

CSE Working Paper

#55

What did they say? Respondent identity, question framing and the measurement of employment

Rosa Abraham, Nishat Anjum, Rahul Lahoti, Hema Swaminathan

September 2023

Centre for Sustainable Employment

cse@azimpremjiuniversity.edu.in

What did they say?

Respondent identity, question framing and the
measurement of employment

Rosa Abraham, Nishat Anjum, Rahul Lahoti, Hema Swaminathan

What did they say?

Respondent identity, question framing and the measurement of employment

Rosa Abraham ¹, Nishat Anjum ², Rahul Lahoti ³, Hema Swaminathan ⁴

Abstract

In developing countries, a precise approach to measuring women's employment remains elusive. Emerging evidence underscores the pivotal role of survey methodology, encompassing respondent selection and question framing, in shaping the assessment of women's employment. Drawing from two labor market experiments in rural India, this study offers insights on the influence of survey design on the measurement of women's employment. The first experiment contrasts self-reported women's and men's employment figures with proxy-reported data from spouses. Women's self-reported workforce participation surpasses proxy-reported estimates by six percentage points, while men's estimates exhibit negligible differences. There are significant differences in the type of employment activities reported by self and proxy for both women and men. These divergences emanate from asymmetric measurement errors, stemming from gender-based norm disparities between spouses, and divergent interpretations of employment. Additionally, information asymmetry between spouses concerning women's marginal activities and disparities in spousal characteristics contribute to these self-proxy differences. The second experiment investigates if framing of questions and recall period has an impact on reporting of labor market outcomes. We find that employing multiple questions to capture weekly employment status yields a 10-percentage-point increase in reported women's workforce participation, but men's participation rate decreases by six percentage points. Furthermore, when a distinct employment query is directed at each day of the preceding week as opposed to a single query for the entire week, reported women's workforce participation increases by seven percentage points, and men's by four percentage points.

¹Azim Premji University, Bangalore. Corresponding author, rosa.abraham@apu.edu.in

²Jawaharlal Nehru University, Delhi

³UNU-WIDER, Helsinki

⁴IIM Bangalore

1 Introduction

"A great deal of information on women exists, but it frequently comes from questions asked of men about their wives, daughters, and sisters, rather than from the women themselves...What women do is perceived as household work and what they talk about is called gossip, while men's work is viewed as the economic base of society and their information is seen as important social communication." - Reiter, 1975

The measurement of women's employment in the Global South faces many challenges. Women tend to be concentrated in the informal sector, in marginal activities, and predominantly engaged in paid and unpaid work. The definitions of what constitutes economic activities are often fuzzy and do not adequately capture the full range of women's responsibilities. Surveys, in most instances, do not speak to women themselves, which alongside gendered social norms pose a challenge to accurately measure women's employment. Employment estimates are then a function of who is being asked, how it is being asked and what is being asked.

Standard methods of data collection may lead to under-reporting and under-valuation of women's work (Dixon-Mueller and Anker, 1988; Greenwood, 2000). The measurement literature (see (Kilic et al., 2022; Koolwal, 2021) for a review) argues that (i) the boundaries of work as conceptualised by labour statistics do not include many productive economic activities that tend to be predominantly performed by women, (ii) the way in which the question of work is framed including the reference period used fails to capture marginal and multiple activities that women engage in, and (iii) the household respondent, who typically gives information for all members, may not present an accurate picture of the extent and type of women's work. Thus, a clear conceptualisation of women's work is missing, compounded by the difficulties of operationalising these concepts.

The burgeoning literature on women's employment in India has typically dwelt on supply side and demand side factors, though recently measurement issues have garnered attention (Hirway, 2010, 2015; Deshmukh et al., 2020; Kapur et al., 2021). Much more has been written about women's paid and unpaid work, definitions of economic activities, and framing of questions with little attention to the role of proxy informants in labour estimates.

In this paper, we use data from an unique survey, the India Working Survey (IWS), conducted in two states of India to examine the impact of methodological variations in survey design on labour market outcomes of men and women.

Specifically, the study investigates the effect of identity of respondent, framing of questions, and reference period on the measurement of employment of men and women and makes several important contributions to the literature. First, we add to the literature on the 'pure' proxy effect in under reporting of work participation i.e. we are able to compare a person's own report with that of their proxies. We are also able to investigate the factors that lead to intra-household disagreements on employment status. To the best of our knowledge, this is the first such study in the Indian context. Second, besides identifying the self-proxy effect on reporting of overall work, it also identifies how this plays out for different kinds of employment - self employment, wage and contributing worker. Third, the study distinguishes between under-reporting and misattribution of employment. A particular activity type may come to be under-reported by a proxy either because they do not recognise the individual to be in any employment or if they perceive them as in a different kind of employment. By asking (self and proxies) about different kinds of employment, we are able to identify the extent of agreement, under-reporting and misattribution in each kind of work. Fourth, in terms of question framing, the study answers, separately, to what extent reference period and detailed questions impact measures of employment and what kinds of work are likely to be mis-measured due to reference period/framing differences.

We find that respondent identity plays a role in measuring employment, but there are differential effects for men and women. There is significant under-reporting of women's work, when reported by men. Women's employment rate declines by nearly six percentage points when men report on behalf of women (proxy-reporting), compared to when women report about themselves (self reporting). However, there is no significant difference in men's employment rate between self and proxy reported estimates. We further disaggregate overall employment into different types of employment - self employment, contributing family work, and casual wage work. We find that, compared to women's self estimates, men tend to over-report women in self-employment while under-reporting women's participation in wage and contributing family work. On the other hand, women are more likely to attribute men as being in wage work, while men are more likely to identify themselves in self employment or contributing family work. Finally, the nature of the experiment allows us to compare responses within households and estimate the extent of intra-household agreements/disagreements. Proxy and self reports are more likely to agree in the case of wage work, although agreement is higher for men. While the highest level of agreement for men is on wage work (67 percent), for women, contributing family work has the highest match on self-proxy reports (51 percent).

With regard to the framing experiment, we find that detailed questions asking specifically

about each kind of employment, rather than a single question about whether they were engaged in employment or not, improves women’s employment rate by 10 percentage points while having no discernible difference for men’s employment rates. Most of this increase in women’s employment comes from an increase in the share of women reporting unpaid family work. Similarly, moving from asking one question about the week to asking one question about each day of the week (single weekly versus single daily) increases women’s employment estimates by about seven percentage points and men’s by four percentage points.

Section 2 reviews the literature on measurement and labour market outcomes. Section 3 discusses the survey instrument used in the India Working Survey with a description of the data and the sample. Section 4 discusses the findings from proxy versus self-reported employment estimates, while Section 5 shares the findings from the framing of questions and their impact on employment estimates. Section 6 concludes.

2 Background

Household surveys typically collect data by eliciting responses from one or more individuals within the household. Measurement error is inherent in this process and refers to the difference between the true (unknown) value of the characteristic being collected and the value recorded in the survey (UN, 2005). Groves (2005) and Biemer and Lyberg (2003) document four main sources of non-sampling errors that arise due to:

- (i) Respondents: Respondents might answer differently from the true value for several reasons; misinterpretation of the questions, social norms that influence reporting, limited domain knowledge, or lack of effort.
- (ii) Questionnaire: Effect of questionnaire design including, the wording and ordering of questions.
- (iii) Data-collection method: Effect of how the questionnaire is administered to the respondent (in-person, phone, online).
- (iv) Interviewers: Enumerator-specific attributes including their skills and biases may affect responses.

In this paper, we examine the first two sources of measurement error, i.e. due to respondents and due to the questionnaire. While thinking about respondent effects, we focus on

whether respondents are reporting for themselves (self-reporting) or other household members (proxy-reporting). It is conceivable that respondent effects will vary with both, the identity of the person providing the information and the identity of the person for whom information is being sought.

2.1 Measurement error: self vs proxy reporting

Nationally-representative household surveys are critical in the Global South given that administrative data may not be readily or consistently accessible. Proxy-reporting or obtaining full information from one member is a standard feature of most national surveys, except for those that deal with specific and sensitive data such as reproductive health or violence. Given this backdrop, there has been some research on how proxy and self reporting impact survey estimates in the developing countries. Across several countries, studies have examined the differences in self and proxy-reporting in employment (Bardasi et al., 2011; Kilic et al., 2022; Kapur et al., 2021), child labour (Galdo et al., 2021; Dillon et al., 2012), land and other asset ownership (Ambler et al., 2021; Kilic et al., 2022), (Twyman et al., 2015), household and agricultural decision making (Ambler et al., 2021; Twyman et al., 2015), and income (Fisher et al., 2010).

Early evidence of the importance of considering respondent identity for measuring labour market outcomes was provided by Bardasi et al. (2011) that examined the impact of proxy responses on measurement of men’s and women’s employment in Tanzania. The authors find that men’s employment is sensitive to respondent selection with women as proxy under-reporting men’s work, while the reporting of women’s work by a male proxy had no significant deviation from self reports. This finding is reinforced by Kilic et al. (2022) in Malawi who show that proxy-reporting is lower than self-reporting for both men and women and across a range of employment activities. Although not examining differences between self and proxy report, the only study in India, Kapur et al. (2021) finds a correlation between gender of respondent and women’s employment rate. When women respond to questions in the household roster, women’s workforce participation rate is likely to be higher than when men respond (Kapur et al., 2021). The divergence between self and proxy could be driven by several factors and can be broadly classified as random measurement error, asymmetric measurement error, and asymmetric information error. We borrow this framework from Ambler et al. (2021) who study disagreements between husbands and wives with respect to asset ownership and household decision-making and adapt it to the context of labour market

outcomes.

Random measurement error refers to self-proxy discrepancies that are unrelated to individual or household attributes. Such errors may arise due to enumerator characteristics or rushed responses. If the measurement error is indeed random, there would be no systematic differences between women's and men's responses.

Asymmetric measurement error, on the other hand, leads to responses that systematically differ between men and women. In our context, asymmetric measurement errors may stem from two sources - differences in the definitional understanding of what constitutes employment and gender norms-based differences in identifying employment. First, the perception of employment for men and women may differ, even though there is complete information on activities performed by either within the household. For instance, men may consider employment to involve only paid labour, while women may also include their work on the family farm which is often not explicitly remunerated. This discrepancy in the understanding of employment leads to consistent under-reporting of women's work by men.

