WPS5192

Policy Research Working Paper 5192

Do Labor Statistics Depend on How and to Whom the Questions Are Asked

Results from a Survey Experiment in Tanzania

Elena Bardasi
Kathleen Beegle
Andrew Dillon
Pieter Serneels

The World Bank
Development Research Group
Poverty and Inequality Team
January 2010

Policy Research Working Paper 5192

Abstract

Labor market statistics are critical for assessing and understanding economic development. In practice, widespread variation exists in how labor statistics are measured in household surveys in low-income countries. Little is known whether these differences have an effect on the labor statistics they produce. This paper analyzes these effects by implementing a survey experiment in Tanzania that varied two key dimensions: the level of detail of the questions and the type of respondent. Significant differences are observed across survey designs with respect to different labor statistics. Labor force participation rates, for example, vary by as much as 10 percentage points across the four survey assignments.

Using a short labor module without screening questions on employment generates lower female labor force participation and lower rates of wage employment for both men and women. Response by proxy rather than self-report yields lower male labor force participation, lower female working hours, and lower employment in agriculture for men. The differences between proxy and self reporting seem to come from information imperfections within the household, especially with the distance in age between respondent and subject playing an important role, while gender and educational differences seem less important.

This paper—a product of the Poverty and Inequality Team, Development Research Group—is part of a larger effort in the department to improve household survey methods. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The author may be contacted at kbeegle@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Produced by the Research Support Team

Do Labor Statistics Depend on How and to Whom the Questions Are Asked? Results from a Survey Experiment in Tanzania

Elena Bardasi
The World Bank

Kathleen Beegle
The World Bank

Andrew Dillon
International Food Policy Research Institute

Pieter Serneels
University of East Anglia

Key words: Labor statistics, Survey design, Tanzania, Field experiment
JEL code: J21, C83, C93
Word count: 8,343

* We would like to thank Economic Development Initiatives, especially Joachim de Weerdt, the supervisory staff, enumerators and data entry teams for thorough work in the field. We also thank Gero Carletto, Louise Fox, Annette Jäckle, David Newhouse, Dominique Van de Walle, and seminar participants at IFPRI, ASSA annual meetings, IZA/World Bank Conference, CSAE University of Oxford, and the World Bank for very useful comments. This work was supported by the World Bank Research Support Budget and PREM-Gender. All views are those of the authors and do not reflect the views of the World Bank or its member countries. Corresponding author is Andrew Dillon a.dillon@cgiar.org

1. Introduction

Labor market statistics are critical for assessing and understanding how an economy functions. They are paramount for identifying sources of income, especially in developing countries and particularly for poor people for whom labor is their most important asset. In practice, how labor market statistics are captured in household surveys varies widely. Recall period, question sequencing, the use of screening questions, and the respondent interviewed vary across surveys, both within and across countries. The wealth of evidence on the quality and reliability of labor statistics in household surveys comes largely from the United States (see Bound, Brown and Mathiowetz., 2001, for a thorough review). In the developing country context, little is known whether these differences have an effect on the labor market statistics produced, and whether they affect the estimates of the economic relationships estimated from the data. It is difficult to extrapolate from the U.S. studies because employment patterns are quite different in East Africa than in the U.S. which may create alternative impacts of measurement error on labor statistics and parameter estimates.[1] In this paper, we provide evidence from a survey experiment in which we varied respondent type (self-reporting versus proxy respondent) and the level of detail of the labor module to study variations in labor force participation, hours, earnings, sectoral distribution, and employment status.

Although the method by which employment data are collected may have a large effect on the labor market statistics, few studies in low-income countries have attempted to

[1] An important part of the U.S. literature focuses on response 'errors' which impact unemployment measurement. In most of this literature (Abowd and Zellner, 1985; Poterba and Summers, 1986), the emphasis focuses on developing procedures to adjust for spurious transitions arising because of classification errors (in a panel setting) rather than on analyzing which aspects of the questionnaire or the interview process originate those errors.

rigorously measure the effect of differences in the specific characteristics of the survey on employment statistics generated from the survey. Moreover, these effects may vary across groups in the population – for example, employment statistics for women and children may be particularly sensitive to the survey method.[2] When the wording of the employment questions emphasizes the main activity in order to define employment status, this may lead to underestimating the number of economically active women because of the large female presence among unpaid agricultural and family workers (Dixon-Mueller and Anker, 1988). Child and teenage work may be similarly underreported.[3]

Different survey methods are expected to yield different measurement errors, and such errors may bias both survey statistics as well as estimates of relationships between measures of employment and other variables. This is true when labor variables are right-hand-side variables and when they are dependent variables. In continuous dependent variables, classic measurement error in the dependent variable does not bias OLS point estimates (although it may reduce precision). However, if the measurement error is introduced by the respondent "actively" trying to guess the true value of a variable of which they do not have exact information (this may be the case of a proxy respondent) it is possible that, in the linear regression framework, biases in the coefficients of interest are introduced – especially if both regressor and outcome variables are measured with error (Hyslop and Imbens, 2001). Moreover, in the case of discrete dependent variables the assumption of classical measurement error is not appropriate; that is, measurement error in discrete dependent variables, such as labor

[2] Guarcello et al. (2009) review the discrepancies in child labor statistics across surveys of different design within a number of low-income countries.

[3] Another important area where questionnaire design can impact the measurement is in regard to labor income from household-owned enterprises. This is not the focus of our paper. De Mel, McKenzie, and Woodruff (2007) implemented a survey experiment on this topic in Sri Lanka.

force participation, biases the point estimates of the coefficients of right-hand-side variables (Hausman, Abrevaya, and Scott-Morton, 1998; Hausman, 2001).

Assessing the internal validity of survey measures, although common in psychological sciences, remains scarce in economics.[4] In this paper we focus on two key characteristics of surveys that potentially influence the labor statistics they generate: the structure of the labor-related questions in the questionnaire and the respondent selected for the interview. Our results show that there are significant differences observed with respect to measures of labor force participation, labor supply, earnings, and the distribution of activity across survey designs. Labor force participation rates, for example, vary by as much as 10 percentage points, across the four survey assignments. Using a labor module with no screening questions generates lower female labor force participation and lower rates of wage employment for both men and women. Response by proxy rather than self-report yields lower male labor force participation, lower female working hours, and lower employment in agriculture for men. Survey design matters to measuring labor and, moreover, valid comparisons across surveys of different design can be compromised by these differences.

The structure of the paper is as follows. In the next section, we discuss briefly the wealth of studies from the U.S. and other high-income countries and the scant evidence from developing countries. In Section 3, we describe the experimental design. The fourth section provides a description of the data collected, while Section 5

[4] In psychological science the sensitivity of results to measurement goes under the label of 'method bias'; see Podsakoff et al (2003) for an overview. With the increasing cross-fertilization of psychology and economics, the latter is increasingly paying attention to the internal validity of measures, especially in the field of experimental economics, where there is a strong interest in the effect of 'framing' or how a question / puzzle / game is presented to a subject (see Tversky and Kahneman, 1981 for early work in this area).

presents our results, exploring the effect of survey design on the statistics generated. The last section concludes.

2. Background and Literature

The specific wording and style of employment questions are posited to have a large influence on labor statistics. This may be particularly relevant in a setting where a significant proportion of individuals are employed in household-owned enterprises or home production and are not directly remunerated in the form of a salary or wage. For example, the standard question “Did you work in the last 7 days?” is hypothesized to systematically undercount persons who work in household enterprise activities without direct wage payments (e.g. unpaid family workers). Likewise, such employment questions may be flawed if applied to settings where employment is highly seasonal or where a significant proportion of workers are casual laborers. Employment in the last seven days or the last month may also yield different statistics than questions about labor force activities over the past year.

Here we discuss some of the literature related to two aspects of survey design which are the primary focus of our experimental design: the structure of questions and the survey respondent. A number of studies have focused on the style of different questions (open vs. closed questions; positive vs. negative statements; etc.) and the effects of their placement in the survey questionnaire, which may produce different responses (for a review see Kalton and Schuman, 1982). Mostly they have confirmed that question-wording effects are important, although the direction of these effects is often unpredictable. Studies have been carried out in the context of the revision of the

employment questions in the U.S. Current Population Survey (CPS) to investigate the concern that irregular, unpaid, and marginal activities may be underreported partly because people do not think of themselves as working. One of the few studies on this topic has been the Respondent Debriefing Study, carried out by the US Census Bureau in 1988 through the use of vignettes. Respondents were asked to classify hypothetical situations in terms of their own understanding of labor force concepts of "work", "job", "business", and so on. The results were quite revealing. Generally, the majority of respondents were able to classify the situations presented consistently with definitions of the CPS. However, for each vignette, large minorities of respondents gave incorrect answers – for example, 38 percent of the respondents included non-work activities under the "work" classification (Campanelli, Rothgeb, and Martin, 1989).[5] An experiment carried out in 1991 to assess the revision of the CPS questionnaire using vignettes and direct screening questions for unreported work confirmed that questionnaire wording and sequence of questions affect the respondent's interpretation of work and, therefore, the employment statistics (Martin and Polivka, 1995). The employment categories most at risk of misreporting were work in the family business or farm, casual employment, and work for commission or work compensated by other means than regular salary/wage payments. Moreover, the use of direct screening questions[6] was found particularly useful to detect

[5] For a discussion of several methodological tools, including hypothetical vignettes and direct screening questions, used to provide diagnostic information to evaluate the effect of questionnaire revisions on reporting of work activities see Esposito et al. (1991).

[6] In the 1991 CPS experiment direct screening questions were asked about individuals who were reported as non-working in the initial employment questions. With specific reference to the individual originally classified as non-working, the question was "In addition to people who have regular jobs, we are also interested in people who may work only a few hours per week. Last week did (NAME) do any work at all even for as little as one hour?"

underreporting of work done in connection with the household business or farm, as well as underreporting of teenage employment.[7]

The 1991 CPS study noted above also pointed to the existence of gender dimensions of these effects. In particular, the revision of the questionnaire, aimed at better capturing unpaid work in a household business or farm, increased the female employment rate. In developing countries, the gender effects may be even more dramatic than in developed countries. Several studies have argued that because female employment patterns tend to include multiple activities in comparison to men, the structure and wording of the questionnaire is likely to produce larger variations in female than in male employment statistics. In particular, concerns about the underreporting and undervaluing of women's work when using the most common methods of employment data collection have been expressed in several studies (Anker, 1983; Dixon-Mueller and Anker, 1988; Charmes, 1998; Mata Greenwood, 2000). In developing countries, women workers tend to have a prominent role in agriculture (especially in subsistence agriculture), often as contributing family members, and in the so-called invisible household economy – that is, in fetching water and wood, carrying out domestic tasks, while providing care to children, the elderly, and persons with disabilities. They generally work on simultaneous (and often seasonal) activities, paid and unpaid, within and outside the System of National Accounts (SNA). They tend to engage in informal sector activities and, because of

[7] Other indirect evidence on the impact of questions is available from comparing the "main activity" question in a household roster reported by the head and the more detailed set of questions in a labor module of the same questionnaire. Of the more than 60 Living Standard Measurement Surveys done by the World Bank, only two (Malawi and Timor Leste) ask both the "main activity" in the household roster *and* collect activity-specific information from each member of the household (such as number of hours in a specific activity in the last 7 days). In Malawi, for women, 33 percent are reported inactive by the "main activity" question, compared to 25 percent from the labor module. This difference can be driven by some combination of the shorter question (and, therefore, different interpretation of work across survey sections) and the proxy response.

assigned cultural roles, may be considered by others and themselves as inactive even when they perform economic activities. In this context, it may be particularly difficult to capture women's work (Mata Greenwood, 2000).

