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

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

December 2023

Abstract

Women and youth are severely underrepresented in the global labor market. We study if undermeasurement of women's and youth's labor market outcomes partially explains existing gender and age employment gaps. To this end, we designed a survey experiment to assess how either using a list of activities survey module (LOA) or enforcing self-reporting affect the measurement of work and employment by gender and age. Using data from 1,008 households and 2,480 working-age respondents in El Salvador, we show that including the LOA increases the average reporting of work by women relative to the men and other women who were not exposed to the LOA module. Although enforced self-reporting has no effects on the average reporting of employment and work, it effectively increases the reporting of these outcomes by young males relative to older males who were also exposed to the same enforcement. In fact, proxy reporting yields lower male working rates than self-reporting. These results have important implications for addressing the undermeasurement of women's and youths' labor outcomes in official statistics, and they also explain to some degree current gender and age employment gaps.

JEL classification: O1, J2, C8
Keywords: work measurement, employment measurement, survey methods, women, youths

[1] We are very thankful to Sylvan Herskowitz and Diego Zardetto for their time and feedback. We also appreciate the comments from participants of the different conferences and seminars where we presented this paper, including the World Bank Research Department, World Bank Data Group Seminar, Statistics Canada's 2022 International Methodology Symposium, European Survey Research Association Conference 2023, and Northwestern GPRL-IPA Methods and Measurement Conference. We acknowledge 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] Corresponding Author. The World Bank. Email: icontreras@worldbank.org.
[3] 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

Statistically, women and youth are severely underrepresented in the global labor market. In 2022, only 43.8 percent of women compared to 67.9 percent of men were employed (ILO, 2022). Youth employment is similarly underrepresented, especially in high-informality settings. Specifically, only 43.7 percent of youth (aged 15 to 24 years) in low-income countries are recorded as employed compared to 71.3 percent of adults (ILO, 2022).[7] Although structural reasons for the wide employment gap between women and men as well as youth and adults exist, including structural change exacerbated by the COVID-19 pandemic and persistent gender norms (Klasen, 2019; Goldin & Mitchell, 2017; Goldin et al., 2017), this gap could also be driven by undermeasurement of women's and youth's labor market outcomes. In fact, existing evidence suggests that survey methods traditionally used to collect labor data may not adequately elicit accurate responses from some groups (Ambler et al. 2021; Bardasi et al. 2011; Arthi et al. 2018; Dillon 2012; Discenza et al. 2021). For example, since women tend to underreport their work activities, official labor data may provide inaccurate measurements of their work (Franck & Olsson, 2014; Muller & Sousa, 2020; Discenza et al., 2021).[8]

At least two explanations exist for why standard survey protocols used to collect data on employment potentially undermeasure women's and youths' labor market outcomes. First, standard labor modules in household surveys are often better at collecting information on formal activities and fail to properly capture informal activities that international guidelines, in fact, classify as work. In addition, respondents

[7] Some of the disparity on employment rates can be explained by youth enrollment in education.

[8] For example, Franck and Olsson (2014) find that women in Malaysia categorize their work activities as housework because social norms uphold that women are responsible for the home while men are accountable for earning income outside of the home. Similarly, Muller and Sousa (2020) use qualitative data to document that Honduran women tend to underreport the informal work that they do because they consider housework and outside employment mutually exclusive. In addition, Discenza et al. (2021) find that women in Sri Lanka are less likely to report their informal work without recovery questions—i.e., questions used to identify individuals who are employed but who were not classified as employed during the questioning—when responding to a labor module.

frequently do not consider their informal activities to be work, such as contributing to the family income by preparing food to sell or helping in a family-owned business. Second, standard data collection protocols for household surveys allow for proxy respondents to provide responses on behalf of other household members that cannot be interviewed. This practice may result in biased reports of labor indicators, particularly when the absent household member works in the informal sector and the proxy respondent may not be aware of the household member's labor activities. This would be the case, for example, for a male household head who reports information (as proxy) on behalf of his female spouse and/or child living within the household.

In this paper, we provide experimental evidence that demonstrates how survey methods can, to some extent, address the underreporting of labor market outcomes among women and youth in El Salvador. This is particularly relevant in a context such as El Salvador where gender and age-based gaps in labor market outcomes are more pronounced compared to economically similar countries, and informal work is prevalent (World Bank, 2023). Moreover, in El Salvador the gender gap in employment was 28 percentage points in 2022, which has remained stable for the past 24 years, and youth experience 2.4 times more unemployment than adults (World Bank, 2023).

We designed and implemented a methodological survey experiment to estimate the impact of two survey methods on reporting employment and work. These approaches included 1) providing concrete and applicable examples of work in a household survey so that respondents could clearly understand what constitutes work and 2) enforcing self-responses from all eligible respondents to participate in the survey. Our findings confirm that women's labor statistics vary when they are given examples that enable them to recognize how they also contribute to the working population. We also provide evidence between females and males that excluding proxy respondents increases the reporting of labor indicators of young men compared to older males.

We designed the survey experiment in different stages. First, we conducted focus group discussions to identify the most common activities performed in rural and peri-urban communities in two regions of El Salvador. In the second stage, we classified the activities described in the focus groups into three categories: agricultural work, production of goods to sell, and provision of services. We then used these categories to design a survey module called "List of Activities" (LOA), which provides region-specific examples of activities that are considered work or employment. From a cognitive perspective, this module is intended to help respondents recall the activities they had engaged in throughout the week before they respond to the standard labor module.[9] In the third stage, we randomly selected 48 communities (enumeration areas, EA) distributed evenly across rural and peri-urban areas of El Salvador, conducted a full household listing for each EA, and randomly selected a total of 1,008 households from these listings. Lastly, between 2 to 4 working-age household members (aged 15 to 64 years) were randomly selected from participating households in order to yield our study population of 2,480 respondents. Since the main hypothesis considers the differential impacts of these survey methods by age and sex, we stratified the study population by these two variables, where youth are defined as individuals between the ages of 15 to 24 years.

To estimate each survey method's causal effect on the respondents' reporting of employment and work, we randomly distributed the 1,008 households and their respective 2,480 working-age household members with equal probability across three groups. The first group was exposed to the LOA module before responding to the standard labor module and proxy respondents were allowed. The second group was not exposed to the LOA but self-responses were required from and all eligible

[9] As we explain in Section 3, the purpose of the LOA is to help respondents identify themselves as conducting doing the activities listed. This would help them recall the activities they engaged in during the past week so that they could report their work and employment information more accurately during the standard labor module. The data provided in the standard labor module are used to measure the main outcomes.

respondents—not proxies—for the standard labor module. The third group (C) was made up of respondents who were not exposed to the LOA and proxy responses were allowed to respond the standard labor module. To ensure that all the effects were driven by exposure to the two survey methods, all other survey protocols were kept exactly the same for all respondents across each of the three experimental groups. For example, the remaining survey modules; the maximum number of visits per household; the incentives; and all other survey protocols were also identical for the three groups.

We find that the average respondent exposed to the LOA module is 4.0 percentage points (6.9 percent of average employment in C) more likely to report being employed and 4.3 percentage points (6.3 percent of the average work in C) more likely to report working. Moreover, enforced self-reporting, rather than proxy reporting, has no statistically significant effect on either employment or work for the average survey experiment participant.

We then explored the survey methods' heterogeneous effects by sex and age. In line with our hypotheses, we find that providing a list of activities (included in the LOA module) is more effective at addressing underreporting bias on labor market outcomes for women relative to men. Women exposed to the LOA module are 8.2 percentage points more likely to report working relative to men who were exposed to the same module. Moreover, the LOA module increases the probability of women reporting to be employed and working by 6.8 and 8.1 percentage points, respectively, compared to the other women who were not exposed to the LOA (in C group). Lastly, and similar to the findings for women in Bardasi et al. (2011), we find no statistically significant differential effect by sex on the enforcement of the self-reporting labor information relative to proxy reporting.

In terms of heterogeneity by age, we document two main results. First, the survey methods by age have no differential effects on the reporting of labor outcomes. We only find that adults exposed to the LOA module are more likely to report working relative to the adults not exposed to the LOA. Second, we find that enforcing self-

reporting increases the share of young males who reported being employed compared to older males by 13.8 percentage points (19.1% of the C outcome mean). This result may reflect that young males are more likely to engage in work activities of which the other household members are not aware or may not consider as work. Moreover, and similar to Bardasi (2011), proxy reporting rather than self-reporting yields lower male working rates. We find no differential effects by age among the subsample of men when we included the LOA module.

Furthermore, we provide suggestive evidence that explains the observed differences in the survey methods' effects by sex and age. Specifically, we assess if the effectiveness of these survey methods can be determined by the prevalence of peers' informal employment or by social norms related to domestic obligations. Our findings suggest that women in communities where the prevailing norm emphasizes a domestic role are more likely to report employment or work when using the list of activities, compared to women in similar communities without using this list. Although our findings provide only suggestive insights, we believe they warrant further exploration, as they can serve as a starting point for new avenues of academic and policy-relevant research.

Our paper contributes to the literature in several important ways. First, we provide evidence on how different survey methods can affect labor market indicators, especially youth employment. The majority of published studies focus primarily on understanding how women underreport their labor outcomes when responding to traditional survey methods and how alternative survey methods improve women's reporting of labor indicators relative to men (Ambler et al. 2021; Arthi et al. 2018; Bardasi et al. 2011; Dillon et al., 2012; Discenza et al. 2021; Franck & Olsson, 2014; Kilic et al., 2021; Kilic et al., 2022; Muller & Sousa, 2020). Yet, there is a lack of evidence of how alternative methods affect the reporting of labor outcomes by other population groups such as youth who are more likely engage in informal activities that may not be

captured in standard labor modules.[10] In this sense, our paper is the first of its kind to rigorously demonstrate how survey methods can increase the reporting of employment and work of young males compared to older adults. In this sense, our study can inform data collection efforts to ensure that women's and youth's labor activities are well documented and represented statistically.

Second, we provide explanations that account for the large employment gap between women and men as well as youth versus adults by using alternative types of measurement to understand the different gaps today. Studies have documented that occupational segregation by sex, which is rooted in gender norms (Franck & Olsson, 2014; Muller & Sousa, 2020) and persists over time, contributes greatly to gender wage gaps (World Bank, 2011; Borrowman & Klasen, 2019). At the same time, the prevalence of informal work to which the respondent is exposed can bias the respondent's reporting of labor market indicators. Despite these findings, studies have yet to demonstrate how norms related to occupational segregation by sex and the incidence of informality influence how effective the different survey methods are. Thus, another novel and noteworthy contribution of our paper is its evidence for how exposure to gender norms related to housework and the prevalence of informality in the respondent's community along with preliminary participation in alternative survey methods can have heterogeneous effects on women's and youth's reporting of labor indicators.

2. Study Design

2.1 Randomized Methodological Survey Experiment

We designed and implemented a methodological survey experiment to test whether different survey methods capture work activities that respondents would not report

[10] 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.

when responding to a standard labor module.[11] In particular, we focus on women and youth's reporting of employment and work compared to men and adults, respectively, within the context of a household survey. We conducted this experiment in four stages described below.

Stage 1: Identify the most common work activities. Between June and July 2022, we facilitated eight focus group discussions among the rural and peri-urban residents of six municipalities[12] that cover two regions of El Salvador. The aim of the discussions was to identify the most common activities the members of these communities perform. We began the discussions by asking the participants the following question, "We would like to talk about the activities in which members of your community most commonly engage. Could you please describe to us the activities that are most prevalent in your community? You can write them down on the paper provided." The activities were grouped into two categories: income- and non-income generating. Within these two categories, the activities were further classified according to who performed them: women, men, or youth. Next, we discussed whether the employment and work-related activities were for pay, performed inside or outside the home, and their duration, among other things. Lastly, we asked respondents to revise and re-categorize the activity lists based on our discussion of which activities can be characterized as employment and which as work.

Stage 2: Design the List of Activities (LOA) module. In a second stage, we followed a frequentist approach and classified all of the activities into three categories: (i) agricultural and aquacultural activities (e.g., crop production, livestock care, fruit cultivation, fish farming, and fishing) to generate income and for personal consumption; (2) producing items to sell (e.g., pastries or jewelry); and (3) provision of

[11] The standard labor module follows the guidelines of the 19th International Conference of Labor Statisticians (ICLS) and collects respondent labor data for the previous 7 days.

[12] The municipalities within the Department of San Salvador are Nejapa, Guazapa and Jiquilisco, and the municipalities from the Department of Usulután are Puerto El Triunfo, San Dionisio, and California. More details on the focus group discussions can be found in Appendix A and on the selection of these municipalities in Section 3.

services (e.g., transportation, laundry, cleaning). Based on the information provided by the respondents, we then designed a List of Activities (LOA) module (see Table A1), which included a list of the most common work activities performed within a given region. The list aimed to provide a sufficient number of examples of income-generating activities that can be classified as work for statistical purposes. The objective of the LOA module was not to collect data to measure the main outcomes of the paper, but to help respondents better recall the activities in which they had engaged during the previous week so that they could report their employment information in the standard labor module more accurately. We then used the latter module as our data source to measure the main outcomes. We discuss this process in more detail in Section 3.

