Wife vs. Husband: Does It Matter Who Answers the Survey?

Adan Silverio-Murillo* Edmundo Molina †

Abstract

Information on household assets is often used to conduct empirical research and to guide public policy. Practitioners use these variables, because they are assumed to be less susceptible to misreporting. To test this assumption, the current study employs data from poor households participating in Mexico's PROGRESA program. Separately, both the wife and the husband were asked the same questions regarding household assets. The study finds the following: (1) There were major discrepancies in the information reported by the spouses. For example, there was disagreement among 24% of the couples as to the possession of a washing machine. (2) The latter result has consequences for identifying families living in poverty. For example, if husbands were to be asked, 10.1% of households would be classified as non-poor, but if wives were to be asked, they would be classified as poor. (3) The discrepancies observed can be partially explained by careless responses given by husbands. This result is robust to a bounding argument for omitted variable bias implemented by Oster (2017). Overall, these findings suggest that survey information on household assets is not free of misreporting, and that who answers the survey matters.

Keywords: Poverty Measurement, Household Survey, Gender.

JEL: C81, C83, D01.

*School of Government, Tecnologico de Monterrey. E-mail: adan.sm@tec.mx

†School of Government, Tecnologico de Monterrey. E-mail: edmundo.molina@tec.mx

1 Introduction

Household surveys are among the main instruments for conducting empirical research and making public policy decisions. Thus, understanding the quality of surveys is an issue of great importance. One of the commonly used variables is information on household assets. This information is used to target programs for the poor, to generate multidimensional measures of poverty, and to test baseline differences in randomized controlled trials, among other uses. Most importantly, such information on household assets is commonly assumed to be free of misreporting problems.

To examine the accuracy of survey data on assets, the current study examined a random sample of 960 couples participating in the social program, PROGRESA in Mexico.[1] Separately, both the wife and the husband were asked questions about the possession of 18 assets. Upon analyzing their responses, discrepancies could be found in the information reported between spouses under every item. The extent of this mismatch ranged from 2.2% to 32.6%. For example, there was disagreement among 21.5% of the couples regarding the possession of a refrigerator.[2]

Following this, we analyzed the extent to which the mismatching observed was random or not. There are many reasons why misreporting may occur when collect-

[1] Mexico implemented the PROGRESA program in 1997, which transfers money to low-income families under the condition that they send their children to school.

[2] In particular, the husband reported having a refrigerator and the wife did not in 9.6% of the households surveyed. Similarly, the husband reported not having a refrigerator and the wife reported having one in 11.9% of households.

ing survey data: (1) individuals simply respond carelessly (careless responses); (2) individuals fail to understand the question or have problems to recall information (cognitive ability); (3) impulsive individuals may have incentives to underreport particular information (lack of self-control); (4) individuals report information that they consider to be socially desirable; (5) the size of the survey can affect how individuals report the information (survey design); and (6) interviewers can affect the quality of information.

In this paper, we test for careless responses, cognitive ability, and the lack of self-control. In the case of careless responses, we look for patterns in the questions regarding assets. In particular, it is well known that annoyed respondents tend to answer surveys in straight vertical lines (Leiner, 2019). Thus, we identify those individuals who answer the questions regarding assets in this way. With respect to the cognitive process, it involves the comprehension of the question and recalling information. To analyze this hypothesis, we use the percentage of correct answers to a cognitive test supplied at the beginning of the interview. The idea is that the percentage of correct answers (a proxy for the cognitive ability of the individuals) may predict the individual's accuracy throughout the rest of the survey. Finally, regarding fraudulent responses, individuals with self-control problems will be more likely to report fraudulent responses. To analyze this hypothesis, we use a test that measures self-control. We do not find evidence that the cognitive process or self-control impacted the discrepancies observed regarding assets. Yet, we find evidence that careless responses given by husbands explained partially the discrep-

ancies observed regarding assets. This result is possibly a consequence of omitted variables. Following the procedures proposed by Oster (2017), the current results were found to be robust to the aforementioned econometric problem.

In addition, the paper analyzes the consequences of such disagreement with the information reported on assets, for the purposes of poverty classification. A key finding here is that who reports the information matters. For example, 10.1% of households would be classified as non-poor if the husband were to be asked, but would be classified as poor if the wife were surveyed.

This paper relates to a body of literature regarding the causes and consequences of discrepancies in the information reported by couples in household surveys. Using data from Malawi, Fisher et al. (2010) found that it is not sufficient to interview only one member of the household to obtain the household income. For example, they found that wife's income provided by both husband and wife would match in only 6% of the households interviewed. Ambler et al. (2019), using data from Bangladesh, found that women are much more likely than men to report women's roles in asset ownership and decision making. Doss et al. (2018), using data from Ghana, Ecuador, and India, analyzed whether wives and husbands provide different responses to questions about the monetary value of their home. They found that wives tend to report lower values than husbands.

There are three main contributions of this paper. First, it contributes to the

literature showing discrepancies in the information on assets collected through household surveys. Second, this paper adds to a growing literature about the causes of misreporting. In particular, it presents evidence that careless responses given by husbands can generate problems in the data regarding assets. Finally, the paper presents evidence that who answers the survey matters for social programs targeting poverty.

The remainder of this paper is organized as follows. Section 2 reviews the related literature and Section 3 introduces the data. Section 4 describes the empirical strategy. Section 5 presents the results, and Section 6 concludes.