In addition to key features of questionnaire design, different surveys adopt different approaches to designating the respondent to the questionnaire. For example, standard surveys in developing countries, like Household Budget Surveys (HBS), Income/Consumption Expenditure Surveys (ICES) and Core Welfare Indicator Questionnaires (CWIQ) typically ask the household head employment questions about all household members. An alternative approach is to ask each household member above a certain age directly. This is done in the Living Standards Measurement Study surveys (LSMS) (Glewwe and Grosh, 2000) and it is also the approach adopted by Labor Force Surveys (LFS). While responses by proxy are accepted for household members who are not available for interview, LFS guidelines warn that proxy respondents may not always provide accurate information and this can cause biases in estimation of employment and jobs (Hussmanns, Mehran, and Verma, 1990). The premise is that self-reporting provides more accurate information than the use of proxy respondents. At the same time, however, requiring all individuals to self-report makes the fieldwork quite burdensome and expensive, creating a trade-off between accuracy of the information and the cost to obtain it. It must be said, however, that although intuitively self-reporting should be superior to proxy response in generating more reliable information, there are few studies that provide evidence on this.[8] Some experimental evidence exists on the reporting of

[8] Experimental studies are especially useful in assessing the 'true effect' of using proxy- vs. self-respondents. Non-experimental studies tend to suffer from the problem of self-selection (Hill, 1987; Moore, 1988) – that is, proxy respondents may be individuals who happen to be at home. These proxy

health events by self- and proxy respondents. In their U.S. study, Mathiowetz and Groves (1985) found that a randomly selected person reports fewer health events for him/herself compared to what he/she reports for other household members, but they are not able to offer a clear explanation of why this is the case.

Most survey experiments that study the effects of using proxy respondents in lieu of self-respondents on employment statistics are from developed countries.[9] Martin and Butcher (1982) compared self-responses with the proxy responses of 1,324 paired adults in the same household from the U.S., 70 percent of which were husband-wife pairs. When comparing the answers of husband and wife, the study found that employment variables (class of worker, occupation, hours worked, etc.) had less than a 10 percent disagreement rate, while higher level of disagreement was obtained for income variables. Moreover, approximately 20 percent of the proxies did not know the income of their spouse. In a similar U.S. survey, Boehm (1989) interviewed pairs of individuals in households to compare proxy and self-reports. Both individuals were interviewed separately about their own status and the other member's status. The study found that self and proxy responses resulted in the same labor force classification 83 percent of the time. However, this study was based on a small sample of 84 individuals from a group of participant volunteers, 42 pairs from 42 households. In general, the little experimental evidence and the non-experimental studies indicate that self-respondents produce higher household and person non-interview rates, but proxies produce higher item non-response rate, especially for wages and income

respondents will typically have different characteristics than those who are absent from the household and those characteristics are generally correlated with the type of information that it is collected.

[9] Bound, Brown and Mathiowetz (2001) examine the implications of measurement error in survey data and provide a detailed review of many studies of labor statistics, mostly U.S., and measurement error in reports of earnings, hours, and unemployment. The validation sources for these studies are usually employer records and administrative records (for taxes and transfer program income), but sometimes it is re-interview and time dairies. Tables 1, 2, 4 and 5 in their study summarize these studies.

variables (Biggs, 1992). The use of proxy respondents may amplify recall errors or affect the reporting of hours of work, especially in the case of irregular or multiple activities (Hussmanns, Mehran, and Verma, 1990). Moreover, the use of proxy respondents is also considered to potentially be a source of sex biases in the form of underestimation of women's participation in economic activity (ILO, 1982).

The survey experiment we conducted and analyze here seeks to inform the method by which labor statistics are collected in household surveys in low-income countries, and, therefore, the information base for analytical work on employment. We hope that this work leads to improvements in the measurement of, among others, labor force participation, the nature of work in terms of type and intensity (particularly work that occurs in household enterprises and farms), the changing patterns in employment over time, and nuanced changes in labor market activity that could be missed when using existing data collection instruments.

3. The Survey Experiment

Whether changes in measurement have an effect on the statistics they produce is, ultimately, an empirical question. We designed and implemented a survey experiment focusing on two key dimensions of labor survey design: the level of detail of the questionnaire (specifically the screening questions to establish employment status) and the type of respondent (Table 1). Households were randomly selected for the survey, and, after being selected, randomly assigned to one of the four survey assignments based on these two dimensions.[10] Eligible persons to respond to the

[10] An alternative approach to randomly assigning households to different surveys to study mis-measurement is to test existing data for abnormalities based on prior distributional assumptions about

individual-level module on employment (or have labor statistics collected for) were all those aged 10 and older.

For the first dimension, we developed a detailed labor module and a short labor module. The short labor module reflects the approach in shorter questionnaires, such as the Core Welfare Indicator Questionnaire (CWIQ). The CWIQ approach was developed by the World Bank as an off-the-shelf survey that country statistical agencies could implement on a large sample in a short time frame. Many countries regularly field CWIQ-type surveys (in many cases they are called by another name, such as Welfare Monitoring Survey), especially with increasing demands to produce sub-regional household survey statistics. This shorter module is often used in generating statistics with a higher frequency, for example with annual regularity, in lieu of multi-topic household surveys which are too demanding to implement on an annual basis. The detailed labor module reflects the approach in longer questionnaires typically used in multipurpose household surveys, that are often modeled on the Living Standards Measurement Study (LSMS) surveys (Glewwe and Grosh, 2000) in developing countries.

In our survey experiment, the detailed module differed from the short module in two ways: in the set of screening questions to determine employment status and in asking about second and third jobs. Here we focus on the former issue. The detailed module starts with three questions to determine employment status: specifically, whether the person has worked for someone outside the household (as an employee), whether s/he has worked on the household farm, and whether s/he has worked in a non-farm

the data – see, for example, Judge and Schechter (2009) who study the first digits of numbers, or Benford's law.

household enterprise (for these three questions the response is yes or no). These questions were asked with respect to the last 7 days (the reference period for identifying those who are "employed" and the set of detailed questions on that employment) and, if reported to not work in the last 7 days, then asked for the last 12 months. In the shorter module, there was only one question to determine employment status with respect to the last 7 days: whether s/he did any type of work (response also yes or no). As in the detailed module, the question was asked twice – for the last 7 days and the last 12 months. The short and detailed employment modules are reported in Appendix 1.

In the second dimension of the experiment, we varied whether questions are asked directly to the subject or asked to a proxy respondent. Response by proxy rather than individuals themselves reflects the common practice to interview an informed household member (often the household head), rather than each individual him or herself. In practice proxy respondents are often used when individuals are away from the household or otherwise unavailable to interview in the time allotted in an enumeration area to conduct interviews. In the survey experiment, the proxy respondent was randomly chosen among household members at least 15 years old.[11] The person selected to be the proxy respondent then reports on up to two other randomly selected household members age 10+. Of course, not all persons selected to be a proxy respondent could actually be interviewed. Also discussed in the next section, in our experiment we find that the proxy respondents selected did not significantly differ from the pool of potential proxy respondents with respect to gender or education. The proxy respondent could be any adult: the head of household,

[11] The Tanzanian CWIQ 2006 data indicate that the average Tanzanian household has between two to three adults who could serve as a proxy with a minimum age of 15. This informed the design of our survey, and, in fact, our sample households had about 2.7 members 15 years and older.

spouse of the head, an older child, or older relative living in the household. In actual implementation of surveys, proxy respondents are not randomly chosen, but are normally selected by interviewers on the basis of availability. In this sense, the experiment did not exactly mimic the actual conditions that result in proxy responses in household surveys. Because we have information about the relationship between the proxy respondent and the individual on whom the proxy responds, we can assess whether there are systematic response patterns that depend on who the proxy is. However, because a typical survey does not generally identify the proxy in the relation to the person for whom the information is collected, we cannot determine what our results imply in terms of potential 'bias' of a typical survey due to the use of proxy respondents.[12]

The four different survey assignments broadly reflect commonly used approaches in practice. The benchmark reference to which the other survey assignments are compared is the detailed self-report questionnaire. This is generally considered to be the "best practice" approach of household surveys. The use of multiple questions is recommended by ILO especially when some categories of workers (especially casual workers, unpaid family workers, apprentices, women engaged in non-market production, workers remunerated in-kind, etc.) may not be able to correctly interpret a question about "any type of work" as referring to their situation (Hussmanns, Mehran, and Verma 1990).

[12] An alternative research design to assess the effect of proxy respondents would have been to interview two members of the household who report on their own labor activities and proxy report on the other. We did not implement such a design because it proved to be too difficult to ensure a proper implementation for a medium to large sample. After consultation with counterparts in Tanzania, we concluded that it would be difficult to assure that proxy and self responses would be independent and would remain unaffected by the knowledge that another household member reports on the same information, given the normally social nature of an interview. The specific concern was that the design (and open communication about this design within the village) would trigger either a coordinated response by household pairs and/or accommodation of response to other's expectations, which would introduce potentially much larger (unobserved) respondent biases.

For those identified as working in the last 7 days, either through the set of three questions (in the detailed module) or through the single question (in the short module), information on the occupation, sector, employer, hours, and wage payments was collected for the main job. These questions are identical across assignments. In the detailed module, this same set of questions was repeated for the second and third job.

For all of the survey assignments, in addition to the labor module, the questionnaire also included four other modules: household roster, assets, dwelling characteristics, land, and consumption expenditures. In the detailed and short questionnaire, the questions followed the same sequence; identical types of questions follow the same phrasing and recall periods are the same. This was to ensure that variations across experiments are not driven by these other survey elements.

From an analytical perspective, we are interested in two general issues. First, we want to assess the effects of the change in survey assignment (presence of screening questions and type of respondent). Second, we want to assess whether these effects vary by the characteristics of the respondents, the person on whom the respondent is reporting (if proxy), and household traits.[13] We estimate the average survey assignment effect for each assignment type, taking the detailed survey with self respondents as the reference group, as follows:

[13] To estimate the average effect of survey assignment (the "treatment effect"), we ideally want to estimate $\Delta = Y_t^1 - Y_t^0$ which is the difference of the outcome variable of interest at time t between two treatments denoted by the superscripts 1 and 0. However, since Δ is unobservable to the researcher because a household does not receive two survey assignments simultaneously, one estimates the survey assignment effect given the observable data, i.e. $TE = E(Y_t^1 \mid T=1) - E(Y_t^0 \mid T=0)$. Since in a properly implemented randomized design the different groups by survey assignment have identical characteristics because the survey assignment was randomly allocated, $E(Y_t^0 \mid T=1) - E(Y_t^0 \mid T=0)$ equals zero and the estimate of the survey assignment effect is unbiased.

$$y_i = \alpha_j + \beta_j T_i^j + \gamma X_i + \lambda V + \varepsilon_{i,j} \quad (1)$$

where y_i are the different labor statistics like labor force participation, labor supply, earnings, and occupational choice for the i^{th} individual, T_i^j is an indicator variable for the j^{th} respective survey assignment for the i^{th} individual, X_i is a vector of individual and household characteristics for i^{th} individual, V is a village fixed effect, and $\varepsilon_{i,j}$ is the stochastic error term which is randomly distributed.