Stage 3: Select households and respondents. The experiment was conducted in two regions within the Departments of San Salvador and Usulután (see map in Figure A1). We selected these regions in order to achieve high variation in the types of activities reported. Region 1, which is located in the Department of San Salvador, includes households located close to the Metropolitan Area of San Salvador, the main economic activities of which are commerce, construction, manufacturing, and sugar cane production. Region 2 is situated in the Department of Usulután, where the main economic activities include coffee production, fishing, and commerce. Using the 2007 Salvadoran Population Census (i.e., the most recent census available), we identified 276 enumeration areas (EAs): 114 in Region 1, and 162 in Region 2. Each of these Eas had at least 30 households and was classified as either rural or peri-urban. From the 276 Eas, we randomly selected 48 that were evenly distributed across the rural and peri-urban areas and then conducted a full household listing for each EA.[13] We then randomly selected 21 households within each EA for a total of 1,008 households to constitute our study sample. Within each household, we randomly chose a minimum of two and up to a maximum of four working-age household members aged of 15 to 64 years, which

[13] An enumeration area (EA) is the smallest statistical sample unit in the census. Larger geographic areas are divided into smaller EA units, which make it easier to organize and facilitate census activities.

yielded a total of 2,480 working-age individuals. Since our main hypothesis consists of the differential impacts of these survey methods by age and sex, we stratified the selection of the working age household members by these two variables, and we defined youth as individuals aged 15 to 24 years.

Stage 4: Randomly assign households to experimental groups. We conducted a stratified randomization at the EA-level to assign the 1,008 households with equal probability to one of the following treatment groups (Figure 1):

T1. *LOA Module group.* A group of households were randomly assigned to participate in the LOA module before responding to the standard labor module. This treatment arm permitted proxy household respondents who provided information about the labor activities of the main respondent when latter was not available for the interview.

T2. *Enforced self-reporting group*. The main respondents who belonged to the households that were randomly assigned to this group were required to respond to the standard labor module (i.e., proxy respondents were not permitted) without previous exposure to the LOA module. We asked the interviewers to visit households up to five times, if necessary, in order to speak directly with the selected household member, and not a proxy.

C. *Control group.* Individuals living in the households randomly assigned to this group were interviewed using the standard labor module without asking to the LOA module. When the selected household members were not available to respond to the standard labor module, proxy respondents were permitted to respond on their behalf.

To avoid contamination across the groups, all individuals within a particular household were assigned to only one treatment status. To ensure that all the effects were driven by being exposed to the different survey methods described above, all other survey protocols were exactly the same for all respondents across the three experimental groups. In addition, we made sure that all respondents completed the same number of additional modules on demographic and socio-economic data at the

individual- and household-level. Further details regarding data collection are provided in Section 3.

3. Data

3.1 Data Collection Activities and Survey Instrument

To collect the relevant data for our experiment, we surveyed the sample of 1,008 households (2,480 household members) between August and October 2022.[14] The interviews were conducted using a Computer-Assisted Personal Interviewing (CAPI) program, in which we embedded the household assignments to the treatment groups.

The survey instrument included up to 18 modules (see Table A2 in the Appendix). The first module collected information from the household roster to identify all household members. After completing the household roster, the CAPI program randomly selected the working-age household members (aged 15 to 64 years) who would participate in the experiment. Each randomly selected member (or the member's proxy respondent for T1 or C) provided information solicited via the three main modules: education, access to technologies, and labor. Participants in T1 also responded the LOA module. We made sure that the three relevant modules were collected at the beginning of the experiment and from all of the randomly selected self- or proxy respondents in the three treatment groups.

3.2 Outcomes

We measure the effects of these alternative survey methods on employment and work as outcome variables. For this paper, the working-age population includes individuals aged 15 to 64 years.

Employment: We use data collected from all study participants via the standard labor module. According to the 19th International Conference of Labor Statisticians (ICLS), being employed is defined as working for pay or profit. Individuals who have engaged

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

in work only for their own consumption in the past week are not considered employed. We define this variable as a dummy indicator that takes the value of 1 if the person engaged in at least one activity that that generates income in the last week, and 0 if the person reports either not engaging in any income generating activity or engaging only in activities intended for his/her own consumption during the past week.[15]

Work: A person considered working must be involved in producing goods or providing services (whether paid or unpaid) in the last week. The activities could be for personal use (e.g., farming produce for household consumption without the intention to sell) or for use by others. We define this variable as a dummy indicator that takes the value of 1 if the person reported engaging in at least one income-generating activity during the previous week (e.g., farming or working for pay in a family- or non-family-run business).

3.3 Summary Statistics

Table 1 reports the summary statistics of the variables used in our analysis.[16] Panels A and B show descriptive statistics of the characteristics of the 1,008 participating households and 2,480 enrolled household members, respectively. As we show in Panel A, the average household included in our sample has 3.7 members, which is consistent with the average household size (3.3 members) in El Salvador (EHPM, 2022). Moreover, for every working-age adult within the household, there are approximately 0.27 dependents—children under the age of 14 or elderly aged 65 years and older. In addition, half of the households in our sample are located in rural areas, which is consistent with the fact that 53 percent of these households have an asset index below the median, and approximately one-third of these households reported experiencing moderate to severe food insecurity. In terms of sources of income, about one-third of the households receive remittances, and about one-third of the household members work by doing non-farming related activities. Lastly, 54 percent of the household

[15] Durazo et al. (2021) elaborate on these definitions of employment and work.
[16] Appendix C includes definitions for each of these variables.

members are engaged in agricultural activities, including growing and harvesting crops, raising livestock and/or fish, and fishing.

Panel B shows that the average participant is 35 years of age; 55 percent of the participants are women; and 34 percent of the respondents have not been married. In terms of human capital, 28 percent of the respondents have completed high school or higher education, and 87 percent can read or write. Among the participants, 88 percent report having access to at least one mobile phone, but only 19 percent have access to the internet at home.

To address how our study sample compares to the average household in the regions surveyed we compare the average characteristics of our respondents to the average Salvadoran respondents of the 2022 Household and Multipurpose Survey (EHPM). Specifically, we compare the variables of the EHPM that were measured similarly and in the same two regions as in our survey. We present this comparison in Table A3. Overall, we find that the two samples are similar in terms of age, sex composition, human capital, and marital status, all of which confirm that the average household members who participated in our experiment are similar to the average adult respondent of the EHPM sample.

3.4 Validity of the Experiments

To support the validity of our identification assumption in the survey experiment, we compare average household and individual characteristics across the three groups in our study. We present these results in Table 2. Columns (1) to (3) present the mean of the control variables, T1 and T2, respectively. Columns (4) through (6) provide *p*-values for t-tests for equality of means between T1 and C (Column [4]), T2 and C (Column [5]), and T1 and T2 (Column [6]). Apart from three of the 15 variables tested, we do not find significant differences in these variables across the treatment and control groups at *p*-values less than 0.1. We find that households in T1 are less likely to work in agriculture (49 percent in T1 versus 56 and 57 percent in C and T2, respectively; $p = 0.066$ and $p = 0.035$). Moreover, T2 respondents are, on average, 1.4 years older than T1 and C

respondents (relative to an average age of 35 years; p = 0.014 and p = 0.018) and are between 3-4 percentage points (pp, out of 89 pp) less likely to know how to read and write (p = 0.066 and p = 0.023). Lastly, T1 respondents are 3 pp less likely than C respondents to have access to a mobile phone (relative to 86 pp in T1 and 89 pp in C). Statistically, however, this is not surprising when testing across many variables and different groups. Overall, the differences in the respondents' average age, reading and writing skills, and access to a mobile phone are relatively small. Nevertheless, to account for these differences, we control for these variables as a robustness check, as we explain in Section 4.

Since households were randomized within EAs, one concern is that intra-cluster contamination could occur, whereby the response of a participant assigned to one treatment arm could affect the response of another respondent assigned to a different treatment arm within the same cluster. To mitigate this issue, we collected all surveys simultaneously within the same EA. In addition, intra-cluster contamination occurs more often when multiple rounds of data are collected. In our case, the EAs were large, and the experiment consisted of only one survey.

Lastly, another concern is that substitute households within each EA would be different compared to original households that were not replaced. As we show in Table A4, despite some statistically significant differences between these two groups, these differences are nevertheless very small. For example, original households had 0.2 household members more than substitute households (a 6 percent 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 percent of original households were in a rural area relative to 61 percent of substitute households, which were located in peri-urban areas. 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

below, 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.

4 Empirical Strategy

Random allocation of households to T1, T2, and C in the survey experiment provides for exogenous variation in the treatment status that allows us to estimate the LOA module's and enforced self-reporting's effects on the reporting of labor market outcomes. We report treatment effect estimates based on the following specification:

$$Y_{ihs} = \alpha_0 + \beta_1 T1_{hs} + \beta_2 T2_{hs} + \Theta X_{ihs} + b_s + \epsilon_{ihs} \qquad (1)$$

where Y_{ihs} is the outcome of interest (employment or work) for individual *I*, in household *h,* in enumeration area *s*. $T1_{hs}$ is a dummy indicating that household *h* in stratum *s* was randomly chosen to respond the LOA module before providing labor information. $T2_{hs}$ consists of a dummy indicating that the members of household *h* were required to complete the labor module themselves (rather than by proxy). X_{ihs} is a vector of control variables. To select the variables included in this vector, we control only for variables that are thought to influence the outcomes of interest (Bruhn & McKenzie, 2009). Since we collected several such variables, we use a Double-Least Absolute Shrinkage and Selection Operator (LASSO) regression approach to select the variables that are more relevant for each outcome (Urminsky et al., 2016). We include all of the individual and household-level characteristics presented in Table 2 in the LASSO specification.[17] To test for the robustness of the estimated coefficients, we also estimate Equation (1) excluding the vector of control variables, controlling for those variables for which we have differences in means across the three treatment groups, and controlling for the variables for which there are differences in means between the original and substitute households. b_s captures stratum fixed effects for the EA, which consist of the

[17] Table A5 presents the list of control variables that LASSO selected for each outcome and model.

interaction between the department (San Salvador or Usulután) and the area of residence (rural or peri-urban).

Lastly, ϵ_{is} is an individual error term. Since treatment status was clustered at the household level, we estimate clustered standard errors. We also take a more agnostic approach to the structure of the standard errors and use the randomization inference approach to estimate standard errors (and the respective *p*-values) (Gerber and Green, 2012). The estimate of $\widehat{\beta_1}$ captures the ITT effect of being exposed to the LOA module before completing the standard labor module. Similarly, $\widehat{\beta_2}$ captures the ITT effect of enforced self-reporting when responding to the standard labor module.

Our main hypothesis is that, by randomly assigning households to either T1 (the LOA module) or T2 (enforced self-reporting), we could determine the differences in the reporting of our outcomes of interest by sex and age. Thus, we estimate heterogeneous effects by sex and age by modifying Equation (1) incorporating an interaction between each treatment indicator Tj_{hs} and a dummy D_{ihs} that specifies whether individual *i* is a woman or a youth, separately. We use one model for sex and one model for age as follows:

$$Y_{ihs} = \alpha_0 + \beta_1 T1_{hs} + \beta_2 T2_{hs} + \beta_3 T1 \times D_{ihs} + \beta_4 T2 \times D_{ihs} + \Theta X_{ihs} + b_s + \epsilon_{ihs} \quad (2)$$

where X_{ihs} includes the indicator D_{ihs} and the list of control variables selected using a double-LASSO procedure for each outcome. All estimations contain EA (stratification variable) fixed effects. We report the estimated coefficients from Equation (2) and the estimations of the total effect of each survey method ($\beta_1 + \beta_3$ for the LOA treatment and $\beta_2 + \beta_4$ for the enforced self-responses) in the Results section.

5 Results

Our pre-analysis plan guided our empirical work. In this sense, we examine the implications of different survey methods on average employment and work. Then, we investigate the differential effects of the survey methods by sex and age. If the LOA

module and self-reporting are effective, then we expect that a higher proportion of women and youth will report work or employment.

5.1. Survey methods' impacts on the measurement of labor market outcomes.

To examine whether the LOA module and the enforced self-reporting survey method increase employment and work on average, we estimate Equation (1) for each of these two outcomes. Table 3, Column (1) shows the different treatments' effects on employment, and Column (2) shows their estimated impacts on the indicator of whether the individual is working. We document two main results. First, the LOA module is effective in increasing this experiment's average participant's reporting. We find that, if participants responded to the LOA module (T1) before the standard labor module, then they were 4.0 percentage points ($p<0.1$) more likely to report being employed compared to the average participant in C. This estimated effect is equivalent to 6.9 percent of the average employment of C. Furthermore, the probability of reporting working is 4.3 percentage points higher when the respondent is exposed to the LOA ($p<0.05$; 6.3 percent of the average work rate for C). Second, we also show that—for the average participant—enforced self-reporting has no statistically significant effect on employment or work relative to proxy reporting.

5.2. Survey methods' heterogeneous effects on labor market outcomes

Since evidence has shown heterogeneity by sex, we subsequently explore the survey methods' heterogenous effects by sex and age. Bardasi et al. (2011) show that response by proxy rather than self-reporting has no effect on female labor statistics but yields substantially lower male employment rates. Yet, evidence of the differential effects of survey methods by the respondent's age is still lacking. To better understand the survey methods' differential impacts by age and sex on employment and work in our setting, we conducted a heterogeneity analysis by sex (female) and age (youth, ages 15 to 24 years) using Specification (2). We report the results in Table 4. Columns (1) and (2) depict the heterogeneity results of each survey method on employment and working

status, respectively, by sex, whereas Columns (3) and (4) present the heterogeneity estimates of each survey method by age.