Second, social norms that assign specific roles to men and women may also influence reporting by sex. Men and women may not conform to these prescribed roles themselves, but their proxies, for various reasons, may report in line with social norms. For example, the reporting of women's employment by men may be driven by perceived social censure associated with a certain activity Jayachandran (2021). In India where women working is often considered a symbol of low social status, men may be averse to reporting their wives being employed. Indeed, Bernhardt et al. (2018) finds that while men themselves are open to their wives being in employment, their 'second-order' beliefs about what society expects results in men disapproving women's employment. It is also that women are mainly recognised as homemakers and thus, their paid roles do not get much attention by proxy respondents (Comblon et al., 2017). For women, the dispersion of their time between home-based employment and child care may result in their employment activity being not reported if it is not perceived as economic activity by the respondent. Women, in keeping with the male breadwinner norm, may hesitate to report men as unemployed and prefer to categorise them as self-employed. Thus, social norms potentially shape both the understanding of employment, and the reporting of such employment by men and women.

The third type of error, asymmetric information error occurs due to different information sets being available to men and women about each other's activities. The information asymmetry could be intentional as spouses strategically hide information about their activities

from each other to keep income private or avoid censure if they are going against social norms. The information asymmetry could also be unintentional due to issues of observations (spouse could be employed in different locations - farm, house etc) or specific domains of responsibility with little intersection across men and women. One would expect that the information asymmetry error to impact proxy report, but not self reports of employment.

We hypothesise that asymmetric measurement error will impact both self and proxy reporting, irrespective of the gender of the respondent, but will vary across employment categories. We expect greater divergence between self and proxy reports in categories like contributing workers, where there are higher chances of differences in perceptions between the spouses as the boundaries between contributing worker and self-employment may be blurred. A social desirability bias may also inform both self and proxy reports – reporting of men as unemployed may be less desirable than reporting them as contributing family workers.

Similarly, information asymmetry is likely to vary across activities. If the work activity is marginal, intermittent, and likely to be performed simultaneously with other activities – taking care of livestock or contributing to the family farm – it is more likely to be missed by the proxy respondent. Work activities that are more visible and structured are less likely to be prone to information asymmetries. Wage work is one example of such activity. The extent of information asymmetry might also vary by couple characteristics or household structure. Couples who have large age or education gaps might share less information and hence, have higher disagreements in reporting of each other’s status. If spouses are employed in the same kind of activities, information asymmetries may be reduced owing to joint participation and shared experiences (Blair et al., 2004). In the empirical analysis of the divergence between self and proxy reports, it is not possible to disentangle between asymmetric measurement error and asymmetric information error. However, we hypothesise on the likely mechanisms behind the divergence and include appropriate control variables in the regression analysis.

2.2 Measurement error: questionnaire design

The framing of questions can have significant impacts on the measurement of employment particularly in economies dominated by agriculture due to its seasonality and inherent multiplicity of activities ((Dixon, 1982)). Given women’s role in agriculture, these issues will have much larger implications for the measurement of women’s employment. To this end, it is important to estimate the impact of detailed probing questions versus short questions on

labour force statistics, especially for women.

Bardasi et al. (2011) estimate differences in labour statistics comparing between a short module questionnaire versus a longer module in the context of Tanzania. The short module consists of one question on work - “Did you do any type of work in the last 7 days?”. The detailed module consists of screening questions, specifying three main groups of economic activity. Comparing employment rates across these two instruments they find that a higher share of women and men reported as working in the short module versus the detailed. However, after re-classifying domestic work as “no work”, women’s employment is about five percentage points lower in the short module than in the detailed module. Benes and Walsh (2018) also emphasize the importance of asking recovery questions in particular in capturing the work of unpaid family helpers who are otherwise reported as unemployed.

A “list of activity” approach has been found to provide higher estimates of women’s employment than a “keyword” questions (questions that contain a typically recognisable keyword about overall employment such as “main activity”, “secondary activity”, “pay or profit”) approach as it allows for capture of women’s part-time and home-based work (Anker, 1983; Langsten and Salen, 2008). A recent study by Deshmukh et al. (2020) in the Indian context reiterate the findings of Anker (1983). Based on questionnaires administered to respondents in selected districts of Delhi National Capital Region, the study finds that asking about primary and secondary activities (analogous to the “keyword” questions identified in Anker (1983) results in a higher share of women being listed as home-makers. In contrast, follow-up questions on the major sources of household income and who contributed to this income resulted in much higher estimates of women’s employment participation with a large share of this increase coming from the reporting of women’s work in caring for livestock.

We compare employment rates across three treatment arms - short daily, short weekly and detailed weekly. The detailed weekly calls out the list of activities while the short weekly and daily use keywords in estimating employment rates. Additionally, we also explore the impact of a change in reference period (from weekly to daily) using “keyword” questions on employment rates. We hypothesise that for women, detailed weekly will generate higher estimates of employment over modules where the activities are not called out. The detailed questions help with both recall of the multiple set of activities women usually undertake and also help them identify their activities as being economically productive and not only home maintenance tasks.

3 Data and Survey Experiments

The survey experiments used in this analysis are embedded within a larger study, the India Working Survey (IWS) that was conducted in early 2020 across two states, Karnataka (southern India) and Rajasthan (western India). The main aim of IWS was to investigate if there was an interaction of social identities (gender, caste, and religion) with various dimensions of the labour market. In addition to the methodological experiments with respect to measuring employment, IWS collected information on diverse domains, including, household living standards, wages, time spent on household production activities, occupational life history, decision-making, social network, and experience of discrimination. The IWS followed a stratified multistage sampling design and intended to survey approximately 4,000 households in each state, which would have provided a state-representative sample. However, the survey was disrupted in March 2020 due to COVID 19, yielding a final (non-representative) sample of 3,646 households and 5,951 individuals (3,371 women and 2,580 men) across the two states. Our analytic sample is restricted to rural areas as 85 percent of the interviewed households are primarily from rural areas. The sample is relatively evenly distributed between the two states.

Two survey experiments, namely the self-proxy experiment and the framing experiment, were conducted as part of the IWS. The self-proxy experiment examines the variation in responses when questions are directed either to the subject directly (self) or to a proxy informant. In many developing countries, including India, household surveys collect information on the entire household from a single household member – usually the head or the individual who is available at the time of enumeration.

IWS followed an alternative strategy where two respondents from each household were interviewed. In every household, one adult man and one adult woman were randomly selected as respondents. The respondents were matched with same sex enumerators and to the extent possible, the interviews were conducted privately and simultaneously. We conducted the self-proxy experiment if the selected respondents were a couple. Each respondent was asked questions about their own and their spouse’s labour market participation. The framing of the questions asking about their own and the spouse’s labour market participation was exactly the same. Specifically, a detailed weekly module with six binary questions to ascertain labour market participation was administered to determine their and their spouse’s employment status. If they or their spouse were employed, then details on the type of activity and hours worked were also collected. Among the 5,951 respondents in the IWS, 3,750 (63 percent)

were spousal pairs. However, due to non-participation by one or both respondents and data recording issues, the final sample for spousal pairs consists of 2,674 observations, including 1,337 husbands and 1,337 wives.

In households with non-couple respondents, we did not collect proxy information for the other respondent. Thus, we cannot investigate how proxy responses differ from self when the proxy is a non-spouse. We expect that spouses are likely to be better informed of each other’s activities than other household members. So our estimates of differences in self and proxy reporting are likely to be underestimates of the self-proxy differences obtained across non-spouse household members.

The framing experiment investigates the impact of question detail and reference period on reported labour market participation. Three different labour modules were developed for the experiment: a short weekly module, a short daily module, and a detailed weekly module. The short weekly and short daily modules resemble those used in labour market surveys conducted by the official statistical apparatus in India. In the short weekly module, respondents were asked a single question: "In the last week, what were the activities you were doing, even if only for an hour?" They were allowed to report multiple activities, but enumerators did not provide a specific list of potential activities. In the short daily module, the framing question was the same, but the reference period was the previous day. The same question was asked for each of the seven preceding days, resulting in a total of seven questions. The weekly labour market status is calculated from these seven daily questions. If they are employed even a single hour or any one day in the previous week, they are deemed as employed according to weekly status.

The detailed weekly module differs from the shorter modules in two ways. First, there was a specific question for each potential employment activity, including, self-employment activities, assistance in family farms or businesses (unpaid), wage or salaried work, paid apprenticeships or internships, and small-scale production of goods or services for sale. Five questions, pertaining to each one of the activities, were asked, with respondents indicating "yes" or "no" for each activity. Second, the module included a recovery question that probed if the respondent missed out on reporting any other income-generating activity that they were involved in. The questions in the detailed module adhere to the recommendations of the International Labour Organization (ILO) for measuring key labour market indicators (Benes and Walsh, 2018) and are similar to those used in the World Bank’s Living Standards Measurement Study (LSMS) surveys. For the full list of questions fielded in the different modules refer to Table A2. Across the three modules, if the respondent participated in an

activity, further details regarding hours worked, industry/sector of work, and income were collected.

The short modules (weekly and daily) were administered only in Karnataka and to randomly selected sets of households based on a household listing exercise. Every household was randomly assigned either the single weekly module or the single daily module. Similar to the main survey, an adult man and an adult woman were selected as respondents in each household and both administered the same module. However, unlike the detailed weekly module, even if the respondents were a spousal pair, they were asked only about their own activities. The survey instrument for these experiments collected details on household demographics, asset ownership but did not include experiences of discrimination, life history calendar, and time spent on household production. In total, 299 and 300 individuals responded to the single weekly and single daily modules respectively.

4 Methods

We assess the impact of survey experiments on labour market outcomes through a two-part investigation. First, we use the self-proxy experiment to analyze the discrepancies in reported labour market outcomes between self and proxy responses for men and women separately, focusing on indicators such as labour force participation, workforce participation, unemployment rate, and hours of work. Additionally, we investigate differences in the type of employment activity being reported by self and proxy, for men and women. To delve deeper into intra-household dynamics, we examine the consistency between self and proxy reporting within couples. Second, we use the information from the three employment modules to analyze the impact of variation in question details and reference period on men and women’s employment status.

Following standard protocols, the study defines an individual as ”employed” if they report engaging in any of the following activities for at least an hour in the last week: business, including farming (own consumption and sale) or other self-employment; contributing work in the household business/farm/livestock; wage or salaried work; paid apprenticeship or internship; and small-scale production of goods and services for sale. Domestic duties, including, household maintenance or care work within the household is not considered employment.