Because questions on hours, earnings, and sector are identical across assignments, variations in statistics across survey assignments are not due to question wording. However, the response to labor force participation determines whether statistics on those other dimensions are collected at all for the individual (in other words, these statistics are conditional on the individual being classified as employed). In the case of self-respondents, the screening questions which differentiate the start of the short and detailed modules entirely explain variations in selection into employment and therefore variations in hours, earnings, and sector statistics. In the case of proxy respondents, variations in statistics for these other outcomes derive from both the quality of the reporting by proxy respondent on a specific variable (how well does the wife know how many hours her husband works?) and the accuracy of the reporting on the employment status (if the husband does not report that his wife works, then he will not be asked her hours). Only the latter is an issue of selection.

4. Data and Context

The survey experiment was implemented in Tanzania, which has different types of labor market surveys, including CWIQs, LFSs and multipurpose household surveys, like the Household Budget Survey (HBS). Figures reported for 2000/1 by the National Bureau of Statistics illustrate the relevance of our research. Whereas the Integrated Labour Force Survey (ILFS) reports labor force participation rates of 90.6 percent for men and 89.5 percent for women (NBS 2003), the Household Budget Survey (HBS) reports 91.1 percent and 82.4 percent respectively for the same year (NBS 2002).[14] The large difference in labor force participation of women between two nation-wide surveys that refer to the same year may reflect genuine differences (attributable to differences in samples or in the timing of the survey implementation), or may reflect differences caused by the use of distinct survey instruments. Our experimental set up intends to shed some light on the role that questionnaire design and respondent type may have on such empirical regularities.

The survey experiment conducted was the Survey of Household Welfare and Labour in Tanzania (SHWALITA), and was implemented by a well-established data collection enterprise, Economic Development Initiatives (EDI). This survey was purposively designed and fielded to study the implications of the alternative survey designs for consumption expenditure measures and labor market indicators. Here we focus on the component that applies to labor market indicators.

[14] The former is obtained from the tables in the appendix of the ILFS 2000/1 report, available at http://www.nbs.go.tz/labourforce/index.htm, while the latter is reported in the 2000/1 HBS report available at http://www.tanzania.go.tz/hbs/Index_FinalReport_HBS.htm

The survey assignments were carefully piloted. A qualitative debriefing with the field supervisors took place at the end of each day during the pilot, in order to solicit their feedback on a range of issues.[15] In addition, a subset of households was selected for qualitative interviews with respondents, in order to see whether wording and structure of the questionnaire could be further improved. During this qualitative interview, respondents were asked open-ended questions to solicit how they thought about the questions, why they chose the responses they did, and how they thought about concepts such as work, household production, and their primary activities.

The field work was conducted from September 2007 to August 2008 in villages and urban areas from 7 districts across Tanzania: one district in the regions of Dodoma, Pwani, Dar es Salaam, Manyara, and Shinyanga region and two districts in the Kagera region. Households were randomly drawn from the listing of villages (urban clusters) and randomly allocated to one of the four survey assignments. The total sample is 1,344 households, with 336 households assigned to each of the four survey assignments. Although the sample of 1,344 is not designed to be nationally representative of Tanzania, the districts were selected to capture variations between urban and rural areas as well as along other socio-economic dimensions to inform survey design related to labor statistics and consumption expenditure for low-income settings.

[15] The feed back focused on nine areas: 1. General impressions of the respondent's comprehension; 2. Question phrasing; 3. Question sequencing; 4. Completeness of lists of question responses; 5. Clarity of interviewer instructions; 6. Completeness of interviewer manual to resolve field problems encountered; 7. Questions that should be restructured for greater clarity and respondent comprehension; 8. Conceptual or cultural difficulties in translating questions to local language; 9. Areas of emphasis for training enumerators. One of the most important parts of the questionnaire to pilot was the selection of proxy and self-reporting respondents. After a day of training, interviewers spent significant time practicing with examples. They appeared to have no trouble in the field selecting proxies or self-reported respondents following the protocols designed for the experiment.

The basic characteristics of the sampled households generally match the nationally representative data from the Household Budget Survey (2006/07) (results not presented here). Household interviews were conducted over a 12-month period, but because of small samples, we do not explore the survey assignment effects across seasons (such as harvest time with peak labor demand and dry seasons with low demand). The random assignment of households is validated when examining a set of household characteristics (Table 2). With the exception of household size and acres owned (driven by a handful of very large households with many children in one group), household traits are not statistically different across survey assignment. Of a total of 96 pairwise comparisons, eight are statistically significant.

Turning to individuals, we classify people on the basis of the survey assignment that they actually received in order to measure their labor statistics. An individual's actual survey assignment is the result of the initial assignment of their household among one of the four survey assignments, whether the individual is selected to be a respondent or a self-report, and whether the self-report/proxy assignment is realized. In the case of self-report modules, up to two persons over age 10 are randomly selected to self-report. If persons randomly selected to self-report are unavailable, an alternative person is selected at random. In the case of proxy assignment, one person in the household over the age of 15 is selected to self-report and to proxy report on up to two random household members. Thus, in the proxy assignment, one household member actually self-reports in addition to reporting on other household members. Therefore, the number of self-reports should be about half the number of proxy reports for households in the proxy assignment. In total there will be more self-reports than proxy reports. Appendix 2 shows the sample sizes by planned assignment and by

actual (realized) assignment. The latter reflects both the availability of someone for a self-report and the number of age-eligible household members. Because the survey experiment highly emphasized the importance of avoiding proxies, the project was fairly successful at completing self-reports when assigned. In about five percent of cases, the team was unable to interview a person selected for self-report (32 out of 638 and 35 out of 636 in Appendix 2). About 10 percent of households have only one adult 10 years and older which means that these households, if assigned to the proxy survey experiment, will actually have no proxy report possible. In that case, the household's one member over age 10 becomes a self-report and there will be no proxy response. Between 12-14 percent of households have one adult 15 and older, which means there can be only self-reports. The results presented in this paper are virtually the same if we exclude the observations where we had to deviate from the planned design.

Comparing across survey assignment of individual respondents, we again observe that household characteristics are very similar across assignments (Table 3). This is consistent with what was observed in Table 2. For individual traits we see some differences between the self-reports and those whose information is reported by proxy. Of the 144 pairwise comparisons, 47 are statistically different. Some of these differences have been introduced by the experiment design which we control for them in the multivariate analysis. For example, the self-reporters are older and more likely to be married, which is consistent with self-reporters including persons selected to be a proxy respondent for other members which requires being 15 or older. In Appendix 3 we report these characteristics for the sample 15 and older and find that age and marital patterns are quite similar across assignments.

5. Results

The presentation of the results of the experiment is divided into two parts. In the first part, we examine differences across the survey assignments for key labor statistics on the individual's main activity: labor force participation, weekly hours, daily earnings, the sector of work, and type of work (employment status). We compare the statistics produced for three comparisons: 1) the short module compared to the detailed module, 2) responses given by proxy compared to self-reported responses, and 3) the short module and proxy respondent interaction compared to the other three groups. The latter is of interest because the combined short-proxy assignment reflects the least expensive approach which is frequently used. We then estimate the survey assignment effects using standard models (OLS, probit, multinomial logit) controlling for observed household and individual characteristics. In principle, we expect the multivariate results to be similar to the first set of results, due to the randomized experimental design, and this is indeed what we find.

In the second part, we analyze whether the characteristics of proxy respondents and/or the difference in characteristics between the proxy respondent and the individual on whom the proxy reports affect the labor statistics. Again, we first examine the mean statistics, and then run multivariate regressions where we control for individual characteristics of the proxy and the interaction between individual characteristics of the proxy and the subject.

Differences in Labor Statistics across Survey Assignment

Table 4 presents the findings, disaggregated by gender, for labor force participation (LFP), weekly hours, and daily earnings. In each case we test for a difference in means across survey assignments using a t-test. Row 1 of Table 4, for instance, reports the LFP rate of men from the short module (82 percent) and from the detailed module (83 percent), and finds that the difference (about 1 percentage point) is not statistically different from zero. LFP is defined following ILO guidelines – that is, having engaged in any work for pay (as wage/salaried worker), profit (as employer, self-employed, own-account worker), or family gain (as paid or unpaid worker in a family farm or family business) in the last 7 days. Domestic duties (unpaid domestic activities) are not counted as labor force participation because they are not included in the SNAs.[16]

We find that there is no difference in LFP rates between the short and detailed module for men, but there is a significant difference for women. LFP is about 7 percentage points lower when reported by the short module. The difference between the proxy and self-reported statistic is significant for both men and women: LFP is lower by 13 and 8 percentage points respectively for the proxy surveys. The short-proxy questionnaire combination yields significantly lower LFP rates than the other designs for both sexes, around 10 percentage points in both cases.

[16] While unpaid domestic activities such as housework, child care giving, and house repair work are outside the boundaries of SNA, other activities such as fetching water and fuel are technically considered SNA activities but are routinely excluded from employment statistics because of a lack of data due to difficulties in collecting data on and valuing these activities.

Reported weekly hours of those who are working are only slightly different between the short and detailed module for men, but hours are lower when reported by a proxy compared to self-reported, and this difference is larger for men than for women (7 fewer hours for men and 4 fewer hours for women). The short-proxy module also generates lower working hours than the other modules, and again the difference is larger in the case of men (almost 5 fewer hours) compared to women (4 fewer hours).

Daily earnings also differ across survey assignments, but following a different pattern. The detailed module produces higher average earnings for women but not men, but differences are not significant between the proxy and self-reported modules.

In Table 5, we turn to the distribution of main activities. Here we classify all subjects into four categories. Labor force participants are in agriculture or other sectors.[17] Those who declared having worked in the past 7 days can report that their main activity in the last week was domestic duties or no work. Participation in domestic duties, while not formally included in a labor force statistic, is commonly collected in surveys. This is usually done by including domestic duties as a possible answer to the question in what sector the main activity is; this approach is also followed in both the short and detailed modules. However, in the detailed module, like in most multi-purpose modules, this question is preceded by the three screening questions on work in the last 7 days, which aim to find out the type of work of the respondent in more detail.

[17] The non-agricultural sectors are too small to leave disaggregated. These include: mining/quarrying/manufacturing/processing, gas/water/electricity, construction, transport, buying and selling, personal services, education/health, and public administration. Buying and selling activities are the most frequently reported of these activities (4-7 percent, depending on the sub-group).