In line with our hypothesis, the LOA module is more effective at addressing underreporting bias on labor market outcomes for women relative to men. Women exposed to the LOA module are 8.2 percentage points more likely to report working (p<0.05) relative to the men exposed to the module. Although the coefficient for the effect on reporting of employment is not statistically significant at the conventional level, the magnitude of the differential effect of the LOA module on women's reporting of employment relative to men is sizeable and close to 6.3 percentage points. Moreover, we also document that the LOA module increases the probability of women reporting to be employed and working by 6.8 and 8.1 percentage points (p<0.05), respectively, compared to other women in C. Lastly, and similar to the findings for women in Bardasi et al. (2011), we find that enforced self-reporting has no statistically significant differential effect by sex in the reporting of labor information relative to proxy reporting.

In terms of heterogeneity by age, we show in Table 4, Columns (3) and (4) that the survey methods have no differential effects on employment and work. We find only that those exposed to the LOA module are 4.3 percentage points (p<0.1) more likely to report working compared to adults not exposed to the list. As we explain below, this result is driven by older women.

We also note that the magnitudes of the estimated coefficients of the differential impact of enforced self-reporting on working status and employment between youth and adults are relatively large, between 6 and 7 percentage points, respectively. These findings, however, are not statistically significant at the conventional level. Since these coefficients might capture both age and sex simultaneously,[18] we separate the samples by sex and estimate the differential impacts of each survey method by age using the

[18] For example, enforced self-reporting might be more relevant for male youth compared to young women.

subsamples of women and men. These results can be found in Table 5. Columns (1) and (2) present the age-related heterogeneity results based on the subsample of women, while Columns (3) and (4) show the heterogeneity impacts by age using the subsample of men.

Taking stock. Our results show that the survey methods' effects on reporting labor market outcomes depend on the respondent's sex and age. Specifically, the LOA module increased the average reporting of employment and work relative to the other two groups. These effects seem to be driven by women assigned to the LOA module relative to men who were also exposed to the LOA module and to other women who did not participate in the LOA. This last result seems to be driven by the subsample of adult women (older than 25 years of age). Although enforced self-reporting has no effects on the average reporting of employment and work, it is effective at increasing the reporting of these outcomes by young males (aged 15 to 24 years) relative to older males also randomly assigned to T2. In fact, proxy reporting yields lower male working rates for older men. Overall, our results indicate the importance of exercising caution when using different survey methods across diverse populations.

6 Mechanisms

In this section, we present further suggestive evidence to explain the survey methods' observed differences in effects by sex and age. Although our findings are suggestive, we believe they warrant further exploration, since they can serve as a starting point for new avenues of academic and policy-relevant research.

6.1 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 percent of women are engaged in informal employment compared to 87.5 percent of men. In El Salvador, the trend is similar: 72 percent of women are engaged in informal work relative to only 66 percent of men (El Salvador Central Bank, 2022). Since measurement of informal work or employment is likely to be biased when using traditional labor

modules or when permitting proxy reporting, we explore whether the impacts of alternative survey methods on the reporting of employment and work differ by the prevalence of informal work to which the individual is exposed.

To test our hypothesis, we first estimate a measure of prevalence of peers' informal work at the community level to which each respondent is exposed. This measure is defined as the share of employment in the informal sector[19] in each EA by sex, excluding the respondent's employment status. By using the average at the EA-level and excluding the respondent, we partially account for potential endogeneity. Within our sample, the average informal employment to which the respondents are exposed at the EA-level is 76 percent; with an average of 70 percent for men and 84 percent for women (Table A6). To ease the interpretation of these results, we standardize this measure, where 1 SD in prevalence measure equals 11 percentage points. We then interact this measure of prevalence of informal employment with each of the treatment indicators and estimate the results separately by sex and by age. In this sense, the result from the estimation should be interpreted as the differential effects of the survey methods by living in a community with certain level of informal employment among peers.

Results from these estimations for the samples of women and men are presented in Table 6. We find that the LOA module increases women's reporting of employment and work relative to other women who did not participate in this module, irrespective of the level of their peers' informal employment in the community. Furthermore, within the sample of men, we observe that the measurement of work and employment using the LOA module is not affected by the exposure to informal employment. We also interact Equation (2) with an indicator of female to confirm our previous results and to

[19] Employment in the informal sector is defined using the characteristics of the entity, enterprise, or employer of the employed population. A person is employed in the informal sector if any of the following conditions are true: a) the employer does not keep accounts or only keeps accounts for personal use; b) the employer is not registered at the national level; c) the employer is unincorporated (business and owner are not separate legal entities); d) the place of work is without fixed premises; and e) the employer has fewer than five employees.

test for differences in the survey methods' effects and prevalence of informal employment by sex. As we show in Figure 1 and Table A7, we find that enforced self-reporting and the LOA module have no differential impacts on women's and men's levels of work and employment exposed to similar levels of informal employment, which confirms our findings for the sample of women reported in Table 6, Columns (1) and (2).

We further explore whether the incidence of informality in the community determines the effects of self-reporting survey method (T2) in the sample of men. Table 7 presents these results. We first document that the incidence of informality in a community affects the effectiveness of using self-reporting survey method in capturing youth employment. Youth assigned to T2 and living in communities with large levels of youth informal employment are 14.2 and 10 percentage points more likely to report being employed and working respectively, than other youth assigned to T2 and exposed to lower levels of informal peer employment. For every 1 SD in the informality measure, young men in T2 living in highly exposed communities are 14.2 percentage points (p-value<0.05) more likely to report being employed and 10 percentage points (p-value<0.10) more likely to report working.

At the same time, in settings with high exposure to informal peer employment, enforced self-reporting increases the reporting of labor outcomes among youth relative to proxy respondents. For an average of 1 SD of peer informal employment, young men who were required to respond themselves to the labor module were more likely to report that they were employed (22 percentage points; p-value<0.05) or that they were working (15.3 percentage points; p-value<0.10) relative to other respondents among who self-reporting was not enforced (C). In line with this result, we show that, in contexts of high incidence of informal youth employment, self-reporting survey method reduces the reporting of employment relative to proxy reporting in the sample of adult men. This, then, suggests that proxy respondents can be biased by the level of informal employment within the community; for, proxy respondents exposed to high youth informal employment end up underreporting the labor indicators of their young

household members, whereas those exposed to high levels of adult informal employment overreport the labor indicators of the adults living in their household.[20]

6.2 Social norms around gendered work activities

Occupational and sectoral separation by sex is remarkably persistent over time (World Bank, 2011) and a major contributor to gender wage gaps. The prevalence of working women and men in particular occupations that classify jobs by type and skill and in sectors of the economy differs greatly (Borrowman & Klasen, 2019). For example, women are more likely to be engaged in a wider range of activities compared to men, such as unpaid domestic work or caregiving responsibilities, which traditional survey methods may not capture effectively. As evidence shows, this gendered separation of work is rooted in social norms: omen frequently classify their work activities as housework to conform to social norms and designate their husband as the primary breadwinner. (Franck & Olsson, 2014).

Considering existing evidence that women underreport their labor market outcomes due to social norms, especially in contexts where women devote a large portion of their time to domestic obligations, we test whether the survey methods, especially the LOA module, offer the female participants to this study a more comprehensive framework of occupations, enhancing their ability to report work activities accurately.

Specifically, we estimate a measure for the respondent's exposure to social norms associated with domestic obligations. This measure consists of the average amount of time in hours that the women—excluding the respondent—in each EA devote to domestic duties, which we obtain from the main survey's time use module.[21] By using the average time at the EA-level and excluding the respondent, we partially

[20] Figure 3 and Table A8 present the results of interacting Specification (1) with the measure of prevalence and age for the sample of men. Overall, our findings confirm the results obtained when separating the sample into groups of young males and old males.

[21] Table A6 in Appendix A documents descriptive statistics on the average number of hours women devote to domestic duties.

account for potential endogeneity. We then interact this average time with each treatment indicator (similar to Equation [2], but where D_{ihs} is the average time in this exercise) and estimate the results separately for the sample of men and women. Thus, we can interpret the estimation result as the socially acceptable amount of time women are expected to spend on housework, determined by the gender norms prevailing within their community.

Results from this heterogeneity analysis are presented in Table 8. Columns (1) to (2) and (3) to (4) present the results for the sample of women and men, respectively. First, we demonstrate that, for every hour that other women from the community dedicate to domestic duties, female respondents are 5.9 percentage points more likely to report being employed compared to other women in similar communities who did not participate in the LOA module. Similarly, women who participated in the LOA module and are subject to higher expectations regarding time spent on housework (more hours) are more likely to report working relative to other women living in communities with a similarly high average time spent on housework but who were not exposed to the LOA module. Second, although the estimated difference in the reporting of work between women exposed to high and low average times spent on housework within their communities is not statistically significant at the conventional level, it is worth noting that the magnitude of the difference is large (8.4 percentage points). This, therefore, indicates that time spent on housework is a potentially relevant mechanism that requires further exploration with a larger sample size. Third, we show that the survey methods have no differential effects in the sample of men. Lastly, to account for a potential reduction of statistical power due to the separation of the sample by sex, we interact Equation (2) with an indicator for female. These results are presented in Figure 4 and Table A9. Similar to our findings when separating the sample by sex, we find that women living in communities where other women devote more time to domestic duties are 6.3 and 6.9 percentage points more likely to report being employed or working, respectively, relative to other women living in similarly exposed communities but who did not participate in the LOA module.

7 Robustness Checks and Additional Results

7.1 Assessing Sensitivity from Selection of Control Variables.

As we discussed in Section 4, we use a double LASSO approach to determine the control variables for our estimations. For each main outcome, we test for the stability of our estimated coefficients after excluding the control variables that LASSO selected. As we show in Tables A10-A12, the estimated coefficients and their statistical significance are stable for most of the results after excluding all control variables selected by LASSO and included in Specification (1).

7.2 Randomization inference.

We take a more agnostic approach to the estimation of standard errors and the respective *p*-values by estimating them using randomization inference. This approach allows us to assign the standard error to a given treatment effect by observing where that treatment effect falls within the distribution of all possible estimated effects from the 1,000 randomizations that we simulated under the assumption of no effects (Blattman et al., 2021). In our setting, randomization inference may be preferable to clustered standard errors because our experimental design is a clustered randomization, and the sample size is small. We present the estimated randomization inference *p*-values in each table below the estimation of the standard errors and draw the same inferences from using either randomization inference or clustering in the estimation of standard errors.

7.3 Addressing differences in household characteristics due to substitution

As mentioned before, we find statistically significant differences in the mean of some baseline variables when we compare original and substitute households. To account for these differences, we add these variables as controls (in addition to the variables selected by LASSO) and summarize the results in Tables A13-A15. Overall, we do find that our results are robust to controlling for the variables for which we find differences in the substitute households.

7.4 Estimating the effects on a placebo variable

To address the concern that participants reported improvements in all labor-market outcomes after being exposed to the two alternative survey methods, we use the variable employment in the formal sector as a placebo outcome. We argue that including the LOA module should not affect the reports of this variable because the activities included in this module are usually performed in the informal sector. Moreover, given the definition of employment in the formal sector, it is unlikely that a proxy respondent would not be aware of that a household member is engage in those types of activities.[22] Considering this, we use Specifications (1) and (2) to estimate the survey methods' effects on measures of formal employment and wages. We show in Tables A16 – A18 in the Appendix that most of the main results are not statistically significant at the conventional levels. If anything, we only observe a positive effect of T2 for the subsample of women when comparing young women versus older women exposed to that treatment.

7.5 Associations between employment and work and characteristics of proxy respondents.

To explore how proxy's characteristics can affect the reporting of employment or work for others household members, we restrict the sample to the groups where proxy respondents were allowed, that is, for those assigned to the LOA module and to the control group. We focus our analysis in three characteristics of the proxy respondents: sex, education level (high school or more), and relationship with the household member for which they are reporting (spouse). Unlike Bardasi et al. (2011), we did not randomly select the proxy respondent. In this sense, our results can only be interpreted as associations between these proxy characteristics and the reporting of the two main outcomes. The results of this analysis are presented in Table A19. We find that proxies who are better educated are more likely to underestimate employment of women and to increase the reporting of both employment and work for men. Moreover, proxies who

[22] We acknowledge that using employment in the formal sector as a placebo outcome for T2 is more nuanced. In this sense, we are more cautious in the interpretation of that result in particular.

are male spouses are more likely to increase the reporting of employment and work for women and female spouses are more likely to increase the reporting of employment of their male spouse. No statistically significant associations were found in the sex of the proxy respondent. In sum, selecting the spouse as proxy respondent may offer some advantages in the reporting for both men and women. However, selecting the most educated proxy respondent may be good for men but can increase the underreporting of labor market outcomes for women.

7.6 Assessing the effects in reporting employment across economic sectors.

According to Bardasi et al. (2011), survey methods can change the reporting of employment and work in different economic sectors. In this sense, we explore if the increase in reporting in employment and work that we observe in our main results is driven by changes in the reporting of employment in one sector relative to no work (extensive margin) or relative to other sectors (i.e., intensive margin). Since the main economic activity in the regions under analysis is agriculture, we use this economic sector as reference and compare the reporting of employment in this sector relative to manufacture, services, and commerce as the additional sectors. These results are presented in Table A20 in the Appendix. We document three main findings. First, the LOA module has no effects on changes in reporting of employment across economic sectors. Second, enforcing self-responses can increase the reporting of employment in agricultural activities relative to being employed in any other economic sector by 3.8 pp. Third, the increase in reporting of employment in agriculture relative to other sectors is driven by women assigned to T2, relative to men treated in T2 (for services and manufacture, by 13 and 15 pp, respectively) and relative to other women in the control group (13.1 pp for services and 15.4 pp in manufacture). In this sense, enforcing self-responses can help to increase the accuracy in the measurement of the economic sectors where respondents are employed, particularly for women.

