Closing the Gaps: The Role of Screening Questions and Self-Reporting in Measuring Women's and Youths' Employment and Work[1]

Ivette Contreras[2] Lelys Dinarte-Diaz[3] Amparo Palacios-Lopez[4]
Valentina Costa[5] Steffanny Romero[6]

April 2025

Abstract

Accurate measurement of labor market outcomes is essential for addressing employment disparities, yet standard survey methods often fail to fully capture the contributions of women and youth, particularly in informal settings. This study investigates whether alternative survey methods can reduce underreporting of women's and youths' labor market outcomes, thereby narrowing the measured gender- and age-based gaps. Using a survey experiment in El Salvador, we compare two innovative approaches—a list of activities survey module and enforced self-responses—to the standard labor module, which follows the guidelines of the 19th International Conference of Labor Statisticians (ICLS) and relies on proxy and self-responses. The findings reveal that the list of activities module increases reported paid and unpaid work, reducing measured gender gaps by 8.2 percentage points. Similarly, enforced self-responses reduce age-related gaps in male paid and unpaid work by 13.9 and 12.3 percentage points, respectively. These results highlight the potential of alternative survey techniques to improve the accuracy of labor market statistics and provide a clearer understanding of employment patterns among underrepresented groups.

JEL classification: O1, J2, C8
Keywords: labor gender gaps measurement, labor age gaps measurement, survey methods, informality.

[1] We are thankful to Damien de Walque, Kathleen Beegle, Aletheia Donald, Jason Kerwin, Devaki Ghose, Jessica Goldberg, Sylvan Herskowitz, Ardyn Nordstrom, Laura Schechter, Niharika Singh, Meiping Sun, and Diego Zardetto for their time and feedback. We also appreciate the comments we received from the participants of the different conferences and seminars where we presented this paper, including the World Bank Research Department and Data Group Seminars, 2022 International Methodology Symposium, European Survey Research Association Conference 2023, Northwestern Global Poverty Research Lab and IPA's Methods and Measurement Conference, ASSA 2024 Annual Meeting, and the Washington Area Labor Economics Symposium (WALES). We are grateful for the support of the International Fund for Agricultural Development and the World Bank Research Support Budget. The Institutional Review Board (IRB) of *Universidad Francisco Gavidia* in El Salvador reviewed and approved this research project's protocol, which is registered under ID #011-2021. This study trial is also included in the AEA RCT Registry under the unique ID #EARCTR-0009541. The findings, interpretations, and conclusions expressed in this report are entirely the authors and do not necessarily represent the views of the World Bank, its affiliated organizations, Executive Directors, or the governments that they represent.
[2] The World Bank. Email: icontreras@worldbank.org.
[3] Corresponding Author. The World Bank, IZA, CESifo, HiCN. Email: ldinartediaz@worldbank.org.
[4] The World Bank. Email: apalacioslopez@worldbank.org.
[5] The World Bank. Email: vcosta@worldbank.org.
[6] The World Bank. Email: sromeroesteban@worldbank.org.

1. Introduction

Accurate labor market indicators are critical for informed policymaking, particularly in low- and middle-income countries where employment disparities between women and men, as well as between youth and adults, are notably large. While these disparities are often attributed to structural factors—exacerbated by the COVID-19 pandemic—and persistent gender norms (Klasen, 2019; Goldin and Mitchell, 2017; Goldin et al., 2017), they may also stem from underlying data quality issues. In particular, the undermeasurement of women's and youths' work activities can distort official labor statistics.

Two key factors contribute to the undermeasurement of women's and youths' labor market outcomes in standard surveys. First, a growing body of evidence suggests that the standard household survey approach used in more than 106 countries is primarily designed to capture formal employment, following the guidelines of the 19th International Conference of Labor Statisticians (ICLS). Therefore, this survey design fails to accurately capture informal and casual work, resulting in an incomplete picture of the global labor market participation of women and youths (Ambler et al., 2021; Bardasi et al., 2011; Arthi et al., 2018; Dillon, 2012; Discenza et al., 2021). Examples of such activities include preparing food for sale or assisting in a family-owned business. Second, despite earlier literature recommended to prioritize self-reporting as best practice (Benes and Walsh, 2018; Blair et al., 2004), many standard survey protocols, including the 19th ICLS guidelines, rely on proxy respondents to provide labor information on behalf of other household members (Desiere and Costa, 2019).[7] As a result, labor indicators may be biased if proxies are unaware of certain activities carried out by family members—particularly in extended households—or if they have incentives to underreport some activities (Blair et al., 2004).

In this paper, we show that alternative household survey methods can help reduce misreporting of women's and young adults' labor market outcomes, thereby providing more accurate measures of underlying gender- and age-based employment gaps. To achieve this, we designed and conducted a household survey experiment in El Salvador to estimate the impact of two alternative household survey methods: a List of Activities (LOA) module administered

[7] A comprehensive review of Labor Force Surveys indicates that in half of these surveys, proxy respondents provide 30% to 50% of the responses (Desiere and Costa, 2019). This reliance on proxies persists primarily because interviewing every household member can be logistically challenging (Thomsen and Villund, 2011).

before the standard labor module and enforced self-responses (ESR) for respondents. We compared these methods against the standard household survey approach, which does not include the list of activities module and allows for proxy responses, to evaluate their effects on the reported employment and work outcomes.[8]

To estimate the causal effects of each survey method on reported employment and work, we randomly assigned 1,008 households—comprising 2,480 working-age members between 15 and 64 years old—to one of three groups, each with equal probability. The first group (LOA) completed the LOA module, which presented a list of work activities, before answering the standard labor module. This module aims to prompt participants to recall any income-generating activities they have undertaken in the past week. Importantly, participants were only asked if they had performed any of the listed activities; they were not explicitly told that these activities constituted work or employment. In this group, proxy responses were allowed.[9] The second group (ESR) completed the standard household labor module, with enforced self-reporting for all eligible respondents, and without the LOA module. The third group (Control) followed the standard household labor module protocol, which excluded the LOA module and permitted proxy responses.

Our experimental design aims to assess the (marginal) effects of altering specific survey features on the reporting of labor market outcomes. For this purpose, we compare the outcomes of participants in the LOA or ESR groups exclusively against those in the control group, rather than comparing them directly to one another. We do not compare LOA against ESR treatments because they differ in multiple features: LOA treatment includes the list of activities module and permits proxy responses, while ESR treatment does not include the list of activities module and enforces self-responses. By comparing each treatment group with the control group, we isolate the marginal effect of the list of activities module (LOA vs. control, both allowing proxy responses) and the marginal effect of enforcing self-responses (ESR vs. control, both using the standard household labor module). To ensure that any observed effects are attributable solely to

[8] As we explain later in the paper, *employment* is defined as engaging in at least one *activity for pay or profit* in the previous week, excluding individuals who worked only for their own use. And *work* is defined as engaging in activities that involved producing goods or providing services *either for their own use or for pay, sale, or profit.*

[9] Some authors have argued that a carefully prepared activity list is easily understood and allows capturing multiple job holdings as well as atypical employment (Langsten and Salen, 2008; Oya, 2013).

the tested survey methods, all other survey protocols—such as the sequence of modules, respondent incentives, and administration procedures—were held constant across all groups.[10]

Our results indicate that including the LOA module leads to a significant increase in respondents' average reported employment and work rates compared to the standard labor module. Specifically, respondents in the LOA group were 4.0 percentage points (pp) more likely to report being employed—equivalent to 6.9% of the average employment rate in the control group (C)—and 4.3 pp (6.3% of the average work rate in C) more likely to report working in unpaid activities that contribute to family income. In contrast, enforced self-reporting in the standard household labor module did not have a statistically significant effect on either employment or work outcomes for the average survey experiment participant.

Next, we analyze how the survey methods affect reported labor outcomes across different demographic groups, focusing on sex and age. First, we find that the LOA module reduces the reported work gap between women and men by 8.2 pp. This result is driven by women being 8.1 pp more likely to report working. Additionally, women responding to the LOA module are 6.8 pp more likely to report being employed compared to women in the control group. Second, the used of enforced self-reporting in the ESR method does not produce any significant difference in reporting between men and women. Third, neither of the alternative survey methods shows significant differential effects by age on reported employment or work. However, when examining age-related effects separately for male and female respondents, we find that enforcing self-responses reduces the employment gap and the work gap between younger and older male respondents in the ESR group by 13.9 pp and 12.3 pp, respectively.

In sum, alternative survey methods improve different labor market gaps: LOA improves gender-based gaps whereas ESR improves age-based gaps for men. Combining these estimated impacts with a cost analysis, we find that the cost of reducing each gender or age gap by one percentage point is comparable across methods, ranging from 1.5% to 2.4% increase in cost.

As further suggestive evidence, we find that requiring women and youth to report their own employment or work activities—rather than relying on proxy respondents—may more

[10] As we note later in the paper, we are unable to evaluate the potential complementarities between the LOA and ESR methods through this experiment. To that end, an arm that combines both LOA and enforced self-responses would be needed.

effectively address misreporting, especially in communities where informal employment among peers is common. In these contexts, individuals tend to have greater awareness of their informal activities, which are often overlooked by standard survey approaches, and direct self-reporting can better capture this information. We also observe that certain characteristics of proxy respondents influence the reported labor market indicators for women and men. For instance, when male spouses serve as proxies, the increase in reported work or employment for their female partners is similar to that reported by female spouses for their male partners. As a result, gender-based gaps in reported employment and work remain unaddressed because the average increase in reporting is equivalent for both sexes. In contrast, proxies with higher levels of education may unintentionally widen these gaps by applying stricter definitions of employment and underreporting informal or non-standard work. Finally, consistent with the findings of Dillon and Mensah (2024), we find that enforcing self-reporting improves the accuracy in identifying the economic sectors where women work, thereby ensuring that their diverse contributions, including informal and non-traditional roles, are accurately represented.

Our paper contributes to existing literature in several significant ways. First, we provide evidence on how alternative survey methods can help address underreporting of women's and youths' engagement in paid and unpaid work activities, and therefore the undermeasurement of their employment participation. Most previous studies focus on how women tend to underreport their labor outcomes when using standard survey methods and how alternative approaches can improve women's reporting relative to men's (Ambler et al., 2021; Arthi et al., 2018; Bardasi et al., 2011; Dillon and Mensah, 2024; Dillon et al., 2012; Discenza et al., 2021; Franck and Olsson, 2014; Kilic et al., 2021; Kilic et al., 2023; Muller and Sousa, 2020). However, no research has yet explored how alternative household survey methods affect reporting among other groups, such as youth, who are often engaged in informal and seasonal activities that standard labor modules fail to capture.[11] In this context, our paper is the first to rigorously demonstrate that an alternative survey method (enforcing self-responses) can improve the reporting of employment and work among young males compared to older male adults. Our findings provide valuable insights for

[11] To our knowledge, only one study shows how response fatigue can have differential effects on the reporting of productive activities by gender and age (Ambler et al., 2021). The authors, however, do not study the impacts of any specific survey method on the reporting of labor outcomes.

enhancing data collection practices to ensure that women's and youths' labor contributions are accurately documented and statistically represented.

Second, while existing evidence shows that gender norms are a key driver of gender wage gaps (World Bank, 2011; Borrowman and Klasen, 2020), no studies have demonstrated that the prevalence of informal work in a respondent's environment can influence their reporting of labor indicators. In this regard, our paper makes a novel and noteworthy contribution by providing evidence on how exposure to informality within the community, combined with the use of alternative survey methods, produces heterogeneous effects on the labor market reporting of women and youth.

Third, this paper contributes with new evidence on the characteristics of proxy respondents and their influence on labor market reporting (Bardasi et al., 2011; Blair et al., 2024; Dillon et al., 2012; Kilic et al., 2021; Klasen, 2019). Researchers or governments interested in collecting data may assume that proxies who are either a spouse or have a high level of education will provide more accurate information about other household members' employment or work. However, our findings suggest the opposite. Selecting proxy respondents with these characteristics does not improve the accuracy of reporting and may perpetuate or even widen gender-based employment and work gaps.