In Table 5 panel A, the previously observed association between lower female LFP and the short module (see Table 4) is shown to result primarily from an increase in domestic duties as women's main activity. In the case of the detailed module, women are more likely to result as "not working" than carrying out domestic duties; the decline in domestic duties is 16 percentage points from short to detailed module. For men, while there is no difference in LFP, we do find a shift from domestic duties to "no work" when moving from the short to the detailed module. This suggests that the additional questions at the start of the labor module act as screening questions, filtering out at least part of the individuals that get classified under domestic duties in the short questionnaire. So, the three work questions in the detailed module (any farm work, any wage work, and any work on a non-farm enterprise) frame the notion of work to exclude domestic duties in the minds of respondents. The results also indicate that the sectoral decomposition of the labor force (by agriculture and other sectors) is not significantly affected by the short or detailed modules for either men or women.

Quite interestingly, if one only relied on the direct question about work (such as "Did you do any type of work in the last seven days?"), the short module generates an employment rate that is actually *higher* than the employment rate derived from the combination of the three screening questions in the detailed module (work on farm, work for household business, work for wages). It is only after the re-classification of those who declared that their main activity was 'domestic duties' that the employment rate from the shorter questionnaire becomes lower. This is seen in the comparison the rates of 'no work' from the short and detailed module in Table 5 Panel A). As we will stress in the conclusions, this result is in contrast with what we expected *a priori* – that a generic or vague question on work would miss people in activities with no

remuneration. But following the protocols commonly used by analysts, where employment is having worked in the last seven days excluding domestic work, we find lower LFP in the shorter questionnaire.

Comparing proxy and self-reports, we find that the sector composition (agriculture and other sectors) is not much affected (Table 5 Panel B). Proxy respondents report lower labor force participation as a result of both a higher frequency of main activity as 'domestic duties' as well as higher percentage of 'no work'. The percentage point declines are larger for agriculture when moving from self-report to proxy. Sixty two percent of men are identified as working in agriculture by self-report compared to 54 percent by proxy (almost 8 percentage point decline), where the participation in other sectors is 26 percent and 21 percent, self-report and proxy report respectively.

Finally, in panel C of Table 5 we compare the short proxy module to the other survey assignments. The combined effect of a short questionnaire and proxy respondents results in lower participation in both agriculture and other sectors for both men and women. The decline is not neutral – it is proportionally larger in 'other sectors' with respect to 'agriculture', so that the short-proxy module produces a higher percentage of participation in agriculture among workers than in 'other sectors' (although these variations are not statistically significant for men).

In Table 6, we explore the distribution of the labor force by employment status. We define four mutually exclusive groups: paid employees (i.e. working for wages or salary), self-employed with employees (hired labor in the household enterprise), self-employed with no employees (often farmers), and unpaid family worker (those

working on the household farm or in the non-farm enterprise but who do not identify themselves as being the main person running/owning the activity). When comparing the different survey assignments, one should keep in mind that the differences here will be driven in part by the inclusion/exclusion of some individuals as being in the labor force (the LFP impacts we observe in Table 4).

Comparing the short and detailed modules, we find that there are significant variations in most categories. Paid employees represent between 14 and 20 percent of working men and between 6 percent and 11 percent of working women, depending on survey assignment. Self-employment without employees is higher for men when using the short module (57 percent) compared to the detailed module (52 percent). By respondent type, while the absolute decline in being self-employed with paid employees for men is small when moving from proxy to self-report (-2.5 percentage points), the relative change is large because of the small base (-50 percent). Across survey assignments, the percentage of unpaid family workers varies dramatically for women and men. When women are directly asked about their type of employment, the percentage of unpaid family workers drops by about 13 percentage points from 73 (when proxy are used) to 60 percent. The impact is also very large for men (33 percent with proxy vs. 21 percent for self-report). Similarly, the use of the short rather than the detailed questionnaire produces a higher percentage of unpaid family workers among women, 67 compared to 62 percent. In this case, the effect is smaller for men and not statistically significant.

A summary of the results of the descriptive analysis is found in Table 7. Deviations from the detailed and self-reported survey design yield lower estimates of LFP and, in

the case of the proxy, lower working hours among those in the labor market. We find higher rates of domestic work as main activity when the short module or proxy module is used. This lowers participation in agriculture, other sectors, as well as the category of 'no work'. Among those in the labor force, the employment status shifts, with fewer paid employees and increases in the share of unpaid family workers.

To extend our analysis, we estimate equation 1 which includes controls for characteristics of the subjects (age, gender, education), household characteristics (household composition, asset holdings) and village-level fixed effects. We identify the effect of survey assignment with separate dummy variables for the short module, the proxy module, and the combined short-proxy module. This approach allows us to identify the separate impact of each survey assignment. The results for LFP, obtained using a probit model, are reported in Columns 1 and 2 of Table 8 for men and women respectively. We find that, in the case of men, it is the use of proxy respondents that drives the lower reporting of LFP – about 12 percentage points – relative to self-reporting, the omitted category. By contrast, in the case of women, the short questionnaire lowers LFP by about 5 percentage points but the effect is not statistically significant. The combination of a short questionnaire and proxy respondents does not appear to add further effects for either men or women. Column 3 and 4 in Table 8 report the results for (log) weekly hours of work using OLS. The short questionnaire is associated with higher hours, possibly reflecting the selection into the hours question – that the shorter questionnaire results in jobs of fewer hours being under-reported. The use of proxy respondents produces lower average working hours for both men (-9 percent) and women (-10 percent) with respect to what is reported by self-respondents – although this difference is only statistically significant

for women. Finally, with respect to daily earnings, the combination short questionnaire and proxy respondents generates particularly low earnings for women, while earnings for men are not affected in any significant way. However, the sample size (those who work for a wage or salary) is small.

Using a multinomial logit, we estimate the survey assignment effects on the allocation across three employment categories: agriculture, other sectors, and the omitted category out of the labor force (combining domestic duties and no work). The results in Table 9 show that proxy respondents report a lower participation of men in both agriculture and other sectors with respect to out of the labor force (no work or domestic duties). For women, the short module generates lower levels of employment in both agriculture and other sectors and a higher rate of out of the labor force, but these results are not statistically significant. Proxy reporting for women lowers women's employment in other sectors but not agriculture (relative to domestic duties and no work).

Following a similar estimation approach, in Table 10 we estimate the survey assignment impact on employment status for those in employment. Controlling for proxy assignment, the short module decreases the probability of men being wage employed compared to unpaid family labor; it also decreases women's probabilities to be in wage employment or self-employed without employees with respect to in unpaid employment. Relying on proxy respondents does not have a statistically significant impact on status in main job for either men or women. The combined use of proxy respondents and short questionnaire produces no additional impact.

Effect of proxy respondent characteristics and of differences in characteristics between proxy and the subject (on whom the proxy reports)

Focusing on the proxy assignment, we investigate whether the responses are affected by the characteristics of the respondent or by the difference in characteristics between the respondent (the proxy) and the person about whom the response is reported (the subject). We focus on differences in age, gender, and schooling. In Table 11a, we examine labor statistics across the quartiles of proxy-subject age differences (proxy age minus subject age). Quartile one (Q1) reflects the smallest age difference in proxy age minus subject age, where the proxy is actually younger than the subject. In quartile four (Q4), there is the largest age difference with the proxy being much older than subject. We compare the proxy-subject age difference for each quartile to the mean age difference, disaggregating the sample by subjects that are less than or equal to the median age and those who are older than the median age. This disaggregation of the sample by the median subject's age controls for potential life cycle effects of the subject's employment, which might otherwise be partially captured in age differences between the proxy and subject. For younger subjects (less than or equal to the median age), reports of LFP, weekly hours worked, and earnings all decrease significantly as the age of the proxy increases relative to that of the subject. For instance, there is a 16 percentage point decline in LFP between the first quartile (proxy much younger than the subject) and fourth quartile (proxy much older than subject). The difference in reported weekly hours decreases by 19 hours between quartile one and four. For subjects older than the median age, we find less variation across the proxy-subject age gap, and less of a clear pattern. The age difference, either

a proxy much older or much younger, matters most for younger workers' labor statistics.

Labor statistics may also be affected by the differences in gender between proxies and subjects. We test whether these differences are significant by comparing male and female reports on the labor activities of women in the household (see Table 11b, columns M-F and column F-F), as well as female and male reporting on men in the household (F-M and M-M). We find that male proxies for women and female proxies for men both report greater LFP and weekly hours worked than same-gendered proxy respondents. A male proxy for a female within the household yield a LFP rate 12 percentage points higher and reports 5 more weekly hours worked than a female proxy. For a female proxy reporting on a male subject, the female proxy also reported a 13 percentage point difference in the LFP rate than a male proxy and 8 hours more worked per week. Given that proxy reporting is associated with lower LFP relative to self-reporting (as shown in Table 4), this implies that same gender proxies are further from self-reported LFP (for example, self-reported male LFP is 87 percent, while M-M proxy report is 65 percent and F-M is 78 percent, which is closer to the self-reported rate, a similar pattern holds for women). If we assume that reports from self-respondents are closer to the "truth", this suggests that opposite gender proxies do better. This is likely to be driven by the importance of the spousal relationship, since in 50 percent of the cases when proxy and respondent are from a different gender, they are spouses.[18] The result suggests that spouses may know better than others in the household about the employment status of their partner, although the differences

[18] Of the proxy responses in the sample, 33 percent of responses are from male proxies reporting for female subjects, 34 percent of responses are from female proxies reporting on male subjects, 20 percent of responses are female proxies reporting on other females within the household, and 13 percent of responses are men reporting on other men in the household.

are still large with respect to self-report (for example, about 10 percentage point difference in male participation in employment between what reported by themselves and what reported by a female proxy – 87 vs. 78 percent).

To consider the effect of differences in schooling, we compare the reports by proxies and subjects of the education gap classified as having any schooling (S) or no schooling (N). The results are reported in Table 11c, where S-N reflects a schooled proxy reporting on non-schooled subject, and so on. For LFP, non-schooled proxy respondents report higher LFP than schooled counterparts, but this difference is statistically significant only for LFP of schooled subjects (although the gap is also there for non-schooled subjects). We do not find differences in hours reporting. Daily earnings are reported as higher by schooled proxies for schooled subjects by 434 TZ shillings than the non-schooled proxy reporting on a schooled subject. We see no statistically significant differences in labor statistics for non-school subjects by the schooling of the proxy respondent. Schooling differences between the proxy and subject matter most for schooled subjects.

6. Conclusions

Despite the importance of household survey instruments as a source for labor statistics, there is a dearth of evidence on the best practices for collecting these statistics in developing countries. The differences in survey design for national labor statistics over time within a country and across countries has serious implications for both measuring labor markets and for research on labor activities. This paper presents a survey experiment focusing on two key aspects of survey design to estimate their

effects on the labor statistics that they generate: the set questions to identify labor force participation and the choice of respondent. With four alternative survey designs, we compare the three assignments with the 'best practice' approach of a detailed questionnaire with self reporting.

Our findings suggest that both types of survey design decisions have statistically significant effects on labor statistics. These effects are largest on the measure of labor force participation, but also exist for weekly hours of work, daily earnings, main activity, and type of work. The effects are distinct for different statistics and remain significant even after controlling for individual, household, and village characteristics. Using the short questionnaire lowers female labor force participation and also affects the distribution of workers across sectors, lowering the share of paid employees among the employed. Response by proxy leads to lower reports of labor force participation and weekly hours for both men and women, and higher reports of female earnings among those employed; it also results in a higher share of unpaid family workers among the employed. Combining a short questionnaire with response by proxy, the least expensive approach to implement, results in lower reported female earnings. The differences between proxy and self reporting seem to come from information imperfections within the household, with the distance in age between respondent and subject playing an important role, while gender and educational differences seem less important.

