



INTERNATIONAL
FOOD POLICY
RESEARCH
INSTITUTE

IFPRI



RESEARCH
PROGRAM ON
Policies,
Institutions,
and Markets

Led by IFPRI

IFPRI Discussion Paper 02050

October 2021

Labor (Mis?)Measurement in Agriculture

Kate Ambler

Sylvan Herskowitz

Mywish K. Maredia

Markets, Trade, and Institutions Division

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

The International Food Policy Research Institute (IFPRI), a CGIAR Research Center established in 1975, provides research-based policy solutions to sustainably reduce poverty and end hunger and malnutrition. IFPRI's strategic research aims to foster a climate-resilient and sustainable food supply; promote healthy diets and nutrition for all; build inclusive and efficient markets, trade systems, and food industries; transform agricultural and rural economies; and strengthen institutions and governance. Gender is integrated in all the Institute's work. Partnerships, communications, capacity strengthening, and data and knowledge management are essential components to translate IFPRI's research from action to impact. The Institute's regional and country programs play a critical role in responding to demand for food policy research and in delivering holistic support for country-led development. IFPRI collaborates with partners around the world.

AUTHORS

Kate Ambler (k.ambler@cgiar.org) is a Research Fellow in the Markets, Trade, and Institutions Division at the International Food Policy Research Institute (IFPRI), Washington, DC.

Sylvan Herskowitz (s.herskowitz@cgiar.org) is a Research Fellow in IFPRI's Markets, Trade, and Institutions Division, Washington, DC.

Mywish K. Maredia (maredia@msu.edu) is a Professor in the Department of Agricultural, Food, and Resource Economics at Michigan State University, Lansing, MI.

Notices

¹ IFPRI Discussion Papers contain preliminary material and research results and are circulated in order to stimulate discussion and critical comment. They have not been subject to a formal external review via IFPRI's Publications Review Committee. Any opinions stated herein are those of the author(s) and are not necessarily representative of or endorsed by IFPRI.

² The boundaries and names shown and the designations used on the map(s) herein do not imply official endorsement or acceptance by the International Food Policy Research Institute (IFPRI) or its partners and contributors.

³ Copyright remains with the authors. The authors are free to proceed, without further IFPRI permission, to publish this paper, or any revised version of it, in outlets such as journals, books, and other publications.

Abstract

Livelihoods are changing rapidly in rural areas. Measuring and categorizing peoples' labor activities in relation to the agricultural sector is important for understanding income earning opportunities and designing effective policy. Conventional data collection methods ask about individuals' main work activities over the past year. Descriptions are recorded in the field, post-coded, and eventually categorized. This approach is costly to collect, fatiguing for respondents, and may create distortions. We show that a more direct approach, asking respondents to categorize their major work activities themselves, provides similar resulting data despite some caveats and lessons for best enumeration practices. We compare these main activities to a series of yes/no questions about participation in a set of specific work tasks. We find a 12% incidence of "missing" work, whereby individuals who reported participation in at least one but did not have any recorded major activities. Looking by sector of work, women and youth are disproportionately more likely to have agricultural contributions "missed," while we find no corresponding bias in undercounting of non-agricultural work. Finally, we test the effect of randomly positioning the task-based questions before the listing of major activities but do not find significant effects on the number or type of activities reported.

Acknowledgements

This project was funded by the CGIAR Research Program on Policies, Institutions, and Markets and 3ie, and was approved by the IFPRI Institutional Review Board. We thank Mike Murphy for his project and data management as well as the team at Innovations for Poverty Action Ghana for project implementation, research assistance, and data collection, particularly Usamatu Salifu, Nicole Gargano, Salifu Amadu, Hassan Moomin, Federica Di Battista, and Madeleen Husselman. Order of authorship is alphabetical.

1. Introduction

Livelihoods and work opportunities in rural areas are rapidly evolving as the preeminence of on-farm agriculture diminishes (Yeboah and Jayne 2018; de Vries, Timmer, and de Vries 2015; McMillan, Rodrik, and Sepúlveda 2017). Accurate measurement of labor activities both in general and with respect to the agricultural sector is important for establishing a clear understanding of rural work opportunities, structural transformation, and for developing effective social programs and policies. Capturing this information systematically and completely can be challenging, particularly in rural settings where a large share of peoples' work is irregular and informal (Fox et al. 2016; Fox and Thomas, 2016; La Porta and Shleifer, 2014). Existing literature examining rural employment has both tried to document levels of participation and to categorize the sector or types of work that people are doing (Dolislager et al. 2020; Mueller and Lee, 2019; Yeboah and Jayne, 2018). A widespread approach in this literature uses a labor module format whereby individuals' main activities are listed for different recall windows along with a brief description.¹ These descriptions are later converted into codes and, eventually, sorted into different types of work categories. This method is arduous, costly, and may be prone to biases or errors.

In this paper we examine the challenges related to measurement and categorization of labor supply in the agricultural sector, both on and off the farm and make three main contributions. First, we explore a promising alternative method of categorizing individuals' main work activities, leveraging respondents' own information about household members' work and its relation to the agricultural sector. We compare this to the conventional, but costlier, method of recording and sorting descriptions of each activity. Second, we ask a series of yes/no questions about every

¹ This labor module format is notably used in most Living Standards Measurement Study surveys, described in Schaffner (2020).

individual's participation in a set of nine specific work tasks, regardless of their reported main activities, and show that for a substantial proportion of the sample, participation in these tasks is "missed" in the recording of main activities. Both agricultural and non-agricultural work have a similar incidence of missing participation, but women and especially youth are more likely to have their agricultural work left out of main activity listings. Finally, by randomly placing these task-based questions either before or after the activity listings, we test the impact of their location on recorded activities but find a precisely estimated null effect.