2 Why Do People Misreport Information?

Philipson and Malani (1999) point out that economists pay much more attention to the consumption of data than to the production of data. They propose the analysis of the data collection process as a principal-agent problem, whereby the investigator is the principal and individuals who provide information are agents. The problem is that agents have preferences (i.e., does the respondent want to tell the truth?), and may experience problems pertaining to information accuracy (i.e., does the respondent know the truth?). This situation is the principal source of erroneous reporting.

Judge and Schechter (2009) proposed that Benford's law can be used as a tool to detect problems in survey data. The idea behind Benford's law is that, in large data sets, numbers with a first digit of 1 are observed more often than those starting with 2, which are, in turn, observed more often than those starting with 3, and so on.[3] Judge and Schechter (2009) analyzed data from nine commonly used datasets including the Matlab Health and Socioeconomics Survey from Bangladesh, PROGRESA data from Mexico, the Living Standards Measurement Survey from Peru, and the Agricultural Resource Management Survey from the United States, among others. Their principal finding is that the data from developing countries are of poor quality, while data from the United States are of better quality. They also found that female and male respondents provide data of similar quality.

However, recent evidence calls into question the quality of data collected through household surveys in the United States. For example, Meyer et al. (2018) examined three large household surveys used in the United States, finding underreport on participation in the food stamps program in the three surveys analyzed. Specifically, the underreport was found to be 23% with the Survey of Income and Program Participation, 35% with the American Community Survey, and 50% with the Current Population Survey. They also found misreporting to be associated with household characteristics (householder age, speaking poor English, and the non-U.S. citizen status). In addition, they uncovered other variables associated with misreporting (being disabled, respondent's education level, and living in a

[3]In particular, Benford's law proposes that: P(First digit is d)= $log_{10}(1 + \frac{1}{d})$, where d goes from 1 to 9.

rural area); yet, the effects were seen as mixed or inconclusive.

The potential reasons behind misreporting can be related to individuals, characteristics of the survey, and interviewers. Regarding the individuals, the literature has identified three problems: careless responses, cognitive problems, and fraudulent responses. In a simplified theoretical model, the individuals decide first if they will provide careless responses or not. If they decided to provide careless responses then we will have misreporting problems. If the individuals decide to provide careful responses, Tourangeau et al. (2000) propose that respondents follow these stages when answering survey questions: comprehension of the question, retrieval of the information needed to answer the question, and a judgment of the information they will report. The first two stages are related to the cognitive process, and the last one is linked to the potential of fraudulent responses.

Leiner (2019) proposes that careless responses can be identified when there are patterns in the answers provided by the individuals. In particular, annoyed respondents typically answer the questions in straight vertical lines, diagonal lines, and a combination of both. For its part, the cognitive process involves the comprehension of the question, the effort involved in answering the survey, and recalling information, all of which are factors that will affect the quality of the data obtained. Finally, regarding fraudulent responses they can be related to the incentives provided to answer the survey and to social desirability. Regarding the first aspect, being given incentives to answer a survey matters. For example, if the survey will

be used to assign respondents' participation in a social program, individuals may have incentives to underreport particular information. Regarding social desirability, it refers to the phenomenon of giving socially desirable answers, whether they are true or not. For example, if having a computer at home is something socially desirable, an individual who may not have a computer at home will answer that she has it to meet what is social desirable.

In addition, the survey design can affect how individuals respond; for example, if the survey is short, individuals may be more focused on the answers, compared to the case of very long surveys. Finally, interviewers themselves can affect the quality of information. For example, in a survey of domestic violence, women would possibly reveal more accurate information to a female than a male interviewer.

Some papers have found empirical evidence related to the aforementioned factors affecting the quality of the information obtained through surveys. For example, Martinelli and Parker (2009), using data from PROGRESA, find evidence that supports the social desirability hypothesis. In particular, they find overreporting on household goods that may carry a certain social "status" (e.g., concrete floor, tap water, toilet). Kilic and Sohnesen (2017) present evidence regarding the survey design. Using an experiment conducted in Malawi, they find that the size of the survey matters, as the same households answered the same questions differently depending on the length of the questionnaire. Finally, using data from Nigeria,

Onwujekwe et al. (2006), found discrepancies in the socio-economic information obtained by different interviewers visiting the same home.

3 Data

The main data source used in this paper is the Survey of Resilience and Social Mobility of participants in Mexico's PROGRESA program (rebranded as *Oportunidades* and then as PROSPERA).[4] The survey is composed of four parts, in the following order: (1) a cognitive test, (2) psychological tests, (3) socioeconomic aspects, and (4) childhood. The questions pertaining to household assets fall under module (3). The question regarding the possession of assets is as follows: "Does your household own any of the following items?" The 18 items then referred to are as follows in this order: television, photographic camera, music device, automobile, sofa, washing machine, refrigerator, gas stove, landline, bicycle, motorcycle, canoe or boat, machinery or work equipment, farm animals, local business, house, apartment or room for rent, other land (apart from home), and savings account or saved money.