These results provide clear evidence that survey design does matter for measuring labor outcomes. By extension, it also matters for the relationships between labor and other variables, since the types of biases introduced may affect the point estimates of

these relationships, as implied by the literature on non-random measurement error. A study of these impacts is a topic we plan to explore in future work.

Of course, in this paper we have focused only on two dimensions of survey design; future work is needed to look at other issues, like, for instance, the wording of questions. But even our (limited) results provide some clear advice for survey design. First, the impacts are not consistently associated with one specific design but differ for different types of individuals. Using a short module rather than a detailed module produces lower LFP for women but not for men; using proxy respondents rather than asking the individual directly strongly impacts (negatively) the LFP of men but not for women. Similarly, the impacts are different for hours or earnings than for LFP. This indicates that the 'best' approach – if it exists – may differ depending on the purpose of the survey (i.e. the type of variables and the type of sample for which information is collected). Second, using a category 'domestic duties' as a possible answer to the 'main sector of activity' can be problematic and may produce ambiguous results, especially in a short employment module. Our experiment, which incorporates the category 'domestic duties' in a very similar way many other surveys do, shows that individuals may classify themselves (or be classified) as 'employed' when responding to a very direct yes/no question about working, but then *de facto* rule themselves out of employment by indicating that their main activity was domestic duties. When using a short employment module an ambiguity remains as to whether these individuals are not fully able to interpret the meaning of 'employment' and/or 'domestic duties' or if they tend to prefer 'domestic duties' over other sectors of employment, even if they actually worked somewhere else as well, because this is where they spent the longest time. With no further information about second job, this

ambiguity is left unresolved. Third, our results underline the importance of staying with the same design if the aim is to make comparison over time. Whatever the preferred design is, changes in design may result in changes in employment statistics even when no actual changes occurred. Finally our findings provide a base for a cost-benefit analysis.[19] If the detailed self-reported module is best practice – as recommended by ILO guidelines and accepted in this paper by adopting it as a 'benchmark' – then we can assess what implications the use of less expensive approaches, like response by proxy or using a shorter module, has. The results suggest that for some indicators it may be acceptable to take short cuts, while for others the price paid in measurement error is probably quite high.

[19] The cost implications of the design are mainly in terms of length of field work (and only very marginally the length of printing questionnaires and the data entry) through more days required in each sample village. Unfortunately we could not measure this directly in this experiment because it was part of a larger survey experiment with a consumption diary component, and we cannot isolate the costs related to the labor module. However, a back-of-the envelope calculation teaches us that to complete self-reports (and track down each adult), a team that otherwise would finish a survey in 2 days in the village will need 1 or 2 more days, implying 33-66% extension of the field work and corresponding increase in the budget.

References

Abowd, J. M. and A. Zellner (1985) "Estimating Gross Labor-Force Flows." *Journal of Business & Economic Statistics* 3(3):254-283.

Anker, R. (1983) "Female Labour Force Participation in Developing Countries: A Critique of Current Definitions and Data Collection Methods." *International Labour Review* 122(6):709-724.

Boehm, L. M. (1989) "Reliability of Proxy Response in the Current Population Survey." in *Proceedings of the Survey Research Methods Section, American Statistical Association*.

Biggs, B. (1992) "Self/Proxy Respondent Rules and Data Quality." Research Paper, Income Research Paper Series, Ottawa: Statistics Canada.

Bound, J., C. Brown and N. Mathiowetz (2001) "Measurement Error in Survey Data." in *Handbook of Econometrics* Vol. 5. ed. J. Heckman and E. Leamer. Amsterdam: North-Holland, Elsevier Science.

Campanelli, P., J. M. Rothgeb, and E. A. Martin (1989) *The Role of Respondent Comprehension and Interviewer Knowledge in CPS Labor Force Classification.* American Statistical Association Proceedings (Survey Research Methods Section).

Charmes, J. (1998) *Women Working in the Informal Sector in Africa: New Methods and New Data.* Paris: Scientific Research Institute for Development and Co-operation.

de Mel, S., D. McKenzie, and C. Woodruff (2007) "Measuring Microenterprise Profits: Don't Ask How the Sausage is Made." *Journal of Development Economics* 88(1):19-31.

Dixon-Mueller, R., and R. Anker (1988) *Assessing Women's Economic Contributions to Development.* Training in Population, Human Resources and Development Planning Paper number 6, Geneva: International Labour Office.

Esposito, J. L., P. C. Campanelli, J. Rothgeb, and A. E. Polivka (1991) "Determining which Questions Are Best: Methodologies for Evaluating Survey Questions." *Proceedings of the American Statistical Association* (Survey Research Methods Section).

Glewwe, P. and M. Grosh (eds) (2000) *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Development Study*. Oxford University Press (for the World Bank).

Guarcello, L., I. Kovrova, S. Lyon, M. Manacorda, and F.C. Rosati (2009) "Towards Consistency in Child Labour Measurement: Assessing the Comparability of Estimates Generated by Different Survey Instruments." Understanding Children's Work Project Draft Working Paper.

Hausman, J. A. (2001) “Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left.” *Journal of Economic Perspectives* 15 (4):57-67.

Hausman, J.A., J. Abrevaya, and F.M. Scott-Morton (1998) “Misclassification of the Dependent Variable in a Discrete Response Setting.” *Journal of Econometrics* 87: 239-269.

Hill, D. H. (1987) “Response Errors in Labor Surveys: Comparisons of Self and Proxy Reports in the Survey of Income and Program Participation (SIPP).” in *Proceedings of the Bureau of Census, Third Annual Research Conference*.

Hussmanns, R., F. Mehran, and V. Verma (1990) *Surveys of Economically Active Population, Employment, Unemployment and Underemployment: An ILO Manual on Concepts and Methods*, ILO: Geneva.

Hyslop, D. R. and G. W. Imbens (2001) “Bias from Classical and Other Forms of Measurement Error.” *Journal of Business & Economic Statistics* 19(4): 475-481.

ILO (1982) Resolution Concerning Statistics of the Economically Active Population, Employment, Unemployment and Underemployment. Adopted by the Thirteenth International Conference of Labour Statisticians, Geneva: International Labour Organization.

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

Kalton, G., and H. Schuman (1982) “The Effect of the Question on Survey Responses: A Review.” *Journal of the Royal Statistical Society* 145(1):42-57.

Martin, E. and A. E. Polivka (1995) “Diagnostics for Redesigning Survey Questionnaires: Measuring Work in the Current Population Survey.” *Public Opinion Quarterly* 59:547-567.

Martin, J. and B. Butcher (1982) “The Quality of Proxy Information: Some Results from a Large Scale Study.” *The Statistician* 31:293-319.

Mata Greenwood, A. (2000) *Incorporating Gender Issues in Labour Statistics*, Geneva: International Labour Office, Bureau of Statistics.

Mathiowetz, N. A. and R. M. Groves (1985) “The Effects of Respondent Rules on Health Survey Reports.” *American Journal of Public Health* 75(6):639-644.

Moore, J. (1988) “Self/Proxy Response Status and Survey Response Quality: A Review of the Literature.” *Journal of Official Statistics* 4:155-172.

National Bureau of Statistics Tanzania (2003) Integrated Labour Force Survey, 2000/01 – Analytical Report, see http://www.nbs.go.tz/labourforce/index.htm

National Bureau of Statistics Tanzania (2002) Household Budget Report, 2000/01, Chapter 5. Dar es Salaam;
http://www.tanzania.go.tz/hbs/Index_FinalReport_HBS.htm

Podsakoff P.M., S.B. MacKenzie, J-Y Lee, and N.P Podsakoff (2003) "Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies." *Journal of Applied Psychology* 88(5):879-903.

Poterba, J. M. and L. H. Summers (1986) "Reporting Errors and Labor Market Dynamics." *Econometrica* 54(6):1319-1338.

Tversky A., and D. Kahneman (1981) "The Framing of Decisions and the Psychology of Choice." *Science* 211(4481):453-458.

Table 1. Four types of survey assignments

	Self-reported	Response by proxy
Detailed module	Group A	Group B
	336 households	336 households
	638 individuals	834 individuals
Short module	Group C	Group D
	336 households	336 households
	636 individuals	837 individuals

Note: See text for a detailed description of the four survey assignments (Groups A-D).

Table 2. Household characteristics, by survey assignment of household

	Household survey assignment			
	Detailed Self-report	Detailed Proxy	Short Self-report	Short Proxy
Head: female (%)	21.7	19.6	19.6	19.0
Head: age	46.5	45.8	45.8	47.7
Head: highest school grade completed	4.6	4.7	4.8	4.7
Head: married (%)	72.3	74.1	70.8	75.0
Household size*	5.5	5.0	5.0	5.3
Adult equivalence household size*	4.0	3.6	3.6	3.9
Share of members less 6 years*	19.3	18.2	17.5	17.1
Share of members 6-15 years	24.9	23.7	23.8	24.0
Number of adults 15+ years	2.8	2.7	2.7	2.8
Concrete/tile flooring (non-earth) (%)	25.0	25.3	24.7	25.9
Main source for lighting is electricity/generator/solar panels (%)	10.4	8.9	10.4	11.3
Owns a mobile telephone (%)	30.1	30.1	28.6	32.5
Bicycle (%)	42.9	39.9	44.3	44.9
Owns any land (%)	78.9	80.1	78.3	81.3
Acres of land owned (including 0s) *	3.3	2.9	2.8	3.3
Month of interview (1=Jan, 12=Dec)	6.0	5.9	5.9	5.9
N	336	336	336	336

Notes: * indicates statistical difference across at least two pairs at 5%. See NBS (2002) for details on the adult equivalence scales. Acres sample excludes 28 households deemed outliers.

Table 3. Household and individuals characteristics, by survey assignment of individual

	Individual survey assignment			
	Detailed Self-report	Detailed Proxy	Short Self-report	Short Proxy
Female*	53.6	52.6	49.6	53.2
Age*	33.9	28.9	34.4	29.5
Higher grade completed	4.6	4.3	4.5	4.3
Married (%)*	55.2	47.7	57.3	45.1
Head: female (%)	19.6	17.5	18.9	16.8
Head: age*	46.0	47.4	46.7	48.2
Head: highest school grade completed	4.6	4.7	4.7	4.8
Head: married (%)*	75.3	80.4	74.8	82.3
Household size*	5.4	5.8	5.2	6.2
Adult equivalence household size*	3.9	4.2	3.8	4.5
Share of members less 6 years*	19.3	18.1	17.7	17.6
Share of members 6-15 years*	25.4	27.9	24.6	28.6
Number of adults 15+ years*	2.8	3.0	2.8	3.2
Concrete/tile flooring (non-earth) (%)	24.3	25.7	24.5	24.6
Main source for lighting is electricity/generator/solar panels (%)	9.4	10.4	10.4	11.8
Owns a mobile telephone (%)*	29.9	32.3	29.8	33.5
Bicycle (%)*	43.2	43.2	45.9	49.6
Owns any land (%)*	79.9	83.2	80.7	83.6
Acres of land owned (including 0s) *	3.3	3.2	3.1	3.7
Any hours collecting firewood last 24 hours (%)*	27.4	22.1	28.7	25.7
Hours collecting firewood last 24 hours (including 0s) *	0.4	0.3	0.4	0.3
Any hours collecting water last 24 hours (%)*	48.7	44.5	50.6	48.9
Hours collecting water last 24 hours (including 0s)	0.4	0.4	0.4	0.4
Month of interview (1=Jan, 12=Dec)	6.1	5.8	5.9	5.8
N	942	530	937	536

Notes: * indicates statistical difference across at least two pairs at 5%. See NBS (2002) for details on the adult equivalence scales.