This paper uses data from a survey with a sample of farmer group members in rural Northern Ghana. The labor module follows a frequently used format whereby enumerators ask about each eligible household member's primary and secondary activities over the previous year as well as their primary activity over the past week.² In the conventional "code-based" approach, each reported activity is described by respondents and recorded by enumerators. Field supervisors then filter these responses into a set of occupation and industry codes and, using these codes, researchers decide how to classify each activity. This process is time consuming, tiring, and susceptible to distortions at each step of the process.

We compare this standard approach with an alternate, "direct question," approach that classifies employment activities by asking respondents themselves to indicate the appropriate categorization, in our case whether they were on- or off-farm, whether the off-farm activities were linked to agriculture, and whether those activities were as employees for a wage or for household led businesses. This approach leverages the respondent's own interpretation of the work, making it less time consuming than providing full descriptions and therefore less likely to both create and

² Respondents were allowed to confer with household members if they were available or to have them respond for themselves, but in practice this happened rarely, less than 20% of the time.

be vulnerable to response fatigue.³ However, it may suffer from respondents' inconsistent interpretation of questions and definitions. We find that these two approaches result in broadly similar characterizations of work types between on-farm, off-farm agricultural, and non-agricultural work activities. However, more than half of self or household employment activities from the survey code approach are considered wage or employee work by respondents themselves in the direct question method and 47% of off-farm agricultural activities identified with survey codes are characterized as non-agricultural by respondents themselves. We see this alternative approach as promising but also requiring extra care in question design and data collection.

We then examine a third method in which respondents are asked a series of yes/no questions about their participation in a set of work-related tasks, both agricultural and non-agricultural. These questions are asked of all eligible household members, regardless of responses to the activities reported in the labor module. Using these questions, we can see how often people are reported to have participated in work tasks that are not captured in the reporting of main activities. Twelve percent of individuals in the sample participated in work tasks but were not reported to have had major activities. Although women and youth's participation in non-agricultural work is similarly likely to be missed as men and older individuals, their agriculture-linked work is significantly more likely to be missed in the activity listing, which may more broadly bias our understanding of participation in agriculture for these groups.

Finally, randomly assigned placement of these task-based questions in the survey, either before or after the listing of main activities, allows us to test whether their location impacts activity reporting. Asking about an individuals' work tasks could remind main respondents about their work and make them more likely to report additional relevant work activities for that individual.

³ In an experiment within this same study, Ambler, Herskowitz, and Maredia (2021) find that response fatigue in labor modules leads to losses in reported labor activities and bias against certain demographic groups.

They could alternatively make respondents feel like this contribution had already been captured and less likely to report that activity. We do not see any average affects in either direction. We do, however, find a modest effect on the reporting of the tasks themselves, in which reporting agricultural activities first reduces recognition of participation in agricultural tasks.

This paper contributes to a growing literature trying to measure rural labor participation. Our results speak to methods of capturing and categorizing work, as well as potential for biases in underreporting the contributions of certain groups. A growing number of studies use the conventional approach referred to in this study to make similar categorizations of work activities with respect to the agriculture sector. Yeboah and Jayne (2018) describe changing patterns across countries and time using full time equivalents. Dolislager et al. (2020) employ a similar categorization with a focus on rural and urban differences. This paper makes a methodological contribution to this literature, proposing a method for reducing survey costs and response burden. While applied here to the agriculture sector, this approach could be adapted to other relevant classifications or sectors as well.

This paper also links to a broader literature studying the most effective ways to collect information on labor supply in household surveys in developing countries. This growing literature includes studies on the use of proxy respondents (Ambler, Herskowitz, and Maredia 2021, Beegle et al. 2012b, Serneels et al. 2016, Bardasi et al. 2011), different ways to ask questions and level of detail (Bardasi et al. 2011, Deininger et al. 2019, Langsten and Salen 2008, Comblon et al. 2015, and Benes and Walsh 2018), screening questions (Martin and Polivka 1995, Serneels et al. 2016, Dillon et al. 2012, and Fox and Pimhidzai 2013), and recall windows (Beegle et al. 2012a, Heath et al. 2020, Das et al. 2012, Deininger et al. 2019, Gaddis et al. 2020), and Arthi et al. 2016). This paper expands on these methodological contributions, testing a less costly method of achieving

similar data objectives, identifying gaps and possible biases in a conventionally used data collection approach, and testing the effects of framing questions on data quality.

This paper proceeds by describing the data collection and coding approach in Section 2. In Section 3 we present the results. We conclude with a brief discussion in Section 4.

2. Data and Data Treatment

The primary source of data in this paper comes from a household survey conducted with a rural sample in Northern Ghana that took place between April and June of 2019, covering 12 districts in four regions. Respondents were members of farmer business organizations organized by the Ghana Agricultural Sector Investment Programme (GASIP) for the purpose of involvement in their agricultural programs.⁴ The full sample includes 3,806 individuals from 1,095 households and 4,358 unique work activities.

2.1 Activity Categorization Approach 1: Survey Codes

The recording of labor activities in the labor module was designed to follow a structure similar to those frequently used in LSMS and other surveys. The module progresses one eligible household member at a time and begins by asking whether that member had worked at all over the last year. If so, enumerators record information about the member's primary and secondary activities in the last 12 months, and their main activity in the last week. For each activity, enumerators recorded a text description of the occupation and industry. Later, these responses are reviewed and categorized by a data collection supervisor using the International Standard Classification of Occupations (ISCO) and the International Standard Industrial Classification (ISIC). Following a similar approach to that used by Yeboah and Jayne (2018), we then attribute each possible occupation and industry code to one of three categories: on-farm agricultural work,

⁴ See Ambler, de Brauw, and Murphy, 2020 for further details and results of the study.

off-farm agricultural work (a business or activity linked directly to an agricultural product or its processing), or not agriculturally linked. When categorizations were ambiguous or conflicting between occupation and industry codes, we examined the codes and the text responses to assign the correct designation. While this approach essentially eliminated cases that could not be assigned, this might be intractable in larger surveys, where these data are less carefully recorded, or with public data where text responses are not available.⁵

Employee status was derived from a separate question asking the respondent to describe their work as working for someone else for pay, self-employed, without pay for a household business, an apprentice, on the household farm (including livestock farm), or as an employer themselves. We reduce this to two categories—self or household employment and employee or wage employment.