The sample design used the list of 5 million households enrolled in the program in May 2009. From these 5 million households, localities with fewer than 45 households were excluded. Thus, the list was reduced to 2.4 million households. From

[4]PROGRESA was the biggest social program in Mexico, providing assistance to over six million families. This program offered cash transfers to families living in poverty on the condition that they send their school-age children to school.

this list, a probabilistic survey of 1,960 households was selected: 850 households in rural and 1,110 in urban areas (Palomar, 2012). From these 1,960 households, 960 households were randomly selected to collect information about wives and husbands.[5] The survey was collected between October and December of 2010.[6]

The 18 assets are presented in Table 1. Columns 1 and 2, respectively, show the percentage of husbands and wives reporting ownership of the assets. In general, there appear to be no major discrepancies between the information reported by the couples. For example, 63.2% of husbands and 65.5% of wives reported possession of a refrigerator inside the home. Column 3 tests for the same mean assets reported by wives and husbands. Among the 18 assets analyzed, statistically, significant differences were found for six of them: bicycle, automobile, photographic camera, land, machinery, and savings. Column 4 shows the percentage of cases containing disagreement in the information reported by the spouses. The disagreement ranges from 2.2% to 32.6%. For example, disagreement can be seen among 21.5% of the couples regarding the possession of a refrigerator; in particular, husbands reported having a refrigerator while their wives did not in 9.6% of households (Column 5). In addition, husbands reported not having a refrigerator while their wives did report having one in 11.9% of households (Column 6). From a methodological

[5] If the respondent was a widow, a single mother, or if the husband worked outside the locality, that house was replaced with another from the selected sample of 1,960 households.

[6] A pilot study was conducted to validate the survey. Regarding the interviewers, each of them received an instructor's manual and training prior to the field survey. The survey involved 100 full-time interviewers who were organized into 25 brigades, with a field supervisor per brigade. The surveys were conducted in households, with an average duration of two hours. The surveys applied to women were obtained on a single occasion. However, in the case of men, an extra field trip was made to some localities due to the fact that in the first field trip it was only possible to survey 15% of the male spouses. (Palomar, 2012).

perspective, the important question is whether the disagreement observed in the information reported by the spouses is random or not. Columns 3 and 4 in Table 1 appear to support the hypothesis that the discrepancies observed are not random.

Table 2 shows that, on average, there are 2.62 items under which the information reported by the spouses does not match within the households. It is observed that only in 14.2% of households does the information reported by the spouses completely coincide. In the rest of the homes, the disagreement in the reported information ranged from one item to a maximum of 12 items.

The current paper analyzes how careless responses, cognitive ability, and self-control (aspects related to the individuals) can affect discrepancies in the information reported. To identify careless individuals, we exploit the characteristic of the module regarding assets in the survey. As mentioned previously, when asking information about the possession of assets the order is as follows: (1) television, (2) photographic camera, (3) music device, (4) automobile, (5) sofa, (6) washing machine, (7) refrigerator, (8) gas stove, (9) landline, (10) bicycle, (11) motorcycle, (12) canoe or boat, (13) machinery or work equipment, (14) farm animals, (15) local business, (16) house, apartment or room for rent, (17) other land (apart from home), and (18) savings account or saved money. On average, the households reported the possession of 4.4 assets with a standard deviation of 2.4. Notice that with high probability there should be at least one of the first nine items in the household. Also, with high probability it sounds implausible that almost all of the

last nine items are present in the household. Thus, we classify an individual as providing careless answers when they: (1) answered the first nine questions of the module regarding assets in a negative way forming a vertical line or (2) answered affirmatively the last nine questions forming a vertical line. Notice that these patterns are highly unlikely and are suspicious of careless responses. Table 2 presents evidence of careless responses by four percent of the husbands. Yet, we fail to find evidence of careless responses by the wives.

Regarding the cognitive process, we use the percentage of correct answers to a cognitive test supplied at the beginning of the interview. The cognitive test used is a Raven test. The Raven test is designed to measure the non-verbal, abstract, and cognitive functioning. It includes a matrix of geometric designs with one piece missing. The interviewed choose one diagram from a set of eight answers. The Raven test used in this survey has 12 questions and it was adopted to the Mexican case by Palomar (2012). Table 2 shows that regarding the percentage of correct answers to the Raven test, husbands responded correctly 36% and wives 34% of the questions, on average.

To test for potential fraudulent responses, we use a test regarding self-control. The hypothesis is that individuals with low self-control engage in deviant behavior. Gottdredson and Hirschi (1990) affirm that in order to get away with rule-breaking conducts, such as cheating, individuals are required to have self-control. The self-control test is based on the Self-Regulatory Questionnaire (SQR) of Brown et al.

(1999), whose test was adapted for the Mexican context by Palomar (2012). The principal components were used to get a measurement of these variables. Then, these results were standardized to a mean of zero and a standard deviation of one. Table 2 presents the results for wives and husbands. It is observed that husbands (0.06) exhibit higher levels of self-control than wives (-.06).

Table 2 also presents information regarding other variables applied as controls in the current study. As can be seen, in the case of age, husbands are typically older than wives, at an average of 48.7 and 44.8 years old, respectively. Regarding education, 64% percent of husbands can read and write. This percentage is a little higher in the case of wives (68%). Regarding speaking an indigenous language, no important difference was observed (5% for both wives and husbands). We also include information regarding suffering a disability (motor, visual, or auditory). It is observed that 7% of husbands suffer a disability and so do 6% of wives. Another variable included as a control was cohabiting (free union),[7] given that the type of marital arrangement could possibly affect the exchange of information within the household. As shown in Table 2, 21% of couples in the surveyed households reported cohabiting.