8 Cost Implications of the Survey Design

As highlighted in Bardasi (2011), different survey methods may have different implementation costs. We use the cost data related to this randomized survey experiment to estimate the cost implications of each treatment arm. The households' members assigned to T1 underwent a marginally longer interview since they participated in the additional LOA module. It took participants an average of 3.1 minutes to complete this module. Each household assigned to T1 had on average, 2.5 members; therefore, the additional interview duration per household is 7.8 minutes. Since the average total cost per interview in C was US$135.4 and the average duration was 111.6 minutes, the cost for the additional time required to complete the LOA module is US$9.44 (US$1.21/minute x 7.8 minutes), which corresponds to 6.9% of the total cost per interview for group C. In light of these calculations, the cost implications of adding this module are negligible.

The cost implications of T2 are not small compared to C. Although the average duration of the survey was similar between enforced self-reporting and proxy reporting, the additional costs of T2 are due to the fact that interviewers had to pay more visits to T2 households to ensure that they located all the respondents. Households in C and T1 needed an average of 1.50 and 1.49 visits, respectively, to collect the necessary information from all the respondents. However, households assigned to T2 required an average of 1.94 visits to complete the interviews. Since the average cost per visit per household was US$90.27, it costs 29.6 percent more to complete an interview in the T2 group compared to C. We do note, however, that these costs are lower than estimates for other settings. For example, Bardasi (2011) calculated that their interviewers needed at least one extra day to track down the respondents, which led to a 50 percent increase in cost.

In sum, researchers, survey practitioners, and policy makers must evaluate the tradeoffs between increased reporting of labor outcomes of youth relative to older adults when self-reporting is enforced versus the additional cost that this survey

method would entail. If nationally representative household surveys permit proxy reporting for cost reasons, then further research is needed to assess the implications of this survey method on welfare indicators in other countries.

9 Conclusion

Labor market policies that address existing inequalities in access to and quality of jobs are needed more than ever, but especially during global crises (UN, 2023). As existing data reveal, a fundamental challenge when drafting such policies is the fact that measurements of female and youth labor market outcomes are biased compared to measurements of these same outcomes for men. In this paper, we argue that part of the gap in employment and work between women and men as well as between young and older adults can be attributed to undermeasurement bias—that is, traditional survey methods might not be sensitive enough to measure these outcomes as accurately for women and youth as they do for men. This paper aims to contribute to the literature on the sensitivity of survey methods by analyzing the data of a randomized survey experiment conducted in urban and peri-urban regions in El Salvador. The study design made it possible to compare employment and work estimates obtained from a traditional labor module, which was embedded into a household survey, by means of two variations in the collection of these outcomes: the LOA module and enforced self-reporting.

In addition, this paper illustrates the effectiveness of our novel survey methods in the context of the prevalence of informal employment and gender norms. We find that the LOA module increases women's reporting of labor outcomes relative to other women who were not exposed to this module, irrespective of the level of peers' informal employment within their communities. On the other hand, exposure to informality can influence the effectiveness of self-reporting survey method (compared to proxy reporting) in the sample of men. We further observe that proxy respondents in communities with high levels of youth informal employment tend to underreport the labor indicators of their young household members. Conversely, those exposed to high

levels of adult informal employment tend to overreport the labor indicators of their adult household members.

Lastly, our results indicate that women living in communities where the traditional division of roles still persist, are more likely to report being employed or working when they use the list of activities compared to other women living in similar communities but have not used this list. In certain cultural or societal contexts, like El Salvador, unpaid care work might not be explicitly acknowledged by women as work, leading to a lack of recognition for the time and effort invested in these activities. This perception can have implications for how the value of unpaid work is understood, addressed, and integrated into discussions on labor, gender roles, and societal contributions. The list of activities can help women to recognize unpaid work as a form of work.

This paper's findings offer lessons for researchers, survey practitioners, and policy makers responsible for producing national statistics. The impact of various survey methods on labor market bias is contingent on the sex and age of the respondent. Moreover, given the differences in the budget implications between using the LOA module or self-reporting survey methods, it is crucial to analyze the trade-offs between collecting data from the actual respondents and the additional cost to do so. In the case of this experiment, enforced self-reporting has a clear impact on the labor indicators of young males that are not captured via traditional household surveys that permit proxy reporting. Yet, this statistical benefit comes at a 30 percent higher implementation cost. Conversely, the LOA module has a significant impact on capturing higher levels of work and employment and its implementation costs are negligible.

References

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.

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.

Carson, R. T., & Groves, T. (2007). Incentive and informational properties of preference questions. *Environmental and resource economics*, *37*, 181-210.

Datta, N. (2019). Willing to pay for security: a discrete choice experiment to analyse labour supply preferences.

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

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.

Eriksson, T., & Kristensen, N. (2014). Wages or fringes? Some evidence on trade-offs and sorting. *Journal of Labor Economics*, *32*(4), 899-928.

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.

Gaddis, I., Oseni, G., Palacios-Lopez, A., & Pieters, J. (2023). Who Is Employed? Evidence from Sub-Saharan Africa on Redefining Employment. *Journal of African Economies*, *32*(2), 151-174.

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.

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

Loomis, J. (2011). What's to know about hypothetical bias in stated preference valuation studies? *Journal of Economic Surveys*, *25*(2), 363-370.

Mas, A., & Pallais, A. (2017). Valuing alternative work arrangements. *American Economic Review*, *107*(12), 3722-3759.

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

Ohnsorge, F., & Yu, S. (2022). *The long shadow of informality: Challenges and policies*. World Bank Publications.

Reuben, E., Wiswall, M., & Zafar, B. (2017). Preferences and biases in educational choices and labour market expectations: Shrinking the black box of gender. *The Economic Journal*, *127*(604), 2153-2186.

Train, K., & Weeks, M. (2005). *Discrete choice models in preference space and willingness-to-pay space* (pp. 1-16). springer Netherlands.

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

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 of Household and Individual Characteristics

Variable	(1) Mean	(2) SD	(3) Min.	(4) Max.
PANEL A. Household Characteristics				
Household size (N)	3.71	1.44	1	14
Dependency ratio (%)	0.27	0.20	0	0.8
Households in rural area (%)	0.50	0.50	0	1
Households with assets index above the median (%)	0.47	0.50	0	1
Moderate to severe food insecurity (%)	0.36	0.48	0	1
Remittances received (%)	0.35	0.48	0	1
Households has a non-farm enterprise (%)	0.33	0.47	0	1
Households engages in agricultural activities (%)	0.54	0.50	0	1
Observations	**1008**			
PANEL B. Individual Characteristics				
Female (%)	0.55	0.50	0	1
Age (years)	35.49	13.52	15	64
High school or higher education (%)	0.28	0.45	0	1
Read and write (%)	0.87	0.33	0	1
Never married (%)	0.34	0.47	0	1
Access to mobile phone (%)	0.88	0.32	0	1
Access to Internet (Wi-Fi) (%)	0.19	0.39	0	1
Observations	**2480**			

Notes: This table shows the average summary statistics of the household- and individual-level variables—Panels A and B, respectively—used in our analysis. Columns (1) and (2) present the mean and standard deviation (SD) for each characteristic of the participants in our sample. Columns (3) and (4) show the minimum and maximum values of the characteristics, respectively. These variables were collected using the following modules: sociodemographic characteristics, education, food security, household characteristics, and assets. Appendix C includes 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. Balance Tests

Variable	(1) Mean C	(2) Mean T1	(3) Mean T2	(4) *P*-value T1 vs. C	(5) *P*-value T2 vs. C	(6) *P*-value T2 vs. T1
PANEL A. Household Characteristics						
Household size (N)	3.723	3.762	3.643	(0.748)	(0.519)	(0.256)
Dependency ratio (%)	0.268	0.253	0.274	(0.274)	(0.665)	(0.212)
Households in rural area (%)	0.500	0.500	0.500	--	--	--
Households with assets index above the median (%)	0.461	0.491	0.461	(0.476)	(1.000)	(0.449)
Moderate to severe food insecurity (%)	0.372	0.354	0.366	(0.637)	(0.883)	(0.722)
Remittances received (%)	0.324	0.369	0.366	(0.213)	(0.201)	(0.934)
Household has a non-farm enterprise (%)	0.333	0.327	0.315	(0.856)	(0.625)	(0.753)
Household engages in agricultural activities (%)	0.560	0.491	0.574	(0.066)	(0.683)	(0.035)
Observations	336	336	336			
PANEL B. Individual Characteristics						
Female (%)	0.549	0.539	0.559	(0.494)	(0.512)	(0.195)
Age (years)	35.006	35.032	36.479	(0.990)	(0.014)	(0.018)
High school or higher education (%)	0.282	0.296	0.275	(0.610)	(0.705)	(0.400)
Read and write (%)	0.878	0.888	0.847	(0.638)	(0.066)	(0.023)
Never married (%)	0.333	0.351	0.320	(0.433)	(0.636)	(0.241)
Access to mobile phone (%)	0.892	0.864	0.888	(0.097)	(0.652)	(0.180)
Access to Internet (Wi-Fi) (%)	0.164	0.193	0.198	(0.241)	(0.110)	(0.721)
Observations	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 (3) present the mean for each characteristic of households and individuals assigned to the control group, T1, and T2, respectively. Columns (4) to (6) 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 C 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. *** p<0.01, ** p<0.05, * p<0.1

Table 3. Survey Methods' Impacts on the Reporting of Labor Market Outcomes

	(1) Employed	(2) Working
T1	0.040*	0.043**
	(0.023)	(0.022)
	[0.087]	[0.045]
T2	-0.001	-0.012
	(0.024)	(0.023)
	[0.958]	[0.601]
Observations	2,480	2,480
Outcome (Control Group) Mean	0.582	0.680

Notes: This table shows the estimated effects of including the List of Activities (LOA) module (T1) or enforcing self-reporting (T2) when reporting employment or work status. We present the 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 control by variables selected using a double-LASSO procedure (Table A5 in the Appendix shows the variables selected). 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 4. Survey Methods' Heterogeneous Effects by Respondent Sex or Age

	(1)	(2)	(3)	(4)
	D = Female		D = Youth	
	Employed	Working	Employed	Working
T1	0.005	-0.001	0.033	0.043*
	(0.030)	(0.027)	(0.027)	(0.025)
	[0.880]	[0.972]	[0.216]	[0.079]
T2	-0.022	-0.025	-0.019	-0.029
	(0.033)	(0.028)	(0.027)	(0.025)
	[0.465]	[0.340]	[0.480]	[0.233]
T1 x D	0.063	0.082**	0.016	-0.007
	(0.043)	(0.041)	(0.052)	(0.052)
	[0.168]	[0.043]	[0.795]	[0.889]
T2 x D	0.042	0.025	0.069	0.064
	(0.046)	(0.044)	(0.051)	(0.050)
	[0.335]	[0.547]	[0.179]	[0.200]
T1 + T1 x D	0.068**	0.081**	0.049	0.036
	(0.033)	(0.033)	(0.045)	(0.045)
	[0.041]	[0.010]	[0.320]	[0.446]
T2 + T2 x D	0.020	-0.000	0.050	0.035
	(0.034)	(0.034)	(0.045)	(0.045)
	[0.551]	[0.996]	[0.246]	[0.404]
Observations	2,480	2,480	2,480	2,480
Outcome (Control Group) Mean	0.582	0.680	0.582	0.680

Notes: This table presents the treatment heterogeneity of including the LOA module (T1) or enforcing self-reporting (T2) by sex and age for reporting 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 (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 A5 in the Appendix contains the variables selected). All estimations include EA (stratification variable) fixed effects. 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. 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 5. Heterogeneity for Employment and Work by Age
Separating the Samples by Sex

	(1)	(2)	(3)	(4)
	Women		Men	
	Employed	Working	Employed	Working
T1	0.046	0.059+	0.015	0.020
	(0.039)	(0.038)	(0.034)	(0.029)
	[0.232]	[0.109]	[0.648]	[0.522]
T2	-0.003	-0.028	-0.050	-0.050+
	(0.039)	(0.039)	(0.036)	(0.030)
	[0.929]	[0.503]	[0.161]	[0.106]
T1 x Youth	0.018	0.010	0.009	-0.031
	(0.071)	(0.073)	(0.075)	(0.070)
	[0.799]	[0.883]	[0.906]	[0.670]
T2 x Youth	0.032	0.044	0.138*	0.123+
	(0.067)	(0.070)	(0.080)	(0.072)
	[0.648]	[0.555]	[0.094]	[0.122]
T1 + T1 x Youth	0.064	0.070	0.023	-0.011
	(0.059)	(0.061)	(0.065)	(0.062)
	[0.285]	[0.261]	[0.707]	[0.861]
T2 + T2 x Youth	0.029	0.016	0.088	0.073
	(0.057)	(0.061)	(0.071)	(0.064)
	[0.633]	[0.807]	[0.241]	[0.310]
Observations	1,361	1,361	1,119	1,119
Outcome (Control Group) Mean	0.461	0.562	0.728	0.823

Notes: This table presents the treatment heterogeneity of including the LOA module (T1) or enforcing self-reporting (T2) 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, 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 A5 in the Appendix for the selected variables). All estimations include EA (stratification variable) fixed effects. 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. 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, + p<0.130