Employment plays a critical role in driving economic development, fostering social inclusion and improving well-being (ILO, 2022). Without precise measurements of employment and labor market participation, policymakers are left to design interventions and allocate resources based on incomplete or misleading information. This paper demonstrates that alternative, cost-effective survey methods can substantially enhance the accuracy of employment and work statistics, offering a more reliable foundation for informed policy decisions, better targeted development programs, and progress toward global equity and prosperity.

2. Study Design

2.1 Randomized Methodological Survey Experiment

We conducted the survey experiment in four stages that we describe below.

Stage 1: Identification of the most common work activities. Between June and July 2022, we facilitated eight focus group discussions among rural and peri-urban residents in six municipalities across two regions of El Salvador. These discussions aimed to identify the most common activities

performed by community members. Participants reported a wide range of activities, which were categorized as either income-generating or non-income-generating and classified by the group typically performing them—women, men, or youth. Additionally, the discussions explored whether these activities were paid or unpaid, conducted inside or outside the home, and their typical duration.

Stage 2: Design of the List of Activities (LOA) module. Based on the focus group findings, we designed the List of Activities module to include the most frequently performed work activities in each region (see Table A1).[12] The purpose of this module was not to collect data for measuring the study's primary outcomes but to provide respondents with examples of income-generating activities that are statistically classified as work. This approach aimed to help respondents more accurately report their work in the standard labor module.

Stage 3: Selection of regions, households, and respondents. We selected two regions in the departments of San Salvador and Usulután to ensure significant variation in reported activities. Region 1, located in San Salvador, includes households near the Metropolitan Area of San Salvador, where commerce, construction, manufacturing, and sugar cane production dominate. Region 2, in Usulután, is characterized by coffee production, fishing, and commerce. Using the 2007 Salvadoran Population Census—the most recent census available—we identified 276 enumeration areas (EAs): 114 in region 1 and 162 in region 2. Each EA had at least 30 households and was classified as rural or peri-urban. We randomly selected 48 EAs and conducted a complete household listing for each. From these, we randomly selected 21 households per EA, yielding a total sample of 1,008 households. Within each household, we randomly chose between two and four working-age members (ages 15–64), resulting in a final sample of 2,480 individuals.

Because our primary hypothesis focuses on differential impacts by sex and age, we stratified the selection of household members by sex (male or female) and age (youths aged 15–24 and adults aged 25–64).

Stage 4: Random assignment of households to experimental groups. The 1,008 households were randomly assigned with equal probability to one of the following three groups (Figure A1):

1. *LOA Module group (LOA).* Households randomly assigned to this experimental group participated in the LOA module *before* responding to the standard labor module. This

[12] A potential concern with the selection of activities is that they might drive the results by gender or age. To avoid this bias, we made sure that the list included a similar number of activities performed by men, women, and youth.

module aims to prompt participants to recall any income-generating activities they have undertaken in the past week. Importantly, participants were only asked in the LOA if they had performed any of the listed activities; they were not explicitly told that these activities constituted work or employment. Proxy respondents were allowed to provide information on behalf of other household members unavailable for the interview.

2. *Enforced self-reporting group (ESR)*. Adults in the households randomly assigned to this experimental group completed the standard labor module without prior exposure to the LOA module. Self-reporting was strictly enforced, with interviewers making up to five attempts to speak directly with the selected household member.
3. *Control group (C)*. Households in this group completed the standard labor module without the LOA module. Proxy respondents were allowed in this group.

Two important clarifications are in order. First, we argue that traditional survey methods, as represented by the control group, may not accurately capture responses from women and youth. We hypothesize that this issue can be addressed by modifying certain survey features. The experimental design allows us to test whether *marginal changes* in survey methods can help reduce gender- and age-based discrepancies in reporting. Specifically, we examine the causal marginal effects of these adjustments by making two key comparisons: the LOA group versus C group and the ESR group versus the C group. Comparing the LOA group to C isolates the impact of including the LOA module while allowing proxy respondents, whereas comparing ESR to C estimates the effect of enforced self-reporting without the LOA module. Second, we do not directly compare LOA and ESR because these groups differ in more than one survey element. LOA allows proxy respondents and includes a list of activities, while ESR does not include the list and requires self-reporting.[13]

To prevent contamination, all individuals within a household were assigned the same treatment status. Furthermore, all respondents completed the same set of additional survey

[13] Including an experimental group that combines both enforced self-responses and the LOA module would have allowed us to explore the potential complementarities between these two survey features and their effects on reported labor market outcomes. However, due to budget constraints, we prioritized examining the *marginal effects* of each feature independently, which aligns with the practical needs of researchers and policymakers, who are more likely to adopt smaller, cost-effective adjustments. However, studying the combined impact of these survey features is a valuable question that warrants further exploration.

modules on demographic and socioeconomic characteristics. This consistency ensures that observed effects are solely attributable to the experimental variations in survey methods.

3. Data

3.1 Data Collection Activities and Survey Instrument

We collected data for this experiment from 1,008 households (comprising 2,480 household members) between August and October 2022. Interviews were conducted using a Computer-Assisted Personal Interviewing (CAPI) program, which embedded household assignments to treatment groups. The survey instrument consisted of up to 18 modules (see Table A2). The first module gathered household roster information to identify all household members. Once the roster was complete, the CAPI program randomly selected working-age members (aged 15 to 64) to participate in the experiment, stratifying them by age and sex. Each randomly selected individual (or the proxy in the LOA or control group, when the selected household member was not available) provided responses to three main modules: education, access to technologies, and labor. Only respondents assigned to the LOA group completed it before the labor module.[14]

3.2 Outcomes

We measure the effects of alternative survey methods on two main outcomes: employment and work. Data for these variables were collected from all participants through the standard labor module,[15] following the definitions of the 19th International Conference of Labor Statisticians.

Employment: This variable is a binary indicator equal to 1 if the individual engaged in at least one activity for pay or profit in the previous week (Durazo et al., 2021). It excludes individuals who worked only for their own use.

Work: This variable is a binary indicator equal to 1 if the individual reported engaging in at least one activity during the previous week that involved producing goods or providing services—either for their own use (e.g., subsistence farming) or for pay, sale, or profit.

3.3 Summary Statistics

Table 1 presents summary statistics for the variables used in our analysis.[16] Panel A shows that the average household in the sample consists of 3.7 members, closely matching the national

[14] See Appendix A for more details on other methodological activities, including enumerator training, replacement protocol, fieldwork, and data quality assurance.

[15] The standard labor module used for this study can be accessed through this link. The data collected using the LOA module was not used to estimate any of the outcomes presented in this paper.

[16] Appendix B presents a description of the variables included in Table 1.

average household size of 3.3 members (EHPM, 2022). On average, for every working-age adult, there are 0.27 dependents (children under 14 or elderly adults aged 65 and older). Half of the sample households are in rural areas, and about one-third report moderate to severe food insecurity. Regarding income sources, approximately one-third of household members are engaged in non-farming activities, while 54% are involved in agriculture, including crop production, livestock raising, and fishing. Panel B provides individual-level characteristics. The average participant is 35 years old, with 55% identifying as women and 34% reporting that they have never married. In terms of human capital, 28% have completed high school or higher education, and 87% can read and write. Most participants (88%) have access to at least one mobile phone, but only 19% have internet access at home.

To validate our identification strategy, we compare average household and individual characteristics across the three study groups. These results are shown in Table 1, Columns (5) through (7). We find a few small differences in the means between the groups. Specifically, households in the LOA group are less likely to engage in agriculture (49% in LOA versus 56% and 57% in the control and ESR groups, respectively; p = 0.066 and p = 0.035). ESR respondents are, on average, 1.4 years older than those in LOA and control groups (relative to an average age of 35; p = 0.014 and p = 0.018) and are 3 to 4 percentage points (out of 88%) less likely to read and write (p = 0.066 and p = 0.023). Lastly, individuals assigned to LOA are 3 percentage points (out of 89%) less likely to have a mobile phone (p = 0.097). To address these small differences, we control for these variables as a robustness check, as we discuss in Section 6.

Since households were randomized within EAs, a potential concern is intra-cluster contamination. To minimize this risk, we conducted all surveys simultaneously within each EA. Additionally, intra-cluster contamination is more likely when multiple survey rounds are conducted within small clusters. In our study, the EAs were relatively large, and data collection occurred in a single round, further mitigating this concern.

4 Empirical Strategy

The random assignment of households to the LOA, ESR, and C groups in the survey experiment provides exogenous variation in treatment status, allowing us to estimate the effects of the LOA module and enforced self-reporting on reported labor market outcomes. Specifically, we use the following specification:

$$Y_{ihs} = \alpha_0 + \beta_1 LOA_{hs} + \beta_2 ESR_{hs} + \Theta X_{ihs} + b_s + \epsilon_{ihs} \qquad (1)$$

where Y_{ihs} is the outcome of interest (employment or work) for individual i, in household h, in stratum (EA) s. LOA_{hs} is a binary indicator denoting that household h in stratum s was randomly assigned to respond the LOA module prior to completing the labor module. ESR_{hs} is a binary indicator that household h's members were required to complete the labor module themselves, rather than through proxy respondents. X_{ihs} is a vector of control variables, selected using a double-LASSO procedure (Urminsky et al., 2016) to identify the variables most relevant for each outcome. The control variables include the individual-level characteristics presented in Table 1.[17] Stratum fixed effects b_s are captured through the interaction of department (San Salvador or Usulután) and area of residence (rural or peri-urban). Lastly, ϵ_{is} is an individual error term. Standard errors are clustered at the household level.

The estimate of $\widehat{\beta_1}$ represents the intention-to-treat (ITT) effect of being exposed to the LOA module before completing the standard labor module (with proxy responses allowed). Similarly, $\widehat{\beta_2}$ captures the ITT effect of enforced self-reporting on the standard labor module, compared to the control condition that permits proxy responses.

Our main hypothesis posits that alternative survey methods can reduce gender- and age-based discrepancies in reported outcomes. To test this, we estimate heterogeneous treatment effects by sex and age. Specifically, we extend Equation (1) by interacting each treatment indicator Tj_{hs} with a binary variable D_{ihs}, indicating whether individual i is female or youth. Separate models are estimated for sex and age, as follows:

$$Y_{ihs} = \alpha_0 + \beta_1 LOA_{hs} + \beta_2 ESR_{hs} + \beta_3 LOA \times D_{ihs} + \beta_4 ESR \times D_{ihs} + \Theta X_{ihs} + b_s + \epsilon_{ihs} \quad (2)$$

where X_{ihs} includes D_{ihs} and the set of control variables identified through double-LASSO. As before, all estimations include stratum fixed effects. In this model, $\widehat{\beta_3}$ ($\widehat{\beta_4}$) captures the impact of the LOA module (enforced self-reporting) on the gender-based or age-based gaps in reported labor market outcomes.

5 Results

5.1. Impacts of Survey Methods on Reporting Labor Market Outcomes

[17] Table A3 presents the list of control variables that LASSO selected for each outcome and model.

Table 2, Columns (1) and (2) shows the effects of the LOA module and enforced self-reporting on employment and work, respectively, using specification (1). Our results indicate that the LOA module significantly reduces underreporting of labor market activity. Specifically, respondents who completed the LOA module before the standard labor module were 4.0 percentage points (pp) more likely to report being employed compared to respondents in the control group (C). This effect, significant at the 10% level, corresponds to 6.9% of the average employment rate in C. Similarly, respondents exposed to the LOA module were 4.3 pp more likely to report working (p<0.05), an increase equal to 6.3% of the average work rate in C. In contrast, enforced self-reporting had no statistically significant effect on reported employment or work outcomes relative to proxy reporting in the control group.

5.2. Heterogeneous Effects of Survey Methods on Reporting Labor Market Outcomes

Figure 1 and Table A4 present the results from Specification (2). Panel A in Figure 1, and Columns (3) and (4) in Table A4, show the heterogeneous impacts of each survey method on employment and working status by sex, while Panel B in Figure 1, and Columns (3) and (4) in Table A4, present the results by age (with youths defined as individuals aged 15–24 years).

Results indicate that the LOA module is particularly effective in reducing underreporting of labor market outcomes for women. Exposure to the LOA module decreases the work gap between women and men by 8.2 pp (p < 0.05). Although the LOA's effect on the employment gap is not statistically significant at conventional levels (p = 0.175), its magnitude—6.3 pp—is sizeable. Moreover, the LOA module increases women's likelihood of reporting employment and work by 6.8 pp and 8.1 pp, respectively (p < 0.05), compared to women in C. In contrast, enforced self-reporting does not show statistically significant differential effects by sex in reporting labor market outcomes, which is consistent with findings from Bardasi et al. (2011).