Finally, we control for the occurrence of natural disasters. Individuals who suffer a natural disaster can be more aware of the assets they lost and have a better measure of them. The data used to measure natural disasters were drawn from the National Center for the Prevention of Disasters, where the information is classified

[7]Applied to those couples who declared not being in a legal marriage

by type of disaster (hydrometeorological, earthquakes, droughts, and others) for all municipalities in Mexico. As shown in Table 2, 15% of the households analyzed suffered some type of natural disaster during the 12 months before conducting the survey.

4 Estimation Strategy

4.1 Identification Strategy

We analyze the effects of careless responses, cognitive ability, and self-control on the differences in the assets as reported by the wife and the husband. Ideally, we would like to estimate the following equations for the husband (i) and the wife (j):

$$Y_{ih} - A_h = \alpha_1 + \beta_{11} {T_{ih}}^1 + \beta_{12} {T_{ih}}^2 + \beta_{13} {T_{ih}}^3 + \theta_1 X_{ih} + e_{ih} \tag{1a}$$

$$Y_{jh} - A_h = \alpha_2 + \beta_{21} {T_{jh}}^1 + \beta_{22} {T_{jh}}^2 + \beta_{23} {T_{jh}}^3 + \theta_2 X_{jh} + e_{jh} \tag{1b}$$

where $Y_{ih}(Y_{jh})$ is the asset reported by individual i(j) in house h. A_h is a dummy variable that measures if the asset is really inside the home. ${T_{ih}}^1$ and ${T_{jh}}^1$ represent a dummy variable that measures if individuals respond in a careless way for the husband and the wife respectively. ${T_{ih}}^2$ and ${T_{jh}}^2$ represent the percentage of correct answers to the cognitive test for the husband and the wife, respectively.

${T_{ih}}^3$ and ${T_{jh}}^3$ represent the measures of self-control for the husband and the wife, respectively. X_{ih} and X_{jh} represent a vector of control variables for the husband and the wife, and e is an error term with mean zero. Unfortunately, we cannot observe A_h, thus we use the following specification:

$$Y_{ijh} = Y_{ih} - A_h\text{- } Y_{jh} + A_h = Y_{ih} \text{ - } Y_{jh} = \alpha_1 - \alpha_2 + \beta_{11}{T_{ih}}^1 - \beta_{21}{T_{jh}}^1 + \beta_{12}{T_{ih}}^2 - \beta_{22}{T_{jh}}^2 + \beta_{13}{T_{ih}}^3 - \beta_{23}{T_{jh}}^3 + \theta_1 X_{ih} - \theta_2 X_{jh} + e_{ijh} \quad (2)$$

Where Y_{ijh} is the difference in the information reported between the husband (i) and the wife (j). In addition, we will include fixed effects at the level of municipality. These fixed effects control for the team that collected the survey in that specific area. Finally, e_{ijh} is an error term with mean zero. Standard errors are clustered at the municipality level. The coefficients of interest are β_{1k} and β_{2k} for k=1,2, and 3. They represent the effects of careless responses, cognitive ability, and self-control on the differences in the assets reported by the wife and the husband. For example, if β_{11} is positive, it implies that, when husbands respond carelessly, they overreport the possession of that asset. On the other hand, when $\beta 11$ is negative, it implies that, when husbands respond carelessly, they subreport the possession of that asset.

An important challenge for this specification is the possibility of omitted variable bias. Although every effort was made to control for variables mentioned in the literature that may affect misreporting, some variables were possibly correlated

with the variables of interest that are not present in the data. If such variables are correlated with the outcomes of interest, then they would be in the error term, e, and it would generate bias in the estimated impacts of the variables of interest. To check for the robustness of these results, the bounding approach proposed by Altonji et al. (2005) and refined by Oster (2017) was used.[8] Altonji et al. (2005) observed that a common approach to evaluating robustness to omitted variable bias is to include additional control variables on the right hand side of the regression. If such additions do not affect the coefficient of interest, then this coefficient can be considered as unlikely to be biased. This strategy implicitly assumes that the selection on observables is informative about the selection on unobservables. Oster points out that it is not only necessary to add controls, but also to observe the movements in the R-squared. Oster formalizes this idea, and provides conditions for bounds and identification. If the bounds exclude zero, then the results from the regressions can be considered to be robust to omitted variable bias (see Appendix A).

[8] A number of recent empirical studies use Oster's bounding methodology; see, for example, Nghiem et al. (2015), and Walther (2018).

5 Results

5.1 Principal Results

Table 3 columns 1 to 18 present the effects of careless responses, cognitive ability, and self-control on differences in the assets as reported by the spouses.[9] Age, literacy, speaking an indigenous language, and disability were controlled for, for both wives and husbands, as was cohabiting, living in a rural area, living in a municipality impacted by a natural disaster in 2010, and municipality fixed effects. The results show that careless responses given by husbands are an important predictor of discrepancies in the information of assets provided by the spouses. A statistically significant effect is found for TV (-0.478), music device (-0.285), automobile (-0.132), refrigerator (-0.296), bicycle (-0.233), work equipment (0.111), and farm animals (-0.293). Notice that when the coefficient is negative it implies that the husband underreports the item with respect to the wife (see equation 2). Finally, there is no evidence that cognitive ability or self-control explain the discrepancies in the information reported by the spouses. In particular, the coefficients associated with these variables are not significant at the 5 percent level.