Table 6. Differential Effects on Employment and Work Status by Prevalence of Informal Employment

	(1)	(2)	(3)	(4)
	Women		Men	
	Employed	Working	Employed	Working
T1	0.063*	0.077**	0.009	0.004
	(0.033)	(0.033)	(0.031)	(0.027)
	[0.054]	[0.018]	[0.759]	[0.873]
T2	0.013	-0.003	-0.028	-0.027
	(0.034)	(0.034)	(0.033)	(0.028)
	[0.700]	[0.933]	[0.368]	[0.310]
T1 x Informal employment in the EA	0.008	0.024	-0.026	0.026
	(0.032)	(0.032)	(0.032)	(0.030)
	[0.821]	[0.461]	[0.389]	[0.326]
T2 x Informal employment in the EA	0.048	0.052+	-0.007	0.027
	(0.034)	(0.034)	(0.035)	(0.030)
	[0.150]	[0.112]	[0.863]	[0.346]
T1 + T1 x Informal employment in the EA	0.070	0.100**	-0.017	0.031
	(0.047)	(0.045)	(0.043)	(0.040)
	[0.137]	[0.034]	[0.688]	[0.425]
T2 + T2 x Informal employment in the EA	0.060	0.049	-0.034	0.000
	(0.048)	(0.047)	(0.047)	(0.041)
	[0.192]	[0.299]	[0.463]	[0.999]
Observations	1,361	1,361	1,119	1,119
Outcome (Control Group) Mean	0.461	0.562	0.728	0.823

Notes: This table presents the treatment heterogeneity of including the LOA module (T1) or enforcing self-reporting (T2) 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 (see more descriptive statistics of this measure in Table A6). 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 (see Table A5 in the Appendix for the selected variables). All estimations include EA (stratification variable) fixed effects. 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. 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 7. Differential Effects on Employment and Work Status by Prevalence of Peers' Informal Employment

Male Respondents Only

	(1)	(2)	(3)	(4)
	Young Men		Adult Men	
	Employed	Working	Employed	Working
T1	-0.012	-0.036	0.010	0.014
	(0.062)	(0.059)	(0.034)	(0.028)
	[0.810]	[0.534]	[0.786]	[0.623]
T2	0.077	0.053	-0.044	-0.046
	(0.064)	(0.058)	(0.036)	(0.031)
	[0.234]	[0.340]	[0.225]	[0.121]
T1 x Informal employment in the EA	0.048	0.082	-0.051	0.008
	(0.056)	(0.055)	(0.036)	(0.034)
	[0.427]	[0.164]	[0.149]	[0.790]
T2 x Informal employment in the EA	0.142**	0.100*	-0.046	0.006
	(0.062)	(0.056)	(0.038)	(0.036)
	[0.031]	[0.061]	[0.232]	[0.868]
T1 + T1 x Informal employment in the EA	0.036	0.046	-0.041	0.022
	(0.084)	(0.081)	(0.049)	(0.044)
	[0.688]	[0.597]	[0.381]	[0.609]
T2 + T2 x Informal employment in the EA	0.220**	0.153*	-0.090*	-0.041
	(0.091)	(0.084)	(0.051)	(0.047)
	[0.016]	[0.077]	[0.075]	[0.343]
Observations	301	301	818	818
Outcome (Control Group) Mean	0.641	0.724	0.760	0.859

Notes: This table presents treatment heterogeneity of including the LOA module (T1) or enforcing self-reporting (T2) by the prevalence of employment in the informal sector. This prevalence measure is estimated as the share of youth between the ages of 15 to 24 years living in the same EA as the respondent and employed in the informal sector. This variable is estimated using the standard labor module. (see the descriptive statistics of this measure in Table A6). We present ITT estimates from Equation (2), where D_{ihs} is the prevalence of youth employment 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 (see Table A5 in the Appendix for the variables selected). All estimations include EA (stratification variable) fixed effects. 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. 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 8. Differential Effects on Employment and Work Status by Exposure to Social Norms Related to Housework

	(1)	(2)	(3)	(4)
	Women		Men	
	Employed	Working	Employed	Working
T1	0.035	-0.019	-0.033	0.004
	(0.084)	(0.083)	(0.087)	(0.076)
	[0.668]	[0.813]	[0.699]	[0.958]
T2	-0.040	-0.052	-0.042	-0.003
	(0.085)	(0.086)	(0.090)	(0.066)
	[0.635]	[0.541]	[0.633]	[0.960]
T1 x Average time (hrs.) spent on housework (women)	0.025	0.084	0.038	0.001
	(0.067)	(0.065)	(0.070)	(0.060)
	[0.705]	[0.209]	[0.583]	[0.989]
T2 x Average time spent on housework (women)	0.048	0.045	0.012	-0.020
	(0.065)	(0.067)	(0.076)	(0.052)
	[0.458]	[0.514]	[0.872]	[0.708]
T1 + T1 x Average time spent on housework (women)	0.059*	0.065*	0.004	0.005
	(0.035)	(0.035)	(0.033)	(0.029)
	[0.088]	[0.048]	[0.898]	[0.869]
T2 + T2 x Average time spent on housework (women)	0.008	-0.007	-0.030	-0.023
	(0.037)	(0.036)	(0.034)	(0.029)
	[0.832]	[0.832]	[0.346]	[0.418]
Observations	1,361	1,361	1,119	1,119
Outcome (Control Group) Mean	0.461	0.562	0.728	0.823

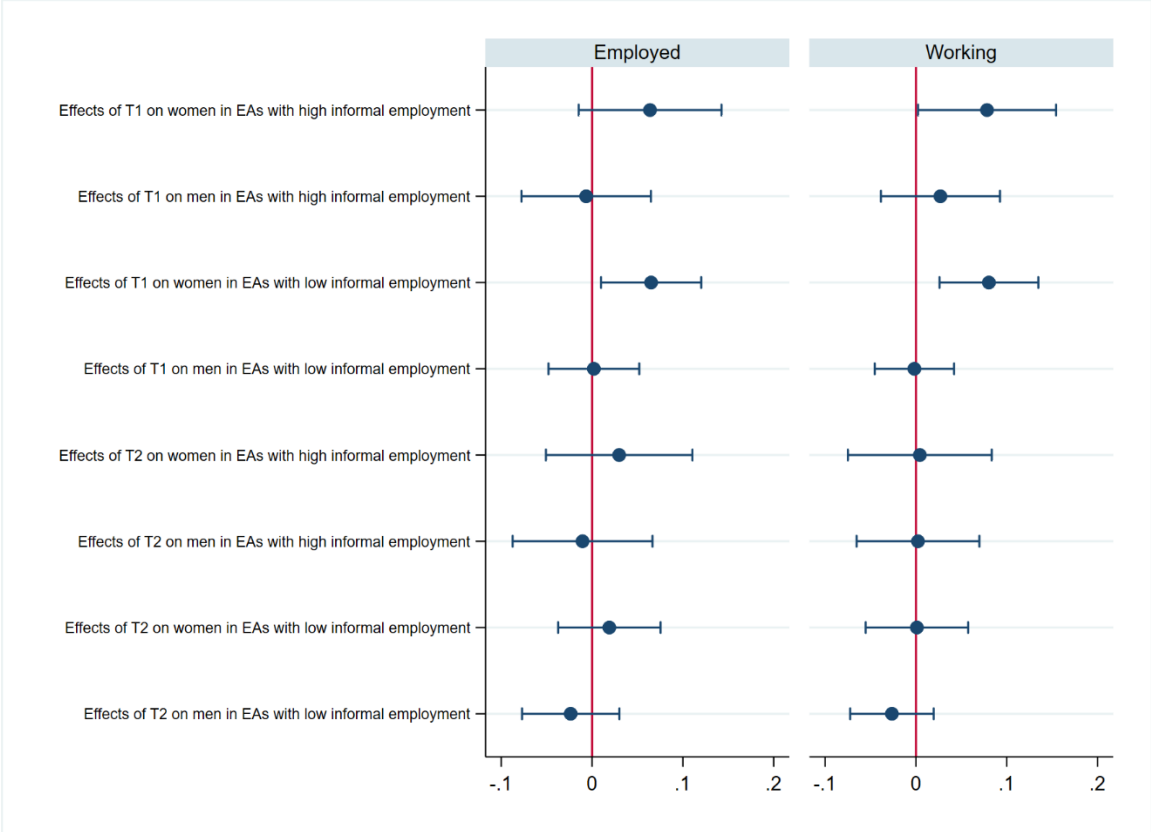
Notes: This table presents treatment heterogeneity of including the LOA module (T1) or enforcing self-reporting (T2) by exposure to social norms related to housework. This exposure measure is estimated as the average time that other women in the EA (excluding the respondent) devoted to housework the previous day (Columns [1] and [2]). This variable is estimated using the Time Use module. We present ITT estimates from Equation (2), where D_{ihs} is the exposure measure (in hours), and the estimations of the total effect of each survey method (Table A6 contains the descriptive statistics of this measure). The vector X_{ihs} includes the variable D_{ihs} and the list of control variables selected using a double-LASSO procedure for each outcome (Table A5 in the Appendix shows the variables selected). All estimations include EA (stratification variable) fixed effects. 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. 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. Experimental Design

Group	Type of labor module	Type of respondent
Control group: 336 households	Standard labor module (following 19th, ICLS)	**Proxy responses accepted** for other HH members if they were not available. If available each HH member (15-64 years old) provides his/her own information
T1: 336 households	Standard + LOA module	**Proxy responses accepted**
T2: 336 households	Standard labor module (following 19th, ICLS)	**Self-responses enforced** for the eligible household members (15-64 years old).

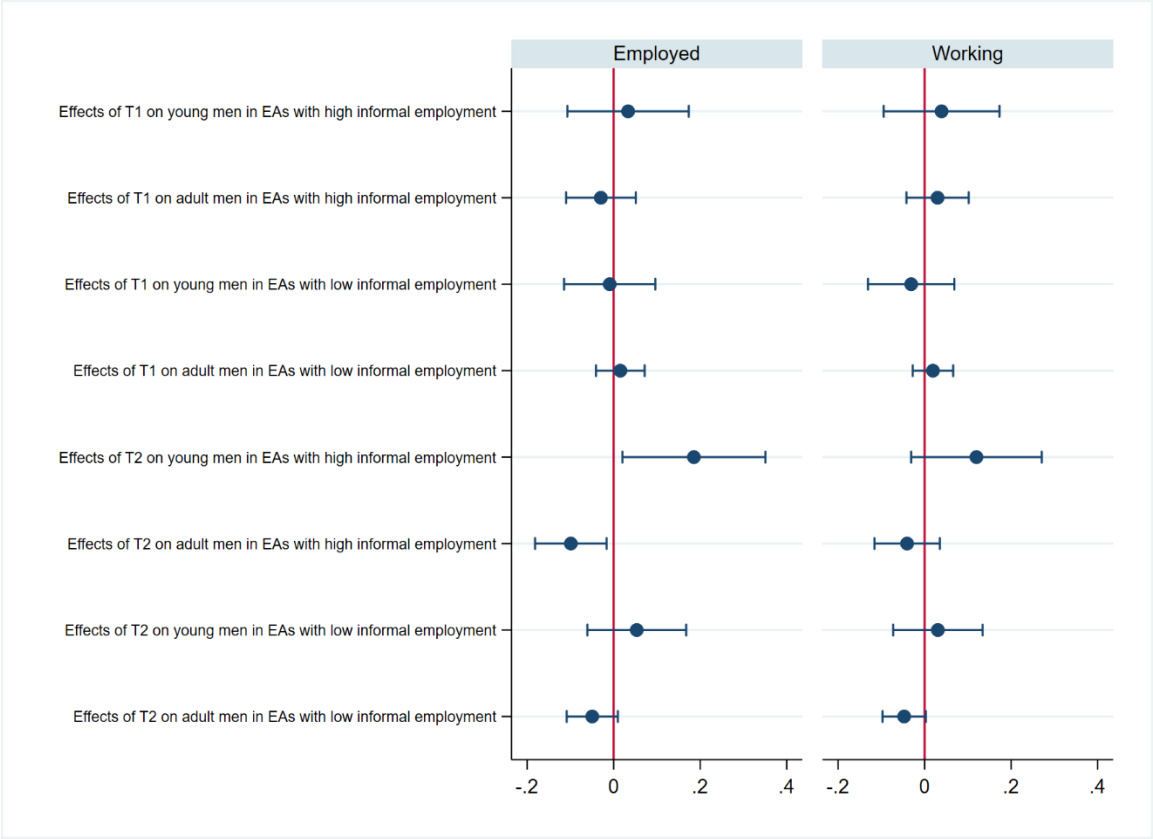
Notes: This figure presents the type of labor module, the type of respondent permitted to report within each experimental group, and the household distributions within each group. The assignments were carried out by means of stratified randomization at the EA level. LOA=List of Activities module; HH=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

Figure 2. Effects on Employment and Work Status by Prevalence of Informal Employment and Sex