Turning to age-related heterogeneity, the results in Figure 1 (and Table A4, Columns (5) and (6)) show no significant differential impacts of survey methods on employment and work between adults and youths. The only notable exception is a 4.3 pp increase (p < 0.1) in the likelihood of older adults reporting work when exposed to the LOA module, compared to other older adults not exposed to the list. As discussed below, this effect is likely driven by older women.

While the estimated coefficients for the differential impacts of enforced self-reporting on employment and work between youths and adults are relatively large (6 to 7 pp), they are not

statistically significant. To better understand these results, we further disaggregate by sex. Figure 2 and Table A5 present separate age-related heterogeneity results for women and men.

Among men, enforced self-reporting reduces the employment and work gaps between young and old respondents by 13.9 pp (p<0.10) and 12.3 pp (p<0.10), respectively. These reductions appear to reflect adjustments in the reporting behavior of both younger and older male respondents. Specifically, older men in the ESR group are less likely to report labor market activity compared to older men in the control group. At the same time, although the estimated coefficients are less precise, younger male respondents appear more likely to report employment and work—by 8.9 pp and 7.3 pp, respectively—compared to their counterparts in the control group. In this way, enforced self-reporting helps improve the reporting of labor market outcomes among both younger and older male respondents, ultimately reducing age-based gaps in these indicators.

6 Robustness Checks

Representativeness of the study sample. We compare the average characteristics of our respondents with those of Salvadoran respondents from the 2022 EHPM living in San Salvador and Usulután. Table A6 shows that the two samples are similar in terms of age, sex composition, human capital, and marital status, confirming that the households participating in our experiment are representative of the broader population.

Assessing sensitivity from selection of control variables. As shown in Tables A7 through A9, our main results remain stable after excluding all LASSO-selected control variables. Similarly, Tables A10 through A12 demonstrate that the results are robust to including variables with significant differences across treatment groups as additional controls in the main specifications.

Randomization inference. To further validate our findings, we estimated p-values using randomization inference. By comparing each treatment effect to the distribution of all possible effects under 1,000 randomizations with no effects, we assign a standard error that reflects the likelihood of observing our results under the null hypothesis (Gerber and Green, 2012; Heß, 2017). These randomization inference p-values, presented below the standard error estimates in Tables 2, A4 and A5, yield nearly identical inferences to those obtained using clustered standard errors.

Addressing differences in household characteristics due to substitution. Some baseline variables, including household size, location, and access to mobile phones or the internet, differ between original and substituted households. These differences are detailed in Appendix A. To ensure the

robustness of our findings, we add these variables as controls—beyond those selected by LASSO—and show in Tables A13 through A15 that the main results remain unchanged.

7 Additional Results

7.1 Prevalence of Local Informality

Women and young men are more likely to engage in informal work in developing countries (ILO 2018). For example, in low- and middle-income countries, 92% of women are engaged in informal employment compared to 87.5% of men. In El Salvador, the gap is wider: 72% of women are engaged in informal work relative to only 56% of men (UN Women, 2023). Standard labor surveys often fail to capture informal work, as respondents may tend to associate "income-generating activities" with formal employment. In contexts with high informal employment, proxies may further bias these measures by overlooking informal activities as legitimate work.

To investigate this, we estimate the prevalence of peers' informal employment at the community level, defined as the proportion of informal employment in each EA, excluding the respondent's status. For clarity, we standardize this measure: one standard deviation (SD) corresponds to 12–15 percentage points of informal employment among women or men, respectively. We then interact this prevalence measure with each treatment indicator, estimating separate results by sex and age.

Table A16 presents the results for women. Exposure to one SD of female peers' informal employment is associated with increases on work reported by women assigned to ESR by 5.2 pp. Similarly, one SD of male peers' informal employment is associated with increases in young men's likelihood of reporting employment and work by 14.2 pp and 10 pp, respectively. By contrast, community informality has no differential effect on the LOA module's impacts. These findings suggest that requiring women and youth to report their own employment—rather than relying on proxies—may be particularly important in communities with high informal employment. In such settings, proxies may fail to recognize informal economic activities as work, leading to underreporting.

7.2 Associations Between Employment, Work, and Proxy Respondents' Characteristics

To explore how proxies' characteristics are associated with the reporting of employment and work of other household members', we limit the sample to groups allowing proxies—those assigned to the LOA module or the control group. Since proxies were not randomly assigned,

these results are best interpreted as associations. We focus on three characteristics: the proxy's sex, education level (high school or higher), and whether the proxy is the respondent's spouse. An indicator for LOA treatment is included to account for exposure to this survey method.

As shown in Table A17, proxies with higher levels of education are more likely to underreport women's employment and overreport men's employment and work. This pattern may reflect a stronger focus on formal activities; more educated proxies might not recognize informal labor—activities that women are more likely to engage in—as employment or work. Meanwhile, proxies who are spouses tend to increase reporting of their partners' employment and work. These effects do not vary by the gender of the partner for whom the proxy is reporting. Consequently, the gender gap in reporting remains unchanged. Although no statistically significant associations are found regarding the sex of the proxy respondent, the magnitude of the effects suggests that women are less likely to report employment or work for other women than for men. In summary, selecting a spouse as the proxy may improve reporting for both men and women, but it does not address the underlying gender-based mismeasurement. In contrast, relying on more educated proxies could exacerbate gender-based reporting gaps.

7.3 Assessing the Effects of Reporting Employment Across Economic Sectors

Finally, we examine whether survey methods influence how employment is reported across economic sectors. Using agriculture as the reference sector, we compare employment reporting in agriculture to both no work (extensive margin) and other sectors (intensive margin). Table A18 shows that, in line with Dillon and Mensah (2024), female respondents in the ESR group are more likely to report employment in agriculture than in other sectors, such as services or manufacturing. This pattern may be due to the perception of agricultural work as informal or irregular, leading proxies to underreport such activities when describing women's employment. Enforcing self-reporting thus improves the accuracy of sector-specific employment data, particularly for women.

8 Cost Implications of the Survey Design

Using cost data from this randomized survey experiment, we estimate the cost implications for each treatment arm. Households assigned to the LOA group experienced slightly longer interviews due to the additional LOA module, which took participants an average of 3.1 minutes to complete. With an average of 2.5 members per LOA household, this resulted in an additional

7.8 minutes per household. Given an average cost of US$135.4 per interview and an average interview duration of 111.6 minutes, the additional time required for the LOA module increased costs by approximately US$9.44 per interview (US$1.21 per minute x 7.8 minutes). Additionally, implementing the LOA module involved conducting and analyzing focus group discussions, which totaled about US$3,000, or US$8.93 per interview when averaged across all LOA interviews. Combined, these costs represented roughly 13.5% of the total per-interview cost for the control group (C). This implies that using the LOA to reduce the gender-based employment or work gap by one percentage point costs between 1.5% and 2.2% more, respectively, compared to the standard module.

ESR method was nominally more expensive than using proxy respondents. The additional costs for ESR were primarily driven by the need for more visits to ensure all respondents were located. On average, households in the C and LOA groups required 1.50 visits to complete interviews, while households in the ESR group needed 1.94 visits. With each visit costing an average of US$90.27, the per-interview cost for ESR households was 29.6% higher than for control households. This translates to an increase in cost of between 2.11% and 2.4% for each one-percentage-point reduction in the age-based employment or work gap. Overall, these costing estimations illustrate that both survey methods can improve the measurement of gender- and age-based labor market outcomes gaps at a similar cost.

9 Conclusion

The need for labor market policies that address inequalities in job access and quality is more pressing than ever, particularly during global crises (UN, 2023). A fundamental challenge in formulating such policies is that measurements of female and youth labor market outcomes often differ from those for men. This paper argues that a portion of the observed gaps in employment and work between women and men, as well as between younger and older adults, may because traditional survey methods used to collect labor market information may lead to underreporting by women and youths.

Our study contributes to the literature on survey sensitivity by analyzing data from a randomized survey experiment conducted in urban and peri-urban areas of El Salvador. This experimental design allowed us to compare employment and work estimates derived from a standard household survey labor module with those obtained using two alternative approaches: the LOA module and enforced self-reporting.

The results of this study provide valuable insights for researchers, survey practitioners, and policymakers tasked with producing national statistics. Specifically, the findings highlight how survey methods can influence labor market bias depending on respondents' sex and age. Additionally, the differing budget implications of the LOA module and enforced self-reporting illustrate the trade-offs between improving data quality and the additional costs involved. In this experiment, enforced self-reporting significantly improved labor indicators for young males, which are often underreported in traditional household surveys that rely on proxies.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT to review the English grammar of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

Abraham, R., & Shrivastava, A. (2022). How Comparable are India's labour market surveys?. *The Indian Journal of Labour Economics*, 65(2), 321-346.

Ambler, K., Herskowitz, S., & Maredia, M. K. (2021). Are we done yet? Response fatigue and rural livelihoods. *Journal of Development Economics*, *153*, 102736.

Arthi, V., Beegle, K., De Weerdt, J., & Palacios-López, A. (2018). Not your average job: Measuring farm labor in Tanzania. *Journal of Development Economics*, *130*, 160-172.

Bardasi, E., Beegle, K., Dillon, A., & Serneels, P. (2011). Do labor statistics depend on how and to whom the questions are asked? Results from a survey experiment in Tanzania. *The World Bank Economic Review*, *25*(3), 418-447.

Benes, E., & Walsh, K. (2018). ILO LFS Pilot Studies experimental field tests: methodology, process and outcomes. *ILO*.

Blair, J., Menon, G., & Bickart, B. (2004). Measurement effects in self vs. proxy response to survey questions: an information-processing perspective. *Measurement errors in surveys*, 145-166.

Bruhn, M., & McKenzie, D. (2009). In pursuit of balance: Randomization in practice in development field experiments. *American economic journal: applied economics*, *1*(4), 200-232.

Borrowman, M., & Klasen, S. (2020). Drivers of gendered sectoral and occupational segregation in developing countries. Feminist Economics, 26(2), 62-94.

Deshmukh, N., Desai, S., Pramanik, S., & Tiwari, D. (2020). Improving measurement of women's work participation: Sensitivity of work participation rates to questionnaire design. *NCAER National Data Innovation Centre Measurement Brief No 2020-1*.

Desiere, S., & Costa, V. (2019). Employment Data in Household Surveys: Taking Stock, Looking Ahead. *Looking Ahead (June 11, 2019). World Bank Policy Research Working Paper*, (8882).

Dillon, A., Bardasi, E., Beegle, K., & Serneels, P. (2012). Explaining variation in child labor statistics. *Journal of Development Economics*, *98*(1), 136-147.

Dillon, A., & Mensah, E. (2024). Respondent biases in agricultural household surveys. *Journal of Development Economics*, *166*, 103198.

Discenza, A., Gaddis, I., Palacios-Lopez, A., Walsh, K. (2021). Measuring Women and Men's Work: Main Findings from a Joint ILO and World Bank Study in Sri Lanka. Washington DC: World Bank.

Durazo, J., Costa, V., Palacios-Lopez, A., & Gaddis, I. (2021). *Employment and Own-Use Production in Household Surveys: A Practical Guide for Measuring Labor (English)*. LSMS Guidebook Washington, D.C.: World Bank Group.

Franck, A. K., & Olsson, J. (2014). Missing women? The under-recording and under-reporting of women's work in Malaysia. *International Labour Review*, *153*(2), 209-221.

Gerber, A. S., & Green, D. P. (2012). Field experiments: Design, analysis, and interpretation. WW Norton.

Goldin, C., Kerr, S. P., Olivetti, C., & Barth, E. (2017). The Expanding Gender Earnings Gap: Evidence from the LEHD-2000 Census. American Economic Review, 107(5), 110-114.

Goldin, C., & Mitchell, J. (2017). The New Life Cycle of Women's Employment: Disappearing Humps, Sagging Middles, Expanding Tops. *Journal of Economic Perspectives*, 31(1), 161-182.

Heß, S. (2017). Randomization inference with Stata: A guide and software. *The Stata Journal*, 17(3), 630-651.