Table 4. Labor statistics by survey assignment and sex

	A.			B.			C.		
	Short	Detailed	Diff	Proxy	Self-rep	Diff	Short Proxy	Other	Diff
Labor force participation (%)									
Men	82.4	83.0	-0.6	74.3	87.3	-13.0***	74.1	84.6	-10.5***
Women	69.9	77.0	-7.2***	68.4	76.5	-8.1***	64.6	75.6	-11.0***
Weekly hours last week									
Men	30.0	27.7	2.3**	24.5	31.3	-6.9***	25.1	29.7	-4.6***
Women	22.3	23.0	-0.8	20.3	24.2	-4.2***	19.4	23.4	-4.0**
Daily earnings (Tshillings)									
Men	541	662	-121	637	580	57	471	628	-157
Women	198	384	-187**	271	306	-35	80	342	-262**

Notes: *** indicates statistical significant mean differences at 1%, ** at 5%, * at 10%. Samples for weekly hours and daily earnings are conditional on any wage work in the last 7 days (they exclude zeros).

Table 5. Sector distribution of the main activity in last 7 days by survey assignment and sex

	Men			Women		
A. Short or Detailed	Short	Detailed	Diff	Short	Detailed	Diff
Main activity^						
Agriculture	58.6	59.0	-0.4	60.1	65.7	-5.6**
Other sectors	23.8	24.0	-0.2	9.6	11.4	-1.8
Domestic Duties	7.9	2.2	5.7***	18.8	2.4	16.4***
No work	9.7	14.8	-5.1***	11.3	20.5	-9.2***
N	723	688		750	784	
Main activity among workers^						
Agriculture	71.1	71.1	0.0	86.1	85.3	0.8
Other sectors	28.9	28.9	0.0	13.7	14.7	-0.9
N	596	571		524	604	
B. Proxy or Self-report	Proxy	Self-rep	Diff	Proxy	Self-rep	Diff
Main activity^						
Agriculture	53.8	61.6	-7.8***	59.8	64.8	-5.1**
Other sectors	20.5	25.7	-5.2**	8.5	11.6	-3.1*
Domestic Duties	7.8	3.6	4.1***	13.5	8.7	4.8***
No work	17.9	9.0	8.9***	18.1	14.8	3.2**
N	502	909		564	970	
Main activity among workers^						
Agriculture	72.4	70.5	1.9	87.3	84.8	2.5
Other sectors	27.6	29.5	-1.9	12.4	15.2	-2.8
N	373	794		386	742	
C. Short proxy or not	Short, Proxy	Other	Diff	Short, Proxy	Other	Diff
Main activity^						
Agriculture	55.4	59.6	-4.2	56.8	64.4	-7.5***
Other sectors	18.7	25.0	-6.3**	7.4	11.2	-3.8**
Domestic Duties	11.6	3.7	7.8***	23.5	7.4	16.0***
No work	14.3	11.7	2.6	11.9	16.9	-5.0**
N	251	1,160		285	1,249	
Main activity among workers^						
Agriculture	74.7	70.4	4.3	88.0	85.2	2.9
Other sectors	25.3	29.6	-4.3	11.4	14.8	-3.4
N	186	981		184	944	

Notes: Other sectors are specifically listed on the questionnaire and include mining/quarrying, manufacturing/ processing, gas/water/electricity, construction, transport, trading, personal services, education/health, public administration, and other *** indicates statistical significant mean differences at 1%, ** at 5%, * at 10%.
^ Within group, the percentages sum to 100.

Table 6. Employment status of those in the labor force by survey assignment and sex

		Men			Women	
A. Short or Detailed	Short	Detailed	Diff	Short	Detailed	Diff
Paid employee	14.1	19.6	-5.5***	6.3	11.4	-5.1***
Self-employed, with employees	3.5	4.2	-0.7	1.0	0.8	0.2
Self-employed, no employees	56.5	52.0	4.5*	25.4	24.8	0.6
Unpaid family worker	25.5	23.5	2.0	66.6	61.9	4.7*
N	596	571		524	604	
B. Proxy or Self-report	Proxy	Self-rep	Diff	Proxy	Self-rep	Diff
Paid employee	14.5	17.9	-3.4*	7.5	9.8	-2.3
Self-employed, with employees	2.1	4.7	-2.5**	0.8	0.9	-0.1
Self-employed, no employees	50.1	56.3	-6.2**	18.6	28.4	-9.8***
Unpaid family worker	33.0	20.6	12.4***	72.5	59.7	12.8***
N	373	794		386	742	
C. Short proxy or not	Short, Proxy	Other	Diff	Short, Proxy	Other	Diff
Paid employee	12.9	17.5	-4.6*	5.4	9.7	-4.3**
Self-employed, with employees	2.2	4.2	-2.0*	1.1	0.8	0.3
Self-employed, no employees	48.9	55.4	-6.5*	17.4	26.6	-9.2***
Unpaid family worker	36.0	22.3	13.7***	76.0	61.9	13.7***
N	184	944		186	981	

Notes: *** indicates statistical significant mean differences at 1%, ** at 5%, * at 10%.

^ Column percentages by group may not sum up to 100 due to missing values.

Table 7. Summary of descriptive findings for men and women

	Short v. Detailed	Proxy v. Self-report	Short proxy v. others
Labor force participation	Lower (women)	-	Lower
Working hours	-	Lower	Lower
Income	Lower	-	Lower
Activity distribution	More domestic duties Less 'no work' Less agric and other sectors (women)	More domestic duties More 'no work' Less agric and other sectors	More domestic duties Less 'no work' (women) Less agric (women) and other sectors
Employment status	Less paid employee (men) More self-empl (men) More unpaid family worker (women)	Less paid employee (men) Less self-employed More unpaid family worker	Less paid employee Less self-employed More unpaid family worker

Table 8. Probit and regression of labor statistics by survey assignment and sex

	(1) Labor force participation		(2) Conditional weekly hours		(3) Conditional daily earnings	
	Men	Women	Men	Women	Men	Women
Short	-0.012	-0.048	0.168***	0.114**	0.131	0.229
	(0.024)	(0.030)	(0.050)	(0.051)	(0.121)	(0.232)
Proxy	-0.124***	-0.039	-0.086	-0.104*	0.239	0.662*
	(0.036)	(0.028)	(0.073)	(0.062)	(0.254)	(0.383)
Short*proxy	0.017	-0.043	-0.040	0.018	-0.090	-1.538***
	(0.043)	(0.039)	(0.114)	(0.084)	(0.398)	(0.399)
N	1,411	1,534	1,166	1,127	199	107

Notes: Standard errors in parentheses. *** indicates statistical significance at 1%, ** at 5%, * at 10%. Other covariates included but not presented are characteristics of subjects (age, gender, education) and households (composition and assets) as well as village fixed effects.
Column 1: Labor force participation estimates are conducted using a probit model; marginal effects are reported. Column 2: OLS estimation on log hours. Column 3: OLS estimation on log daily earnings.

Table 9. Multinomial logit of main activity in last 7 days by survey assignment and sex

	Men		Women	
	Agriculture	Other Sectors	Agriculture	Other Sectors
Short	0.860	0.927	0.750	0.675
	(0.194)	(0.197)	(0.139)	(0.171)
Proxy	0.386***	0.449***	0.822	0.661*
	(0.101)	(0.132)	(0.159)	(0.146)
Short*proxy	1.292	1.064	0.774	0.741
	(0.495)	(0.397)	(0.206)	(0.283)
N	1,411		1,534	

Notes: Other covariates included but not presented are characteristics of subjects (age, sex, education) and households (composition and assets) as well as village fixed effects. The multinomial logit model is used where the three categories include agriculture, other sectors and the omitted category (domestic work and no work). Relative risk ratios are reported. See Table 6 note for explanation of other sectors. *** indicates statistical significance at 1%, ** at 5%, * at 10%.

Table 10. Multinomial logit of employment status in main job in the last 7 days by survey assignment and sex

	Men			Women		
	Wage employee	Self-employed with employees	Self-employed, no employees	Wage employee	Self-employed with employees	Self-employed, no employees
Short	0.393***	0.594	1.014	0.318***	0.715	0.957
	(0.093)	(0.234)	(0.511)	(0.103)	(0.273)	(0.441)
Proxy	0.560	0.414	0.243	0.439	0.585	2.085
	(0.234)	(0.243)	(0.533)	(0.444)	(0.794)	(3.359)
Short*proxy	0.766	0.976	0.497	0.657	0.838	0.644
	(0.194)	(0.399)	(0.241)	(0.144)	(0.232)	(0.230)
N	1,233			1,279		

Notes: Other covariates included but not presented are characteristics of subjects (age, gender, education) and households (composition and assets) as well as village fixed effects. The multinomial logit model is used where the three categories include wage employee, self-employed with employees, self employed without employees, and the omitted category (unpaid family worker). Relative risk ratios are reported. *** indicates statistical significance at 1%, ** at 5%, * at 10%.

Table 11a. Labor statistics by proxy-subject characteristics – difference in age

	Mean	Proxy-subject age difference quartiles			
		Q1: proxy younger than subject	Q2	Q3	Q4: proxy older than subject
Subjects <= median age					
Labor force participation (%)	60.6	71.4	71.4***	60.3	55.6**
Weekly hours last week	14.5	33.0***	19.6***	13.8	12.2***
Daily earnings (Tshillings)	204	0	769***	87	54*
N	523	7	105	199	232
Subjects > median age					
Labor force participation (%)	82.2	81.8	79.6	87.7*	85.7
Weekly hours last week	30.0	30.9	28.1*	32.4	27.3
Daily earnings (Tshillings)	692	927*	481	461	391
N	523	258	157	73	35

Notes: *** indicates statistical significant at 1%, ** at 5%, and * at 10%. Labor statistics are disaggregated by the quartile of age differences between proxies and the person on whom they report (subject). The ttest conducted is between the mean of the disaggregated labor statistic in the given group versus the mean of the labor statistics in all other groups. The smaller sample size in this table is due to restricting the sample to only proxy responses. The median subject age is 24. The quartile intervals are as follows: Q1< -5 years difference, -5 years ≤ Q2 < 7 years, 7 years ≤ Q3 < 25 years, and Q4 ≥ 25 years.