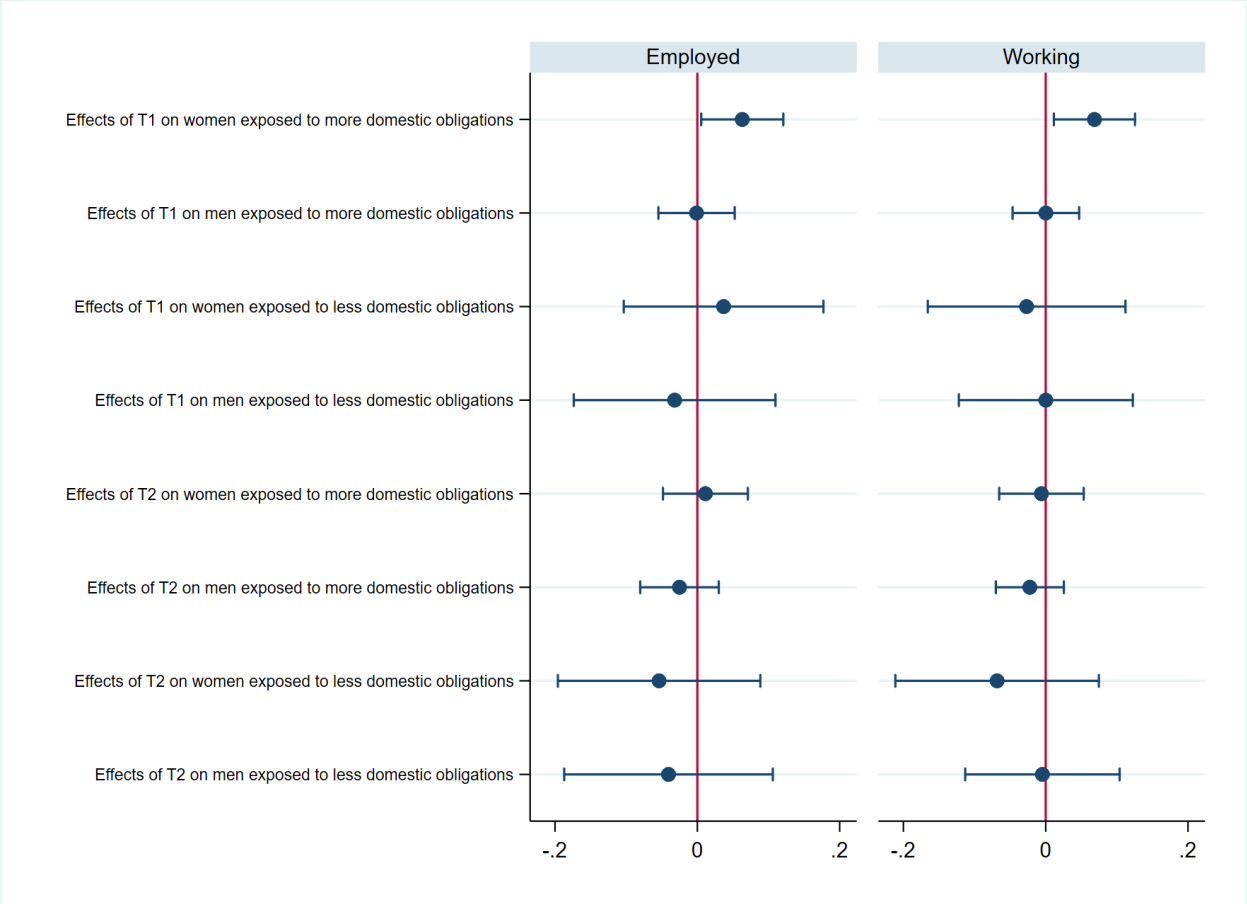
Notes: This figure presents the treatment heterogeneity of including the LOA module (T1) or enforcing self-reporting (T2) on the reporting of employment and work status by the prevalence of informal employment. This prevalence measure is estimated as the share of individuals living in the same EA as 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 (see Table A6). We report the total effects of T1 and T2 for each subsample relative to the group of individuals with similar exposure in the control group. For example, we present the total effect of T1 on women living in communities with a high prevalence of informal employment relative to other women living in communities with a similar prevalence in the control group. These estimated effects were obtained from a model that includes a triple interaction between treatment, sex, and the prevalence measure (see the notes of Table A7 in the Appendix for more details). Each line represents a 90-percent confidence interval.

Figure 3. Effects on Employment and Work Status by Prevalence of Informal Employment and Age
Male Respondents Only



Notes: This figure presents the treatment heterogeneity of including the LOA module (T1) or enforcing self-reporting (T2) on the reporting of employment and working status by the prevalence of informal employment. This prevalence measure is estimated as the share of individuals 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 (Table A6). We report the total effects of T1 and T2 for each subsample relative to the group of individuals with similar exposure in the control group. For example, we present the total effect of T1 on women living in communities with a high prevalence of informal employment relative to other women living in communities with a similar prevalence in the control group. These estimated effects were obtained from a model that includes a triple interaction between treatment, sex, and the prevalence measure (see the notes of Table A8 in the Appendix for more details). Each line represents a 90-percent confidence interval.

Figure 4. Effects on Employment and Work Status by Exposure to Social Norms Related to Housework



Notes: This figure presents the treatment heterogeneity of including the LOA module (T1) or enforcing self-reporting (T2) on the reporting of employment and work status by exposure to social norms related to housework. This exposure measure is estimated as the average time (hours) that other women in the EA (excluding the respondent) devoted to housework the previous day (Columns [1] and [2]). We report the total effects of T1 and T2 for each subsample relative to the group of individuals with similar exposure in the control group. For example, we present the total effect of T1 on women exposed to strong norms related to housework relative to other women with similar exposure in the control group. These estimated effects were obtained from a model that includes a triple interaction between treatment, sex, and the exposure measure (see the notes of Table A9 in the Appendix for more details). Each Line represents a 90-percent confidence interval.

Appendix Tables and Figures

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, caring for the sick or elderly for one or more hours?	... provide transportation services such as taxi, Uber, Mototaxi, pickup truck, 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	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 T1 (LOA module).

Table A2. Modules Included in the Survey

#	Module	Level	Respondent	Could a proxy provide responses?
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	
19	Contact	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 we could collect information from a proxy respondent for each of the modules at the individual level. For example, data for module Time Use was collected only if the individual was available at the time of the survey and self-reported the data directly. If the household was assigned to the T1 or C, we then collected self-reported data on time use only from those respondents who were available at the time of the survey.

Table A3. Mean Characteristics of the Study Participants vs. Salvadorians

Variable	(1) Study Sample	(2) EHPM (Usulután and San Salvador)	(3) *P*-val. Diff.	(4) Study Sample Usulután	(5) EHPM Usulután	(6) *P*-val. Diff.	(7) Study Sample San Salvador	(9) EHPM San Salvador	(9) *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	

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) 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.

Table A4. Original and Substitute Household Characteristic Means

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*-value for the test of the differences in means between the two groups.

Table A5. Control Variables Selected by LASSO 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 3	Employed (Column 1)	√	√	√	√		
	Working (Column 2)	√	√	√	√		
Table 4	Employed (Column 1)	√	√	√	√	√	
	Working (Column 2)	√	√	√	√		
	Employed (Column 3)	√	√	√		√	
	Working (Column 4)	√	√	√		√	
Table 5	Employed (Column 1)			√			
	Working (Column 2)						
	Employed (Column 3)	√			√	√	
	Working (Column 4)	√			√		
Table 6	Employed (Column 1)						
	Working (Column 2)						
	Employed (Column 3)	√			√		
	Working (Column 4)	√			√		
Table 7	Employed (Column 1)				√		√
	Working (Column 2)				√		√
	Employed (Column 3)	√				√	√
	Working (Column 4)	√					
Table 8	Employed (Column 1)						
	Working (Column 2)						
	Employed (Column 3)	√			√		
	Working (Column 4)	√			√		
Table A7	Employed (Column 1)	√	√	√	√	√	
	Working (Column 2)	√	√	√	√		
Table A8	Employed (Column 1)	√			√	√	
	Working (Column 2)	√			√		
Table A9	Employed (Column 1)	√	√	√	√	√	
	Working (Column 2)	√	√	√	√		

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

Table A6. Summary Statistics for Variables Used for the Mechanisms Analysis

	Mean (1)	SD (2)	Median (3)	Min. (4)	Max. (5)	N (6)
Informal employment in the EA (share of women, SD)	0.00	1.00	0.12	-2.75	1.51	2480
Informal employment in the EA (share of men, SD)	0.00	1.00	0.05	-3.54	2.02	2480
Average time spent on housework (women, hours)	1.14	0.48	1.05	0.22	2.80	2480

Notes: This table reports summary statistics for the variables *average informal employment* and *women's average time (hours) spent on housework*, both measured at the enumeration area (EA)-level. The variable *prevalence of employment in the informal sector* is estimated in SD as the share of individuals living in the same EA as the respondent and employed in the informal sector. This variable is estimated using the standard labor module based on the sample of employed individuals. An individual is working in the informal sector if the firm (production unit) or employer meets any of the following criteria: a) does not keep accounts or only keeps accounts for personal use; b) is not registered at the national level; c) is unincorporated (business and owner are not separate legal entities); d) place of work is without a fixed premises; or e) the firm has fewer than five employees. The variable *women's average time spent on housework* is estimated as the average time in hours that other women in the EA (excluding the female respondent) devoted to housework the previous day.

Table A7. Differential Effects on Employment and Work Status by Prevalence of Informal Employment

	(1) Employed	(2) Working
Informal employment in the EA	-0.016	-0.042**
	(0.023)	(0.020)
Female	-0.297***	-0.292***
	(0.031)	(0.030)
T1	0.002	-0.002
	(0.030)	(0.027)
T2	-0.024	-0.026
	(0.033)	(0.028)
T1 x Female	0.063	0.082**
	(0.043)	(0.041)
T2 x Female	0.043	0.027
	(0.046)	(0.043)
Informal employment in the EA x Female	0.024	0.058*
	(0.037)	(0.034)
T1 x Informal employment in the EA	-0.008	0.029
	(0.031)	(0.028)
T2 x Informal employment in the EA	0.013	0.029
	(0.034)	(0.029)
T1 x Informal employment in the EA x Female	0.007	-0.031
	(0.048)	(0.046)
T2 x Informal employment in the EA x Female	-0.002	-0.025
	(0.051)	(0.047)
Treatment Effects		
Effects of T1 on men in EAs with low informal employment	0.002	-0.002
Effects of T1 on men in EAs with high informal employment	-0.006	0.027
Effects of T1 on women in EAs with low informal employment	0.065*	0.080**
Effects of T1 on women in EAs with high informal employment	0.064	0.078*
Effects of T2 on men in EAs with low informal employment	-0.024	-0.026
Effects of T2 on men in EAs with high informal employment	-0.010	0.002
Effects of T2 on women in EAs with low informal employment	0.019	0.001
Effects of T2 on women in EAs with high informal employment	0.030	0.004
Observations	2,480	2,480
Outcome (Control Group) Mean	0.582	0.680

Notes: This table presents the treatment heterogeneity of including the LOA module (T1) or enforcing self-reporting (T2) by the prevalence of informal employment. This prevalence measure is estimated as the share of individuals living in the same EA as the respondent and employed in the informal sector. We present estimated coefficients from a model that includes a triple interaction between treatment, sex, and the exposure measure. We control for a vector of variables (X_{ihs}) that includes sex, the prevalence measure, double interactions, and the list of control variables selected using a double-LASSO procedure for each outcome. Under *Treatment Effects* we report the total effects of T1 and T2 for on subsample relative to the group of individuals with similar exposure in the control group. For example, we present the total effect of T1 on women in EAs with high exposure to informal employment relative to other women with similar exposure in the control group (see Figure 2 for a graphic representation of these total 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. *** p<0.01, ** p<0.05, * p<0.1

Table A8. Differential Effects on Employment and Work Status by Prevalence of Informal Employment

Male Respondents Only

	(1) Employed	(2) Working
Informal employment in the EA	0.029	-0.010
	(0.028)	(0.026)
Youth	-0.082	-0.058
	(0.056)	(0.050)
T1	0.016	0.019
	(0.034)	(0.028)
T2	-0.050	-0.047
	(0.036)	(0.031)
T1 x Youth	-0.025	-0.050
	(0.073)	(0.067)
T2 x Youth	0.103	0.078
	(0.078)	(0.069)
Informal employment in the EA x Youth	-0.100**	-0.060
	(0.050)	(0.047)
T1 x Informal employment in the EA	-0.045	0.011
	(0.036)	(0.034)
T2 x Informal employment in the EA	-0.049	0.007
	(0.038)	(0.035)
T1 x Informal employment in the EA x Youth	0.087	0.060
	(0.067)	(0.064)
T2 x Informal employment in the EA x Youth	0.182**	0.082
	(0.079)	(0.070)
Treatment Effects		
Effects of T1 on adult men in EAs with low informal employment	0.016	0.019
Effects of T1 on adult men in EAs with high informal employment	-0.029	0.030
Effects of T1 on young men in EAs with low informal employment	-0.009	-0.031
Effects of T1 on young men in EAs with high informal employment	0.033	0.039
Effects of T2 on adult men in EAs with low informal employment	-0.050	-0.047
Effects of T2 on adult men in EAs with high informal employment	-0.099**	-0.040
Effects of T2 on young men in EAs with low informal employment	0.053	0.031
Effects of T2 on young men in EAs with high informal employment	0.186*	0.120
Observations	1,119	1,119
Outcome (Control Group) Mean	0.728	0.823