4.1 Self-proxy survey experiment

The self-proxy design, where both partners provide information about their own and their spouse’s employment status, enables us to identify causal differences between self and proxy reporting. We estimate the following equation while controlling for individual (self and proxy), household, and interviewer characteristics:

$$y_{ie} = \alpha + \beta P_{ie} + \lambda X_{ie} + \phi_e + \epsilon \quad (1)$$

Here, y_i represents various labour statistics (e.g., labour force participation, hours of work, and type of activity) for the i^{th} individual interviewed by enumerator e . P_i is an indicator variable denoting proxy reporting of labour market outcomes, where P_i equals 1 when the outcome is proxy-reported and 0 when self-reported. The coefficient β captures the difference between self-reported and proxy-reported labour market outcomes. X_i denotes a vector of individual and household characteristics for the i^{th} individual in household h . These control variables encompass factors such as respondent characteristics (age, education), characteristics of the person being reported on (age, education and major activity for the majority of the year) and household attributes (asset ownership, social group). ϕ_e represents enumerator fixed effects, which control for idiosyncratic impacts attributed to individual enumerators. The term ϵ accounts for the stochastic error term, which is randomly distributed across households.

There is the possibility of omitted variable bias as unobserved characteristics of the person providing information may influence the outcomes of interest. To address this concern, we employ an individual fixed-effects regression model:

$$y_i = \alpha + \beta P_{ie} + \phi_e + \gamma_i + \epsilon \quad (2)$$

In this model, the included variables are similar to those in equation 1, but instead of individual and household characteristics, we introduce individual fixed-effects γ_i . These fixed effects capture the impact of all time-invariant observed and unobserved characteristics of the respondent.

In addition, we explore the determinants of disagreement between spouses on overall employment status and with respect to the specific employment category. For this analysis,

we consider only the subsample where the individuals report themselves as being employed. We estimate the following regression equation:

$$D_i = \alpha_0 + \beta X_i + \epsilon_i \tag{3}$$

Here, the dependent variable D_i is binary indicating disagreement between spouses. The variable D_i takes a value of 1 when the proxy respondent disagrees with their spouse's self-report of being employed and a value of zero when the proxy and self agree on being employed. This type of disagreement is referred to as under reporting based on the assumption that self-reports regarding employment are more accurate. Thus, the under reporting is from the point of view of the individual who is being reported on by the proxy.

Equation 3 is also estimated for disagreements between spouses where individuals do not report themselves as working but the spouses identify them as employed. Similar to the previous specification, the variable D_i takes a value of 1 when the spouse (proxy respondent) disagrees with the partner's self-report status of not working and reports them in employment, and a value of zero when the proxy agrees with the spouse's reporting of not being engaged in any economic activity. In this specification, the disagreement is referred to as over reporting, based on the premise that self reports are likely to be more accurate (or at least, less prone to error). Separate regressions for over and under reporting are conducted for each activity type (self-employment, contributing worker, and wage). All regression equations, 1-3, are estimated separately for both men and women.

As discussed earlier, reporting can be subject to random error, asymmetric measurement error, or asymmetric information error. We expect responses to be systematically different between men and women and across activities due to the measurement and information error. This is not the case random errors where responses ought not to differ by sex. Although it is not possible to disentangle the source of error in the responses, we include variables in our regressions that can serve as proxies for the underlying mechanisms at work. We discuss these variables briefly here.

To capture gendered norms that affect reporting of employment, we include a couple of proxy variables. The first variable relates to the employment status of the husband's mother. Our expectation is that if his mother was employed, then he is more likely to be aware of employment related activities of his wife as he has been exposed to women who have stepped beyond the prescriptive "caregiver or homemaker" role. A second proxy for gender norms is

the spousal difference in time spent in household maintenance (cooking, cleaning, child and elderly care, and fetching water). It is reasonable to expect that a low difference in time spent is indicative of greater sharing of domestic duties and thus, a willingness on part of both men and women to step outside of their gendered roles.

Mismatches in reporting may arise due to self and proxy not having access to the same kind of information about each other’s activities (Ambler et al., 2021). Self and proxy respondents may not share the same spaces (husbands may leave for work while wives work in the household farm or business), or if the work is marginal or performed intermittently, under-reporting may occur. To account for this, we include select correlates that proxy for attributes of the employment activities of the individuals. For the person being reported about, we include their self-reported activity status that they were engaged in for the majority of the year. This can capture whether the weekly activity that is being reported is marginal. To assess the possibility of asymmetric information error that may arise when work is less visible since it is marginal or intermittent, we include the number of self-reported hours of work spent in that employment in an average week. The regressions control for spousal differences in age and educational attainment that may also contribute to informational asymmetry. Additionally, the estimations control for household-level attributes including, an asset index reflecting the household’s asset ownership profile, the social group categorization (SC/ST, OBC, or General category), and state dummies accounting for state-specific differences.

4.2 Framing experiment

The framing experiment aims to ascertain the impact of level of detail and reference period of labour market questions on the reporting of labour market status. Three different labour modules were administered to three sets of randomly selected households to do this. We explore the mean differences in labour market outcomes and in the reported activity status across the three modules.

Our balance tests on the three samples indicate that there might be some differences across the samples. To account for these, we employ a regression analysis, controlling for the characteristics that differ significantly across the samples - majority activity status over the previous year and proportion of female (refer table 8) The regression model is as follows:

$$y_i = \alpha + \beta L_i + \lambda X_i + \epsilon \tag{4}$$

Here, y_i represents various labour market outcomes (e.g., labour force participation, work force participation, unemployment rate, and hours of work) for the i^{th} individual. L_i is an indicator variable denoting the labour module administered to individual i . The coefficient β compares the outcome reported in detailed weekly and single daily module with the single weekly module. X_i denotes a vector of individual and household characteristics for the i^{th} individual in household h . These control variables encompass factors such as respondent characteristics (age, education, and major activity for the majority of the year) and household attributes (asset ownership, social group).

5 Self-Proxy Experiment Results

The sample for the self-proxy experiment are households where the randomly selected respondents are a spousal pair, and have been administered the detailed weekly module. We discuss the characteristics of this sample, present the differences in self-proxy reporting in overall employment and activity type using both, descriptive statistics and regression analysis. Finally, we investigate the correlates of disagreement on employment status.

5.1 Spousal respondent sample

Sample characteristics are presented in Table 1. The average age of men in the sample is 43 years, while women, at an average age of 37 years are relatively younger. Educational attainment is expectedly gendered with women at greater disadvantage. Almost half the sample of women are illiterate, while about a quarter of men are illiterate. A significant proportion of both men and women have education below the secondary level, with 67 percent of men and 78 percent of women falling into this category.

The majority of the sample consists of Hindus, accounting for approximately 95 percent of respondents.⁵ About 49 percent of respondents belong to the Other Backward Class (OBC), 23 percent belong to the Scheduled Caste (SC), 13 percent belong to the Scheduled Tribes (ST), and the remaining 15 percent belong to other categories.⁶

⁵The survey encountered challenges in obtaining representation from Muslims due to the government’s proposed Citizenship Amendment Bill, which sparked protests and increased religious tensions. As a result, refusals among Muslim households were higher than anticipated.

⁶The survey collected religious affiliation and caste membership at the individual level, and the proportion of individuals reporting a different religion or caste from their spouse is negligible.

The activity that individuals are involved in for majority of the time during the last year differs substantially by gender. Women are most likely to be involved in household work (46 percent) while men are most likely to be in self-employment (53 percent). The incidence of salaried wage is low overall, but substantially higher for men (12 percent) than women (4 percent). Women’s main employment activity is being a contributing worker, which is unpaid work on family farm or enterprise.

To ensure the generalizability of our findings, we compared the characteristics of the spousal pairs in this sample with all the married men and women in the overall sample. The analysis reveals no significant differences between the individuals in the spousal sample and their counterparts in the overall sample. This suggests that the findings from the spousal sample can be broadly extended to other married individuals in the overall sample. Further details can be found in Appendix A1.

5.2 Employment reporting by self and proxy

The parameters for comparison between self and proxy reports are labour force participation rates (LFPR), workforce participation rates (WPR), unemployment rates (UR), and average working hours (Table 2).

For women, several proxy-reported labour market statistics are significantly lower compared to self-reported numbers. While 70 percent of women report themselves as being part of the labour force, only 64 percent of women are reported to be part of the labour force by their husbands, indicating a significant discrepancy of 5.7 percentage points. Similar differences exist in the workforce participation rate for women. Although not statistically significant, the proportion of women reporting themselves as unemployed is higher than when reported by their husbands.⁷

Regarding unconditional working hours, which account for average working hours of all individuals regardless of their employment status, women’s self-reports show significantly higher average working hours compared to proxy reports. However, when the analysis is limited to those reported as working, there is no statistically significant difference between self and proxy reports. This suggests that the differences between self and proxy reports for women occur primarily at the extensive margin, i.e., in identifying women as employed or

⁷An individual is identified as unemployed if they did not engage in any of the work activities in the week and responded in the affirmative to the question of either seeking work or being available for work in the last week.

Table 1: Self-proxy experiment sample

Characteristics	Gender of Respondent	
	Women	Men
Age (years)	37	43
<i>Education (%)</i>		
Illiterate	46	25
Primary/Middle	35	44
Secondary+	19	31
<i>Religion (%)</i>		
Hindu	98	95
Others	2	5
<i>Social group (%)</i>		
Scheduled Caste	26	26
Scheduled Tribe	15	14
Other Backward Caste	49	50
General	10	10
<i>Activity type (%)</i>		
Self employed	11	53
Contributing worker	20	1
Salaried	4	12
Casual wage	17	25
Unemployed	0	1
Household work	46	2
<i>State (%)</i>		
Karnataka	52	52
Rajasthan	48	48
N	1,134	1,134

Note: The sample is only for rural areas. Activity type is defined as the activity the individual is doing for majority of the time (more than six months) in the last year.

not, rather than at the intensive margin, i.e., the number of hours spent on paid work.

In contrast, no statistically significant differences exist between self-reported and proxy-reported labour market outcomes for men. Proxy reports indicate slightly higher labour and workforce participation rates for men compared to self-reports, but these differences are not significant.⁸

⁸Typically, convergence between self and proxy reports is taken as an indicator of reporting accuracy (Blair et al., 2004). In the absence of a third source of validation data to triangulate the different reports, we consider the self-reported estimates to be the benchmark. Hence, under-reporting and over-reporting are

Table 2: Difference in self and proxy reported labour market outcomes

	Women			Men		
	Self	Proxy	Difference	Self	Proxy	Difference
Labourforce participation rate	69.5	63.8	5.7***	79.7	81.5	-1.8
Workforce participation rate	63.2	57.9	5.4***	76.9	78.7	-1.8
Unemployment rate	9.0	6.0	3.0	3.5	2.8	0.7
Hours (unconditional)	20.2	17.7	2.5***	32.3	32.7	-0.4
Hours (conditional)	32.5	31.2	1.3	42.4	41.9	0.5

Note: *, **, *** indicates significance at 10% 5% and 1% level of significance for two-sided t-tests. Unconditional hours worked correspond to average weekly hours worked, averaged across all responses, irrespective of reported employment status (if not working, then they are assumed to have zero hours). Conditional hours worked corresponds to average weekly hours worked conditional on reporting as being employed. Number of observations - 1,134 men and women.

