1$), the likelihood in (18) reduces to 1 for all a and b .

The form for VarExp in (13) depends on the first two moments of λ_{hi} through the form

$$\text{VarExp}(a, b, \tau_h) = \frac{\kappa_1 \text{Var}[\lambda_{hi}]}{\kappa_1 \text{Var}[\lambda_{hi}] + \kappa_2 (\text{E}[\lambda_{hi}]^2 + \text{Var}[\lambda_{hi}]) + \kappa_3 \text{E}[\lambda_{hi}] + \kappa_4},$$

where

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

The assumption that $\lambda_h \sim \text{Dirichlet}(\tau_h, \dots, \tau_h)$, implies

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

We use the posterior means of the observed λ_{hi} to generate maximum likelihood estimates for a, b and then VarExp. 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. To simplify notation, we relabel the observed members of household h according to $j = 1, \dots, s_h^o$ so that the j^{th} observed member of household h is that for which $I_{hi}^o = j$. Then, $\lambda_h^o[j, k]$ represents the k^{th} posterior draw of for observed member j , and $\bar{\lambda}_h^o[j]$ represents the posterior mean based on the 500 draws. For the k^{th} posterior draw, the absolute deviation from the posterior means for household h is defined as

$$\Delta_h^k = \sum_{j=1}^{s_h^o} |\lambda_h^o[j, k] - \bar{\lambda}_h^o[j]|.$$

Households are then ordered according to the values of Δ_h^k , to yield the order statistics $\Delta_h^{(k)}$ for which $\Delta_h^{(k)} < \Delta_h^{(k+1)}$. The k^{th} estimate of a, b , and VarExp are based on $\Delta_h^{(k)}$ for each household, with τ_h taken as its posterior mean.