Table 11b. Labor statistics by proxy-subject characteristics – difference in sex

	Mean	Proxy-subject gender interactions					
	(1)	M-F (2)	F-F (3)	Diff.	F-M (4)	M-M (5)	Diff.
Labor force participation (%)	71.2	73.1	60.7	12.4***	77.7	65.2	12.5***
Weekly hours last week	22.1	22.0	16.7	5.4***	26.5	18.9	7.6***
Daily earnings (Tshillings)	444	307	213	94	740	360	382
N	1,066	350	214		367	135	

Notes: *** indicates statistical significant at 1%, ** at 5%, and * at 10%. Labor statistics are disaggregated by proxy-subject gender interactions (M-F indicates a male proxy who reports on a female subject, and so on). The ttest conducted is between M-F and F-F in Columns (2) and (3), and F-M and M-M in Columns (4) and (5). The smaller sample size in this table is due to restricting the sample to only proxy responses.

Table 11c. Labor statistics by proxy-subject characteristics – difference in education

	Mean	Proxy-subject education interactions					
	(1)	S-S (2)	N-S (3)	Diff.	S-N (4)	N-N (5)	Diff.
Labor force participation (%)	71.2	67.3	78.8	-11.5***	76.5	78.9	2.4
Weekly hours last week	22.1	21.7	22.3	0.4	24.4	23.8	0.6
Daily earnings (Tshillings)	444	579	145	434	133	472	339
N	1,066	679	179		132	76	

Notes: *** indicates statistical significant at 1%, ** at 5%, and * at 10%. Labor statistics are disaggregated by proxy-subject interactions in the education level (S-N indicates that the proxy went to school, while the subject received no education, and so on). The ttest conducted is between S-S and N-S in Columns (2) and (3) and S-N and N-N in Columns (4) and (5). The smaller sample size in this table is due to restricting the sample to only proxy responses.

Appendix 1: Short and detailed labor modules

SHORT MODULE

1. Did [NAME] do any type of work in the last seven days? Even if for 1 hour.	YES...1 (»3) NO....2	
2. Did [NAME] do any type of work in the last 12 months?	YES...1 (»12) NO....2 (»12)	
3. What is [NAME]'s primary occupation in [NAME]'s main job? MAIN OCCUPATION IN THE LAST 7 DAYS	a. OCCUPATION	b. OCCUPATION CODE (TO BE FILLED IN BY SUPERVISOR)
4. In what sector is this main activity? READ ALL RESPONSES	AGRICULTURE. . . . . .1 MINING/QUARRYING. .2 MANUFACTURING/ PROCESSING. . . . . ..3 GAS/WATER/ ELECTRICITY . . . . . .4 CONSTRUCTION. 5	TRANSPORT. . . . . 6 BUYING AND SELLING . . . . . .7 PERSONAL SERVICES. . . . . 8 EDUCATION/HEALTH . . . . . 9 PUBLIC ADMINISTRATION. .10 DOMESTIC DUTIES . . . . . 11 OTHER, SPECIFY. 12
5. For how many hours did [NAME] work in the last 7 days in [NAME]'s main job? IF DID NOT WORK ENTER 0	HOURS	
6. What is [NAME]'s employment status in [NAME]'s main job? READ ALL RESPONSES	PAID EMPLOYEE . . . ..1 SELF-EMPLOYED WITH EMPLOYEES . . . . . 2 (>>12) SELF-EMPLOYED, NO EMPLOYEES. . . . . .3 (>>12)	UNPAID FAMILY WORKER. . .4 (>>12) DOMESTIC EMPLOYEE. . . . . .5 OTHER, SPECIFY. 6
7. Who is [NAME]'s employer in [NAME]'s main job? READ ALL RESPONSES	GOVERNMENT. . . . . .1 NGO . . . . . .2 COOPERATIVE. . . . . 3 INTERNATIONAL ORGANIZATION. 4	PRIVATE SECTOR. . . . . 5 HOUSEHOLD. . . . . .6 OTHER, SPECIFY. 8
8. Did [NAME] receive wages, salary, cash payments or other in kind payments from this employer for this work?	YES...1 NO....2 (»12)	
9. How much was [NAME]'s last payment? IF NONE ENTER 0	a. AMOUNT (in TSH)	b. TIME UNIT HOUR . . . . . .1 DAY . . . . . 2 WEEK . . . . . .3 MONTH . . . . . .4 YEAR 5

10. Did [NAME] receive any in-kind payment or regular allowance for the work [NAME] performed for this employer? IF NONE ENTER 0	a. IN KIND ITEM ALLOWANCE . . . . . 1 HOUSING . . . . . 2 TRANSPORT . . . . . 3 PRODUCE . . . . . 4 ANIMALS. . . . . 5 MEALS. . . . . 6 CLOTHING. . . . . 7 MEDICATION . . . . . 8 OTHER, DURABLE GOODS SPECIFY. . . . . 9 OTHER, NON-DURABLE GOODS SPECIFY . . 10	b. VALUE (ESTIMATE AMOUNT IN TSH) c. TIME UNIT HOUR . . . . 1 DAY. . . . . 2 WEEK . . . . 3 MONTH. . . 4 YEAR 5
11. Did [NAME] receive any other in-kind payment or regular allowance for the work [NAME] performed for this employer? IF NONE ENTER 0	a. IN KIND ITEM ALLOWANCE . . . . . 1 HOUSING . . . . . 2 TRANSPORT . . . . . 3 PRODUCE . . . . . 4 ANIMALS. . . . . 5 MEALS. . . . . 6 CLOTHING. . . . . 7 MEDICATION . . . . . 8 OTHER, DURABLE GOODS SPECIFY. . . . . 9 OTHER, NON-DURABLE GOODS SPECIFY . . 10	b. VALUE (ESTIMATE AMOUNT IN TSH) c. TIME UNIT HOUR . . . . 1 DAY. . . . . 2 WEEK . . . . 3 MONTH. . . 4 YEAR 5
12. Now I would like to ask about activities that you or [NAME] does for the household. How much time in the last 24 hours did you or [NAME] spend on any of the following activities for the household? IF NONE ENTER 0		
a. COLLECTING FIREWOOD?	HOURS MINUTES	
b. FETCHING WATER?	HOURS MINUTES	
13. Does you or [NAME] usually do any of the following activities?		
a. WASHING CLOTHES?	YES...1 NO....2	
b. COOKING?	YES...1 NO....2	
c. TAKING CARE OF SICK/ELDERLY?	YES...1 NO....2	

DETAILED MODULE

1. During the past 7 days, has [NAME] worked for someone who is not a member of your household, for example, an enterprise, company, the government or any other individual?	YES...1 (»3) NO....2	
2. At any time during the past 12 months, has [NAME] worked for someone who is not a member of your household, for example, an enterprise, company, the government or any other individual?	YES...1 NO....2	
3. During the past 7 days, has [NAME] worked on a farm owned, borrowed or rented by a member of your household, whether in cultivating crops or in other farm maintenance tasks, or have you cared for livestock belonging to a member of your household?	YES...1 (»5) NO....2	
4. At any time during the last 12 months has [NAME] worked on a farm owned, borrowed or rented by a member of your household, whether in cultivating crops or in other farm maintenance tasks, or have you cared for livestock belonging to a member of your household?	YES...1 NO....2	
5. During the past 7 days, has [NAME] worked *on your own account or in a business enterprise* belonging to you or someone in your household, for example, as a trader, shop-keeper, barber, dressmaker, carpenter or taxi driver?	YES...1 (»7) NO....2	
6. At any time during the last 12 months, has [NAME] worked on your own account or in a business enterprise belonging to you or someone in your household, for example, as a trader, shop-keeper, barber, dressmaker, carpenter or taxi driver?	YES...1 NO....2	
7. CHECK THE ANSWERS TO QUESTIONS 1, 3 AND 7. (WORKED IN LAST 7 DAYS)	ANY YES..1 ALL NO...2 (»37)	
8. What is [NAME]'s primary occupation in [NAME]'s main job? (MAIN OCCUPATION IN THE LAST 7 DAYS)	a. OCCUPATION	b. OCCUPATION CODE (TO BE FILLED IN BY SUPERVISOR)
9. In what sector is this main activity?	AGRICULTURE. . . . . .1 MINING/QUARRYING. .2 MANUFACTURING/ PROCESSING. . . . . ..3 GAS/WATER/ ELECTRICITY . . . . . .4 CONSTRUCTION. 5	TRANSPORT. . . . . 6 BUYING AND SELLING . . . . . .7 PERSONAL SERVICES. . . . . 8 EDUCATION/HEALTH . . . . . 9 PUBLIC ADMINISTRATION. .10 DOMESTIC DUTIES . . . . . 11 OTHER, SPECIFY. 12
10. For how many hours did [NAME] work in the last 7 days in [NAME]'s main job? (IF NOT WORKED, ENTER 0)	HOURS	

11. What is [NAME]'s employment status in [NAME]'s main job? READ ALL RESPONSES	PAID EMPLOYEE . . . . .1 SELF-EMPLOYED WITH EMPLOYEES . . 2 (>>17) SELF-EMPLOYED, NO EMPLOYEES. . . . .3 (>>17)	UNPAID FAMILY WORKER. .4 (>>17) DOMESTIC EMPLOYEE. . . . .5 OTHER, SPECIFY. 6
12. Who is [NAME]'s employer in [NAME]'s main job? READ ALL RESPONSES	GOVERNMENT. . . . . .1 NGO . . . . . .2 COOPERATIVE. . . . . 3 INTERNATIONAL ORGANIZATION. 4	PRIVATE SECTOR. . . . . 5 HOUSEHOLD. . . . . .6 OTHER, SPECIFY 8
13. Did [NAME] receive wages, salary, cash payments or other in kind payments from this employer for this work?	YES...1 NO....2 (»17)	
14. How much was [NAME]'s last payment?	a. AMOUNT (in TSH)	b. TIME UNIT HOUR . . . . . .1 DAY . . . . . 2 WEEK . . . . . .3 MONTH . . . . . .4 YEAR 5
15. Did [NAME] receive any in-kind payment or regular allowance for the work [NAME] performed for this employer?	a. IN KIND ITEM ALLOWANCE . . . . . 1 HOUSING . . . . . 2 TRANSPORT . . . . . 3 PRODUCE . . . . . 4 ANIMALS. . . . . 5 MEALS. . . . . 6 CLOTHING. . . . . 7 MEDICATION . . . . . 8 OTHER, DURABLE GOODS SPECIFY. . . . . 9 OTHER, NON-DURABLE GOODS SPECIFY . . 10	b. VALUE (ESTIMATE AMOUNT IN TSH) c. TIME UNIT HOUR . . . . 1 DAY. . . . . 2 WEEK . . . . 3 MONTH. . . 4 YEAR 5
16. Did [NAME] receive any other in-kind payment or regular allowance for the work [NAME] performed for this employer?	a. IN KIND ITEM ALLOWANCE . . . . . 1 HOUSING . . . . . 2 TRANSPORT . . . . . 3 PRODUCE . . . . . 4 ANIMALS. . . . . 5 MEALS. . . . . 6 CLOTHING. . . . . 7 MEDICATION . . . . . 8 OTHER, DURABLE GOODS SPECIFY. . . . . 9 OTHER, NON-DURABLE GOODS SPECIFY . . 10	b. VALUE (ESTIMATE AMOUNT IN TSH) c. TIME UNIT HOUR . . . . 1 DAY. . . . . 2 WEEK . . . . 3 MONTH. . . 4 YEAR 5