5.3 Activity distribution by self and proxy reports

In this subsection, we analyze the differences in activities as reported by individuals and their proxies, focusing on gender differences (Table 3). An individual may be self-employed, or contributing to the family farm or business or in casual or salaried wage work. The individual’s primary activity is taken as the activity where the individual spends the maximum time (measured in hours) during the week.

Women’s activity distribution shows significant differences when reported by themselves versus their proxies. According to self-reports, approximately 21 percent of women identify as self-employed. However, when their husbands report on their behalf, this proportion increases to 33 percent. Conversely, women are less likely to be reported as contributing family workers or working for wages by their husbands compared to their own reporting. While 55 percent of women self-report as contributing workers, only 49 percent are classified as such by their husbands, reflecting a substantial difference of 6 percentage points. This contrast in reporting between husbands and wives regarding women’s self-employment and contributing work may stem from varying perceptions of these activities. The distinction lies in the fact that self-employment involves earning direct income, whereas contributing work entails contributing to the household farm or enterprise without receiving direct payment. It is possible that women acknowledge the lack of payment and classify themselves as contributing workers, whereas men may not perceive it the same way and categorize women as self-employed. Additionally, husbands tend to under-report women working for wages compared to women’s

used taking self-reports as the baseline.

own reports by around 7 percentage points (18 percent versus 25 percent), most of which is driven by differences in casual wage labour in agriculture.

Although there are no significant differences in the reports of men's overall employment between husbands and wives, there are notable variations in the types of activities identified. Compared to what their wives report, men are more likely to classify themselves as self-employed and contributing workers, but less likely to classify themselves as wage workers. The differences in reporting are most pronounced in the case of wage work, particularly in casual employment. While 26 percent of men self-report as wage workers, their wives classify 37 percent of them as such. Men are 6 percentage points more likely to categorize themselves as self-employed and 4 percentage points more likely to identify as contributing workers compared to their wives' reports.

It is notable that these differences in classification between self and proxy reports predominantly arise in the agricultural sector. With the exception of casual non-agricultural work for men, all other significant disparities in self-proxy reports are related to agricultural activities.

In summary, the analysis reveals substantial differences in the classification of employment activities between self and proxy reports, particularly with respect to gender. These findings shed light on the distinct perceptions and categorizations of work within households.

Having understood the extent of differences in the classification of activities, in the next step we investigate the nature of these differences. Table 4 compares the self-reported status (rows) against the proxy-reported status (column) for women and men in separate panels. So the first row shows how women who report themselves to be self-employed are classified by their husbands. The diagonal elements of each panel represent the degree of agreement between the self and proxy reports. For example, 31 percent of women who report themselves as self-employed are also classified as self-employed by their husbands. Overall there are large differences in how women and men are misclassified by proxies as compared to self-reports. The disagreements in classification reflect differences in perceptions, differences in understanding of the various categories, social norms, and information asymmetry.

The degree of agreement on activity type is highest for women who classify themselves as out of labour force (65 percent) and least for those who classify themselves as unemployed (16 percent). About one in five women who classify themselves as doing any employment activity are likely to be classified as not working by their husbands. Self-employment and contributing work are the two activities most likely to be mixed up for women. Thirty-five percent of

Table 3: Differences in self and proxy reports of employment activity distribution

	Women			Men		
	Self	Proxy	Difference	Self	Proxy	Difference
<i>Self Employed</i>	20.5	33.3	-12.8***	56.8	50.7	6.1***
Own account worker agriculture	13.7	23.8	-10.1***	37.6	31.3	6.3**
Own account worker non-agriculture	6.1	7.8	-1.7	12.4	11.1	1.3
Employer agriculture	0.7	1.2	-0.5	5.2	6.6	-1.4***
Employer non-agriculture	0.00	0.5	-0.5	1.6	1.7	-0.1
<i>Contributing Work (CW)</i>	54.7	48.6	6.1	19.5	13.7	4.2
CW agriculture	52.3	46.0	6.3***	17.2	11.8	5.4*
CW non-agriculture	2.4	2.6	-0.2	0.3	1.1	-0.8
<i>Wage Work</i>	24.8	18.2	6.7	25.7	36.5	-10.8
Salaried	4.7	4.7	0.0	9.2	12.6	-3.4
Casual agriculture	15.8	9.8	6.0***	7.0	9.2	-2.2**
Casual non-agriculture	4.3	3.7	0.7	9.5	14.7	-5.2*
N	717	656		872	892	

Note: *, **, *** indicates significance at 10% 5% and 1% level of significance. Own account workers are defined as self-employed individuals who run their own enterprises or farm without any hired help. Employers are defined as self-employed workers who have hired help for their enterprise or farm. The sample is restricted to those who report or are reported as in the workforce. Salaried included paid interns.

women who report themselves as self-employed are reported as contributing workers by their husbands, and 19 percent of women who report as contributing workers are reported as self-employed by their husbands. Interestingly wage work, which one might assume to have more clarity in terms of definition and visibility, shows substantive disagreements. Less than half of husbands agree with their wives on them being wage workers. About half of women reporting as unemployed are reported as out of the labour force by their husbands. About 18 percent of women who report themselves as out of the labour force are viewed as contributing workers by their husbands.

Even for men for whom employment activities tend to be major and more visible, the disagreements between self and proxy reporting on activity type are substantial. Agreements are highest for wage work (67 percent) and self-employment (55 percent). Sixteen percent of men who report themselves as self-employed are reported as contributing workers and 41 percent of male contributing workers are reported as self-employed by their wives, reflecting the confusion in the interpretation of these two categories. More than half of men who report as unemployed or out of the labour force are reported as working in an employment activity by their wives. This might be due to social desirability bias by women who do not want to report their husbands lack of economic activity or men hiding from their wives that they are

unemployed or out of the labour force.

Table 4: Self and proxy activity-wise match

Self reported	Proxy reported				
	Self employed	Contributing worker	Wage work	Unemployed	Out of labour force
<i>Women</i>					
Self employed	31.0	35.4	6.2	1.8	25.7
Contributing worker	18.5	51.2	2.4	5.1	22.8
Wage work	17.1	14.1	47.1	3.5	18.2
Unemployed	10.1	20.3	4.4	15.9	49.3
Out of labour force	5.3	17.6	4.1	8.2	64.8
<i>Men</i>					
Self employed	54.9	15.7	13.5	1.7	14.3
Contributing worker	40.8	16.9	23.9	1.9	16.4
Wage work	10.8	7.8	67.2	3.0	11.2
Unemployed	22.6	12.9	19.4	19.4	25.8
Out of labour force	26.9	8.7	26.5	3.5	34.4

Note: The table reports the mismatch in the activity distribution reported by self and proxy. The rows are the distribution reported by self-reports and the columns are the corresponding reports by proxies. The sum of each row adds up to 100.

5.4 Regression results

Regression results corroborate the descriptive analyses that proxies significantly underreport women’s labour market outcomes, but there are no statistically significant differences in self-proxy reporting about men’s labour market outcomes. Table 5 presents the results of estimation of equation 1 and 2 for work force participation rate. Proxy (husband’s) report of women’s workforce participation is five percentage points lower than that reported by women themselves. This result holds after we add individual, household controls, and enumerator fixed effects. These results are robust to the inclusion of individual fixed effects.

Proxies (wives) report two percentage points higher workforce participation for men as compared to their own reporting. But this difference is not statistically significant across any of the models. Proxies also underreport women’s labour force participation but not men’s labour force participation. No statistically significant difference exists between the unemployment rate reported by self and proxy respondents (Table A3). Reinforcing the descriptive analysis, the weekly hours of work (unconditional) are underreported for women by proxies, but this divergence between self and proxy does not carry over to the conditional weekly hours of work.

Further, we estimate a fixed effects model for each employment activity, controlling for all

Table 5: Differences in self-proxy reporting for workforce participation

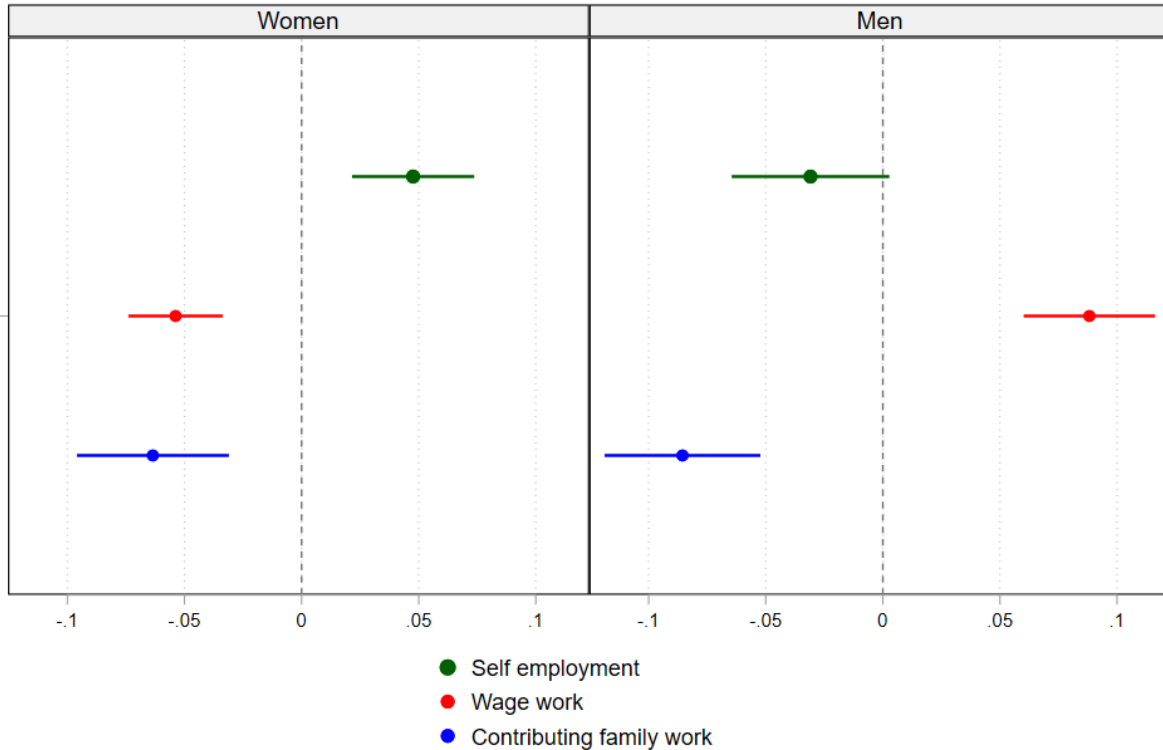
	Women			Men		
	(1)	(2)	(3)	(4)	(5)	(6)
Proxy	-0.05*** (0.02)	-0.05*** (0.02)	-0.05*** (0.02)	0.02 (0.02)	0.02 (0.02)	0.02 (0.02)
Individual controls	No	Yes	No	No	Yes	No
Household controls	No	Yes	No	No	Yes	No
Enumerator fixed effects	No	Yes	Yes	No	Yes	Yes
Individual fixed effects	No	No	Yes	No	No	Yes
N	2,268	2,268	2,268	2,268	2,268	2,268