Why do we observe effects of careless responses but not cognitive ability or self-control in the differences in the information reported? Cognitive ability involves the process of understanding the survey and remembering information. In particular, people failing to understand a survey where they are asked about the

[9] We do not present the coefficient associated with careless responses given by wives because we do not find evidence that women behave in this way (see Table 2).

possession of assets within the home seems unlikely. It also sounds unlikely that people may have trouble remembering the possession of an asset within the home. So, it makes sense not to find evidence of cognitive ability in the discrepancies of the information reported on assets. With regard to self-control, when participating in PROGRESA, impatient spouses were expected to have an incentive to under-report information so as to facilitate their continuation on the program. However, it is possible that this concern was partially mitigated by the way in which the questionnaire was designed and the information collected. The questionnaire used for the purposes of the current paper differed in many ways from the questionnaire used by the program. In particular, this questionnaire emphasizes questions related to psychological factors and the questionnaire used by PROGRESA places the emphasis on socioeconomic questions. In addition, the information was collected in the name of a private university, rather than in the name of the PROGRESA program. These factors possibly affected the impact of self-control on the differences in information reported by the spouses.

One potential concern with the effects of careless responses given by husbands on the differences in the assets reported is that these results can be biased as a consequence of omitted variables. To address the problem of omitted variables, a bounding methodology was implemented following Oster (2017). Oster proposes a method to test the robustness of results under the assumption that the relationship between the observables and the treatment is informative of the relationship between the unobservables and the treatment. This assumption allows to yield some

bounds for the coefficient of interest (see Section 4, Empirical Strategy). The results are presented in Table 4. To check the robustness using this methodology, we only include those results that were statistically significant at the 5 percent level. Column (1) presents a summary of the results from Table 3. Columns (2) and (3) display a solution for the coefficients that would have been obtained if we had assumed that the observables were at least as important as the unobservables ($\delta = 1$ and $\delta =$ -1) for the corresponding assumption on R_{max}. We observe the following bounds for television (-0.479, -0.477), music device (-0.301, -0.271), automobile (-0.135, -0.130), refrigerator (-0.320, -0.270), bicycle (-0.239, -0.227), and farm animals (-0.319, -0.270). To sum up, the effect of careless responses given by husbands emerged as robust to omitted variable bias when applying Oster's methodology.

5.2 Effects of Disagreements in Information on the Classification of Households in Poverty

The previous results confirm that discrepancies in the reported information on assets are not random, but do the differences in the information reported between spouses have consequences for poverty identification? To answer this question, the Simple Poverty Scorecard Poverty-Assessment Tool (Scorecard) was used. This is an index developed by Schreiner (2017) and used by Innovations for Poverty Action (IPA) to identify families in poverty. It has been developed for more than 45 countries. This measure was employed because it is transparent in the sense that

the variables used for its construction have been made public (Schreiner, 2017). The index uses 11 socioeconomic indicators[10] to estimate the likelihood of a household having a consumption below a given poverty line.

From the data, it is possible to recover six out of the socioeconomic indicators proposed by the Scorecard[11]. Regarding the other five items, the results were presented for two cases: (1) the value of zero (not having the item) was assigned for both the wife and the husband, raising the probability of the households to be classified as poor; and (2) a maximum value per item for both the wife and the husband was assigned, decreasing the probability of the households being classified as poor. Then, the effectiveness of this measure in targeting families living in poverty was measured.[12] Table 5, Panel A presents the results for case (1). The results show differences in the targeting depending on who answers the survey. For example, in 10.1% of the cases the household will be classified as non-poor if the husband is asked the questions, but will be classified as poor if the wife is surveyed. In a similar way, in 8.1% of the cases a household will be classified as poor if questions are posed to the husband, but as a non-poor if posed to the wife. Panel B presents the results for case (2). Here, it can be seen that 3.7% of households will be classified

[10]The items referenced are: number of household members aged 17 or younger, education level of male head of household, flooring material, availability of a kitchen sink for washing dishes, gas stove or microwave, piped water, washing machine, number of fans, car, computer, and mobile phone.

[11]For the purposes of the current study, the items referenced were household members aged 12 or younger, education of the male head/ spouse, flooring material, microwave, washing machine, and owning a car.

[12]The Scorecard ranges from 0 points (higher probability of being poor) to 100 points (lower probability of being poor). It was assumed that a family is considered poor when it has a score of 34 or less.

as non-poor if the questions are directed to the husband, but as poor if directed to the wife. In addition, in 4.0% of cases a household will be classified as poor if the husband is asked the questions, but as non-poor if the wife is asked the questions.[13]

6 Summary and Concluding Remarks

This paper analyzes the reliability of survey information regarding household assets. A unique data set is used, with the same questions posed to both husbands and wives participating in the Mexican social program, PROGRESA. The analysis revealed key discrepancies in the information reported between the spouses.

There are many reasons why misreporting may occur when collecting survey data: (1) individuals simply respond carelessly (careless responses); (2) individuals do not understand the question or have problems to recall information (cognitive ability); (3) impulsive individuals may have incentives to underreport particular information (lack of self-control); (4) individuals report information that they consider to be socially desirable; (5) the size of the survey can affect how individuals report the information (survey design); and (6) interviewers can affect the quality of information. In this paper, we test for careless responses, cognitive ability, and the lack of self-control. The results show that careless responses provided

[13]It should be noted that this was an adaptation of the Scorecard Poverty-Assessment Tool (Scorecard). The results presented here act as a point of reference to illustrate how the information reported by spouses may affect the classification of households with regard to poverty; this, however, does not reflect the accuracy of the Scorecard.

by husbands predict disagreement in the information reported on assets. Using a bounding methodology proposed by Oster (2017), this finding was found to be robust to a problem of omitted variable bias.