This code-based approach has the potential to offer a high level of detail on peoples' activities if done carefully and completely, using data that is commonly collected on large household surveys. However, distortions may enter at each stage in the data recording and categorizing process, and lead to substantial aggregate errors. The process is also costly in terms of time and energy spent by the data collection team, researchers, and respondents themselves that may be excessive if the researcher is most interested in a pre-identified set of categories.

2.2 Activity Categorization Approach 2: Direct Questions

Given the high costs of the code-based approach, we designed new questions to categorize work activities intended to leverage the respondents' own information about work being done by household members. In addition to the descriptions recorded as part of the code-based process,

⁵ Ultimately, 44 out of 3,850 individuals had a work activity that we could not categorize and were excluded from the analysis sample.

enumerators asked respondents to categorize each reported activity as: 1) Household farm work, 2) Household business, 3) Wage work – on farm, or 4) Wage work – off farm.

We assume that any on-farm work is linked to agriculture, but activities classified as household businesses or off-farm wage work were followed with a question asking whether this activity was linked to agriculture. Employee status is considered to be for the household or oneself for responses one and two and for wage or as an employee for responses three and four.

2.3 Agriculture and Non-Agriculture Linked Tasks

Finally, we also incorporated an unconditional set of nine questions for every household member about their involvement in a set of work tasks. These did not rely on responses to the major reported activities. Seven yes/no response questions were asked about individuals' participation over the last twelve months in work tasks linked to agricultural: pre-harvest or harvest work, post-harvest work, and livestock-related work on household farms, agriculture-linked household business, wage/salary work on non-household farms, wage/salary work on post-harvest activities, and wage/salary work linked to off-farm agricultural activities. They were additionally asked two broad questions about participation in non-agricultural work either related to household business or for a wage/salary. See Appendix A for the full phrasing of questions. We note that we could have divided these agricultural questions into on or off-farm work. For the purpose of comparing these responses with those recorded in the listed activities, we opted for the simpler distinction of agricultural versus non-agricultural work as we were unsure that respondents' would consider on and off-farm work on the same crop as distinct activities. We discuss this further in next section when presenting the results.

These questions were incorporated as an alternative way to capture the extensive margin of work participation. Because they are not characterized as being “primary” or “secondary” main

activities, it is possible that they reflect lower intensity levels than those recorded in the activity listing. However, they may still constitute meaningful labor contributions and be particularly vulnerable to underreporting by the household's primary respondent.

In addition to their descriptive value, we hypothesize that inclusion of these task-based questions could impact reported activities in two ways. They could reduce reported activities, either because of response fatigue caused by the additional questions or because respondents become less likely to report as full activities work that is similar to the tasks just recorded.⁶ Or conversely, these questions could increase reporting of activities by priming or prodding respondents to think about specific work tasks, thereby reducing the likelihood that these contributions are forgotten. We therefore randomize, at the household level, whether these task-based questions were asked before or after the listing of activities for each member.

3. Results

3.1. Comparing Survey Code and Direct Question Activity Categorizations

First, we contrast recorded levels of participation in different types of work activities using the two main methods of activity categorization: code based and direct questions. With individuals reporting up to three unique activities, participation in a type of work is indicated if any of their three activities falls in that category. Table 1 reports mean participation rates for the different types of work indicated at the start of each row using the categorization approach indicated at the top of each column. We see high levels of consistency between methods for all on-farm activities. Total participation rates in off-farm agricultural activity are similar, 11.5% and 9% using the codes and direct question approaches, respectively. However, differences are bigger when splitting these totals by self-employment versus wage employment. This is also true for non-agricultural work

⁶ Ambler, Herskowitz, Maredia (2020) shows evidence from the same study, that response fatigue based on randomized position in the household roster reduced the number of work activities reported for individual members.

activities which show general agreement about total levels of participation, 22.3% and 26.9% respectively, but less agreement when splitting by form of employment.⁷ With the caveat of disagreements over form of employment, this general consistency between methods is suggestive that the direct method may afford a plausible option, lower-cost alternative to the conventional code-based approach.⁸

Exploring this disagreement in employment status further, we analyze agreement in categorizations at the activity level in Table 2. Panel A shows a cross tabulation of whether activities are classified as self-employment/family business or as wage earners/employees using the two methods. As detailed above, the code-based method and the direct method use slightly different questions to make this distinction. The first uses a broader set of potential employment types and the second comes directly from the reported activity categorization. Panel A suggests that there is a high level of agreement, with 85.2% of activities described as self-employment with the code approach similarly categorized with the direct question approach. Agreement is 89.1% for activities initially categorized as wage employment using the survey code approach. However, Panel B removes on-farm work, and we see a sharp fall among activities categorized as self-employment to just 48.1% agreement with the direct question approach. This pattern explains a large share of the tension observed in Table 1 when disaggregating activities both by sector and form of employment.

Table 3 presents a similar cross-tabulation with the categories of on-farm work, off-farm agricultural work, and non-agricultural work. While on-farm agricultural work identified in the survey code approach is corroborated 96% of the time by the direct method and non-agricultural

⁷ A disaggregated version of Table 1 by gender is presented in Appendix Table A1 and shows similar patterns.

⁸ Appendix Table A2 also shows these categorizations using the Ghana Living Standards Measurement Study survey focusing on the rural sample. We find broadly similar levels of participation across work categorizations, which is encouraging for broader relevance of our analyses.

work activities are corroborated 93% of the time, off-farm agricultural work is only corroborated 50% of the time with 94% of these disagreements categorized as being non-agricultural in the direct question approach. Without an external benchmark of “truth” to resolve this conflict, we cannot say with certainty which approach is more accurate, although the level of specificity in the initial categorizations in the code approach may reflect deeper reflection than the quick generalizations of the direct question approach. At the same time, the codes were not created with the purpose of categorizing activities as related to agriculture or not, which leaves room for ambiguity in the recorded responses or distortions along the initial response to final categorization chain. Code-based classification in the future should carefully consider the context being studied to ensure activities are being appropriately classified while the direct question approach may need to provide more detailed definitions, characterizations, or examples of what is meant by agriculturally-linked work activities. Ultimately, the fact that most disagreement occurs between off-farm agricultural work and non-agricultural work highlights the difficulty in understanding off-farm employment in agricultural value chains.