Notes: Dependent variable is 1 if reports/reported as working, 0 otherwise. Independent variable, proxy, takes value 1 if proxy reported and 0 if self-reported. Individual and household controls include respondent and proxy education, age and age-squared, respondent’s major activity status in the year, household social group, landowning status, household size, share of dependents, age and education difference between respondent and proxy. The reported number of observations (N) belong to a pooled sample where the individual appears twice: once with self reported employment status and then proxy reported employment status. Hence, the number of observations is double that of number of individuals. Robust standard errors are reported in parenthesis. *, **, *** indicates significance at 10% 5% and 1% level of significance.

individual-specific attributes, and, as before, introduce an independent variable that captures whether a reported status is self or proxy reported (Table A4). The coefficient estimate gives the extent to which a particular activity type is under/over reported when reported by the proxy. Men over report women in self employment by 5 percentage points and under report their participation status in casual wage and contributing family work by a similar extent. Placed in the context of the overall workforce participation model, this reiterates the findings from table 3 that for women, under-reporting of work is also combined with misattribution of kinds of employment.

While men’s overall employment estimates was not significantly different between self and proxy reports (Table 2), within each activity type there is some mismatch. Wives under-report their husbands’ self-employment and contributing family work, while their wage employment is over-reported.

Figure 1: Estimates for differences in self and proxy reporting by employment activity



Note: The coefficients correspond to the estimates for the proxy variable from equation 2, estimated for each employment activity. The dependent variable is 1 if reports/reported as employed in that activity, 0 otherwise. Independent variable (proxy) is 1 if reported by proxy, 0 if self-reported. Individual fixed effects and enumerator fixed effects are included as controls. Standard errors are robust.

5.5 Correlates of divergence between self and proxy reports

We now turn to examining the correlates of divergence between the self and proxy reports. for overall and each employment activity. The estimation was specified in Equation 3, where the dependent variable takes the value one if the individual reports themselves as working but their spouse disagrees; and zero if the individual reports themselves as working and so does the spouse. This is labelled as under reporting as we are working under the assumption that the self-report is the true value or will be closer to the true value relative to the proxy report.

Table 6: Correlates of under-reporting, women

	(1) Overall	(2) Self employment	(3) Contributing worker	(4) Wage
Wife's attributes				
Age	-0.01 (0.02) (0.00)	0.09 (0.07) (0.00)	-0.03 (0.02) (0.00)	-0.04 (0.07) (0.00)
Education (Base:Illiterate)				
Primary/Middle	0.03 (0.62)	-0.13 (-0.53)	-0.12 (1.73)	-0.00 (-0.04)
Secondary+	0.05 (0.72)	0.16 (0.65)	0.15 (1.45)	-0.06 (-0.26)
Husband's attributes				
Age	-0.01 (0.01)	-0.05 (0.03)	0.02** (0.01)	0.02 (0.02)
Education (Base:Illiterate)				
Primary/Middle	-0.16*** (0.05)	0.13 (0.23)	-0.15* (0.08)	0.02 (0.13)
Secondary+	-0.23*** (0.06)	0.05 (0.25)	-0.21** (0.10)	-0.00 (0.18)
Asymmetric information covariates				
Woman's major activity(Base:OOWF)				
Self Employed	-0.11 (0.07)	-0.19 (0.27)	0.06 (0.10)	-0.59* (0.31)
Contributing work	-0.06 (0.06)	-0.41 (0.37)	0.08 (0.08)	-0.34 (0.24)
Wage	-0.16** (0.06)	0.03 (0.34)	0.10 (0.09)	-0.45*** (0.17)
Spouses in same activity	-0.12*** (0.04)	-0.03 (0.15)	-0.18*** (0.06)	-0.22 (0.15)
Total weekly hours worked	-0.00 (0.00)	0.01* (0.01)	-0.00 (0.00)	-0.00 (0.00)
Age difference	0.00 (0.01)	0.02 (0.06)	-0.03* (0.01)	-0.03 (0.03)
Educational difference	0.02*** (0.01)	-0.02 (0.03)	0.02*** (0.01)	0.04** (0.02)
Asymmetric measurement covariates				
Difference in time on domestic work (wife - husband)	0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	0.00 (0.00)
Husband's mother employed	-0.11** (0.04)	0.21 (0.21)	-0.17*** (0.06)	-0.10 (0.14)
Household attributes				
Asset tercile(Base:Poorest)				
Middle	-0.06 (0.05)	0.03 (0.29)	-0.11 (0.07)	0.01 (0.13)
Richest	-0.02 (0.05)	0.30 (0.27)	-0.08 (0.08)	0.09 (0.17)
Social group(Base:General/OBC)				
SC/ST	-0.00 (0.05) (0.62)	0.21 (0.25) (0.62)	-0.15** (0.07) (0.68)	0.23 (0.14) (0.72)
Observations	716	143	487	185

This sample is conditioned on women who report as being in that employment. Dependent variable is 1 if the husband disagrees and reports her as not employed, and 0 if he agrees that she is employed.

Other controls include age-squared, number of adult females and dependents (children below 6 years and adults above 65 years), enumerator, PSU and state fixed effects

Robust standard errors in parentheses. *, **, *** indicates significance at 10% 5% and 1% levels.

In the reporting of overall employment for women (column 1), variables that proxy for asymmetric measurement and information error are significant correlates of under-reporting. The woman's primary employment status over the year, and specifically being in wage work, significantly reduces the likelihood of her employment being under reported by their spouses. Therefore, asymmetric information error could explain some of the divergences in reporting; if the woman is primarily a wage worker for the majority of the year, her work is likely to be less marginal/intermittent and more observable, and hence, less likely to be under reported. Not surprisingly, spouses being employed in the same activity is negatively associated with under reporting. Presumably, the common activity facilitates greater information sharing between spouses. On the flip side, educational difference between spouses had a significant impact on increasing likelihood of under reporting. A similar relation was observed by Kapur et al. (2021) in explaining the self-proxy difference in select cities in North India.

Asymmetric measurement errors also explain under reporting to a certain extent. Women whose mother-in-laws were reported as employed (by their husbands) were less likely to have their own employment under reported. All other individual-specific attributes including education and age have no significant association with the likelihood of under reporting. Interestingly, the spouse's attributes have some association with under-reporting. Husbands with higher levels of education are less likely to under report their wives' employment.

For self-employment, neither asymmetric information error nor asymmetric measurement error had a significant impact on explaining reporting differences. Note that self employment is more likely to be over reported rather than under reported, and so under reporting of self-employment only accounts for a small share of the sample. For contributing work, asymmetric measurement error explains some of the under reporting. Specifically, if the husband reported that their mothers were in employment, they were less likely to under report their wife's employment. Similar to overall employment, widening educational difference between the spousal pair increased the probability of under reporting.

For wage work, unlike the other employment types, the woman's predominant activity for the majority of the year mattered. If the woman reported herself as principally employed in self-employment or wage work for the year, they were less likely to be under reported. This is potentially a result of two process at work – reduced asymmetric information error since the woman's primary activity during the year is an economic activity. This also suggests that social norms in that situation do not dissuade men from recognising women's work.

Besides the woman's major activity status of the year, none of her other attributes

including her age, educational attainment influenced the likelihood of her work being under-reported, both overall as well as across different kinds of employment. Interestingly, the under-reporting was also unaffected by the hours of work women spent in that activity.

Interestingly, for wage work, indicators that capture the role of norms did not come into play. Husband's reporting of their mother's employment or the time spent in household reproduction did not have any significant impact on under-reporting of wage work. Therefore the major source of under reporting of wage work was asymmetric information whereas for unpaid work, asymmetric measurement errors also contribute to the mismatch. For self-employment, neither explained under reporting. However, since for self-employment, the major direction of reporting error was towards over reporting (A4), we examine the correlates of over reporting of women's self employment by their husbands (Table 7).

To understand correlates of over reporting, we estimate a similar regression as in 6. As discussed previously, the dependent variable takes the value 1 if the woman is reported as employed by the husband, while she herself reports as not employed. It is 0 if both the woman and her husband agree that she is not in employment. When men report women as self-employed, only in 10 percent of these cases, women are reporting themselves as not in employment. Therefore, the over-reporting of self employment for women is essentially misattribution of the kind of employment women are undertaking, in particular of their unpaid contributing family work (Table 3).

Table 7: Correlates of over-reporting, women

	Overall	Self employment	Contributing worker	Wage
Wife's attributes				
Age	-0.02 (0.02)	0.01 (0.01)	-0.03* (0.02)	0.00 (0.01)
Education (Base:Illiterate)				
Primary/Middle	0.08 (0.07)	-0.02 (0.03)	-0.01 (0.05)	0.02 (0.02)
Secondary+	-0.01 (0.09)	-0.05 (0.04)	-0.00 (0.06)	0.06** (0.03)
Husband's attributes				
Age	0.02 (0.01)	0.00 (0.01)	0.03** (0.01)	0.01 (0.00)
Education (Base:Illiterate)				
Primary/Middle	0.11 (0.07)	0.04 (0.03)	0.07 (0.05)	0.04* (0.02)
Secondary+	-0.01 (0.08)	0.04 (0.04)	0.00 (0.05)	-0.01 (0.02)
Asymmetric information covariates				
Woman's major activity(Base:OOWF)				
Self Employed	0.08 (0.20)	0.26*** (0.08)	0.18** (0.07)	0.03 (0.02)
Contributing work	0.23** (0.11)	0.13*** (0.03)	0.18** (0.08)	-0.01 (0.02)
Wage	0.20** (0.08)	0.10*** (0.03)	0.04 (0.05)	0.17*** (0.04)
Spouses in same activity	-0.14** (0.06)	-0.02 (0.02)	-0.07* (0.04)	0.01 (0.02)
Age difference	-0.01 (0.02)	-0.01 (0.01)	-0.03** (0.02)	-0.00 (0.01)
Educational difference	0.00 (0.01)	0.00 (0.00)	-0.00 (0.01)	0.00 (0.00)
Asymmetric measurement covariates				
Difference in time on domestic work (wife - husband)				
	0.00 (0.00)	0.00 (0.00)	-0.00 (0.00)	0.00 (0.00)
Husband's mother employed				
	0.08 (0.07)	0.05** (0.02)	0.07 (0.05)	0.02 (0.02)
Household attributes				
Asset tercile (Base:Poorest)				
Middle	-0.04 (0.08)	0.02 (0.03)	-0.01 (0.04)	0.00 (0.02)
Richest	-0.06 (0.09)	0.03 (0.03)	-0.06 (0.05)	-0.02 (0.02)
Social group (Base: General/OBC)				
SC/ST	-0.05 (0.07)	-0.01 (0.03)	-0.08* (0.04)	-0.01 (0.02)
Observations	372	855	552	806

This sample is conditioned on women who report as not being in that employment. Dependent variable is 1 if the husband disagrees and reports her as employed, and 0 if he agrees that she is not employed.