The key implications of this paper point to the need for practitioners to be careful when using data regarding household assets. For example, when using data to identify individuals living in poverty, policy makers recognize that individuals have incentives to underreport their income. As a consequence, they use variables such as the possession of assets to proxy households' real income, assuming that these variables are less susceptible to misreporting. Yet, this paper presents evidence that contradicts this assumption.

References

Altonji, J. G., T. E. Elder, and C. R. Taber (2005): "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 113, 151–184.

Ambler, K., C. Doss, C. Kieran, and S. Passarelli (2019): "He Says, She Says: Spousal Disagreement in Survey Measures of Bargaining Power," *Economic Development and Cultural Change*, 0, null.

Brown, J., W. Miller, and L. Lawendowski (1999): "The Self-Regulation Questionnaire," *Innovations in clinical practice: A source book*, 17, 281–292.

Doss, C. R., Z. Catanzarite, W. Baah-Boateng, H. Swaminathan, C. D. Deere, L. Boakye-Yiadom, and J. Suchitra (2018): "Do men and women estimate property values differently?" *World Development*, 107, 75 – 86.

Fisher, M., J. J. Reimer, and E. R. Carr (2010): "Who Should be Interviewed in Surveys of Household Income?" *World Development*, 38, 966 – 973.

Gottdredson, M. and T. Hirschi (1990): *A General Theory of Crime*, Stanford University Press.

Judge, G. and L. Schechter (2009): "Detecting Problems in Survey Data Using Benford's Law," *The Journal of Human Resources*, 44, 1–24.

Kilic, T. and T. P. Sohnesen (2017): "Same Question But Different Answer: Experimental Evidence on Questionnaire Design's Impact on Poverty Measured by Proxies," *Review of Income and Wealth.*

Leiner, D. J. (2019): "Too Fast, too Straight, too Weird: Non-Reactive Indicators for Meaningless Data in Internet Surveys," *Survey Research Methods*, 13, 229–248.

Martinelli, C. and S. W. Parker (2009): "Deception and Misreporting in a Social Program," *Journal of the European Economic Association*, 7, 886–908.

Meyer, B. D., N. Mittag, and R. M. Goerge (2018): "Errors in Survey Reporting and Imputation and their Effects on Estimates of Food Stamp Program Participation," *NBER Working Paper Series.*

Nghiem, H. S., H. T. Nguyen, R. Khanam, and L. B. Connelly (2015): "Does School Type Affect Cognitive and Non-cognitive Development in Children? Evidence from Australian Primary Schools," *Labour Economics*, 33, 55 – 65.

Onwujekwe, O., K. Hanson, and J. Fox-Rushby (2006): "Some indicators of socio-economic status may not be reliable and use of indices with these data could worsen equity," *Health Economics*, 15, 639–644.

Oster, E. (2017): "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business & Economic Statistics*, 0, 1–18.

Palomar, J. (2012): "Identificación de un Modelo de Resiliencia y su Relación con la Movilidad Social en Beneficiarios del Programa Oportunidades," *México, D.F. Universidad Iberoamericana*.

Philipson, T. and A. Malani (1999): "Measurement errors: A principal investigator-agent approach," *Journal of Econometrics*, 91, 273 – 298.

Schreiner, M. (2017): "Progress Out of Poverty Index Mexico 2014," *Mexico: IPA*.

Tourangeau, R., L. J. Rips, and K. Rasinski (2000): *The Psychology of Survey Response*, Cambridge University Press.

Walther, S. (2018): "Noncooperative decision making in the household: Evidence from Malawi," *Journal of Development Economics*, 134, 428 – 442.

7 Tables and Figures

Table 1: Assets' Descriptive Statistics

	Husband's report possession of (%): (1)	Wife's report possession of (%): (2)	P-Value (3)	Percentage that do not match (%) (4)	Husband: Yes Wife: No (5)	Husband: No Wife: Yes (6)
Music device	59.6	57.8	0.363	32.6	17.2	15.4
Bicycle	42.6	36.9	0.001***	30.1	18.0	12.1
Farm animals	30.0	30.3	0.848	25.8	12.8	13.0
Washing machine	42.4	43.3	0.596	24.0	11.6	12.4
Gas stove	19.9	22.0	0.220	22.8	10.4	12.4
Refrigerator	63.2	65.5	0.123	21.5	9.6	11.9
Sofa	23.7	23.5	0.941	19.3	9.7	9.6
Landline	15.8	17.1	0.259	14.0	6.3	7.7
Automobile	18.9	15.9	0.013**	13.5	8.2	5.3
Photographic camera	8.5	6.5	0.061*	10.9	6.4	4.5
Other land (apart from home)	8.1	6.0	0.056*	10.5	6.3	4.2
Television	91.2	90.7	0.578	8.5	4.5	4.0
Machinery or work equipment	5.6	3.2	0.006***	7.5	5.0	2.5
House, apartment or room to rent	4.4	3.6	0.325	7.0	3.9	3.1
Motorcycle	4.9	5.4	0.500	5.8	2.6	3.2
Savings	1.8	3.6	0.011**	4.8	1.5	3.3
Local business	2.6	2.9	0.612	3.7	1.7	2.0
Canoe or boat	1.9	1.7	0.513	2.2	1.3	0.9