ILO (2018). Women and men in the informal economy: a statistical picture (third edition). Geneva: International Labor Office.

ILO (2022). World Employment and Social Outlook. Trends 2022. Geneva: International Labour Office.

Kilic, T., Van den Broeck, G., Koolwal, G., & Moylan, H. (2023). Are You Being Asked? Impacts of Respondent Selection on Measuring Employment in Malawi. *Journal of African Economies*, 32(5), 495-522.

Kilic, T., Moylan, H., & Koolwal, G. (2021). Getting the (Gender-Disaggregated) lay of the land: Impact of survey respondent selection on measuring land ownership and rights. *World Development*, 146, 105545.

Klasen, S. (2019). What Explains Uneven Female Labor Force Participation Levels and Trends in Developing Countries?. *The World Bank Research Observer*, 34(2), 161-197.

Koolwal, G. B. (2021). Improving the measurement of rural women's employment: Global momentum and survey priorities. *World Development*, 147, 105627.

Langsten, R., & Salen, R. (2008). Two approaches to measuring women's work in developing countries: A comparison of survey data from Egypt. *Population and Development Review*, *34*(2), 283-305.

Mehta, B. S., Awasthi, I. C., Mehta, B. S., & Awasthi, I. C. (2019). Measurement of Women's Work. *Women and Labour Market Dynamics: New Insights and Evidences*, 33-54.

Muller, M., & Sousa, L. D. (2020). *"She Helps Me All the Time": Underestimating Women's Economic Engagement in Rural Honduras*. The World Bank.

ONEC (2012). Encuesta de Hogares de Propósitos Múltiples - Preliminar 2022 [Data set]. October 2023 - Oficina Nacional de Estadística y Censos. https://onec.bcr.gob.sv/Repositorio_archivos/

Oya, C. (2013). Rural Wage Employment in Africa: Methodological Issues and Emerging Evidence. *Review of African Political Economy*, *40*(136), 251-273.

UN Women (2023). Empoderamiento económico de las mujeres. URL: https://elsalvador.unwomen.org/es/nuestro-trabajo/empoderamiento-economico Accessed on October 25, 2023.

Urminsky, O., Hansen, C., & Chernozhukov, V. (2016). Using double-lasso regression for principled variable selection. *Available at SSRN 2733374*.

World Bank (2011). *World Development Report 2012: Gender Equality and Development.* Washington, DC: World Bank.

World Bank (2023). Gender Data Portal. URL: https://genderdata.worldbank.org/indicators/sl-uem-zs/?gender=total. Accessed on October 25, 2023.

Tables and Figures

Table 1. Descriptive Statistics and Balance Tests

Variable	(1) Mean All	(2) Mean C	(3) Mean LOA	(4) Mean ESR	(5) *P*-value LOA vs. C	(6) *P*-value ESR vs. C	(7) *P*-value ESR vs. LOA
PANEL A. Household Characteristics							
Household size (N)	3.71	3.723	3.762	3.643	(0.748)	(0.519)	(0.256)
Dependency ratio (%)	0.27	0.268	0.253	0.274	(0.274)	(0.665)	(0.212)
Households in rural area (%)	0.50	0.500	0.500	0.500	--	--	--
Households with assets index above the median (%)	0.47	0.461	0.491	0.461	(0.476)	(1.000)	(0.449)
Moderate to severe food insecurity (%)	0.36	0.372	0.354	0.366	(0.637)	(0.883)	(0.722)
Remittances received (%)	0.35	0.324	0.369	0.366	(0.213)	(0.201)	(0.934)
Household has a non-farm enterprise (%)	0.33	0.333	0.327	0.315	(0.856)	(0.625)	(0.753)
Household engages in agricultural activities (%)	0.54	0.560	0.491	0.574	(0.066)	(0.683)	(0.035)
Observations	1008	336	336	336			
PANEL B. Individual Characteristics							
Female (%)	0.55	0.549	0.539	0.559	(0.494)	(0.512)	(0.195)
Age (years)	35.49	35.006	35.032	36.479	(0.990)	(0.014)	(0.018)
High school or higher education (%)	0.28	0.282	0.296	0.275	(0.610)	(0.705)	(0.400)
Read and write (%)	0.87	0.878	0.888	0.847	(0.638)	(0.066)	(0.023)
Never married (%)	0.34	0.333	0.351	0.320	(0.433)	(0.636)	(0.241)
Access to mobile phone (%)	0.88	0.892	0.864	0.888	(0.097)	(0.652)	(0.180)
Access to Internet (Wi-Fi) (%)	0.19	0.164	0.193	0.198	(0.241)	(0.110)	(0.721)
Observations	2480	830	854	796			

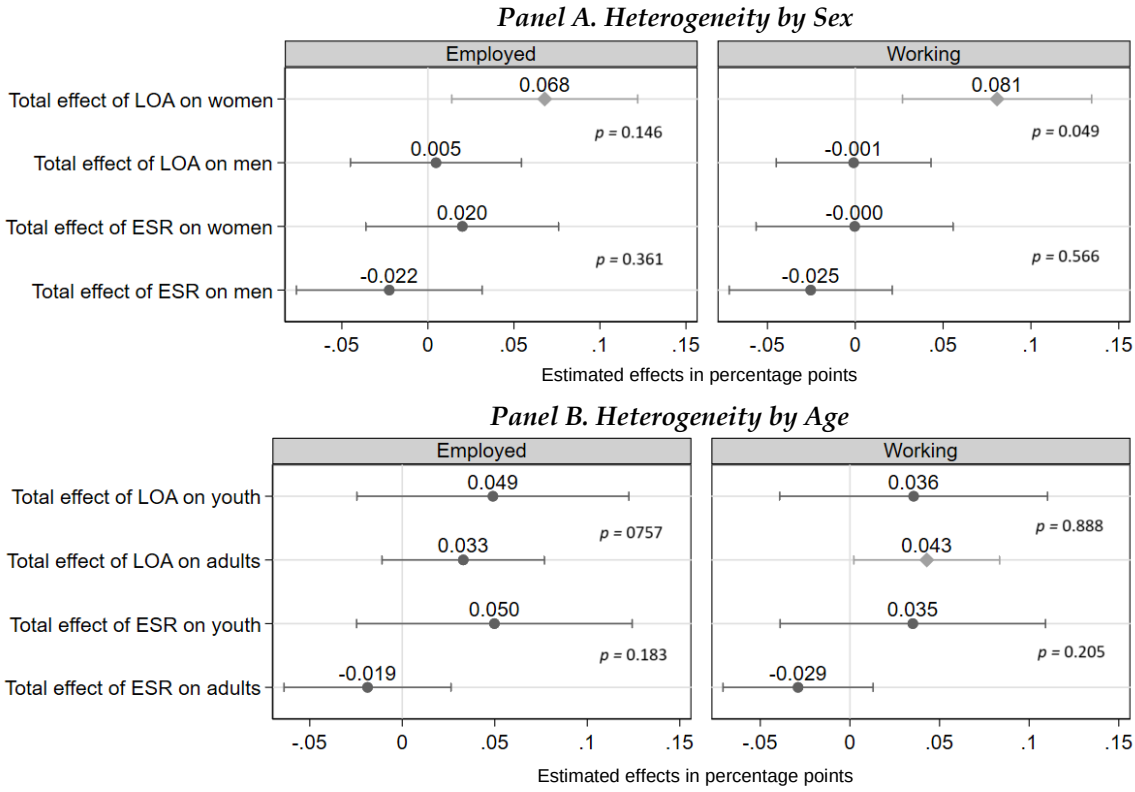
Notes: This table shows the average characteristics of the household- and individual-level variables—Panels A and B, respectively—by treatment status. Columns (1) to (4) present the mean for each characteristic of households and individuals participating in the study or assigned to C, LOA, and ESR, respectively. Columns (5) to (7) present the *p*-value associated with the hypothesis of the mean values across pairs of groups that are the same. For the estimation of *p*-values, we control for enumeration area (EA) fixed effects (stratification variable) and estimate standard errors at the EA level in Panel A and at the household level in Panel B. The variables presented in this table were collected using the following modules: sociodemographic characteristics, education, food security, household characteristics, and assets. Appendix B contains definitions for each of the variables. All indices are estimated using Anderson's (2008) approach of inverse covariance weighting and take a value between 0 and 1. All variables are dummies except when the unit of measurement is indicated in parentheses.

Table 2. Survey Methods' Impacts on the Reporting of Labor Market Outcomes

	(1)	(2)
	Employed	Working
LOA	0.040*	0.043**
	(0.023)	(0.022)
	[0.078]	[0.047]
ESR	-0.001	-0.012
	(0.024)	(0.023)
	[0.972]	[0.601]
Observations	2,480	2,480
Outcome (Control Group) Mean	0.571	0.670

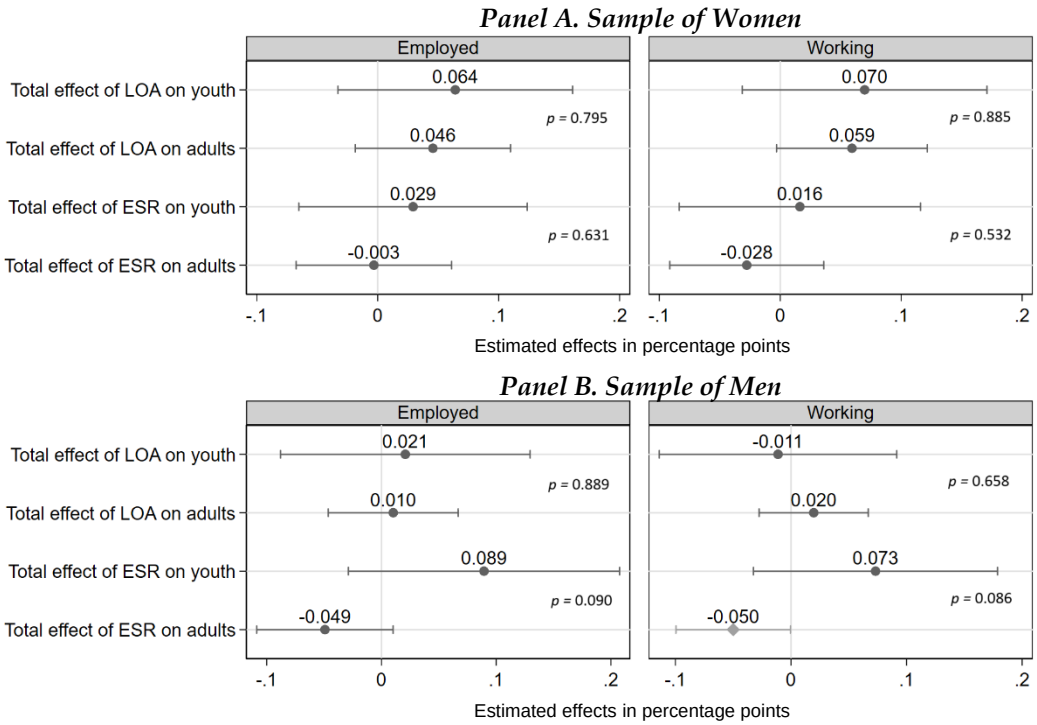
Notes: This table presents the main effects of including the LOA module (LOA) or enforcing self-reporting (ESR) in a household survey on the reporting of employment or work status using specification (1). The vector X_{ihs} includes the list of control variables selected using a double-LASSO procedure for each outcome (Table A3 contains the variables selected). All estimations include strata fixed effects. "Outcome (Control Group) Mean" refers to the mean of the control group for each outcome. Standard errors clustered at the household level are presented in parentheses, and the randomization inference p-values are shown in brackets. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Figure 1. Survey Methods' Heterogeneous Impacts on the Reporting of Labor Market Outcomes by Sex and Age



Notes: This figure presents the heterogeneous impacts of including the LOA module (LOA) or enforcing self-reporting (ESR) on the reporting of employment or work status estimated using specification (2). Results from the estimation are presented in Table A4 in the Appendix. Youth are defined as individuals aged 15 to 24, and adults are defined as individuals aged 25 to 64. All estimated effects are expressed in percentage points and represent the total impact of each survey method on the respective subgroup (women, men, youths, and adults). The *p*-values indicate differences in the effect of each survey method between the two subgroups (i.e., the effect of LOA on women vs men or the effect of ESR on youth vs adults). Solid lines represent confidence intervals at 10%.