3.2 Task-Based Labor Contributions and “Missing” Activities

Next, we use the task-based questions to examine another approach to capturing labor contributions. As discussed earlier, these questions were asked about every household member in the labor module, regardless of what was or was not reported as work activities. Notably, the placement of these binary questions was randomized to occur before or after recording individuals’ main activities. Table 4 examines the responses to these questions and explores whether certain work may be “missing” from individuals’ reported activities. Panel A focuses on agriculture-linked tasks while Panel B looks at non-agricultural work tasks.

The first column reports the incidence of individuals reported to have participated in different tasks anytime over the last 12 months in the full sample. In Panel A, the highest levels of participation were in on-farm work with over 85% of individuals participating in these activities both pre- and post-harvest. Over 50% of people were reported to have helped with livestock activities. Overall, individuals engaged in an average of over 3 out of 7 possible distinct agriculture-linked tasks.

A motivation for including these questions was to see how complete the different activity categorization approaches are in documenting agricultural participation. Columns (2) and (3) therefore restrict the sample to individuals who did not have any reported agriculture-linked activities using the code and direct-question based approaches respectively. Although the incidence of participation falls considerably for all types of tasks, they are far from zero. More than sixty percent of people without reported agriculture-linked activities have still participated in at least one agricultural task, with an average of more than two tasks. Focusing on people who did not have any agriculture-linked activities but *did* have at least one agriculture-linked task identified in columns (4) and (5), we see that these people look very similar to those in the full sample, participating in just over three tasks on average. As discussed, it could be that these contributions are not as substantive as main activities and intentionally excluded, that respondents do not perceive them to be as consequential (potentially suffering from the respondents' own biases), or are simply skipped as a result of forgetfulness or fatigue. Regardless of reason, and for both categorization methods, activity-based approaches are missing the agriculture-linked contributions of more than 15% of the sample.

Although the tasks for non-agricultural work were much more limited than those focused on agriculture, we conduct a similar analysis for missing non-agricultural work in Panel B. As

anticipated, participation in these tasks is much lower than for most of the agricultural-tasks, but almost a third of the sample participates in either a non-agricultural business or for a non-agricultural wage task, and approximately 20% of individuals without non-agricultural activities are reported as participating in non-agricultural tasks. The code-based approach appears to miss the non-agricultural work tasks of nearly 18% of individuals whereas the direct approach misses the tasks of 14% of the sample.

As noted, in the previous section, we could have divided the agriculture-linked tasks into those occurring on and off-farm. We are skeptical however that respondents would necessarily view these as distinct activities and so we opted for the joint and clearer distinction between agricultural and non-agricultural work. Still, we show this further division in Appendix Table A3 where we observe a very high share of missing off-farm agricultural work of more than 40%. However, we suspect that this is the result of these off-farm tasks being included in activities primarily categorized as being on-farm.

Table 5 explores missed work further by examining whether certain types of individuals are more or less likely to have “missing” activities. We first define an indicator for any “missed” work activities, independent of type, that is equal to one if participation in any work-related task was reported but no activities were recorded. We then code a binary indicator for individuals who *did* participate in agriculture-linked tasks but *did not* have any reported activities linked to agriculture and construct a similar indicator for non-agricultural work. Using a regression approach, we can then see if age and gender are significant predictors of missing work activities in column (1). Columns (2) and (3) use the codes-based approach to activity categorization for missing agriculture and non-agriculture linked work activities, whereas columns (4) and (5) use the direct question approach. We additionally control for whether the individual was conferred

with while conducting the labor module, household fixed effects, and relationship to primary respondent fixed effects (including being the primary respondent).

The first column suggests strong patterns of missed labor contributions for both women and younger age groups. Almost 12% of the sample was reported to have participated in some form of work task but was not reported as having had any major activities over the past year. Focusing on agricultural-linked work activities and task-based questions in columns (2) and (4), both categorization approaches retain similarly strong patterns of omitted agricultural work activities for women and youth. In column (2) using the code approach, we see that women are 3.4 percentage points more likely to have their agriculture engagement not reported relative to a mean incidence of 15.5%. The direct question approach appears to perform slightly worse, with women 4.5 percentage points more likely to have their activities missed relative to a mean incidence of 16.4%. While these relationships for women are substantial, reflecting 22-27% of the mean, youth have an even higher incidence of missed activities. Youth age 14-24 are nearly 20 percentage points more likely to have these activities missed relative to those aged 35-59 regardless of categorization methods with mean incidence in the sample of 15-16%. Even those age 25-34 are 12 percentage points more likely to have their agriculture activities missed than the older, omitted group of individuals age 23-59. However, these age-based differences do not exist with either approach for non-agricultural tasks in columns (3) and (5). The codes-based approach does still show a marginally significant relationship for women, though the magnitude is roughly 30% lower than for ag-linked work and there is no significant relationship at all when using the direct-question approach. While these contrasts between agriculture and non-agriculture linked work are notable, we hesitate to over-emphasize these differences both due to the unequal number of task-based questions asked in the surveys by sector and because the setting for our sample is likely to be

heavily skewed towards agricultural work relative to other populations and settings.

Together, these observations suggest that a large amount of work participation captured in the tasks are missed when using an activity listing approach. The contributions of women and youth are particularly likely to be under reported. Although incidence of missing activities is similar for agricultural and non-agricultural work, agricultural work, regardless of categorization approach, continues to suffer from these biases by age and gender.