Source: Survey of Resilience and Social Mobility (Progresa-Oportunidades Program).* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Summary Statistics

	N	Mean	SD	Min	Max
Number of assets with mismatch	957	2.62	1.98	0.00	12.00
Careless responses (husband): 1 Yes 0 No	957	0.04	0.20	0.00	1.00
Careless responses (wife): 1 Yes 0 No	957	0.00	0.00	0.00	0.00
Percentage of correct answers to the Raven test (husband)	957	0.36	0.24	0.00	1.00
Percentage of correct answers to the Raven test (wife)	957	0.34	0.24	0.00	1.00
StandardizedSelf-control (husband)	931	0.06	1.01	-3.82	1.26
Standardized Self-control (wife)	934	-0.06	0.99	-4.25	1.26
Age of husband	957	48.71	13.03	19.00	107.00
Age of wife	954	44.81	12.19	19.00	99.00
Can read and write (husband): 1 Yes 0 No	956	0.64	0.48	0.00	1.00
Can read and write (wife): 1 Yes 0 No	956	0.68	0.47	0.00	1.00
Speak some indigenous language (husband): 1 Yes 0 No	952	0.05	0.21	0.00	1.00
Speak some indigenous language (wife): 1 Yes 0 No	954	0.05	0.21	0.00	1.00
Motor, visual, or auditory disability (husband): 1 Yes 0 No	957	0.07	0.26	0.00	1.00
Motor, visual, or auditory disability (wife): 1 Yes 0 No	957	0.06	0.23	0.00	1.00
Free union: 1 Yes 0 No	954	0.21	0.41	0.00	1.00
Rural:1 Yes 0 No	957	0.48	0.50	0.00	1.00
Suffered a natural disaster in 2010: 1 Yes 0 No	957	0.15	0.36	0.00	1.00

Source: Survey of Resilience and Social Mobility (Progresa-Oportunidades Program)

Table 3: OLS Estimates: Effects of Careless, Cognitive, and Self-Control on the Differences in Assets

Dependent variable: Asset with mismatch	Television	Camera	Music device	Automobile	Sofa	Washing machine	Refrigerator	Gas stove	Landline	Bicycle	Motorcycle	Canoe or boat	Work equipment	Farm animals	Local business	Apartment to rent	Other land	Savings
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Careless responses (husband): 1 Yes 0 No	-0.478***	0.018	-0.285***	-0.132**	-0.109	-0.109	-0.296***	-0.023	-0.045	-0.233**	-0.009	-0.022	0.111*	-0.293***	-0.026	-0.042	-0.021	-0.038
	(0.097)	(0.024)	(0.088)	(0.054)	(0.070)	(0.072)	(0.103)	(0.052)	(0.064)	(0.088)	(0.023)	(0.015)	(0.064)	(0.087)	(0.027)	(0.058)	(0.068)	(0.039)
Husband's percentage of correct answers	-0.009	0.006	0.009	0.026*	-0.021	-0.031	-0.010	-0.016	-0.014	-0.033	-0.001	-0.010	-0.002	0.015	-0.008	-0.001	0.011	0.012
to the Raven test (standardized)	(0.012)	(0.014)	(0.020)	(0.015)	(0.018)	(0.023)	(0.014)	(0.022)	(0.018)	(0.023)	(0.010)	(0.007)	(0.012)	(0.021)	(0.008)	(0.009)	(0.013)	(0.010)
Wife's percentage of correct answers	-0.001	0.005	-0.006	-0.014	0.005	-0.015	0.033*	0.030	-0.016	0.021	0.003	0.001	0.008	0.014	0.006	0.000	0.011	-0.003
to the Raven test (standardized)	(0.011)	(0.014)	(0.022)	(0.013)	(0.019)	(0.016)	(0.018)	(0.020)	(0.017)	(0.025)	(0.009)	(0.006)	(0.008)	(0.018)	(0.009)	(0.009)	(0.015)	(0.009)
Husband's Self-control (standardized)	-0.007	-0.012	-0.005	0.006	-0.016	0.017	0.013	0.003	0.013	0.009	-0.000	0.002	-0.013	0.024	0.004	0.008	0.027*	0.017*
	(0.012)	(0.015)	(0.024)	(0.016)	(0.017)	(0.019)	(0.022)	(0.018)	(0.017)	(0.021)	(0.010)	(0.008)	(0.011)	(0.022)	(0.007)	(0.012)	(0.014)	(0.009)
Wife's Self-control (standardized)	0.006	-0.004	-0.014	-0.007	-0.010	-0.009	0.016	0.016	0.000	-0.015	-0.005	0.004	0.001	-0.035*	-0.000	0.003	0.023*	0.004
	(0.008)	(0.013)	(0.023)	(0.016)	(0.019)	(0.025)	(0.017)	(0.018)	(0.015)	(0.017)	(0.008)	(0.005)	(0.009)	(0.021)	(0.005)	(0.012)	(0.014)	(0.008)
				(0.087)														
Other controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.16	0.10	0.13	0.15	0.08	0.15	0.12	0.09	0.08	0.10	0.10	0.10	0.16	0.15	0.09	0.23	0.13	0.10
Observations	891	889	889	888	888	889	890	887	888	890	889	889	887	890	889	888	886	885