Figure 2. Heterogeneity for Employment and Work by Age
Samples Separated by Sex



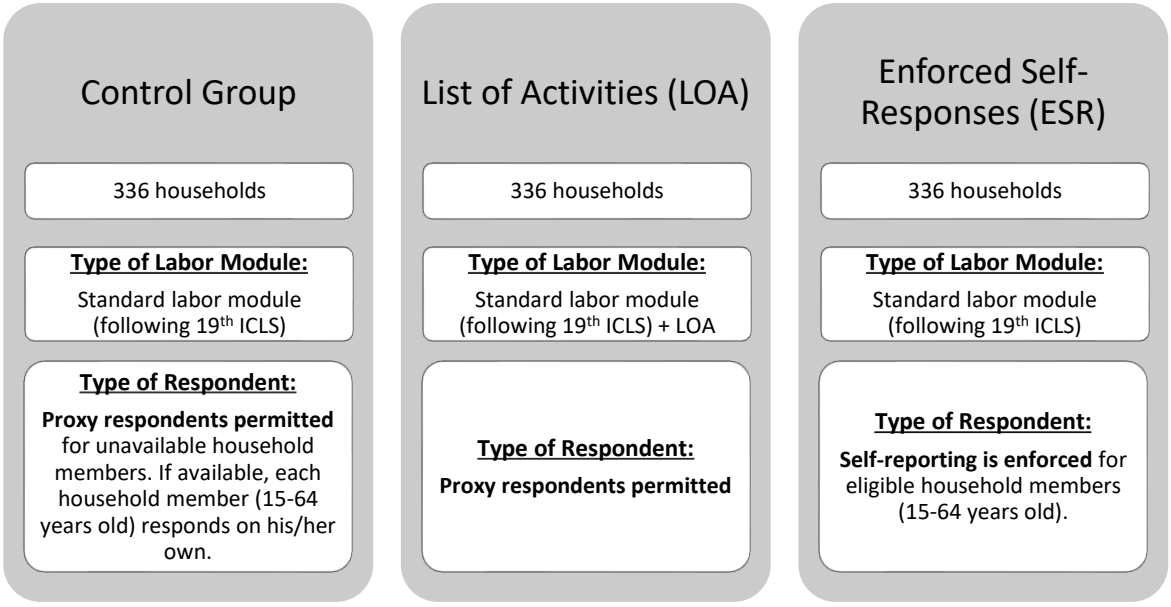
Notes: This figure presents the heterogeneous impacts of including the LOA module (LOA) or enforcing self-reporting (ESR) on the reporting of employment or work status estimated using specification (2). Panel A presents the results for the sample of women and Panel B shows the results for the sample of men. Results from the estimation are presented in Table A5 in the Appendix. Youth are defined as individuals aged 15 to 24, and adults are defined as individuals aged 25 to 64. All estimated effects are expressed in percentage points and represent the total impact of each survey method on the respective subgroup (youths and adults). The *p*-values indicate differences in the effect of each survey method between the two subgroups (i.e., the effect of LOA on female youth vs female adults or the effect of ESR on male youth vs male adults). Solid lines represent confidence intervals at 10%.

Supplementary Appendix for

"Closing the Gaps: The Role of Screening Questions and Self-Reporting in Measuring Women's and Youths' Employment and Work"

1. Figures and Tables

Figure A1. Experimental Design



Notes: This figure presents the type of labor module, the type of respondent permitted to report within each experimental group, and the household distributions within each group. The assignments were carried out using stratified randomization at the enumeration area level. LOA=List of Activities module.

Table A1. List of Activities Module

During the past week, from Monday [DATE] to Sunday [DATE], did [NAME]…											
Agricultural activities		**Production of items for sale**			**Provision of services**						
1.	**2.**	**3.**	**4.**	**5.**	**6.**	**7.**	**8.**	**9.**	**10.**	**11**	**12**
... do any agricultural work, care for your livestock, or fish for one or more hours?	... help a family member with agricultural work, take care of the family's livestock, or fish for one or more hours?	... sell items such as clothing, cell phones, shoes, jewelry, etc. for an hour or more?	... make items to sell, such as tamales, jelly, food, jewelry, etc. for one or more hours?	... sell homemade items such as tamales, jelly, food, jewelry, etc. for one or more hours?	... provide services such as hairstyling, repairs, or masonry, injecting medicines, or caring for the sick or elderly for one or more hours?	... provide transportation services such as taxi, Uber, Mototaxi, pickup truck, or minibus for one or more hours?	... provide home delivery services for one or more hours?	... cook, launder, or perform other services for people for one or more hours?	... help in a non-agricultural family business for one or more hours?	… engage (or be willing to engage) in activities that generate any type of income?	What other activities not mentioned here did [NAME] engage in to generate income?
YES...1 NO....2	YES...1 NO....2	YES...1 NO....2	YES...1 NO....2	YES...1 NO....2	YES...1 NO....2	YES...1 NO....2	YES...1 NO....2	YES...1 NO....2	YES...1 NO....2	YES...1 NO....2	

Notes: This table shows the List of Activities (LOA) module that was included in the survey instrument used to collect data from all randomly selected working-age individuals living in the households assigned to the LOA module.

Table A2. Modules Included in the Survey

#	Module	Level	Respondent	Proxy Permitted?
0	Cover	Household	Main respondent	Yes
1	Household Roster	Individual	Main respondent	Yes
2	Education and Technology	Individual	Individuals 15-64 years old	Yes
3	**List of Activities**	**Individual**	**Individuals 15-64 years old**	**Yes**
4	**Labor**	**Individual**	**Individuals 15-64 years old**	**Yes**
5	Skills and Work Readiness	Individual	Individuals 15-64 years old	No
6	Aspirations	Individual	Individuals 15-64 years old	No
7	Time Use	Individual	Individuals 15-64 years old	No
8	Social Norms	Individual	Individuals 15-64 years old	No
9	Social Desirability	Individual	Individuals 15-64 years old	No
10	Discrete Choice Experiment	Individual	Literate individuals 15-64 years old	No
11	Risk Aversion	Individual	Individuals 15-64 years old	No
12	Housing	Household	Main respondent	
13	Household Enterprises Flap	Household	Main respondent	
14	Household Enterprises	Enterprise	Main respondent	
15	Food Security	Household	Main respondent	
16	Agriculture	Household	Main respondent	
17	Assets	Household	Main respondent	
18	Other Income	Household	Main respondent	

Notes: This table shows the modules included in the survey, along with the corresponding question level and respondent type. We also report whether proxy respondents were permitted for each of the modules at the individual level. For example, data for module Time Use was collected only if the individual him/herself was available at the time of the survey and self-reported the data directly. If the household was assigned to the LOA or C, then we then collected self-reported data on time use only from those respondents who were available at the time of the survey.

Table A3. LASSO-Selected Control Variables for Each Outcome and Model

# Table	Outcome	Access to Mobile Phone	Sex (=1 if Female)	Education Level: High school or Higher	Marital Status (=1 if never married)	Can Read or Write	Age
Table 2	Employed (Column 1)	√	√	√	√		
	Working (Column 2)	√	√	√	√		
Table A4	Employed (Column 1)	√	√	√	√	√	
	Working (Column 2)	√	√	√	√		
	Employed (Column 3)	√	√	√		√	
	Working (Column 4)	√	√	√		√	
Table A5	Employed (Column 1)			√			
	Working (Column 2)						
	Employed (Column 3)	√				√	
	Working (Column 4)	√			√		

Notes: This table reports the control variables selected using the LASSO approach for each outcome and model (table) used in our estimations.

Table A4. Survey Methods' Heterogeneous Impacts on the Reporting of Labor Market Outcomes

	(1)	(2)	(3)	(4)
	Heterogeneous Effects			
	D = Female		D = Youth	
	Employed	Working	Employed	Working
LOA	0.005	-0.001	0.033	0.043*
	(0.030)	(0.027)	(0.027)	(0.025)
	[0.863]	[0.971]	[0.215]	[0.062]
ESR	-0.022	-0.025	-0.019	-0.029
	(0.033)	(0.028)	(0.027)	(0.025)
	[0.493]	[0.353]	[0.481]	[0.244]
LOA x D	0.063	0.082**	0.016	-0.007
	(0.043)	(0.041)	(0.052)	(0.052)
	[0.175]	[0.059]	[0.752]	[0.879]
ESR x D	0.042	0.025	0.069	0.064
	(0.046)	(0.044)	(0.051)	(0.050)
	[0.339]	[0.541]	[0.195]	[0.239]
LOA + LOA x D	0.068**	0.081**	0.049	0.036
	(0.033)	(0.033)	(0.045)	(0.045)
	[0.051]	[0.014]	[0.256]	[0.415]
ESR + ESR x D	0.020	-0.000	0.050	0.035
	(0.034)	(0.034)	(0.045)	(0.045)
	[0.567]	[0.993]	[0.291]	[0.442]
Observations	2,480	2,480	2,480	2,480
Outcome (D=0) Mean	0.728	0.823	0.627	0.732
Outcome (Control Group) Mean	0.571	0.670	0.571	0.670

Notes: This table presents the heterogeneous impacts of including the LOA module (LOA) or enforcing self-reporting (ESR) in a household survey on the reporting of employment or work status using specification (2). The indicator D_{ihs} in the heterogeneity analysis takes the value of 1 if the respondent is a woman (Columns [1] and [2]) and if the respondent is a youth aged 15 to 24 years (Columns [3] and [4]). The vector X_{ihs} includes the list of control variables selected using a double-LASSO procedure for each outcome (Table A3 contains the variables selected). All estimations include strata fixed effects. Standard errors clustered at the household level are presented in parentheses, and the randomization inference p-values are shown in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table A5. Heterogeneity for Employment and Work by Age
Samples Separated by Sex

	(1)	(2)	(3)	(4)
	Women		Men	
	Employed	Working	Employed	Working
LOA	0.046	0.059*	0.010	0.020
	(0.039)	(0.038)	(0.034)	(0.029)
	[0.214]	[0.103]	[0.773]	[0.485]
ESR	-0.003	-0.028	-0.049	-0.050*
	(0.039)	(0.039)	(0.036)	(0.030)
	[0.947]	[0.453]	[0.187]	[0.095]
LOA x Youth	0.018	0.010	0.011	-0.031
	(0.071)	(0.073)	(0.076)	(0.070)
	[0.805]	[0.881]	[0.885]	[0.645]
ESR x Youth	0.032	0.044	0.139*	0.123*
	(0.067)	(0.070)	(0.082)	(0.072)
	[0.603]	[0.542]	[0.105]	[0.104]
LOA + LOA x Youth	0.064	0.070	0.021	-0.011
	(0.059)	(0.061)	(0.066)	(0.062)
	[0.265]	[0.252]	[0.740]	[0.857]
ESR + ESR x Youth	0.029	0.016	0.089	0.073
	(0.057)	(0.061)	(0.072)	(0.064)
	[0.588]	[0.786]	[0.227]	[0.273]
Observations	1,361	1,361	1,119	1,119
Outcome (Control Group) Mean	0.436	0.537	0.735	0.832

Notes: This table presents the treatment heterogeneity of including the LOA module (LOA) or enforcing self-reporting (ESR) by age for the reporting of employment or work status. We separate the samples by sex. We present the ITT estimates from Equation (2), where D_{ihs} takes the value of 1 if the respondent is a youth between the ages of 15 to 24 years, and the total effect of each survey method. The vector X_{ihs} includes the indicator D_{ihs} and the list of control variables selected using a double-LASSO procedure for each outcome (see Table A3 in the Appendix for the selected variables). All estimations include strata fixed effects. The control mean refers to the mean of the control group for each outcome. Standard errors clustered at the household level are presented in parentheses, and the randomization inference p-values are shown in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table A6. Mean Characteristics of the Study Participants vs. General Population of Salvadorans