3.3 Experimental Task Priming and Recorded Work Activities

Finally, we wanted to test whether the order of these task-based questions would impact reporting of agriculture-linked activities. Respondents might feel that having already reported these tasks, they did not need to provide additional details on this work in the form of recorded activities, especially as greater response fatigue sets in. Or conversely, using these task-based questions before the listing of activities could prime respondents and jog their memories about the labor contributions of their family members and increase reported activities. To test these effects, we randomized (at the household level) whether these task-based questions were asked before or after the activity participation questions.

In Table 6, we regress the total number of jobs reported for each individual and binary individual-level indicators for participation in agricultural or non-agricultural activities (based on both the codes and direct-question approaches) on whether the task-based questions were asked before (as opposed to after) the reporting of primary, secondary, and recent activities. The regressions additionally control for individuals' gender, age group, whether they were conferred with on these questions, and community fixed effects (household fixed effects cannot be used because the treatment is at the household level). We do not find meaningful impacts of these leading task-based questions on any of our outcomes. Conversely, we do see marginally significant

but small in magnitude increases (less than 5% of mean) of total tasks reported when they are conducted before, instead of after, the listing of activities. We do not however see any impact on any activity being listed suggesting that if there is any impact of activity questions on tasks, it is likely modest.

4. Conclusion

In this paper we study a number of methodological questions related to the documentation and categorization of work activities in household surveys, with a focus on the agricultural sector. In our sample of rural Ghanaians, we find that a direct approach of activity categorization, which may be less burdensome than a traditional approach based on occupation and industry codes, is promising. However, substantial disagreement between the two methods on the classification of off-farm activities suggests that further work is needed to ensure the accuracy of both approaches. Switching to a reliance on respondents' own characterizations increases the need for training of enumerators and clear explanation and reflection at the time of data collection.

We also document that task-focused questions pick up information about productive activities that are not captured in standard questions about primary and secondary activities, and work contributions documented in these task questions are frequently not captured in the listing of major work activities. This "missing" work is disproportionately common in the agricultural sector for women and youth. Existing work relying exclusively on labor supply captured in activity listings may be systematically biased against the contributions of women and youth. Finally, despite the possibility of improving the accuracy recall about individuals' work activities, asking task-based questions before listing activities does not affect the number or type of activities reported.

All of our results point to the importance of thoughtful survey design that reflects the objectives of researchers and policy-makers along with the paramount importance of careful training and enumeration during the data collection itself to ensure high quality and complete data.

References

- Ambler, Kate, Alan de Brauw, and Mike Murphy (2020). “Increasing the adoption of conservation agriculture: A framed field experiment in Northern Ghana,” IFPRI Discussion Paper 01932.
- Ambler, Kate, Sylvan Herskowitz, and Mywish Maredia (2020) “Are we done yet? Response fatigue and rural livelihoods,” IFPRI Discussion Paper 01980.
- Arthi, Vellore, Kathleen Beegle, Joachim De Weerd, and Amparo Palacios-Lopez (2016) Not your average job: Measuring farm labor in Tanzania: The World Bank.
- Bardasi, Elena, Kathleen Beegle, Andrew Dillon, and Pieter Serneels (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: 418–447.
- Beegle, Kathleen, Calogero Carletto, and Kristen Himelein (2012a) “Reliability of recall in agricultural data,” *Journal of Development Economics*, 98: 34–41.
- Beegle, Kathleen, Joachim De Weerd, Jed Friedman, and John Gibson (2012b) “Methods of household consumption measurement through surveys: Experimental results from Tanzania,” *Journal of Development Economics*, 98: 3–18.
- Benes, Elisa and Kieran Walsh (2018) “Measuring employment in Labour Force Surveys: Main findings from the ILO LFS pilot studies,” ILO.
- Comblon, Virginie, Anne-Sophie Robilliard et al. (2015) “Are female employment statistics more sensitive than male ones to questionnaire design? Evidence from Cameroon, Mali and Senegal,” Technical Report.
- Das, Jishnu, Jeffrey Hammer, and Carolina Sánchez-Paramo (2012) “The impact of recall periods on reported morbidity and health seeking behavior,” *Journal of Development Economics*, 98: 76–88.
- De Vries, Gaaitzen, Marcel Timmer and Klaas de Vries (2015). “Structural transformation in Africa: Static gains, dynamic losses.”, *Journal of Development Studies* 51(6): 674–88
- Deininger, Klaus, Calogero Carletto, Sara Savastano, and James Muwonge, “Can diaries help in improving agricultural production statistics? Evidence from Uganda,” *Journal of Development Economics*, 98: 42–50.
- Dillon, Andrew, Elena Bardasi, Kathleen Beegle, and Pieter Serneels (2012) “Explaining Variation in Child Labor Statistics,” *Journal of Development Economics*, 98: 136–147.