Other controls include age-squared, number of adult females and dependents (children below 6 years and adults above 65 years), enumerator, PSU and state fixed effects

Robust standard errors in parentheses. *, **, *** indicates significance at 10% 5% and 1% levels.

We use the same correlates as the under reporting model. Overall over reporting increases if the woman is predominantly employed as a contributing family worker or wage worker for the majority of the year.

For self employment, the major sources of over-reporting emerges from women who are predominantly in contributing family work being over-reported as self-employed. This points towards the misattribution that we saw earlier 4. Therefore, women who were in wage work or contributing family work for the majority of the year were being reported as self-employed by their husbands. Further, both variables capturing norms were significant correlates of over reporting of self-employment - if husbands reported their mothers as employed, they were more likely to over report. Large differences in time spent in household production by the wife and husband indicating more skewed distribution of care work and adherence to norms contributed to higher likelihood of over reporting.

Similarly, the over reporting of contributing work is largely for those women who are primarily in self-employment or unpaid family work, again indicating the misattribution of work. Wage work was over reported among women who were primarily wage workers.

We undertook a similar analysis of the correlates of under reporting for men's employment (Appendix table A5). For men's employment (overall and across employment types), their predominant activity during the year was the only significant correlate of under reporting. If men were in self-employment or family work for the majority of the year, they were less likely to be under reported. Similarly, higher number of hours in employment in that week reduced the probability of under reporting, a factor that had no significant influence on the under reporting of women's employment. Not surprisingly, the proxy, i.e. the wife's attributes or covariates of asymmetric measurement, had any bearing on the under reporting of men's employment. We find similar results in the case of over reporting of men's work - major activity status is the main correlate of likelihood of over reporting, across employment types (see Table A6 in Appendix).

6 Framing Experiment Results

In this section, we discuss the results of the framing experiment that was conducted to investigate the impact of employment questions on reporting. First, we discuss the characteristics of the sample across treatment arms. Next, using descriptive and regressions analysis, we compare the employment outcomes from the three arms of the experiment.

Most sample characteristics are not statistically different between the single weekly and the other two arms of the experiment (detailed weekly and single daily) (Table 8). Only the yearly activity status of women varies significantly across the experimental arms. To

Table 8: Sample characteristics by experimental arm

	Single Weekly (1)	Detailed Weekly (2)	Single Daily (3)	Difference (1-2) (4)	Difference (1-3) (5)
Women (%)	55	57	62	-2	5*
<i>Average age (in years)</i>					
Men	41	40	40	1	1
Women	41	39	41	2	0
<i>Men's education</i>					
Illiterate	42	37	42	5	0
Primary/Middle	22	25	20	-3	2
Secondary & above	36	38	38	-2	-2
<i>Women's education</i>					
Illiterate	44	50	49	-6	-5
Primary/middle	29	26	20	3	9
Secondary & above	27	24	31	3	-4
<i>Social group (%)</i>					
Scheduled Caste/Tribe	37	42	46	-5	-9
Other Backward Caste	54	51	48	3	6
General	8	8	6	0	2
<i>Men's yearly activity status (%)</i>					
Self employed	42	44	44	-2	-2
Contributing worker	3	2	3	1	0
Wage work	37	40	41	-3	-4
Out of workforce	18	14	12	4	6
<i>Women's yearly activity status (%)</i>					
Self employed	7	9	8	-2	-1
Contributing worker	19	15	12	4	7**
Wage work	27	34	38	-7	-10**
Out of workforce	46	42	43	4	3
N	327	2,415	300		

Note: Column 4 refers to the difference between sample characteristic for the single weekly and detailed weekly. Column 5 refers to the difference between sample characteristic for single weekly and single daily modules. The significance stars are from t-test of these differences. *, **, *** indicates significance at 1% 5% and 10% level of significance.

account for these differences, we control for yearly activity status in the regression analysis.

6.1 Employment outcomes

Table 9 presents the reported employment outcomes from each of the three arms of the experiment separately, for women and men. In the single weekly arm, an individual is considered part of the workforce if they respond in the affirmative to the single employment question. In the detailed weekly, if an individual responds in the affirmative to any of the multiple questions on various employment activities, then they are considered as part of the workforce. In the single daily arm, if an individual reports as being employed in any of the days of the last week for at least an hour, they are included as part of the workforce.

Table 9: Employment estimates by experimental arm

	Single weekly (1)	Detailed weekly (2)	Single Daily (3)	Difference (1-2) (4)	Difference (1-3) (5)
Women					
Labourforce participation rate	48.3	63.2	57.3	-15***	-9.0
Workforce participation rate	48.3	58.2	55.7	-9.9***	-7.4
Unemployment rate	1.15	7.9	5.21	-5.0***	0.0
Avg. daily hours of work (unconditional)	4.8	3.4	3.3	1.4	-1.1
Avg. daily hours of work (conditional)	9.9	5.9	5.9	4.0***	0.0
Men					
Labourforce participation rate	78.2	73.8	83.5	4.3	-5.2
Workforce participation rate	78.2	71.8	82.6	6.4*	-4.4
Unemployment rate	0.0	2.7	4.7	-2.0***	0.0
Avg. daily hours of work (unconditional)	8.4	5.0	5.8	3.4***	2.1***
Avg. daily hours of work (conditional)	10.8	7.0	7.2	3.8***	0.1
N	327	2,415	300		

Note: Daily hours of work (unconditional) correspond to average daily hours worked, averaged across all responses, irrespective of reported employment status (if not working, then they are assumed to have zero hours). Daily hours of work (conditional) corresponds to average daily hours worked conditional on reporting as being employed. Column 4 refers to the difference between sample characteristic for the single weekly and detailed weekly. Column 5 refers to the difference between sample characteristic for single weekly and single daily modules. The significance stars are from t-test of these differences. *, **, *** indicates significance at 1% 5% and 10% level of significance.

Asking multiple questions about employment activities increases the reported labour and work force participation rate of women significantly. When responding to multiple questions on weekly employment activities, women’s LFPR is 63 percent, while this is only 48 percent when a single weekly question is asked. For men, asking multiple questions leads to a decline in workforce participation rate from 78 percent to 74 percent. The average hours of work are higher when a single weekly question is asked as opposed to when multiple questions are asked for both women and men.

Shorter reference period - daily vs. weekly - increases the reported labour force and work force participation rates for both women and men. These differences are not significant,

partly due to the small sample size of these experimental arms. Women’s labour force participation increases by 9 percentage points and men’s increases by 5 percentage points when asked one question about each of the previous seven days as compared to one question about the previous week.

As the balance table 8 shows, there is a clear difference in the size and composition of the sample between the three experiment arms. Given this we first match the sample size between the main and experiment arms of the survey by randomly drawing a random sample subset of individuals from the main survey equivalent in size to the experiment arms. This is pooled with the experiment data. We run a simple linear probability model with employment estimate as the dependent variable, and as controls we introduce dummy variables to account for the source of the estimate (single weekly, single daily, detailed weekly). We estimate bootstrapped coefficients for the dummy variables. We draw multiple sub-samples from the main survey and estimate the linear probability model with bootstrapped coefficients for the dummy variables accounting for the source of estimate. The final coefficient reported is the average of the bootstrapped coefficients. We also do a similar linear probability model with controls to account for the difference in the sample composition.

For this sample, we estimate bootstrap coefficients and standard errors. In order to account for variation in the sample composition between the two surveys, we introduce controls for individual and household characteristics that vary significantly between the samples. Table 10 provides the estimates of the coefficient for the source dummy where the variable takes the value one if the estimate comes from the main survey, otherwise zero. Without controls, we find that women’s employment estimate increases significantly when we move from single weekly to detailed weekly while for men there is no significant change. Since the samples were unbalanced with a significant difference in the distribution of women in terms of their employment, we include employment type as a control. Interestingly, while women’s employment increases when moving from single to detailed weekly, for men there is a decrease. Therefore, asking detailed questions to elude employment information results in a lower share of men being reported as employed. ⁹.

6.2 Activity reporting

Finally, we find that the higher employment estimates from asking detailed questions about the week is a result of higher reporting of contributing family work.

⁹refer appendix for note on power calculation

Table 10: Employment estimates for activities by experimental arm

	(1)	(2)	(3)	(4)
<i>Women</i>				
Labour Module (Base:Single weekly)				
Detailed weekly	0.10*** (0.04)	0.07 (0.05)	0.09** (0.00)	0.07*** (0.00)
Single daily	0.07 (0.05)	0.07 (0.06)	0.07** (0.00)	0.05*** (0.00)
N	1,739	1,190	544	312
<i>Men</i>				
Labour Module (Base:Single weekly)				
Detailed weekly	-0.06* (0.04)	-0.06 (0.04)	-0.06 (0.00)	-0.06 (0.00)
Single daily	0.04 (0.05)	0.03 (0.06)	0.04 (0.00)	0.03 (0.00)
N	1,302	884	383	253
Individual controls	No	Yes	No	Yes
Bootstrap standard errors	No	No	Yes	Yes

Notes: Columns (1)-(4) show regression results for workforce participation on questionnaire source. Dependent variable is 1 if individual is in the workforce, 0 otherwise. Each estimate represents the coefficient on the categorical variable indicating the source of the employment estimate. Controls include individual's age, education, activity status for the majority of the year, social group and household assets. *, **, *** indicates significance at 1% 5% and 10% level of significance.