Panel A. Comparison of study sample and the Salvadoran Household Survey (EHPM)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variable	Study Sample	EHPM Usulután and San Salvador	*P*-val. Diff.	Study Sample Usulután	EHPM Usulután	*P*-val. Diff.	Study Sample San Salvador	EHPM San Salvador	*P*-val. Diff.
Age (years)	35.487	35.081	(0.836)	35.658	35.748	(0.976)	35.325	34.457	(0.733)
Female (%)	0.549	0.541	(0.916)	0.556	0.536	(0.850)	0.542	0.547	(0.960)
High school or higher education	0.285	0.281	(0.951)	0.249	0.221	(0.757)	0.319	0.337	(0.843)
Read or write	0.871	0.878	(0.886)	0.846	0.824	(0.791)	0.895	0.928	(0.501)
Never married	0.335	0.295	(0.536)	0.330	0.258	(0.451)	0.340	0.329	(0.902)
Access to mobile phone	0.983	0.986	(0.856)	0.981	0.977	(0.903)	0.984	0.994	(0.518)
Access to internet (Wi-Fi)	0.185	0.156	(0.581)	0.168	0.125	(0.556)	0.202	0.186	(0.828)
Observations	2,480	2,212		1,210	804		1,270	1,408	

Panel B. Comparison of subgroups of study sample and EHPM	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Self-respondents			**Proxy-respondents**			**Adult**		
	Study Sample	EHPM	*P*-val. Diff.	Study Sample	EHPM	*P*-val. Diff.	Study Sample	EHPM	*P*-val. Diff.
Age (years)	36.539	40.451	(0.194)	32.871	33.219	(0.946)	41.280	41.767	(0.854)
Female (%)	0.621	0.657	(0.741)	0.370	0.467	(0.609)	0.550	0.555	(0.966)
High school or higher education	0.278	0.364	(0.445)	0.301	0.393	(0.623)	0.259	0.405	(0.205)
Read or write	0.870	0.910	(0.554)	0.875	0.937	(0.502)	0.841	0.906	(0.345)
Never married	0.288	0.183	(0.243)	0.453	0.465	(0.949)	0.205	0.204	(0.988)
Access to mobile phone	0.979	0.980	(0.975)	0.992	0.990	(0.971)	0.983	0.984	(0.978)
Access to internet (Wi-Fi)	0.190	0.335	(0.187)	0.173	0.408	(0.210)	0.181	0.378	(0.084)*
Observations	1,769	16,220		711	23,593		1,817	29,371	

	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
		Youth			**Men**			**Women**	
	Study Sample	EHPM	*P*-val. Diff.	Study Sample	EHPM	*P*-val. Diff.	Study Sample	EHPM	*P*-val. Diff.

Age (years)	19.614	19.766	(0.890)	35.580	35.563	(0.997)	35.411	36.655	(0.739)
Female (%)	0.546	0.514	(0.867)	--	--	--	--	--	--
High school or higher education	0.356	0.309	(0.792)	0.282	0.392	(0.451)	0.287	0.372	(0.518)
Read or write	0.953	0.983	(0.555)	0.869	0.936	(0.354)	0.874	0.917	(0.561)
Never married	0.691	0.778	(0.582)	0.309	0.391	(0.574)	0.356	0.316	(0.748)
Access to mobile phone	0.982	0.993	(0.733)	0.983	0.985	(0.963)	0.982	0.987	(0.873)
Access to internet (Wi-Fi)	0.196	0.378	(0.328)	0.181	0.365	(0.202)	0.188	0.389	(0.128)
Observations	663	10,442		1,119	18,123		1,361	21,690	

Notes: This table compares the average characteristics of the individuals in our sample and individuals in El Salvador. The 2022 Household and Multipurpose Survey (EHPM) provided the data for individuals in El Salvador that we compared to data that we measured similarly in our survey. Columns (1) to (3) in Panel A compare the two full samples and Columns (4) to (9) compare the samples by the Departments of San Salvador and Usulután. We restricted the EHPM sample to working-age household members (aged 15 to 64 years) in rural areas. Panel B presents comparison of mean characteristics between subgroups of the study sample and the EHPM. Subgroups included are self-respondents, proxy respondents, adults, youth, women, and men.

Table A7. Survey Methods' Impacts on the Measurement of Labor Market Outcomes, Excluding Controls

	(1) Employed	(2) Working
LOA	0.037	0.040*
	(0.023)	(0.022)
	[0.148]	[0.086]
ESR	-0.004	-0.014
	(0.025)	(0.023)
	[0.879]	[0.585]
Observations	2,480	2,480
Outcome (Control Group) Mean	0.571	0.670

Notes: This table shows the estimated effects of including the LOA module (LOA) or enforcing self-reporting (ESR) on reporting employment or work status. We present estimated coefficients from Equation (1) but exclude the vector of control variables. Each column is a separate dependent variable as defined in Section 3. The control mean refers to the mean of the control group for each outcome. All estimations include strata fixed effects only. Standard errors clustered at the household level are presented in parentheses, and the randomization inference *p*-values are shown in brackets. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A8. Survey Methods' Heterogeneous Effects by Respondent's Sex and Age, Excluding Controls

	(1)	(2)	(3)	(4)
	D = Female		D = Youth	
	Employed	Working	Employed	Working
LOA	-0.000	-0.008	0.023	0.033
	(0.031)	(0.027)	(0.027)	(0.025)
	[0.983]	[0.774]	[0.380]	[0.165]
ESR	-0.020	-0.020	-0.033	-0.042
	(0.034)	(0.029)	(0.028)	(0.026)
	[0.561]	[0.477]	[0.252]	[0.111]
LOA x D	0.064	0.084**	0.053	0.028
	(0.044)	(0.042)	(0.054)	(0.054)
	[0.185]	[0.055]	[0.314]	[0.593]
ESR x D	0.033	0.016	0.095*	0.089*
	(0.047)	(0.044)	(0.054)	(0.054)
	[0.473]	[0.694]	[0.094]	[0.128]
LOA + LOA x D	0.063*	0.075**	0.076	0.061
	(0.033)	(0.032)	(0.046)	(0.047)
	[0.065]	[0.022]	[0.104]	[0.193]
ESR + ESR x D	0.013	-0.005	0.062	0.047
	(0.034)	(0.034)	(0.048)	(0.048)
	[0.707]	[0.884]	[0.214]	[0.347]
Observations	2,480	2,480	2,480	2,480
Outcome (D=0) Mean	0.728	0.823	0.627	0.732
Outcome (Control Group) Mean	0.571	0.670	0.571	0.670

Notes: This table presents the heterogeneous impacts of including the LOA module (LOA) or enforcing self-reporting (ESR) in a household survey on the reporting of employment or work status using specification (2) but excluding the vector of control variables. The indicator D_{ihs} in the heterogeneity analysis takes the value of 1 if the respondent is a woman (Columns [1] and [2]) and if the respondent is a youth aged 15 to 24 years (Columns [3] and [4]). All estimations include strata fixed effects. "Outcome (Control Group) Mean" refers to the mean of the control group for each outcome. "Outcome (D=0) Mean" refers to the mean of the reference subgroup. For example, men for columns (1) and (2) and adults for columns (3) and (4). Standard errors clustered at the household level are presented in parentheses, and the randomization inference p-values are shown in brackets. *** p<0.01, ** p<0.05, * p<0.

Table A9. Heterogeneity for Employment and Work by Age, Excluding Controls

	(1)	(2)	(3)	(4)
	Women		Men	
	Employed	Working	Employed	Working
LOA	0.045	0.059	0.009	0.014
	(0.039)	(0.038)	(0.035)	(0.029)
	[0.216]	[0.103]	[0.812]	[0.619]
ESR	-0.006	-0.028	-0.053	-0.048
	(0.039)	(0.039)	(0.037)	(0.031)
	[0.897]	[0.453]	[0.175]	[0.131]
LOA x Youth	0.026	0.010	0.012	-0.032
	(0.071)	(0.073)	(0.077)	(0.072)
	[0.719]	[0.881]	[0.881]	[0.636]
ESR x Youth	0.035	0.044	0.155*	0.135*
	(0.068)	(0.070)	(0.083)	(0.074)
	[0.579]	[0.542]	[0.081]	[0.095]
LOA + LOA x Youth	0.071	0.070	0.021	-0.018
	(0.058)	(0.061)	(0.067)	(0.064)
	[0.204]	[0.252]	[0.739]	[0.779]
ESR + ESR x Youth	0.029	0.016	0.102	0.088
	(0.057)	(0.061)	(0.073)	(0.066)
	[0.586]	[0.786]	[0.178]	[0.215]
Observations	1,361	1,361	1,119	1,119
Outcome (Control Group) Mean	0.436	0.537	0.735	0.832

Notes: This table presents the treatment heterogeneity of including the LOA module (LOA) or enforcing self-reporting (ESR) by age for the reporting of employment or work status. We separate the samples by sex. We present the ITT estimates from Equation (2), where D_{ihs} takes the value of 1 if the respondent is a youth between the ages of 15 to 24 years, and the total effect of each survey method. All estimations include strata fixed effects. We exclude the vector of control variables for this robustness check. The control mean refers to the mean of the control group for each outcome. Standard errors clustered at the household level are presented in parentheses, and the randomization inference p-values are shown in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table A10. Survey Methods' Impacts on the Measurement of Labor Market Outcomes Including Control Variables with Imbalance at Baseline

	(1) Employed	(2) Working
LOA	0.039*	0.047**
	(0.023)	(0.022)
	[0.093]	[0.033]
ESR	-0.002	-0.015
	(0.024)	(0.023)
	[0.940]	[0.482]
Observations	2,480	2,480
Outcome (Control Group) Mean	0.571	0.670

Notes: This table shows the estimated effects of including the LOA module (LOA) or enforcing self-reporting (ESR) on the reporting of employment or work status. We present estimated coefficients from Equation (1) controlling by variables that were imbalanced across groups at baseline (respondent's age, reading and writing skills and access to mobile phone) as a robustness test for Table 2. Each column is a separate dependent variable as defined in Section 3. All estimations include strata fixed effects. "Outcome (Control Group) Mean" refers to the mean of the control group for each outcome. Standard errors clustered at the household level are presented in parentheses, and the randomization inference *p*-values are shown in brackets. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A11. Survey Methods' Heterogeneous Effects by Respondent's Sex and Age, Including Control Variables with Imbalance at Baseline

	(1)	(2)	(3)	(4)
	D = Female		D = Youth	
	Employed	Working	Employed	Working
LOA	0.006	0.006	0.034	0.049**
	(0.030)	(0.027)	(0.027)	(0.024)
	[0.828]	[0.816]	[0.201]	[0.036]
ESR	-0.024	-0.027	-0.016	-0.029
	(0.033)	(0.028)	(0.027)	(0.026)
	[0.468]	[0.338]	[0.520]	[0.245]
LOA x D	0.060	0.075*	0.016	-0.010
	(0.043)	(0.041)	(0.052)	(0.052)
	[0.193]	[0.081]	[0.735]	[0.827]
ESR x D	0.040	0.021	0.066	0.063
	(0.046)	(0.043)	(0.052)	(0.051)
	[0.362]	[0.611]	[0.210]	[0.242]
LOA + LOA x D	0.066**	0.081**	0.050	0.039
	(0.033)	(0.032)	(0.045)	(0.045)
	[0.055]	[0.012]	[0.241]	[0.359]
ESR + ESR x D	0.016	-0.006	0.050	0.034
	(0.034)	(0.034)	(0.045)	(0.045)
	[0.641]	[0.851]	[0.290]	[0.462]
Observations	2,480	2,480	2,480	2,480
Outcome (D=0) Mean	0.728	0.823	0.627	0.732
Outcome (Control Group) Mean	0.571	0.670	0.571	0.670

Notes: This table presents the treatment heterogeneity of including the LOA module (LOA) or enforcing self-reporting (ESR) by sex and age for the reporting of employment or work status. We present the ITT estimates from Equation (2), where D_{ihs} takes the value of 1 if the respondent is a youth between the ages of 15 to 24 years, as well as the total effect of each survey method. In the vector of control variables, we include variables that were imbalanced across groups at baseline (respondent's age, reading and writing skills and access to mobile phone) as a robustness test of Table A4 and Figure 1. All estimations include strata fixed effects. "Outcome (Control Group) Mean" refers to the mean of the control group for each outcome. "Outcome (D=0) Mean" refers to the mean of the reference subgroup. For example, men for columns (1) and (2) and adults for columns (3) and (4). Standard errors clustered at the household level are presented in parentheses, and the randomization inference p-values are shown in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table A12. Heterogeneity for Employment and Work by Age, Including Control Variables with Imbalance at Baseline