Notes: This table presents the treatment heterogeneity of including the LOA module (T1) or enforcing self-reporting (T2) by the prevalence of informal employment. This prevalence measure is estimated as the share of men living in the same EA as the respondent and employed in the informal sector. We present estimated coefficients from a model that includes a triple interaction between treatment, sex, and the exposure measure. We control for a vector of variables (X_{ihs}) that includes sex, the prevalence measure, double interactions, and the list of control variables selected using a double-LASSO procedure for each outcome. Under *Treatment Effects* we report the total effects of T1 and T2 for each subsample relative to the group of individuals with similar exposure in the control group (see Figure 3 for a graphic representation of these total 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. *** p<0.01, ** p<0.05, * p<0.1

Table A9. Differential Effects on Employment and Work Status by Exposure to Social Norms Related to Housework

	(1) Employed	(2) Working
Average time spent on housework (hrs.)	-0.044	-0.001
	(0.048)	(0.036)
Female	-0.268***	-0.207**
	(0.088)	(0.080)
T1	-0.032	0.000
	(0.086)	(0.074)
T2	-0.041	-0.005
	(0.089)	(0.066)
T1 x Female	0.069	-0.027
	(0.123)	(0.114)
T2 x Female	-0.013	-0.064
	(0.124)	(0.110)
Average time spent on housework x Female	-0.026	-0.074
	(0.072)	(0.065)
T1 x Average time spent on housework	0.031	0.000
	(0.069)	(0.060)
T2 x Average time spent on housework	0.015	-0.018
	(0.074)	(0.052)
T1 x Average time spent on housework x Female	-0.005	0.095
	(0.099)	(0.092)
T2 x Average time spent on housework x Female	0.050	0.080
	(0.100)	(0.087)
Treatment Effects		
Effects of T1 on men exposed to less housework	-0.032	0.000
Effects of T1 on men exposed to more housework	-0.001	0.000
Effects of T1 on women exposed to less housework	0.037	-0.027
Effects of T1 on women exposed to more housework	0.063*	0.069*
Effects of T2 on men exposed to less housework	-0.041	-0.005
Effects of T2 on men exposed to more housework	-0.025	-0.022
Effects of T2 on women exposed to less housework	-0.054	-0.068
Effects of T2 on women exposed to more housework	0.011	-0.006
Observations	2,480	2,480
Outcome (Control Group) Mean	0.582	0.680

Notes: This table presents the treatment heterogeneity of including the LOA module (T1) or enforcing self-reporting (T2) by the exposure to social norms for housework. This exposure measure is estimated as the average time (hours) that other women in the EA devoted to housework the previous day. This variable is estimated using the Time Use module. We present estimated coefficients from a model that includes a triple interaction between treatment, sex, and the exposure measure. We control for a vector of variables (X_{ihs}) that includes sex, the exposure measure, double interactions, and the list of control variables selected using a double-LASSO procedure for each outcome. Under *Treatment Effects* we report the total effects of T1 and T2 for each subsample relative to the group of individuals with similar exposure in the control group (see Figure 4 for a graphic representation of these total 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. *** p<0.01, ** p<0.05, * p<0.1

Table A10. The Survey Methods' Impacts on the Measurement of Labor Market Outcomes, Excluding Controls

	(1) Employed	(2) Working
T1	0.037	0.040*
	(0.023)	(0.022)
	[0.143]	[0.089]
T2	-0.004	-0.014
	(0.025)	(0.023)
	[0.892]	[0.564]
Observations	2,480	2,480
Outcome (Control Group) Mean	0.582	0.680

Notes: This table shows the estimated effects of including the LOA module (T1) or enforcing self-reporting (T2) on reporting 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 only 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 A11. The Survey Methods' Heterogeneous Effects by Respondent Sex and Age, Excluding Controls

	(1)	(2)	(3)	(4)
	D = Female		D = Youth	
	Employed	Working	Employed	Working
T1	-0.000	-0.008	0.023	0.033
	(0.031)	(0.027)	(0.027)	(0.025)
	[0.985]	[0.783]	[0.400]	[0.200]
T2	-0.020	-0.020	-0.033	-0.042
	(0.034)	(0.029)	(0.028)	(0.026)
	[0.537]	[0.467]	[0.246]	[0.117]
T1 x D	0.064	0.084**	0.053	0.028
	(0.044)	(0.042)	(0.054)	(0.054)
	[0.158]	[0.049]	[0.348]	[0.613]
T2 x D	0.033	0.016	0.095*	0.089*
	(0.047)	(0.044)	(0.054)	(0.054)
	[0.465]	[0.709]	[0.083]	[0.095]
T1 + T1 x D	0.063*	0.075**	0.076	0.061
	(0.033)	(0.032)	(0.046)	(0.047)
	[0.062]	[0.018]	[0.116]	[0.207]
T2 + T2 x D	0.013	-0.005	0.062	0.047
	(0.034)	(0.034)	(0.048)	(0.048)
	[0.687]	[0.918]	[0.197]	[0.314]
Observations	2,480	2,480	2,480	2,480
Outcome (Control Group) Mean	0.582	0.680	0.582	0.680

Notes: This table presents the treatment heterogeneity of including the LOA module (T1) or enforcing self-reporting (T2) by sex and age for the reporting of employment or work status. We present the ITT estimates of β_1 - β_4 based on the following equation: $Y_{ihs} = \alpha_0 + \beta_1 T1_{hs} + \beta_2 T2_{hs} + \beta_3 T1 \times D_{ihs} + \beta_4 T2 \times D_{ihs} + \Theta D_{ihs} + b_s + \epsilon_{ihs}$. We also include estimations of the total effect of each survey method ($\beta_1 + \beta_3$ for the LOA module, and $\beta_2 + \beta_4$ for proxy-reporting) at the end of each table. 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 (Columns [3] and [4]). All estimations include strata fixed effects. 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. Standard errors clustered at the household level (randomization 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, Excluding Controls

	(1)	(2)	(3)	(4)
	Women		Men	
	Employed	Working	Employed	Working
T1	0.045	0.059	0.009	0.014
	(0.039)	(0.038)	(0.035)	(0.029)
	[0.245]	[0.109]	[0.793]	[0.644]
T2	-0.006	-0.028	-0.053	-0.048
	(0.039)	(0.039)	(0.037)	(0.031)
	[0.885]	[0.503]	[0.143]	[0.133]
T1 x Youth	0.026	0.010	0.012	-0.032
	(0.071)	(0.073)	(0.077)	(0.072)
	[0.712]	[0.883]	[0.877]	[0.681]
T2 x Youth	0.035	0.044	0.155*	0.135*
	(0.068)	(0.070)	(0.083)	(0.074)
	[0.617]	[0.555]	[0.075]	[0.103]
T1 + T1 x Youth	0.071	0.070	0.021	-0.018
	(0.058)	(0.061)	(0.067)	(0.064)
	[0.237]	[0.261]	[0.749]	[0.771]
T2 + T2 x Youth	0.029	0.016	0.102	0.088
	(0.057)	(0.061)	(0.073)	(0.066)
	[0.617]	[0.807]	[0.197]	[0.248]
Observations	1,361	1,361	1,119	1,119
Outcome (Control Group) Mean	0.461	0.562	0.728	0.823

Notes: This table presents the treatment heterogeneity of including the LOA module (T1) or enforcing self-reporting (T2) by age for the reporting of employment or work status. We present the ITT estimates of β_1 - β_4 based on the following equation: $Y_{ihs} = \alpha_0 + \beta_1 T1_{hs} + \beta_2 T2_{hs} + \beta_3 T1 \times D_{ihs} + \beta_4 T2 \times D_{ihs} + \Theta D_{ihs} + b_s + \epsilon_{ihs}$. We also include estimations of the total effect of each survey method ($\beta_1 + \beta_3$ for the LOA module, and $\beta_2 + \beta_4$ for proxy-reporting) at the end of each table. The indicator D_{ihs} takes the value of 1 if the respondent is a youth between the ages of 15 to 24 (Columns [1] to [4]). All estimations include strata fixed effects. 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. Standard errors clustered at the household level (randomization 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. The Survey Methods' Impacts on the Measurement of Labor Market Outcomes, Addressing the Substitution Protocol's Effects

	(1) Employed	(2) Working
T1	0.040*	0.044**
	(0.023)	(0.022)
	[0.090]	[0.043]
T2	-0.001	-0.011
	(0.024)	(0.023)
	[0.948]	[0.620]
Observations	2,480	2,480
Outcome (Control Group) Mean	0.582	0.680

Notes: This table shows the estimated effects of including the LOA module (T1) or enforcing self-reporting (T2) in 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 (Table A5 in the Appendix shows the selected variables). 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. The Survey Methods' Heterogeneous Effects by Respondent Sex and Age, Addressing the Substitution Protocol's Effects

	(1)	(2)	(3)	(4)
	D = Female		D = Youth	
	Employed	Working	Employed	Working
T1	0.005	-0.000	0.033	0.044*
	(0.030)	(0.027)	(0.027)	(0.025)
	[0.883]	[0.991]	[0.215]	[0.068]
T2	-0.023	-0.024	-0.019	-0.028
	(0.033)	(0.028)	(0.028)	(0.025)
	[0.460]	[0.369]	[0.479]	[0.252]
T1 x D	0.063	0.081*	0.016	-0.007
	(0.044)	(0.042)	(0.052)	(0.052)
	[0.168]	[0.043]	[0.794]	[0.888]
T2 x D	0.043	0.024	0.069	0.063
	(0.047)	(0.044)	(0.052)	(0.051)
	[0.333]	[0.557]	[0.179]	[0.202]
T1 + T1 x D	0.068**	0.081**	0.049	0.036
	(0.033)	(0.033)	(0.045)	(0.045)
	[0.042]	[0.010]	[0.321]	[0.441]
T2 + T2 x D	0.020	-0.000	0.050	0.036
	(0.034)	(0.034)	(0.045)	(0.045)
	[0.551]	[0.998]	[0.246]	[0.400]
Observations	2,480	2,480	2,480	2,480
Outcome (Control Group) Mean	0.582	0.680	0.582	0.680

Notes: This table presents the treatment heterogeneity of including the LOA module (T1) or enforcing self-reporting (T2) 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 (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 A5 in the Appendix shows the variables selected). All estimations include EA (stratification variable) fixed effects. 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. 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
T1	0.052	0.069*	0.013	0.019
	(0.039)	(0.038)	(0.034)	(0.029)
	[0.173]	[0.063]	[0.685]	[0.531]
T2	0.001	-0.022	-0.053	-0.052*
	(0.039)	(0.039)	(0.036)	(0.030)
	[0.984]	[0.602]	[0.134]	[0.092]
T1 x Youth	0.018	0.009	0.009	-0.031
	(0.071)	(0.073)	(0.075)	(0.070)
	[0.813]	[0.905]	[0.907]	[0.674]
T2 x Youth	0.031	0.041	0.139*	0.123*
	(0.068)	(0.070)	(0.080)	(0.072)
	[0.670]	[0.571]	[0.093]	[0.123]
T1 + T1 x Youth	0.071	0.077	0.022	-0.012
	(0.059)	(0.062)	(0.065)	(0.063)
	[0.249]	[0.219]	[0.721]	[0.849]
T2 + T2 x Youth	0.032	0.019	0.085	0.072
	(0.057)	(0.060)	(0.071)	(0.064)
	[0.597]	[0.767]	[0.261]	[0.315]
Observations	1,361	1,361	1,119	1,119
Outcome (Control Group) Mean	0.461	0.562	0.728	0.823

Notes: This table presents the treatment heterogeneity of including the LOA module (T1) or enforcing self-reporting (T2) 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, 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 A5 in the Appendix for the outcome variables). All estimations include EA (stratification variable) fixed effects. 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. 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. The Survey Methods' Impacts on a Placebo Outcome

	(1) Employed in the formal sector
T1	0.005
	(0.017)
	[0.782]
T2	0.016
	(0.017)
	[0.325]
Observations	2,480
Outcome (Control Group) Mean	0.140

Notes: This table shows the estimated effects of including the LOA module (T1) or enforcing self-reporting (T2) in the reporting of employment in the formal sector. We present estimated coefficients from Equation (1). All estimations include strata fixed effects and are controlled by variables selected using a double-LASSO procedure. 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. The Survey Methods' Heterogeneous Effects by Respondent Sex and Age on a Placebo Outcome

	(1)	(2)
	Employed in the formal sector	
	D = Female	D = Young
T1	-0.003	-0.008
	(0.030)	(0.021)
	[0.935]	[0.685]
T2	0.004	0.011
	(0.030)	(0.020)
	[0.742]	[0.205]
T1 x D	0.013	0.047
	(0.034)	(0.036)
	[0.530]	[0.196]
T2 x D	0.022	0.015
	(0.035)	(0.037)
	[0.890]	[0.567]
T1 + T1 x D	0.011	0.039
	(0.017)	(0.031)
	[0.513]	[0.682]
T2 + T2 x D	0.026	0.026
	(0.018)	(0.032)
	[0.127]	[0.381]
Observations	2,480	2,480
Outcome (Control Group) Mean	0.140	0.140

Notes: This table presents the treatment heterogeneity of including the LOA module (T1) or enforcing self-reporting (T2) by sex and age for the reporting of employment in the formal sector. 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 (Column [1]) and if the respondent is a youth between the ages of 15 to 24 (Column [2]). 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

Table A18. Heterogeneity for Employment in the Formal Sector by Age on a Placebo Outcome