- Dolislager, Michael, Thomas Reardon, Aslihan Arslan, Louise Fox, Saweda Liverpool-Tasie, Christine Sauer, and David L Tschirley (2020) “Youth and adult agrifood system employment in developing regions: Rural (peri-urban to hinterland) vs. urban,” *The Journal of Development Studies*, 57(4): 571–593.
- Fox, Louise and Obert Pimhidzai (2013) *Different dreams, same bed: Collecting, using, and interpreting employment statistics in Sub-Saharan Africa—The Case of Uganda*. The World Bank.
- Fox, Louise., Lemma W. Senbet, and Witness Simbanegavi (2016). “Youth employment in Sub-Saharan Africa: challenges, constraints and opportunities,” *Journal of African Economies*, 25: i3-i15.
- Fox, Louise and Alun Thomas (2016). “Africa's got work to do: A diagnostic of youth employment challenges in Sub-Saharan Africa,” *Journal of African Economies*, 25: i16-i36.
- Gaddis, Isis, Gbemisola Oseni, Amparo Palacios-Lopez, and Janneke Pieters (2020) “Measuring Farm Labor: Survey Experimental Evidence from Ghana,” *The World Bank Economic Review*, DOI: <http://dx.doi.org/10.1093/wber/lhaa012>.
- Heath, Rachel, Ghazala Mansuri, Bob Rijkers, William Hutchins Seitz, and Dhiraj Sharma (2020) “Measuring employment: Experimental evidence from Urban Ghana,” *World Bank Policy Research Working Paper*.
- La Porta, Rafael and Andrei Shleifer (2014) “Informality and Development,” *Journal of Economic Perspectives*, 28 (3): 109–126.
- Langsten, Ray and Rania Salen (2008) “Two approaches to measuring women’s work in developing countries: A comparison of survey data from Egypt,” *Population and Development Review*, 34: 283–305.
- Martin, Elizabeth and Anne E Polivka (1995) “Diagnostics for redesigning survey questionnaires: Measuring work in the Current Population Survey,” *Public Opinion Quarterly*, 59: 547–567.
- McMillan, Margaret, Dani Rodrik, and Claudia Sepúlveda, eds. 2017. *Structural change, fundamentals, and growth: A framework and case studies*. Washington, DC, U.S.A.: International Food Policy Research Institute.
- Mueller, Valerie and Hak Lim Lee (2019). *Can Migration be a Conduit for Transformative Youth Employment?* In Mueller, V and Thurlow, J. (eds): *Youth and jobs in rural Africa: Beyond stylized facts*. Oxford, U.K.: Oxford University Press.
- Schaffner, Julie Anderson (2000) “Employment,” in Margaret Grosh and Paul Glewwe eds. *Designing household survey questionnaires for developing countries: The World Bank*, Chap. 9, pp. 135-141.

Serneels, Pieter, Kathleen Beegle, and Andrew Dillon (2016) Do returns to education depend on how and who you ask?, Policy Research Working Papers: The World Bank

Yeboah, Felix Kwame and Thomas S Jayne (2018) “Africa’s evolving employment trends,” *The Journal of Development Studies*, 54: 803–832.

Appendix A: Task-Based Work Questions

Respondents were asked to report participation in work tasks about all household members in the labor module regardless of what work related activities were listed for that individual. The first seven were explicitly linked to agriculture while the last two were targeting non-agricultural work tasks. The phrasing of questions was: “In the last 12 months did [NAME]...”

1. ...work on this household's farm in pre-harvest and harvest activities, including land preparation, planting, weeding, other pre-harvest activities, and harvesting?
2. ...work on this household's farm in post-harvest activities, including processing, storage, or transporting to market? Examples of processing include drying, shelling, cleaning, and others.
3. ...do any work related to household livestock including either care of animals or activities related to animal byproducts? [excluding livestock or livestock byproducts]
4. ...work in an agricultural-related household business, run by them or someone else in this household? This could include selling crops, livestock or animal byproducts at market on a regular basis.
5. ...work for a wage, salary, commission or any payment in kind, from work on others' farms in pre-harvest and harvest activities, including land preparation, planting, weeding, other pre-harvest activities, and harvesting?
6. ...work for a wage, salary, commission or any payment in kind, in post-harvest activities, including processing and storage? Examples of processing include drying, shelling, cleaning, and others.
7. ...work for a wage, salary, commission or any payment in kind, in off-farm agricultural activities?
8. ...work in non-agricultural-related household business, run by them or someone else in this household?
9. ...work for a wage, salary, commission or any payment in kind, in off-farm non-agricultural activities?

1 Main Paper Tables

Table 1: Participation in Ag and Non-Ag Activities by Categorization Approach

Work Type	Approach	
	Codes (1)	Direct (2)
<i>On-Farm</i>		
Self	0.729	0.724
Wage	0.013	0.013
Total	0.733	0.729
<i>Off-Farm Ag</i>		
Self	0.110	0.051
Wage	0.005	0.039
Total	0.115	0.090
<i>Non-Ag</i>		
Self	0.160	0.096
Wage	0.063	0.175
Total	0.223	0.269
N	3,806	3,806

Notes: Table shows participation at the individual level in different types of work activities as reported in their primary activity over the last year, secondary activity over the last year, or primary activity in the last seven days. The first three columns use a common approach of using industry and sector codes to categorize work along with relationship to business owner. The direct question approach directly asks respondents to define their reported work activities to make their own categorizations with respect to the agricultural sector.

Table 2: Correspondence Matrix of Employment Status by Method

		Direct Question		
		Self-Employed	Wage/Employee	Total
Survey Code	Self-Employed	3,447 85.2%	599 14.8%	4,046
	Wage/Employee	34 10.9%	278 89.1%	312
	Total	3,481 79.9%	877 20.1%	4,358 100.0%

		Direct Question		
		Self-Employed	Wage/Employee	Total
Survey Code	Self-Employed	501 48.1%	541 51.9%	1,042
	Wage/Employee	34 10.9%	278 89.1%	312
	Total	535 39.5%	819 60.5%	1,354

Notes: 85.5% overall agreement of activity employment characterization in Panel A. Agreement falls to 57.5% in Panel B after removing on-farm self/household activities. Self-employed includes work for a household farm or business. The Survey Code approach used a question with a set of five different relationships with employers whereas the Direct Question approach asked the respondent to make this binary distinction immediately. The gray shaded boxes show the incidence and share of agreement between the two methods.

Table 3: Correspondence Matrix of Work Tasks Categorizations by Method

		Direct Question			Total
		On-Farm	Off-Farm	Non-Ag	
Survey Code	On-Farm	2,947 96.46%	69 2.26%	39 1.28%	3,055
	Off-Farm	11 2.48%	223 50.23%	210 47.3%	444
	Non-Ag	5 0.58%	59 6.87%	795 92.55%	859
	Total	2,963 67.99%	351 8.05%	1,044 23.96%	4,358

Notes: This table shows the categorization of 4,358 individual tasks following two distinct approaches. Whole number values are counts of the frequency of that type of task as categorized by both methods. Percentages can be understood as row percentages, or the percent of tasks from the initial categorization stated at the left of the row that were categorized using the direct question approach to be in the group indicated in the column header. The gray shaded boxes show the incidence and share of agreement between the two methods.