7 Conclusion

Accurate information on employment is important for policy makers. However, as this paper shows, women's employment estimates are subject to several types of errors. Who is asked the survey questions has a significant impact on the reported level of women's labor force participation rate. Labour surveys typically approach the household head to collect information on employment and other demographic characteristics of all members of the household. We demonstrate that there significant under reporting when men report about women's work. For men, there is no such difference between men's self reported employment estimates and the reports of them by their spouses. We find that the under-reporting of women's work is a result of biases regarding women's expected status as well information asymmetry between spouses.

Table 11: Distribution of activities by different labour modules

	Single Weekly (1)	Detailed weekly (2)	Single daily (3)	ttest (1-2)	ttest (1-3)
<i>Women</i>					
Self-employed	8.9	11.5	6.5	-2.6	2.4
Contributing worker	20.4	24.8	18.4	-4.4	2
Wage worker	22.1	21.9	38.4	0.2	-16.3***
Out of workforce	48.5	41.8	36.8	6.7	11.7**
N	167	1,033	185		
<i>Men</i>					
Self-employed	42.4	34.0	38.3	8.4*	4.1
Contributing worker	7.6	20.7	4.4	-13.1***	3.2
Wage worker	36.4	17.1	46.1	19.3***	-9.7
Out of workforce	13.6	28.2	11.3	-14.5***	2.3
N	132	1,369	115		

Notes: Columns 4 and 5 denote two sided t-tests for differences between single weekly vs detailed weekly and single weekly vs single daily respectively. *, **, *** indicates significance at 1% 5% and 10% level of significance.

The framing and recall period also have a significant impact on reported levels of labor force participation. A single question about women's employment is likely to miss many employed women. Rather, calling out each kind of employment activity leads to significant increase in reported estimates of women's employment. Also asking separate questions about each day of the week leads to higher reporting of women's work as compared to asking a single question for the entire week.

National statistical agencies need to be mindful of who and how they ask labor force questions as the reported levels change. Asking multiple people in the household about their own labor market status when feasible as is done in time use surveys in India or World Bank's Living Standards Measurement Survey (LSMS) or increasingly in other surveys is one option to be considered. In the Indian context, national statistical agencies should change its practice of asking a single question about employment to multiple questions that call out different activities to better capture women's work. This change is relatively easy, cheap and very effective. It will also help align national surveys better with ILO's recommended method of capturing labor market outcomes.

References

- Ambler, K., Doss, C., Kieran, C., and Passarelli, S. (2021). He says, she says: Spousal disagreement in survey measures of bargaining power. *Economic Development and Cultural Change*, 69(2):765–788.
- Anker, R. (1983). Female labour force participation in developing countries: A critique of current definitions and data collection methods. *International Labour Review*, 122:709.
- Bardasi, E., Beegle, K., Dillon, A., and Serneels, P. (2011). Do labor statistics depend on how and to whom the questions are asked? Results from a survey experiment in Tanzania. *The World Bank Economic Review*, 25(3):418–447.
- Benes, E. and Walsh, K. (2018). Measuring employment in labour force surveys: Main findings from the ILO LFS pilot studies. *ILO*.
- Bernhardt, A., Field, E., Pande, R., Rigol, N., Schaner, S., and Troyer-Moore, C. (2018). Male social status and women’s work. *AEA Papers and Proceedings*, 108:363–67.
- Biemer, P. P. and Lyberg, L. E. (2003). *Introduction to survey quality*. John Wiley & Sons.
- Blair, J., Menon, G., and Bickart, B. (2004). Measurement effects in self vs. proxy response to survey questions: an information-processing perspective. *Measurement errors in surveys*, pages 145–166.
- Comblon, V., Robilliard, A.-S., and Roubaud, F. (2017). *Gender analysis of labour market outcomes in Sub-Saharan Africa: Recent evidence from Cameroon and Mali*. UN Women.
- Deshmukh, N., Desai, S., Pramanik, S., and Tiwari, D. (2020). Improving the measurement of women’s work: Sensitivity of work participation rates to survey design. *NCAER National Data Innovation Centre Measurement Brief*, 1.
- Dillon, A., Bardasi, E., Beegle, K., and Serneels, P. (2012). Explaining variation in child labor statistics. *Journal of Development Economics*, 98(1):136–147.
- Dixon, R. B. (1982). Women in agriculture: Counting the labor force in developing countries. *Population and Development Review*, pages 539–566.
- Dixon-Mueller, R. B. and Anker, R. (1988). Assessing women’s economic contributions to development. *ILO*.
- Fisher, M., Reimer, J. J., and Carr, E. R. (2010). Who should be interviewed in surveys of household income? *World Development*, 38(7):966–973.
- Galdo, J., Dammert, A. C., and Abebaw, D. (2021). Gender bias in agricultural child labor: Evidence from survey design experiments. *The World Bank Economic Review*, 35(4):872–891.
- Greenwood, A. M. (2000). The hours that we work: the data we need, the data we get. *Bulletin of Labour Statistics, Geneva*, (1).

- Groves, R. M. (2005). *Survey errors and survey costs*. John Wiley & Sons.
- Hirway, I. (2010). Understanding poverty: Insights emerging from time use of the poor. In *Unpaid work and the Economy*, pages 22–57. Springer.
- Hirway, I. (2015). Unpaid work and the economy: linkages and their implications. *Indian Journal of Labour Economics*, 58(1):1–21.
- Jayachandran, S. (2021). Social norms as a barrier to women’s employment in developing countries. *IMF Economic Review*, (69).
- Kapur, D., Vaishnav, M., and Verley, D. (2021). What is women’s work? Reflections from four north Indian urban clusters. *Urbanisation*, 6:S20–S39.
- Kilic, T., Van den Broeck, G., Koolwal, G., and Moylan, H. (2022). Are you being asked? impacts of respondent selection on measuring employment in malawi. *Journal of African Economies*.
- Koolwal, G. B. (2021). Improving the measurement of rural women’s employment: Global momentum and survey priorities. *World Development*, 147:105627.
- Langsten, R. and Salen, R. (2008). Two approaches to measuring women’s work in developing countries: A comparison of survey data from Egypt. *Population and Development Review*, 34(2):283–305.
- Twyman, J., Useche, P., and Deere, C. D. (2015). Gendered perceptions of land ownership and agricultural decision-making in Ecuador: who are the farm managers? *Land Economics*, 91(3):479–500.
- UN (2005). Household sample surveys in developing and transition countries. *Studies in Methods, Series F., No. 96*.

Appendix

Table A1: Sample characteristics for spousal and non spousal pairs (18 years and above)

Characteristics	Men		Difference	Women		Difference
	Spousal pair	Non spousal pair		Spousal pair	Non spousal pair	
Age	42.33	32.56	-9.77***	36.77	39.8	3.02***
	0.3	0.48	0.57	0.25	0.45	0.51
Education status						
below secondary	0.67	0.44	-0.23***	0.78	0.76	-0.03
	0.01	0.02	0.02	0.01	0.01	0.02
above secondary	0.33	0.56	0.23***	0.22	0.24	0.03
	0.01	0.02	0.02	0.01	0.01	0.02
Religion						
Hindu	0.95	0.93	-0.02	0.97	0.96	-0.01
	0.01	0.95	0.01	0	0.01	0.01
Non-hindu	0.05	0.07	0.02	0.03	0.04	0.01
	0.01	0.01	0.01	0	0.01	0.01
Social group						
SC	0.24	0.24	0	0.25	0.26	0.01
	0.01	0.01	0.02	0.01	0.01	0.02
ST	0.13	0.12	-0.01	0.14	0.12	-0.02
	0.01	0.01	0.01	0.01	0.01	0.01
OBC	0.49	0.47	-0.01	0.49	0.48	-0.01
	0.01	0.02	0.02	0.01	0.02	0.02

Table A2: Detailed weekly module: blocks and questions

Detailed Weekly (Main Survey Questions)	
Core Block	
Employment category	Employment Questions
Self employment	Last week, did you do any kind of business, farming or other self-employed activity to generate income, even if only for one hour?
Unpaid family helper	Last week, did you assist without pay in a business/farm/livestock of a household or family member even if only for one hour?
Wage work (casual/salaried)	Last week, did you work for a wage, salary, commission or any payment in kind, including doing paid domestic work, even if only for one hour?
Apprentice/Intern	In the last week, did you work for pay as an apprentice, intern or trainee even if only for one hour?
Small scale production	Last week, did you engage in small scale production of goods or services at home that were exchanged for cash or kind even if only for one hour?
Unpaid volunteer	Last week, did you work as an unpaid volunteer or do any kind of unpaid social work even if only for one hour?
Recovery block	
Overall recovery	Did you miss out reporting any work activities that led to you earning an income, or helping household members with an activity that generates an income even if only for one hour?
Short absence block	
Duration of absence	Even though you did not work last week, when do you expect to go back to work?
Type of work on return	If you do work in general for wages/profit, what do you normally do?
Underemployment block	
Desire for more work	Would you have wanted to do more work for pay or profit in the last week?
Hours underemployed	Did you have the time to do more work in the last week for pay or profit (in addition to the work you were already doing)? How many hours in the week would you have had time to do more work?
Unemployment block	
Search for jobs	In last week did you look for work either through employment exchanges, intermediaries, friends or relatives, or apply for work with prospective employers?
Willingness to work	In the last week, although you did not look for work, were you willing to work if work was available?
For self-consumption block	
Goods produced	In the last week, did you engage in any production of goods for own/household consumption, even if for only an hour?

Table A3: Difference between self and proxy reports for other labour market outcomes

	Women			Men		
	(1)	(2)	(3)	(4)	(5)	(6)
Labour force participation	-0.06*** (0.02)	-0.06*** (0.02)	-0.06*** (0.02)	0.02 (0.02)	0.02 (0.02)	0.02 (0.02)
Unemployment rate	0.00 (0.01)	0.00 (0.01)	0.01 (0.02)	-0.00 (0.02)	-0.00 (0.01)	0.00 (0.01)
Weekly hours of work (unconditional)	-2.36*** (0.85)	-2.36*** (0.77)	-2.36*** (0.66)	0.05 (1.10)	0.01 (1.02)	0.05 (0.92)
Weekly hours of work (conditional)	-1.16 (1.07)	-0.77 (1.03)	-1.03 (1.00)	-0.70 (1.05)	-0.74 (1.05)	-1.26 (0.98)
Individual Controls	No	Yes	No	No	Yes	No
Household Controls	No	Yes	No	No	Yes	No
Individual Fixed Effects	No	No	Yes	No	No	Yes
N	2,268	2,268	2,268	2,268	2,268	2,268

Note: Dependent variable is 1 if reports/reported as working, 0 otherwise. Independent variable proxy takes value 1 if employment is reported by proxy, 0 if self-reported. Controls include respondent and proxy education, age and age-squared, respondent's major activity status in the year, household social group, landowning status, household size, share of dependents, age and education difference between respondent and proxy, and enumerator fixed effects. *, **, *** indicates significance at 10% 5% and 1% level of significance.