	(1)	(2)	(3)	(4)
	Women		Men	
	Employed	Working	Employed	Working
LOA	0.052	0.074**	0.010	0.025
	(0.039)	(0.037)	(0.034)	(0.029)
	[0.151]	[0.043]	[0.786]	[0.376]
ESR	0.006	-0.018	-0.046	-0.044
	(0.039)	(0.038)	(0.036)	(0.030)
	[0.880]	[0.590]	[0.219]	[0.136]
LOA x Youth	0.022	0.006	0.011	-0.036
	(0.071)	(0.073)	(0.076)	(0.070)
	[0.766]	[0.912]	[0.876]	[0.600]
ESR x Youth	0.029	0.039	0.135	0.119*
	(0.067)	(0.070)	(0.082)	(0.071)
	[0.640]	[0.599]	[0.118]	[0.118]
LOA + LOA x Youth	0.074	0.080	0.020	-0.011
	(0.059)	(0.061)	(0.067)	(0.063)
	[0.192]	[0.181]	[0.742]	[0.866]
ESR + ESR x Youth	0.035	0.021	0.089	0.074
	(0.057)	(0.061)	(0.073)	(0.064)
	[0.520]	[0.745]	[0.230]	[0.277]
Observations	1,361	1,361	1,119	1,119
Outcome (Control Group) Mean	0.436	0.537	0.735	0.832

Notes: This table presents the treatment heterogeneity of including the LOA module (LOA) or enforcing self-reporting (ESR) by age for the reporting of employment or work status. We separate the samples by sex. We present the ITT estimates from Equation (2), where D_{ihs} takes the value of 1 if the respondent is a youth between the ages of 15 to 24 years, as well as the total effect of each survey method. In the vector of control variables, we include variables that were imbalanced across groups at baseline (respondent's age, reading and writing skills and access to mobile phone) as a robustness test of Table A5 and Figure 2. All estimations include strata fixed effects. "Outcome (Control Group) Mean" refers to the mean of the control group for each outcome. Standard errors clustered at the household level are presented in parentheses, and the randomization inference p-values are shown in brackets. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A13. Survey Methods' Impacts on the Measurement of Labor Market Outcomes, Addressing the Substitution Protocol's Effects

	(1) Employed	(2) Working
LOA	0.040*	0.044**
	(0.023)	(0.022)
	[0.081]	[0.043]
ESR	-0.001	-0.011
	(0.024)	(0.023)
	[0.966]	[0.620]
Observations	2,480	2,480
Outcome (Control Group) Mean	0.571	0.670

Notes: This table shows the estimated effects of including the LOA module (LOA) or enforcing self-reporting (ESR) on the reporting of employment or work status. We present estimated coefficients from Equation (1). Each column is a separate dependent variable as defined in Section 3. The control mean refers to the mean of the control group for each outcome. All estimations include strata fixed effects and are controlled by variables selected using a double-LASSO procedure. "Outcome (Control Group) Mean" refers to the mean of the control group for each outcome. Standard errors clustered at the household level are presented in parentheses, and the randomization inference p-values are shown in brackets. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A14. Survey Methods' Heterogeneous Effects by Respondent's Sex and Age, Addressing the Substitution Protocol's Effects

	(1)	(2)	(3)	(4)
	D = Female		D = Youth	
	Employed	Working	Employed	Working
LOA	0.005	-0.000	0.033	0.044*
	(0.030)	(0.027)	(0.027)	(0.025)
	[0.866]	[0.993]	[0.215]	[0.059]
ESR	-0.023	-0.024	-0.019	-0.028
	(0.033)	(0.028)	(0.028)	(0.025)
	[0.492]	[0.375]	[0.481]	[0.262]
LOA x D	0.063	0.081*	0.016	-0.007
	(0.044)	(0.042)	(0.052)	(0.052)
	[0.174]	[0.059]	[0.753]	[0.879]
ESR x D	0.043	0.024	0.069	0.063
	(0.047)	(0.044)	(0.052)	(0.051)
	[0.339]	[0.568]	[0.195]	[0.241]
LOA + LOA x D	0.068**	0.081**	0.049	0.036
	(0.033)	(0.033)	(0.045)	(0.045)
	[0.051]	[0.014]	[0.256]	[0.405]
ESR + ESR x D	0.020	-0.000	0.050	0.036
	(0.034)	(0.034)	(0.045)	(0.045)
	[0.567]	[0.995]	[0.291]	[0.435]
Observations	2,480	2,480	2,480	2,480
Outcome (D=0) Mean	0.728	0.823	0.627	0.732
Outcome (Control Group) Mean	0.571	0.670	0.571	0.670

Notes: This table presents the treatment heterogeneity of including the LOA module (LOA) or enforcing self-reporting (ESR) by sex and age for the reporting of employment or work status. We present the ITT estimates from Equation (2) and the total effect of each survey method. The indicator D_{ihs} takes the value of 1 if the respondent is a woman (Columns [1] and [2]) and if the respondent is a youth between the ages of 15 to 24 years (Columns [3] and [4]). The vector X_{ihs} includes the indicator D_{ihs} and the list of control variables selected using a double-LASSO procedure for each outcome (Table A3 in the Appendix shows the variables selected). All estimations include strata fixed effects. Each column is a separate dependent variable as defined in Section 3. "Outcome (Control Group) Mean" refers to the mean of the control group for each outcome. "Outcome (D=0) Mean" refers to the mean of the reference subgroup. For example, men for columns (1) and (2) and adults for columns (3) and (4). Standard errors clustered at the household level are presented in parentheses, and the randomization inference p-values are shown in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table A15. Heterogeneity for Employment and Work by Age, Addressing the Substitution Protocol's Effects

	(1)	(2)	(3)	(4)
	Women		Men	
	Employed	Working	Employed	Working
LOA	0.052	0.069*	0.009	0.019
	(0.039)	(0.038)	(0.034)	(0.029)
	[0.154]	[0.058]	[0.806]	[0.501]
ESR	0.001	-0.022	-0.052	-0.052*
	(0.039)	(0.039)	(0.036)	(0.030)
	[0.987]	[0.551]	[0.161]	[0.084]
LOA x Youth	0.018	0.009	0.010	-0.031
	(0.071)	(0.073)	(0.076)	(0.070)
	[0.799]	[0.894]	[0.886]	[0.647]
ESR x Youth	0.031	0.041	0.139*	0.123*
	(0.068)	(0.070)	(0.080)	(0.072)
	[0.621]	[0.578]	[0.107]	[0.104]
LOA + LOA x Youth	0.071	0.077	0.019	-0.012
	(0.059)	(0.062)	(0.066)	(0.063)
	[0.223]	[0.198]	[0.751]	[0.847]
ESR + ESR x Youth	0.032	0.019	0.087	0.072
	(0.057)	(0.060)	(0.072)	(0.064)
	[0.552]	[0.755]	[0.235]	[0.293]
Observations	1,361	1,361	1,119	1,119
Outcome (Control Group) Mean	0.436	0.537	0.735	0.832

Notes: This table presents the treatment heterogeneity of including the LOA module (LOA) or enforcing self-reporting (ESR) by age for the reporting of employment or work status. We separate the samples by sex. We present the ITT estimates from Equation (2), where D_{ihs} takes the value of 1 if the respondent is a youth between the ages of 15 to 24 years, as well as the total effect of each survey method. The vector X_{ihs} includes the indicator D_{ihs} and the list of control variables selected using a double-LASSO procedure for each outcome (see Table A3 in the Appendix for the list of variables selected by LASSO). All estimations include strata fixed effects. Each column is a separate dependent variable as defined in Section 3. "Outcome (Control Group) Mean" refers to the mean of the control group for each outcome. Standard errors clustered at the household level are presented in parentheses, and the randomization inference p-values are shown in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table A16. Survey Methods' Differential Effects on Employment and Work Status by Prevalence of Informal Employment

	(1)	(2)	(3)	(4)
	Women		Young Men	
	Employed	Working	Employed	Working
LOA	0.063*	0.077**	-0.012	-0.036
	(0.033)	(0.033)	(0.062)	(0.059)
	[0.059]	[0.017]	[0.843]	[0.569]
ESR	0.013	-0.003	0.077	0.053
	(0.034)	(0.034)	(0.064)	(0.058)
	[0.667]	[0.922]	[0.228]	[0.394]
LOA x Informal employment in the EA	0.008	0.024	0.048	0.082
	(0.032)	(0.032)	(0.056)	(0.055)
	[0.837]	[0.459]	[0.417]	[0.157]
ESR x Informal employment in the EA	0.048	0.052	0.142**	0.100*
	(0.034)	(0.034)	(0.062)	(0.056)
	[0.145]	[0.108]	[0.026]	[0.078]
LOA + LOA x Informal employment in the EA	0.070	0.100**	0.036	0.046
	(0.047)	(0.045)	(0.084)	(0.081)
	[0.124]	[0.026]	[0.704]	[0.591]
ESR + ESR x Informal employment in the EA	0.060	0.049	0.220**	0.153*
	(0.048)	(0.047)	(0.091)	(0.084)
	[0.181]	[0.285]	[0.012]	[0.067]
Observations	1,361	1,361	301	301
Outcome (Control Group) Mean	0.436	0.537	0.620	0.717

Notes: This table presents the treatment heterogeneity of including the LOA module (LOA) or enforcing self-reporting (ESR) by the prevalence of employment in the informal sector. This prevalence measure is estimated as the share of individuals (women in Columns [1] and [2] and men in Columns [3] and [4]) living in the same EA as the respondent and employed in the informal sector. For our estimations, we use informal employment in SD. On average, 1 SD = 11 percentage points of informal employment. This variable is estimated using the standard labor module. We present ITT estimates from Equation (2), where D_{ihs} is the prevalence measure in SD, and the estimations of each survey method's total effect. The vector X_{ihs} includes the variable D_{ihs} and the list of control variables selected using a double-LASSO procedure for each outcome. All estimations include strata fixed effects. Each column is a separate dependent variable as defined in Section 3. "Outcome (Control Group) Mean" refers to the mean of the control group for each outcome. Standard errors clustered at the household level are presented in parentheses, and the randomization inference p-values are shown in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table A17. Participation in Employment and Work by the Proxy Respondent's Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Employed						Working					
	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men
LOA	0.032	0.002	0.051	0.005	0.034	0.005	0.048	-0.014	0.062	-0.012	0.050	-0.013
	(0.062)	(0.040)	(0.062)	(0.040)	(0.062)	(0.040)	(0.063)	(0.036)	(0.063)	(0.035)	(0.062)	(0.035)
	[0.650]	[0.956]	[0.450]	[0.907]	[0.625]	[0.919]	[0.484]	[0.715]	[0.356]	[0.773]	[0.467]	[0.744]
Proxy is female	-0.090	-0.022					-0.089	-0.019				
	(0.067)	(0.081)					(0.066)	(0.072)				
Proxy has high school education or higher			-0.143*	0.076*					-0.086	0.067*		
			(0.073)	(0.044)					(0.077)	(0.038)		
Proxy is spouse					0.169**	0.120*					0.151**	0.074
					(0.077)	(0.065)					(0.076)	(0.051)
Observations	261	448	261	448	261	448	261	448	261	448	261	448
Outcome (Control Group) Mean	0.476	0.749	0.476	0.749	0.476	0.749	0.524	0.835	0.524	0.835	0.524	0.835