	(1)	(2)
	Employed in the formal sector	
	Women	Men
T1	-0.008	-0.013
	(0.020)	(0.035)
	[0.697]	[0.704]
T2	0.004	0.019
	(0.021)	(0.036)
	[0.843]	[0.587]
T1 x Youth	0.045	0.038
	(0.035)	(0.068)
	[0.199]	[0.572]
T2 x Youth	0.064*	-0.043
	(0.037)	(0.074)
	[0.088]	[0.565]
T1 + T1 x Youth	0.037	0.025
	(0.028)	(0.059)
	[0.188]	[0.659]
T2 + T2 x Youth	0.068**	-0.024
	(0.032)	(0.064)
	[0.037]	[0.693]
Observations	1,361	1,119
Outcome (Control Group) Mean	0.071	0.224

Notes: This table presents the treatment heterogeneity of including the LOA module (T1) or enforcing self-reporting (T2) by age for the reporting of employment in the formal sector. 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, 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. 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

Table A19. Participation in Employment and Work by Proxy 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
T1	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)
	[0.648]	[0.963]	[0.442]	[0.902]	[0.633]	[0.902]	[0.481]	[0.675]	[0.351]	[0.735]	[0.445]	[0.711]
Proxy is woman	-0.090	-0.022					-0.089	-0.019				
	(0.067)	(0.081)					(0.066)	(0.072)				
Proxy has high school or more			-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 Mean	0.490	0.743	0.490	0.743	0.490	0.743	0.544	0.821	0.544	0.821	0.544	0.821

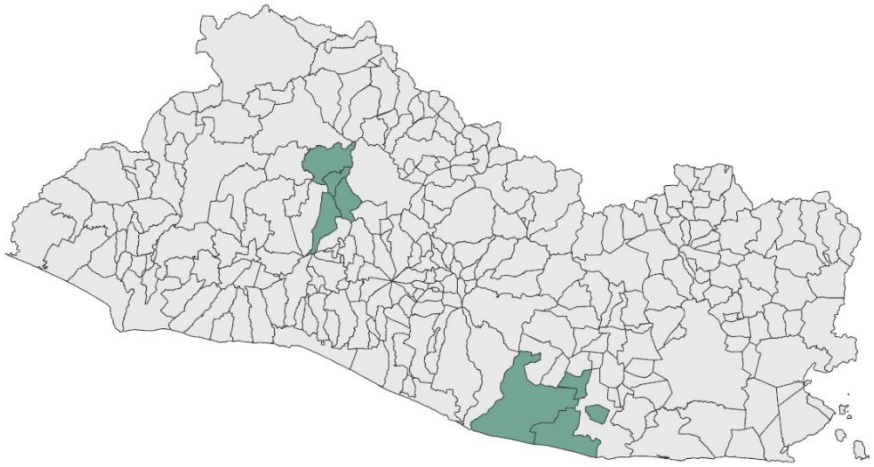
Notes: This table presents the associations between the characteristics of a proxy respondent and the probability of reporting employment and work status. The sample is restricted to participants assigned to T1 and C groups. We present ITT estimates from the following Equation $Y_{ihs} = \alpha_0 + \beta_1 T1_{hs} + \theta X_{ihs} + b_s + \epsilon_{ihs}$. 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. 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 A20. The Survey Methods' Effects on Employment Across Economic Sectors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Agriculture versus No employment		Agriculture versus Other Sectors		Agriculture versus Services		Agriculture versus Manufacture		Agriculture versus Commerce	
T1	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)
	[0.700]	[0.701]	[0.386]	[0.253]	[0.876]	[0.831]	[0.577]	[0.778]	[0.489]	[0.140]
T2	-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)
	[0.898]	[0.386]	[0.101]	[0.570]	[0.182]	[0.959]	[0.179]	[0.978]	[0.194]	[0.494]
T1 x Female		0.051		-0.039		-0.008		0.024		-0.110
		(0.063)		(0.043)		(0.069)		(0.094)		(0.068)
		[0.439]		[0.362]		[0.900]		[0.793]		[0.124]
T2 x Female		0.072		0.044		0.129 *		0.153+		0.021
		(0.064)		(0.044)		(0.072)		(0.102)		(0.071)
		[0.253]		[0.321]		[0.081]		[0.100]		[0.792]
T1 + T1 x Female		0.029		-0.004		0.001		0.038		-0.034
		(0.029)		(0.028)		(0.052)		(0.080)		(0.051)
		[0.300]		[0.885]		[0.988]		[0.625]		[0.486]
T2 + T2 x Female		0.022		0.062**		0.131**		0.154**		0.059
		(0.031)		(0.032)		(0.057)		(0.085)		(0.056)
		[0.453]		[0.034]		[0.018]		[0.040]		[0.246]
Observations	1,085	1,085	1,530	1,530	885	885	535	535	696	696
Outcome (Control Group) Mean	0.268	0.268	0.192	0.192	0.331	0.331	0.548	0.548	0.421	0.421

Notes: This table presents the treatment effects of including the LOA module (T1) or enforcing self-reporting (T2) by sex for the extensive and intensive margins in the reporting of employment in agricultural sector. Columns (1) and (2) present the results on an indicator that takes the value of 1 if the individual es employed in agriculture and 0 if unemployed. Columns (3) and (4) show the results on an indicator equal to 1 if the individual reports being employed in agriculture and 0 if employed in other sectors (manufacture, services, and industry). The outcome in columns (5) – (10) is measured as an indicator that takes the value of 1 if the individual reports being employed in agriculture and 0 if employed 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, +p<0.102

Figure A1. Municipalities Selected for the Survey Experiment



Notes: This figure shows the distribution of the municipalities selected for the survey experiment. The municipalities to the left of center are Aguilares, El Paisnal, Nejapa, and Guazapa; they belong to the Department of San Salvador. The municipalities to the right of center are Jiquilisco, Puerto El Triunfo, San Dionisio, and California; they are part of the Department of Usulután.

Appendix A. Qualitative Study

This appendix provides further information on the methods used in the qualitative component of the study and their main results.

A1.1. Focus Group Discussions

The main goal of this qualitative study is twofold: (1) to create a list of activities considered work in rural and peri-urban communities in El Salvador, and (2) to understand the barriers that women and young people face in entering the labor market and obtaining formal employment.

We conducted eight focus groups between June 13 and July 8, 2022 in six Salvadoran municipalities located in the Department of San Salvador (Nejapa and Guazapa) and the Department of Usulután (Jiquilisco, Puerto El Triunfo, San Dionisio, and California). The focus groups were stratified by age and gender. Four of the focus groups included youth between the ages of 18 and 29 (two with men and two with women), and the four other focus groups included adults aged 29 years and older (two with women and two mixed). A total of 58 people attended the focus groups, 40 women and 18 men within their respective municipalities.

Table B1. Focus Group Participants by Gender and Municipality

Municipality	**Men**	**Women**	**Total**
Nejapa (young men)	7	0	7
Nejapa (adults)	2	7	9
Guazapa (young women)	0	8	8
Guazapa (adult women)	0	8	8
Jiquilisco (young men)	6	0	6
Puerto El Triunfo (young women)	0	8	8
San Dionisio (mixed adults)	3	4	7
California (adult women)	0	5	5
Total	**18**	**40**	**58**

A1.2. Approach

The focus groups facilitators were local qualitative experts and trained by the research team using a technical guide that included all of the focus group questions. During the semi-structured focus groups, the facilitator asked the participants open-ended questions that could lead to additional questions and responses.

Special care was taken to preserve participant anonymity and freedom to consent. Before starting each of the focus group sessions, a team member talked with the potential participant and read a consent statement that included the objective of the focus group, how participation was voluntary, and how the information shared would be used based on the IRB's approval of the experiment design. A facilitator conducted each of the discussions with the support of two team members who were responsible for listing the main work activities that the participants identified during the discussion on a whiteboard. The discussions lasted up to one and a half hours.

A1.3. Focus Group Questions

The semi-structured technical guide helped the facilitators lead the focus group discussions. The guide included two main components. The first component included questions to elicit a list of general activities that the members of the community do. From there, the group identified which activities are considered work and which employment. The discussions started with the following question, "*Think about a typical day in your life. Which activities do you perform*?", followed by "*Now, I would like us to talk about things that members of the community generally do. Could you describe to the group what some of the most common activities are within your community*?" After identifying all the common activities, the facilitator asked: "*Which of these activities is considered work*?" Using different colored markers, the participants highlighted which the activities would be classified as work. After identifying the work activities, the facilitator asked about the heterogeneity of the activities in the following questions: "*Do both men and women perform these activities*? *Or do the men and women in your community do different things*?", and "*Do young people also perform these activities*?"

The second component of the focus group aimed to identify the participants' main barriers to finding stable employment within the community. In this section of the discussion, the facilitator started by asking whether people in the community have difficulty finding stable jobs. All of the focus groups participants agreed that it was difficult to find stable employment. Afterwards, the facilitator asked: "*What are the main obstacles to fostering stable employment within your community*?" The facilitators also explored the heterogeneity in the job barriers by asking the following: "*Do women face different obstacles*? *What job characteristics would entice women to*

work or enable them to work?" and "*Do young people face different barriers than women and men*? *What job characteristics would entice young people to work or enable them to work*?"

A1.4. Main Results

Extensive list of work activities. Participants identified the list of activities, which we summarize in Table AP1. We then created the List of Activities (LOA) module based on the different activities mentioned, and this list later helped respondents identify their work activities more accurately.

Heterogeneity of the work activities. Men primarily performed some of the work activities (e.g., transportation services), while mainly women did others (e.g., make items such as tamales, jelly, food, jewelry, etc. to sell). The focus groups reported that youth primarily helped in different activities.

Main barriers to obtaining stable employment. We found that the most relevant barriers to stable employment reported in the focus groups were: (1) difficulty fulfilling job requirements, (2) lack of childcare services, and (3) insecurity about the commute to work. While adult men reported the first and third barriers, adult women emphasized the second obstacle. Young men reported the first barrier, and young women described struggling with all three barriers.

Table AP1. List of Activities Identified in the Focus Groups

List of Activities	FG 1 - Nejapa Young Men	FG 2 - Nejapa Adults	FG 3 – Guazapa Young Women	FG 4 - Guazapa Adult Women	FG 5 - Jiquilisco Young Men	FG 6 - Puerto El Triunfo Young Women	FG 7 - San Dionisio Mixed Adults	FG 8 - California Adult Women
Care for farm, livestock, or fishing	4	11	4	8	9	5	6	6
Help a family member with agricultural work, take care of the family's livestock and/or fishing	0	0	0	0	1	0	0	0
Sell items such as clothing, cell phones, shoes, jewelry	0	1	1	2	1	1	0	0
Make items to sell such as tamales, jelly, food, jewelry	3	2	3	4	3	0	8	2
Sell homemade items such as tamales, jelly, food, jewelry	2	4	1	4	4	4	3	1
Provide hair and beauty services, repairs or masonry; administer medication; care for the sick or elderly	4	3	2	3	2	4	2	1
Provide transportation services via taxi, Uber, Mototaxi, pickup truck, minibus	0	1	1	1	1	0	1	0
Provide home delivery services	1	0	0	0	1	0	1	0
Cook, launder clothing, or perform other services for individuals	5	4	4	6	4	5	4	5
Help in a non-farm family business	1	2	1	1	3	1	0	0

Appendix B. 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.[23]

We included an additional substitution protocol for the T2 group, which enforce self-responses. Interviewers were asked to visit each T2 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 T2 were replaced primarily because selected household members were hospitalized or visiting family in other municipalities.

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 percent of the surveys and asked two trained monitors to audit the recordings. The monitors listened to these recordings daily and recorded 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.

[23] The main reasons for substitution were 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 T2 household member (0.7%).

Appendix C. Definitions of the Variables Included in the Analysis

C1. 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:

Question: Does your household own any of these items in working conditions?	
1. Radio	YES...1 NO....0
2. Sound system	YES...1 NO....0
3. TV	YES...1

	NO....0
4. Video cassette or DVD player	YES...1 NO....0
5. Refrigerator	YES...1 NO....0
6. Washing machine	YES...1 NO....0
7. Blender	YES...1 NO....0
8. Fan	YES...1 NO....0
9. Computer	YES...1 NO....0
10. Sewing machine	YES...1 NO....0
11. Vehicle, car, or boat (for household use)	YES...1 NO....0
12. Iron	YES...1 NO....0
13. Microwave oven	YES...1 NO....0
14. Video games (Nintendo)	YES...1 NO....0
15. Air conditioning	YES...1 NO....0
16. Well or cistern	YES...1 NO....0
17. Motorcycle	YES...1 NO....0
18. Tablet	YES...1 NO....0

Households that experienced moderate to severe food insecurity: According to the Food Insecurity Experience Scale (FIES),[24] 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

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

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....0
2. … anyone in your household been unable to eat healthy and nutritious foods?	YES...1 NO....0
3. … anyone in your household eaten only a few kinds of foods?	YES...1 NO....0
4. … anyone in your household had to skip a meal?	YES...1 NO....0
5. … anyone in your household eaten less than you thought he/she should?	YES...1 NO....0
6. … your household run out of food?	YES...1 NO....0
7. … anyone in your household been hungry but did not eat?	YES...1 NO....0
8. … anyone in your household gone hungry for the entire day?	YES...1 NO....0

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. Does any member of your household own or have access to any land that is used for crop cultivation during the agricultural season?	YES...1 NO....0
2. Has your household owned any livestock in the past 12 months?	YES...1 NO....0
3. Have you or any other member of your household been involved in fish-related activities (i.e., catching or raising fish) in the last 12 months?	YES...1 NO....0

C2. Individual Characteristics

Female: Denotes the sex of the individual and takes the binary value of 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.