Table A4: Estimates for differences in self and proxy reporting by employment activity

	Women			Men		
	(1)	(2)	(3)	(4)	(5)	(6)
Self Employment	0.05*** (0.01)	0.05*** (0.01)	0.05*** (0.01)	-0.03 (0.02)	-0.03 (0.02)	-0.03* (0.02)
Wage	-0.05*** (0.01)	-0.05*** (0.01)	-0.05*** (0.01)	0.09*** (0.02)	0.09*** (0.02)	0.09*** (0.01)
Contributing work	-0.06*** (0.02)	-0.06*** (0.02)	-0.06*** (0.02)	-0.09*** (0.02)	-0.09*** (0.02)	-0.09*** (0.02)
Individual controls	No	Yes	No	No	Yes	No
Household controls	No	Yes	No	No	Yes	No
Enumerator fixed effects	No	Yes	Yes	No	Yes	Yes
Individual fixed effects	No	No	Yes	No	No	Yes
N	2,268	2,266	2,268	2,268	2,266	2,268

Notes: Dependent variable is 1 if reports/reported as working in that activity, 0 otherwise. Independent variable proxy is 1 if reported by proxy, 0 if self-reported. Individual and household controls include respondent and proxy education, age and age-squared, respondent's major activity status in the year, household social group, landowning status, household size, share of dependents, age and education difference between respondent and proxy. The reported number of observations (N) belong to a pooled sample where the individual appears twice: once with their self reported employment status and then their proxy reported employment status. Hence, the number of observations is double that of number of individuals. Robust standard errors are reported in parenthesis. *, **, *** indicates significance at 10% 5% and 1% level of significance.

Table A5: Correlates of under-reporting, men

	Overall	Self employment	Contributing worker	Wage
Husband's attributes				
Age	-0.01 (0.02)	-0.01 (0.02)	-0.00 (0.04)	-0.11* (0.06)
Education (Base:Illiterate)				
Primary/Middle	-0.04 (0.04)	-0.11 (0.08)	-0.27** (0.13)	0.01 (0.12)
Secondary+	-0.01 (0.05)	-0.07 (0.10)	-0.21 (0.19)	-0.01 (0.15)
Wife's attributes				
Age	0.00 (0.01)	0.01 (0.01)	-0.00 (0.03)	0.04 (0.04)
Education (Base:Illiterate)				
Primary/Middle	-0.01 (0.03)	-0.07 (0.08)	0.06 (0.11)	-0.17 (0.10)
Secondary+	-0.02 (0.05)	-0.01 (0.09)	0.02 (0.14)	-0.20 (0.13)
Asymmetric information covariates				
Man's major activity (Base:OOWF)				
SE	-0.18 (0.12)	-0.04 (0.20)	0.05 (0.24)	-0.42 (0.48)
Unpaid	-0.25* (0.14)	0.72*** (0.27)	-0.34 (0.36)	
Wage	-0.15 (0.12)	0.31 (0.21)	-0.03 (0.25)	-0.52 (0.48)
Total weekly hours worked	-0.00 (0.00)	-0.01*** (0.00)	-0.00 (0.00)	-0.00 (0.00)
Asymmetric measurement covariates				
Spousal diff in domestic work				
	-0.00 (0.00)	-0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Age difference	0.00 (0.01)	0.02 (0.02)	-0.00 (0.04)	0.05 (0.05)
Educational difference	-0.00 (0.01)	0.00 (0.01)	0.02 (0.02)	-0.00 (0.01)
Household attributes Asset tercile (Base:Poorest)				
Middle	0.06* (0.04)	0.11 (0.08)	-0.05 (0.13)	-0.00 (0.10)
Richest	0.04 (0.04)	0.03 (0.08)	-0.15 (0.15)	0.15 (0.17)
Social group (Base: General/OBC)				
SC/ST	-0.00 (0.04)	-0.09 (0.09)	0.11 (0.13)	0.10 (0.12)
Karnataka	0.14 (0.45)	0.99 (0.78)	1.43 (0.99)	0.52 (0.46)
Observations	744	432	295	226

Robust standard errors in parentheses. *, **, *** indicates significance at 10% 5% and 1% levels.

Table A6: Correlates of over-reporting, men

	Overall	Self employment	Contributing worker	Wage
Husband's attributes				
Age	0.09 (0.05)	0.00 (0.03)	-0.01 (0.01)	0.02 (0.01)
Education (Base:Illiterate)				
Primary/Middle	-0.01 (0.15)	-0.05 (0.06)	0.04 (0.04)	0.05 (0.05)
Secondary+	-0.02 (0.13)	-0.13** (0.07)	0.06 (0.05)	0.02 (0.05)
Wife's attributes				
Proxy age	-0.02 (0.02)	-0.01 (0.02)	-0.01 (0.01)	-0.01 (0.01)
Education (Base:Illiterate)				
Primary/Middle	0.00 (0.13)	0.05 (0.05)	-0.04 (0.04)	-0.01 (0.05)
Secondary+	-0.10 (0.17)	0.01 (0.07)	-0.10** (0.05)	0.04 (0.06)
Asymmetric information covariates				
Man's major activity				
SE	0.23 (0.17)	0.39*** (0.08)	0.02 (0.06)	-0.09 (0.05)
Contributing work	-0.98*** (0.28)	0.31** (0.15)	-0.14 (0.19)	-0.11 (0.14)
Wage	-0.02 (0.14)	-0.08 (0.08)	-0.03 (0.06)	0.21*** (0.06)
Asymmetric measurement covariates				
Difference in time in domestic work (wife - husband)				
	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Age difference	-0.04 (0.03)	-0.00 (0.02)	-0.02 (0.01)	-0.01 (0.01)
Educational difference	-0.00 (0.01) (0.07) (0.05)	0.01 (0.01) (0.02) (0.03)	-0.01 (0.00) (0.02) (0.02)	0.01 (0.01) (0.02) (0.02)
Household attributes				
Asset tercile (Base:Poorest)				
Middle	0.00 (0.12)	0.06 (0.05)	-0.04 (0.04)	0.03 (0.04)
Richest	0.02 (0.14)	0.04 (0.06)	0.08* (0.05)	0.01 (0.05)
Social group (Base:General/OBC)				
SC/ST	-0.06 (0.16)	0.03 (0.06)	0.04 (0.04)	-0.05 (0.05)
Karnataka	0.13 (0.77)	-0.45 (0.50)	-0.13 (0.39)	-0.13 (0.33)
Observations	234	546	683	752

Robust standard errors in parentheses. *, **, *** indicates significance at 10% 5% and 1% levels.

Table A7: Impact on employment estimates by experimental arm

	(1)	(2)	(3)	(4)
Labourforce participation rate (LFPR)				
<i>Women</i>				
Labour module (Base:single weekly)				
Detailed weekly	0.15*** (0.04)	0.14*** (0.04)	0.10*** (0.00)	0.12*** (0.00)
Single daily	0.09* (0.05)	0.09 (0.06)	0.09*** (0.00)	0.07*** (0.00)
N	1,739	1,190	547	306
<i>Men</i>				
Labour module (Base:single weekly)				
Detailed weekly	-0.04 (0.04)	-0.04 (0.03)	-0.01*** (0.00)	-0.04*** (0.00)
Single daily	0.05 (0.05)	0.03 (0.04)	0.05*** (0.00)	0.02*** (0.00)
N	1,302	884	380	249
Unemployment rate (UR)				
<i>Women</i>				
Labour module (Base:single weekly)				
Detailed weekly	0.07*** (0.01)	0.06** (0.03)	0.10** (0.00)	0.12*** (0.00)
Single daily	0.04 (0.02)	0.03 (0.04)	0.09*** (0.00)	0.07*** (0.00)
N	1,062	711	296	170
<i>Men</i>				
Labour module (Base:single weekly)				
Detailed weekly	0.03** (0.01)	0.02*** (0.01)	0.06*** (0.00)	0.06*** (0.00)
Single daily	0.05** (0.02)	0.05* (0.03)	0.04*** (0.00)	0.03*** (0.00)
N	979	657	321	203
Individual controls	No	Yes	No	Yes
Bootstrap standard errors	No	No	Yes	Yes

Notes: Dependent variable is 1 if individual is in the workforce, 0 otherwise. Each estimate represents the coefficient on the categorical variable indicating the source of the employment estimate. Controls include individual's age, education, activity status for the majority of the year, social group and household assets. *, **, *** indicates significance at 1% 5% and 10% level of significance.

Table A8: Activity wise employment estimates by experimental arm

	Women				Men			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Self employment								
Detailed weekly	0.02	0.04	-0.02***	-0.004***	-0.02	0.06	-0.07***	-0.02***
	0.04	0.05	0	0	0.05	0.05	0	0
Single daily	-0.06	-0.02	-0.06***	-0.03***	-0.01	0.08	-0.01***	0.05***
	0.05	0.05	0	0	0.07	0.06	0	0
N	984	647	290	159	944	637	313	209
Wage work								
Detailed weekly	-0.05	0.01	0.04***	-0.04***	-0.17***	-0.22***	-0.18***	-0.2***
	0.06	0.06	0	0	0.05	0.04	0	0
Single daily	0.13*	0.07	0.10***	0.02***	0.04	-0.04	0.05***	-0.003***
	0.07	0.07	0	0	0.07	0.06	0	0
N	990	652	521	299	957	645	406	275
Contributing family worker								
Detailed weekly	0.03	-0.03	0.06***	0.05***	0.20***	0.17***	0.30***	0.28***
	0.05	0.06	0	0	0.03	0.03	0	0
Single daily	-0.07	-0.02	-0.06***	0.01***	-0.03	-0.03	-0.03***	-0.02***
	0.07	0.07	0	0	0.04	0.04	0	0
N	984	647	290	159	944	637	313	209
Individual controls	No	Yes	No	Yes	No	Yes	No	Yes
Bootstrapped SE	No	No	Yes	Yes	No	No	Yes	Yes

Note: Columns (1)-(4) (women) and (5)-(8) (men) show regression results for workforce participation (by activity) on questionnaire source. Dependent variable takes the value 1 if individual reports working in self employment, wage or CFW and 0 otherwise (conditional on being in the workforce). Controls include individual's age, education, activity status for the majority of the year, social group and household assets. *, **, *** indicates significance at 1% 5% and 10% level of significance.