Other controls are age, ability to read and write, speaking an indigenous language and having a motor, visual or auditory disability for both wives and husbands. Also, living in free union, living in a rural area, and living in a municipality impacted by a natural disaster in 2010. Note: Clustered standard errors displayed in parenthesis at the municipality level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Bounding Methodology: Effects of Careless on the Differences in Assets

		$R_{max} = 1.3\tilde{R}$		Robust
		β for $\delta = 1$	β for $\delta = -1$	
Television	-0.478^{***}	[-0.479,	-0.477]	Yes
Music device	-0.285^{***}	[-0.301,	-0.271]	Yes
Automobile	-0.132^{**}	[-0.135,	-0.130]	Yes
Refrigerator	-0.296^{***}	[-0.320,	-0.270]	Yes
Bicycle	-0.233^{**}	[-0.239,	-0.227]	Yes
Farm animals	-0.293^{***}	[-0.319,	-0.270]	Yes

Note: Intervals in brackets are the bounds. Other controls include for the wife and the husband: age, can read and write, speaking an indigenous language, Raven test questions, self-control, and having a motor, visual or auditory disability . In addition: free union, living in a municipality suffering a natural disaster in 2010, rural, and municipality fixed effects.

Table 5: Classification of Poverty Based on Who Answer the Survey: Wife vs. Husband

Panel A

		Wife	
		Poor	Non-poor
Husband	Poor	9.7%	8.1%
	Non-poor	10.1%	72.1%

Panel B

		Wife	
		Poor	Non-poor
Husband	Poor	88.1%	4.0%
	Non-poor	3.7%	4.2%

The classification of poverty is estimated using the Simple Poverty Scorecard Poverty- Assessment Tool (Scorecard). The index uses 11 socioeconomic indicators. It is possible to recover 6 out of the socioeconomic indicators. Regarding the other indicators, Panel A assumes value of zero (not having the item) for the wife and for the husband and Panel B assumes the maximum value per item for the wife and for the husband.

8 Appendix A

Following the notation in Oster, the full model has the form:

$$\text{Y}= \beta \text{ T} + X_1 + X_2 + \epsilon.$$

where T is the variable of interest, X_1 contains the *observed* control variables multiplied by their coefficients, i.e., $X_1 = \sum_{j=1}^{J_o} X_j^o \gamma_j^o$, and X_2 contains all of the *unobserved* variables multiplied by their coefficients, i.e., $X_2 = \sum_{j=1}^{J_u} X_j^u \gamma_j^u$. Finally, ϵ is a random error that represents the measurement error in Y, and is uncorrelated with X_1, X_2 and T. Oster suggests the following approach to account for omitted variable bias:

(1) Regress Y on T and report the parameter on T, denoted by β^0, and the R-squared coefficient, denoted by R^0.

(2) Regress Y on T and X_1, and report the parameter on T, denoted by $\tilde{\beta}$, and the R-squared coefficient, denoted by $\tilde{R}$.

(3) Define R_{max} as the overall R-squared of the model, that is, the R-squared that would be obtained from a regression of Y on both observables (T, X_1) and unobservables (X_2).

(4) Define δ to be a parameter that ensures equality $\frac{Cov(T,X_2)}{Var(X_2)} = \delta \frac{Cov(T,X_1)}{Var(X_1)}$. In

other words, this relationship formalizes Altonji et al. (2005)'s idea that the magnitude and sign of the relationship between T and X_1 provides some information about the magnitude and sign of the relationship between T and X_2. For example, if $-1 \leq \delta \leq 1$, then the variable of interest (T) would be no more correlated with the unobservables (X_2) than it would be with observables (X_1). The case $0 \leq \delta \leq 1$ has a similar interpretation, with the additional assumption that the relationship between T and X_1 has the same sign as the relationship between T and X_2.

Oster shows that $\beta^* \approx \tilde{\beta} - \delta\frac{\left(\beta^0-\tilde{\beta}\right)\left(R_{max}-\tilde{R}\right)}{\left(\tilde{R}-R^0\right)}$ is a consistent estimator of the effect of T on Y, β. It should be noted that this is a close approximation to the consistent estimator and is used to present a particular intuition regarding the methodology. The complete approximation is presented in Oster (2017).

In order to estimate β^*, estimates of δ and R_{max} are required. Oster proposes assumptions for δ and R_{max} that allows one to determine whether β^* is different from zero. Oster proposes that $R_{max} = min\{1.3\tilde{R}, 1\}$, where the $\tilde{R}$ is as defined above. The cut-off value of 1.3 is derived from a sample of papers that have used randomized controlled trials and nonrandomized data, and were published in the *American Economic Review*, *Quarterly Journal of Economics*, *The Journal of Political Economy*, and *Econometrica* from 2008-2010. She determined that using this cut-off allowed 90% of the randomized and 50% of the nonrandomized results to continue being statistically significant.

After determining the value of R_{max}, Oster suggests that β^* be calculated for all the following ranges of δ: $0 \leq \delta \leq 1$. In addition, the current paper presents the results for δ: $-1 \leq \delta \leq 0$. This allows for the construction of the set: $[\tilde{\beta}, \beta^*]$. If this set excludes zero, the results from the controlled regressions can be considered to be robust to omitted variable bias. In other words, the results indicate that $\beta^* \neq 0$.