Table 4: Participation in Agriculture and Non-Agriculture Linked Work Tasks

<i>Panel A:</i>					
<i>Agriculture Linked Tasks</i>	Full Sample	No Ag Activities		“Missed”	
		Codes	Direct	Codes	Direct
	(1)	(2)	(3)	(4)	(5)
On-Farm Pre-Harvest	0.866	0.576	0.582	0.908	0.916
On-Farm Post-Harvest	0.859	0.570	0.572	0.897	0.900
Livestock	0.513	0.308	0.316	0.485	0.498
Ag-Linked Business	0.227	0.100	0.108	0.158	0.170
Off-Farm Pre-Harvest	0.268	0.201	0.200	0.317	0.314
Off-Farm Post-Harvest	0.257	0.211	0.210	0.332	0.331
Off-Farm Other Ag	0.118	0.075	0.076	0.118	0.120
Any Ag Tasks	0.899	0.635	0.636	1.000	1.000
Total Ag Tasks	3.109	2.040	2.066	3.214	3.248
<i>N</i>	3806	920	970	584	617
<i>Panel B:</i>					
<i>Non-Ag Linked Tasks</i>	Full Sample	No Non-Ag Activities		“Missed”	
		Codes	Direct	Codes	Direct
	(1)	(2)	(3)	(4)	(5)
Non-Agricultural Business	0.184	0.147	0.120	0.642	0.617
Non-Agricultural Wage	0.204	0.124	0.107	0.541	0.552
Any Non-Ag Tasks	0.324	0.229	0.194	1.000	1.000
Total Non-Ag Tasks	0.388	0.271	0.227	1.183	1.169
<i>N</i>	3806	2958	2782	678	540

Notes: This table shows reported rates of participation in different agriculture (Panel A) and non-agriculture (Panel B) linked tasks. Participation in these tasks was asked about all eligible household members in the labor module, independent of whether any primary (or secondary etc.) activities were reported to indicate such participation. Columns two and three focus on sub-samples of people who were reported as not having any agriculture-linked activities in Panel A or no non-agriculture activities in Panel B, by the two different methods of activity categorization. Columns (4) and (5) focus on sub-samples of individuals who did not have any of these activities identified but who *were* reported to have contributed to at least one of these tasks. The method of this categorization is indicated at the top of columns (2)-(5).

Table 5: Predictors of Missing Ag and Non-Ag Activities - by Categorization Approaches

	Any (1)	Codes		Direct	
		Ag (2)	Non-Ag (3)	Ag (4)	Non-Ag (5)
Female	0.033*** (0.010)	0.034*** (0.012)	0.023* (0.013)	0.045*** (0.012)	-0.005 (0.012)
Age 14-24	0.192*** (0.020)	0.202*** (0.023)	-0.004 (0.018)	0.195*** (0.023)	0.011 (0.017)
Age 25-34	0.087*** (0.015)	0.121*** (0.017)	-0.001 (0.017)	0.121*** (0.018)	-0.005 (0.015)
Reported Own Labor	-0.045*** (0.015)	-0.088*** (0.016)	0.035** (0.017)	-0.095*** (0.016)	0.006 (0.016)
<i>N</i>	3720	3720	3720	3720	3720
Mean Y	0.118	0.155	0.177	0.164	0.141
Relationship FEs	Yes	Yes	Yes	Yes	Yes
HH FEs	Yes	Yes	Yes	Yes	Yes
R2	0.428	0.471	0.475	0.488	0.469

Notes: This table shows regression results of an indicator of “missing” activities on individuals’ gender, age group, and self-reporting, while controlling for relation to primary respondent and household fixed effects. The first column uses an indicator for individuals who had no activities reported but did participate in tasks of any type. The other four columns define missing with respect to agricultural or non-agricultural work, using the two different categorization approaches as indicated above each column. 86 individuals in households where they are the only members in the labor module are dropped from this regression after including household fixed effects.

Table 6: Effect of Task Question Location on Reported Activities and Tasks

	Total Jobs (1)	Ag-Linked (0/1)		Non-Ag (0/1)		Task Questions	
		Codes (2)	Direct (3)	Codes (4)	Direct (5)	Total (6)	Any (7)
Tasks before Activities	0.007 (0.027)	-0.006 (0.016)	0.005 (0.017)	0.005 (0.016)	0.008 (0.017)	0.164* (0.091)	0.022 (0.013)
Female	-0.012 (0.023)	-0.057*** (0.014)	-0.067*** (0.014)	-0.038*** (0.014)	0.008 (0.014)	-0.085 (0.054)	-0.026*** (0.010)
Age 14-24	-0.440*** (0.031)	-0.278*** (0.020)	-0.277*** (0.020)	-0.029 (0.018)	-0.071*** (0.018)	-0.457*** (0.081)	-0.069*** (0.014)
Age 25-34	-0.152*** (0.026)	-0.123*** (0.015)	-0.129*** (0.016)	0.042*** (0.016)	0.035** (0.017)	-0.213*** (0.065)	-0.011 (0.010)
Reported Own Labor	0.276*** (0.025)	0.166*** (0.014)	0.174*** (0.014)	0.010 (0.016)	0.032* (0.017)	0.651*** (0.060)	0.088*** (0.009)
<i>N</i>	3799	3799	3799	3799	3799	3799	3799
Mean Y	1.145	0.759	0.745	0.223	0.269	3.496	0.914
Community FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.176	0.214	0.224	0.067	0.072	0.227	0.093

Notes: This table shows regressions estimating the effect of the task based questions being asked before the reporting of major labor activities on the total number of unique job activities (0-3) and participation in activities linked to agricultural work (0/1) and activities not linked to agricultural work (0/1) by the two categorization methods as indicated in the column headings. Whether the task questions were asked before or after the listing of activities was randomly assigned at the household level. Regressions include controls for female, age group, self-reporting of labor, and community fixed effects. Standard errors are clustered at the household level, the level of the treatment. Seven observations are dropped due to one community having a single family (with no variation in location of task questions).