Notes: This table presents the associations between the proxy respondent's characteristics and the probability of the proxy reporting employment and work status. The sample is restricted to participants assigned to the LOA and control group. We present ITT estimates from the following equation: $Y_{ihs} = \alpha_0 + \beta_1 T1_{hs} + \Theta X_{ihs} + b_s + \epsilon_{ihs}$. The vector X_{ihs} includes the list of control variables selected using a double-LASSO procedure for each outcome. All estimations include EA (stratification variable) fixed effects. "Outcome (Control Group) Mean" refers to the mean of the control group for each outcome. Standard errors clustered at the household level are presented in parentheses, and the randomization inference p-values are shown in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table A18. Survey Methods' Effects on Employment Across Economic Sectors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Agriculture vs. No Working		Agriculture vs. Other Sectors		Agriculture vs. Services		Agriculture vs. Manufacture		Agriculture vs. Commerce	
LOA	0.010	-0.022	0.018	0.036	0.006	0.009	0.021	0.014	0.025	0.075
	(0.028)	(0.056)	(0.022)	(0.033)	(0.036)	(0.047)	(0.043)	(0.050)	(0.040)	(0.054)
	[0.708]	[0.682]	[0.369]	[0.228]	[0.858]	[0.815]	[0.591]	[0.778]	[0.480]	[0.139]
ESR	-0.003	-0.050	0.038*	0.018	0.047	0.002	0.061	0.001	0.048	0.038
	(0.030)	(0.058)	(0.024)	(0.033)	(0.037)	(0.046)	(0.046)	(0.054)	(0.043)	(0.055)
	[0.906]	[0.382]	[0.087]	[0.571]	[0.175]	[0.956]	[0.165]	[0.979]	[0.212]	[0.503]
LOA x Female		0.051		-0.039		-0.008		0.024		-0.110
		(0.063)		(0.043)		(0.069)		(0.094)		(0.068)
		[0.419]		[0.329]		[0.903]		[0.796]		[0.110]
ESR x Female		0.072		0.044		0.129 *		0.153*		0.021
		(0.064)		(0.044)		(0.072)		(0.102)		(0.071)
		[0.272]		[0.317]		[0.069]		[0.103]		[0.775]
LOA + LOA x Female		0.029		-0.004		0.001		0.038		-0.034
		(0.029)		(0.028)		(0.052)		(0.080)		(0.051)
		[0.306]		[0.910]		[0.984]		[0.607]		[0.496]
ESR + ESR x Female		0.022		0.062**		0.131**		0.154**		0.059
		(0.031)		(0.032)		(0.057)		(0.085)		(0.056)
		[0.440]		[0.034]		[0.022]		[0.036]		[0.271]
Observations	1,085	1,085	1,530	1,530	885	885	535	535	696	696
Outcome (Control Group) Mean	0.241	0.241	0.174	0.174	0.314	0.314	0.515	0.515	0.380	0.380

Notes: This table presents the treatment effects of including the LOA module (LOA) or enforcing self-reporting (ESR) by sex for the extensive and intensive margins in the reporting of employment in the agricultural sector. Columns (1) and (2) present the results on an indicator that takes the value of 1 if the individual is working in agriculture and 0 if the individual is not working. Columns (3) and (4) show the results for an indicator equal to 1 if the individual reports being working in agriculture and 0 if the individual is working in other sectors (e.g., manufacture, services, and industry). The outcome in Columns (5) through (10) is measured as an indicator that takes the value of 1 if the individual reports being working in agriculture and 0 if the individual is working in services (Columns [5] and [6]), manufacture (Columns [7] and [8]), or commerce (Columns [9] and [10]). The vector X_{ihs} includes the indicator D_{ihs} and the list of control variables selected using a double-LASSO procedure for each outcome. All estimations include EA (stratification variable) fixed effects. Standard errors clustered at the household level are presented in parentheses, and the randomization inference p-values are shown in brackets. *** p<0.01, ** p<0.05, * p<0.1

Appendix A. Methodological Activities

Enumerator training and fieldwork. Before commencing data collection, supervisors and interviewers underwent a two-week training session. All of the interviewers and supervisors had previous experience using Survey Solutions and conducting household surveys. Following the training, the interviewers carried out mock interviews in order to become more familiar with the questionnaire and interviewing techniques. After the mock interviews, interviewers participated in a 1-day pilot with non-participating households within the EAs included in the experiment. All pilot interviews were recorded and later audited by two trained monitors who provided each of the interviewers with written feedback. The field coordinators then met with each interviewer and discussed the feedback. After the individual meetings, the team led a debrief session with all of the interviewers to address any remaining questions and concerns.

Six teams comprised of one supervisor and three interviewers each conducted the survey between August and October 2022. The teams operated in a roving manner, interviewing all selected households in each EA before moving on to the next. All of the households in each treatment status were interviewed following the same protocol: upon arriving at a selected household, the interviewer introduced herself to the first working-age person who opened the door, read aloud the consent statement, and, after the working-age adult agreed to continue, began the interview. All households that completed the interview received an in-kind incentive valued at USD$5.00.

Household substitution protocols. We randomly selected 21 additional substitute households for each EA in case participating household members or proxies did not show for the interview. Interviewers were asked to visit all households up to 5 times to obtain personal or parental consent (for youth) from household members to participate in the survey. If enumerators were unable to obtain consent after 5 visits, then they had to inform their supervisors, who would try to obtain consent one additional time. If neither of these approaches worked, then the household was replaced. In total, 22.6% of the households were replaced by substitutes.[1]

We included an additional substitution protocol for the ESR group, which enforced self-reporting. Interviewers were asked to visit each ESR household up to 5 times to talk directly with the selected household member. If the interviewer was not able to interview the selected household member after 5 visits, then the supervisor also visited the household and documented the reason why the selected household member was not available. In this case, a randomly selected replacement household within the EA replaced the entire household. Only 2.1% of the households in ESR were replaced due primarily to the fact that selected household members were hospitalized or visiting family in other municipalities.

A possible concern is that replacement households within each EA would differ from the original households that were not replaced. As shown in Table A1a, the two groups are very similar with few statistically significant differences between them, however, these differences are nevertheless very small. For example, original households had 0.2 household members more than substitute

[1] The main reasons for substitution were as follows: household members refused to participate (18.6%); households could not be found (1.0%); households were not eligible to participate in the survey because they did not report having at least 2 household members aged 15 to 64 years (2.3%); and the interviewer was unable to interview the designated ESR household member (0.7%).

households (a 6% difference). We estimated similarly small statistically significant differences for the dependency ratio as well as access to a mobile phone and/or Internet. The only large statistically significant difference between original and substitute households occurred for the share of households in rural areas: only 47% of the original households were in a rural area relative to 61% of the substitute households. However, as we explain in Section 4, we stratify the random assignment of households by EA (which are either rural or peri-urban). Thus, this concern should be addressed by including EA fixed effects as we did in our main specification. Overall, as we explain in the main text, we formally address this potential concern by including the variables for which there are differences between the two groups as controls in our main model and test the robustness of our results.

Data quality assurance. We performed extensive monitoring throughout the fieldwork to ensure data quality. Field coordinators and the team supervisors monitored data collection on site. They also visited the field teams at random times. In addition, we activated Survey Solutions audio recording functionality for 25% of the surveys and asked two trained monitors to audit the recordings. The monitors listened to these recordings daily and logged their observations in a structured questionnaire on interviewer performance. Lastly, once the interview was completed and uploaded to the server, project managers reviewed the data to verify that it was complete. Our field coordinators also checked for additional errors and produced an error file, which was communicated back to the respective field interviewers. These crosschecks were performed daily throughout the duration of the survey.

Table A1a. Means of Original and Substitute Household Characteristics

Variable	(1) Original Household	(2) Substitute Household	(3) *P*-val. Difference
PANEL A. Household Characteristics			
Household size (N)	3.756	3.543	0.052
Dependency ratio (%)	0.259	0.287	0.071
Households in rural area (%)	0.470	0.606	0.000
Households with an assets index above the median (%)	0.482	0.434	0.215
Moderate to severe food insecurity (%)	0.372	0.335	0.307
Remittances received (%)	0.366	0.308	0.110
Household has a non-farm enterprise (%)	0.333	0.299	0.337
Household works in agriculture (%)	0.545	0.529	0.679
Observations	787	221	
PANEL B. Individual Characteristics			
Female (%)	0.546	0.559	0.612
Age (years)	35.513	35.387	0.853
High school or higher education (%)	0.284	0.285	0.967
Read or write (%)	0.876	0.854	0.201
Never married (%)	0.332	0.345	0.587
Access to mobile phone (%)	0.887	0.856	0.055
Access to internet (Wi-Fi) (%)	0.177	0.218	0.036

Observations	1,979	501

Notes: This table compares the mean of the characteristics of the original (Column [1]) and substitute (Column [2]) households included in our sample. Column (3) presents the *p*-values for the test of the differences in means between the two groups.

Appendix B. Definitions of the Variables Included in the Analysis

B1. Household Characteristics

Household size: The number of individuals who normally live and eat their meals together in the household, excluding the domestic servants, other workers such as gardeners, and guests who visit temporarily. Those who normally live at home, but who were absent at the time of the interview because of academic studies, business, visiting friends or relatives, travel for pleasure, hospitalization, etc. are considered household members as long as they have lived in the household for at least 6 of the past 12 months.

Dependency ratio: The percentage of household members who are considered dependents. This ratio is calculated as the sum of dependents aged 0 to 14 years plus individuals over the age of 65 divided by the total number of working-age household members (15 to 64 years).

Households in rural areas: The percentage of households located in rural areas, which includes a binary indicator that takes the value of 1 for households in rural areas, and 0 for those in peri-urban areas. Due to our methodological design, the samples are evenly distributed across rural and peri-urban areas.

Households with an assets index higher than the median: The percentage of households that have an assets index above the median value within the sample. Following Anderson (2008), we construct a standardized index using inverse covariance weighting. We utilize the mean and standard deviation of the control group for the standardization. The items used to measure this assets index are radio, TV, video cassette or DVD player, refrigerator, washing machine, blender, fan, computer, sewing machine, vehicle (car or boat), iron, microwave oven, videogames console, air conditioning, well or cistern, motorcycle, tablet.

Households that experienced moderate to severe food insecurity: According to the Food Insecurity Experience Scale (FIES),[2] moderate to severe food insecurity prevalence refers to a range of food security conditions experienced by the households that have difficulty accessing enough safe and nutritious food for their members' normal growth and development and that fail to enjoy an active and healthy life due to a lack of money or other resources. This variable takes the value of 1 if the household falls within the range of moderate to severe food insecurity, and 0 otherwise.

The questions used to construct the FIES are as follows:

Question: Due to a lack of money or resources, in the past 30 days has	
1. … anyone in your household worried about not having enough food to eat?	YES...1 NO....2

[2] More information on the FIES can be found at FAO (2023).

2. … anyone in your HH been unable to eat healthy and nutritious foods?	YES...1 NO....2
3. … anyone in your household eaten only a few kinds of foods?	YES...1 NO....2
4. … anyone in your household had to skip a meal?	YES...1 NO....2
5. … anyone in your household eaten less than you thought he/she should?	YES...1 NO....2
6. … your household run out of food?	YES...1 NO....2
7. … anyone in your household been hungry but did not eat?	YES...1 NO....2
8. … anyone in your household gone hungry for the entire day?	YES...1 NO....2

Households that received remittances: The proportion of households that have received international remittances, which include cash transfers or gifts from individuals living abroad, such as relatives and friends, within the last 12 months.

Household with a non-farm enterprise: The proportion of households that operate at least one non-farm enterprise (NFE). An NFE can encompass a wide range of income-generating activities, including small businesses, workshops, retail ventures, services, manufacturing, or any other commercial endeavors that do not primarily involve agricultural production.

Household working on agricultural activities: The percentage of households engaged in agricultural, livestock, or fishing activities. This variable is assigned a value of 1 if the household responds affirmatively to at least one of the following questions:

1. Do any members of your household own or have access to land that is used for crop cultivation during the agricultural season?	YES...1 NO....2
2. Has your household owned any livestock in the past 12 months?	YES...1 NO....2
3. Have you or any other member of your household been involved in fishing activities (i.e., catching or raising fish) in the last 12 months?	YES...1 NO....2

B2. Individual Characteristics

Female: Denotes the sex of the individual, = 1 if the person is female, or 0 otherwise.

Age: An individual's chronological age measured in years.

High school or higher education: Indicates whether the individual has completed at least a high school education or has received a higher education.

Read or write: Captures literacy status and takes the value of 1 if the respondent can read or write, or 0 if the respondent is illiterate.

Never married: Identifies marital status and more specifically signifies that the person has never entered a legally recognized marriage or civil partnership.

Access to mobile phone: This variable takes the value of 1 if the individual has access to a mobile phone, or 0 otherwise.

Access to Internet (Wi-Fi): This variable takes the value of 1 if the individual has access to the Internet at home, or 0 otherwise.