2 Appendix Tables

Table A.1: Participation in Ag and Non-Ag Activities by Categorization Approach

Work Type	Industry Code Approach			Direct Question Approach		
	(1) Male	(2) Female	(3) Total	(4) Male	(5) Female	(6) Total
<i>On-Farm</i>						
Self	0.776	0.692	0.729	0.775	0.684	0.724
Wage	0.024	0.005	0.013	0.021	0.007	0.013
Total	0.785	0.693	0.733	0.783	0.687	0.729
<i>Off-Farm Ag</i>						
Self	0.042	0.163	0.110	0.017	0.078	0.051
Wage	0.005	0.005	0.005	0.024	0.051	0.039
Total	0.047	0.168	0.115	0.041	0.129	0.090
<i>Non-Ag</i>						
Self	0.140	0.176	0.160	0.057	0.127	0.096
Wage	0.104	0.031	0.063	0.209	0.148	0.175
Total	0.244	0.206	0.223	0.265	0.272	0.269
N	1,675	2,131	3,806	1,675	2,131	3,806

Notes: This table shows participation at the individual level in different types of work activities as reported in their primary activity over the last year, secondary activity over the last year, or primary activity in the last seven days. All cells show mean levels of participation for the sample indicated at the top of the column. The first three columns use a common approach of using industry and sector codes to categorize work along with relationship to business owner. The direct question approach directly asks respondents to define their reported work activities to make their own categorizations with respect to the agricultural sector.

Table A.2: Participation in Activity Types with Code Approach in Rural Ghana LSMS Sample

Work Type		(1)	(2)	(3)
		Male	Female	Total
On-Farm	Self	0.680	0.628	0.652
	Wage	0.026	0.014	0.019
Off-Farm	Self	0.045	0.178	0.118
	Wage	0.022	0.018	0.020
Maybe Ag	Self	0.026	0.064	0.047
	Wage	0.012	0.006	0.009
Non-Ag	Self	0.076	0.047	0.060
	Wage	0.152	0.040	0.091
N		21,412	25,760	47,172

Notes: Table shows participation at the individual level in different types of work activities as reported in their primary activity over the last year, secondary activity over the last year, or primary activity in the last seven days with data from the rural sample of the Ghana LSMS in 2018. Industry codes and sector codes are used to categorize each reported activity with respect to agricultural work although some work activities remain ambiguous, where sufficient level of details over the activities or outputs are not available in the data. Cells show means of participation rates for each activity type in the group indicated at the top of each column.

Table A.3: Participation in Agriculture and Non-Agriculture Linked Work Tasks

<i>Panel A:</i>	Full Sample	No On-Farm Activities		“Missed”	
		Codes	Direct	Codes	Direct
<i>On-Farm Agriculture Linked Tasks</i>	(1)	(2)	(3)	(4)	(5)
On-Farm Pre-Harvest	0.866	0.577	0.581	0.945	0.943
On-Farm Post-Harvest	0.859	0.570	0.574	0.934	0.932
Livestock	0.513	0.306	0.308	0.502	0.501
Any On-Farm Ag Tasks	0.887	0.611	0.616	1.000	1.000
Total On-Farm Ag Tasks	2.238	1.454	1.464	2.381	2.376
<i>N</i>	3806	1015	1031	620	635
<i>Panel B:</i>	Full Sample	No Off-Farm Activities		“Missed”	
		Codes	Direct	Codes	Direct
<i>Off-Farm Agriculture Linked Tasks</i>	(1)	(2)	(3)	(4)	(5)
Ag-Linked Business	0.227	0.203	0.200	0.434	0.434
Off-Farm Pre-Harvest	0.268	0.271	0.267	0.579	0.580
Off-Farm Post-Harvest	0.257	0.259	0.252	0.554	0.547
Off-Farm Other Ag	0.118	0.112	0.110	0.239	0.239
Any Off-Farm Ag Tasks	0.491	0.468	0.461	1.000	1.000
Total Off-Farm Ag Tasks	0.870	0.845	0.830	1.805	1.800
<i>N</i>	3806	3369	3463	1577	1597
<i>Panel C:</i>	Full Sample	No Non-Ag Activities		“Missed”	
		Codes	Direct	Codes	Direct
<i>Non-Ag Linked Tasks</i>	(1)	(2)	(3)	(4)	(5)
Non-Agricultural Business	0.184	0.147	0.120	0.642	0.617
Non-Agricultural Wage	0.204	0.124	0.107	0.541	0.552
Any Non-Ag Tasks	0.324	0.229	0.194	1.000	1.000
Total Non-Ag Tasks	0.388	0.271	0.227	1.183	1.169
<i>N</i>	3806	2958	2782	678	540

Notes: This table shows reported rates of participation in different on-farm agriculture (Panel A), off-farm agriculture (Panel B), and non-agriculture (Panel C) linked tasks. Participation in these tasks was asked about all eligible household members in the labor module, independent of whether any primary (or secondary etc.) activities were reported to indicate such participation. Columns two and three focus on sub-samples of people who were reported as not having any activities linked to on-farm agriculture in Panel A, off-farm agriculture in Panel B, and non-agriculture in Panel C, by the two different methods of activity categorization. Columns (4) and (5) focus on sub-samples of individuals who did not have any of these activities identified but who *were* reported to have contributed to at least one of these tasks. The method of this categorization is indicated at the top of columns (2)-(5).

ALL IFPRI DISCUSSION PAPERS

All discussion papers are available [here](#)

They can be downloaded free of charge

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

www.ifpri.org

IFPRI HEADQUARTERS

1201 Eye Street, NW
Washington, DC 20005 USA
Tel.: +1-202-862-5600
Fax: +1-202-862-5606
Email: ifpri@cgiar.org