17. Did [NAME] have a second job or economic activity in the last seven days?	YES...1 NO....2 (»37)	
18. What is [NAME]'s primary occupation in [NAME]'s second job in the last 7 days?	a. OCCUPATION	b. OCCUPATION CODE (TO BE FILLED IN BY SUPERVISOR)
19. In what sector is this secondary activity? READ ALL RESPONSES	AGRICULTURE. . . . . .1 MINING/QUARRYING. .2 MANUFACTURING/ PROCESSING. . . . . ..3 GAS/WATER/ ELECTRICITY . . . . . .4 CONSTRUCTION. 5	TRANSPORT. . . . . 6 BUYING AND SELLING . . . . . .7 PERSONAL SERVICES. . . . . 8 EDUCATION/HEALTH . . . . . 9 PUBLIC ADMINISTRATION. .10 DOMESTIC DUTIES . . . . . 11 OTHER, SPECIFY. 12
20. For how many hours did [NAME] work in the last 7 days in [NAME]'s second job? IF NOT WORKED ENTER 0	HOURS	
21. What is [NAME]'s employment status in [NAME]'s second job? READ ALL RESPONSES	PAID EMPLOYEE . . . . .1 SELF-EMPLOYED WITH EMPLOYEES. . . . . .2 (>>27) SELF-EMPLOYED, NO EMPLOYEES. . . . . .3 (>>27)	UNPAID FAMILY WORKER. . 4 (>>27) DOMESTIC EMPLOYEE. . . . . 5 OTHER, SPECIFY. 6
22. Who is [NAME]'s employer in [NAME]'s second job? READ ALL RESPONSES	GOVERNMENT. . . . . .1 NGO . . . . . .2 COOPERATIVE. . . . . 3 INTERNATIONAL ORGANIZATION. 4	PRIVATE SECTOR. . . . . 5 HOUSEHOLD. . . . . .6 OTHER, SPECIFY. 8
23. Did [NAME] receive wages, salary, cash payments or other in kind payments from this employer for this work?	YES...1 NO....2 (»27)	
24. How much was [NAME]'s last payment?	a. AMOUNT (in TSH)	b. TIME UNIT HOUR . . . . . .1 DAY . . . . . 2 WEEK . . . . . .3 MONTH . . . . . .4 YEAR 5

25. Did [NAME] receive any in-kind payment or regular allowance for the work [NAME] performed for this employer? IF NONE ENTER 0	a. IN KIND ITEM ALLOWANCE . . . . . 1 HOUSING . . . . . 2 TRANSPORT . . . . . 3 PRODUCE . . . . . 4 ANIMALS. . . . . 5 MEALS. . . . . 6 CLOTHING. . . . . 7 MEDICATION . . . . . 8 OTHER, DURABLE GOODS SPECIFY. . . . . 9 OTHER, NON-DURABLE GOODS SPECIFY . . 10	b. VALUE (ESTIMATE AMOUNT IN TSH) c. TIME UNIT HOUR . . . . 1 DAY. . . . . 2 WEEK . . . . 3 MONTH. . . 4 YEAR 5
26. Did [NAME] receive any other in-kind payment or regular allowance for the work [NAME] performed for this employer? IF NONE ENTER 0	a. IN KIND ITEM ALLOWANCE . . . . . 1 HOUSING . . . . . 2 TRANSPORT . . . . . 3 PRODUCE . . . . . 4 ANIMALS. . . . . 5 MEALS. . . . . 6 CLOTHING. . . . . 7 MEDICATION . . . . . 8 OTHER, DURABLE GOODS SPECIFY. . . . . 9 OTHER, NON-DURABLE GOODS SPECIFY . . 10	b. VALUE (ESTIMATE AMOUNT IN TSH) c. TIME UNIT HOUR . . . . 1 DAY. . . . . 2 WEEK . . . . 3 MONTH. . . 4 YEAR 5
27. Did [NAME] have a third job or economic activity in the last seven days?	YES...1 NO....2 (»37)	
28. What is [NAME]'s primary occupation in [NAME]'s third job?	a. OCCUPATION	b. OCCUPATION CODE (TO BE FILLED IN BY SUPERVISOR)
29. In what sector is this third activity? READ ALL RESPONSES	AGRICULTURE. . . . . .1 MINING/QUARRYING. .2 MANUFACTURING/ PROCESSING. . . . . ..3 GAS/WATER/ ELECTRICITY . . . . . .4 CONSTRUCTION. 5	TRANSPORT. . . . . 6 BUYING AND SELLING . . . . . .7 PERSONAL SERVICES. . . . . 8 EDUCATION/HEALTH . . . . . 9 PUBLIC ADMINISTRATION. .10 DOMESTIC DUTIES . . . . . 11 OTHER, SPECIFY. 12
30. For how many hours did [NAME] work in the last 7 days in [NAME]'s third job?	HOURS	
31. What is [NAME]'s employment status in [NAME]'s second job? READ ALL RESPONSES	PAID EMPLOYEE . . . . .1 SELF-EMPLOYED WITH EMPLOYEES. . . . . .2 (>>37) SELF-EMPLOYED, NO	UNPAID FAMILY WORKER. . 4 (>>37) DOMESTIC EMPLOYEE. . . . . 5 OTHER, SPECIFY. 6

	EMPLOYEES. . . . . .3 (>>37)	
32. Who is [NAME]'s employer in [NAME]'s second job? READ ALL RESPONSES	GOVERNMENT. . . . . .1 NGO . . . . . .2 COOPERATIVE. . . . . 3 INTERNATIONAL ORGANIZATION. 4	PRIVATE SECTOR. . . . . 5 HOUSEHOLD. . . . . .6 OTHER, SPECIFY. 8
33. Did [NAME] receive wages, salary, cash payments or other in kind payments from this employer for this work?	YES...1 NO....2 (»37)	
34. How much was [NAME]'s last payment?	a. AMOUNT (in TSH)	b. TIME UNIT HOUR . . . . . .1 DAY . . . . . 2 WEEK . . . . . .3 MONTH . . . . . .4 YEAR 5
35. Did [NAME] receive any in-kind payment or regular allowance for the work [NAME] performed for this employer? IF NONE ENTER 0	a. IN KIND ITEM ALLOWANCE . . . . . 1 HOUSING . . . . . 2 TRANSPORT . . . . . 3 PRODUCE . . . . . 4 ANIMALS. . . . . 5 MEALS. . . . . 6 CLOTHING. . . . . 7 MEDICATION . . . . . 8 OTHER, DURABLE GOODS SPECIFY. . . . . 9 OTHER, NON-DURABLE GOODS SPECIFY . . 10	b. VALUE (ESTIMATE AMOUNT IN TSH) c. TIME UNIT HOUR . . . . 1 DAY. . . . . 2 WEEK . . . . 3 MONTH. . . 4 YEAR 5
36. Did [NAME] receive any other in-kind payment or regular allowance for the work [NAME] performed for this employer? IF NONE ENTER 0	a. IN KIND ITEM ALLOWANCE . . . . . 1 HOUSING . . . . . 2 TRANSPORT . . . . . 3 PRODUCE . . . . . 4 ANIMALS. . . . . 5 MEALS. . . . . 6 CLOTHING. . . . . 7 MEDICATION . . . . . 8 OTHER, DURABLE GOODS SPECIFY. . . . . 9 OTHER, NON-DURABLE GOODS SPECIFY . . 10	b. VALUE (ESTIMATE AMOUNT IN TSH) c. TIME UNIT HOUR . . . . 1 DAY. . . . . 2 WEEK . . . . 3 MONTH. . . 4 YEAR 5
37. Now I would like to ask about you or [NAME]'s activities that you or [NAME] does for the household. How much time in the last 24 hours did you or [NAME]		

spend on any of the following activities for the household?	
a. COLLECTING FIREWOOD?	HOURS MINUTES
b. FETCHING WATER?	HOURS MINUTES
38. Does [NAME] usually do any of the following activities?	
a. WASHING CLOTHES?	YES...1 NO....2
b. COOKING?	YES...1 NO....2
c. TAKING CARE OF SICK/ELDERLY?	YES...1 NO....2

Appendix 2: Planned and actual survey assignments

	Household survey assignment				
	Detailed self-reported	Detailed proxy response	Short self-reported	Short proxy	Total
Households					
Number (planned = actual)	336	336	336	336	1344
Percent with one adult 15+	14.0	12.2	14.6	11.9	
Percent with one member 10+	9.8	9.2	10.7	10.7	
Planned individual assignment, if every household has at least 3 members over 10 years of age, and at least one member over 15 years.^					
Detailed self-reported	672	336	0	0	1008
Detailed proxy response	0	672	0	0	672
Short self-reported	0	0	672	336	1008
Short proxy planned	0	0	0	672	672
Planned individual assignment, given assumption about household composition^ #					
Detailed self-reported	672	336	0	0	1008
Detailed proxy response	0	504	0	0	504
Short self-reported	0	0	672	336	1008
Short proxy planned	0	0	0	504	504
Actual individual assignment					
Detailed self-reported	606	336	0	0	942
Detailed proxy response	32	498	0	0	530
Short self-reported	0	0	601	336	937
Short proxy	0	0	35	501	536
Total actual number of individuals					2,945

^ Assuming that each household has at least 2 persons age 10+ to be randomly selected for self-report.

Assuming that each household has one member 15+ and an average of 2.5 household members 10+years per household. Thus, there are 1.5 *336 other members to be reported on by proxy.

Appendix 3. Household and individuals characteristics, by survey assignment of individual, among individuals 15+ years

	Individual survey assignment			
	Detailed Self-report	Detailed Proxy	Short Self-report	Short Proxy
Female	53.6	52.3	50.1	52.7
Age*	36.6	34.8	37.0	35.8
Higher grade completed	4.8	4.8	4.7	4.9
Married (%)	62.0	64.5	63.8	61.3
Head: female (%)	19.2	16.1	18.2	16.2
Head: age*	45.7	47.5	46.4	48.2
Head: highest school grade completed	4.7	4.8	4.8	4.9
Head: married (%)*	75.4	81.9	75.1	82.3
Household size*	5.3	5.7	5.1	6.0
Adult equivalence household size*	4.2	4.6	4.1	4.8
Share of members less 6 years*	19.5	18.7	18.0	17.7
Share of members 6-15 years	23.4	22.6	22.4	23.8
Number of adults 15+ years*	2.8	3.2	2.8	3.3
Concrete/tile flooring (non-earth) (%)	25.3	28.1	25.2	25.3
Main source for lighting is electricity/generator/solar panels (%)*	9.8	12.0	10.8	12.9
Owns a mobile telephone (%)*	31.1	34.4	30.6	36.0
Bicycle (%)*	42.2	43.1	45.8	49.4
Owns any land (%)*	79.5	81.4	80.6	83.5
Acres of land owned (including 0s) *	3.2	3.1	3.1	3.4
Any hours collecting firewood last 24 hours (%)*	27.5	18.9	27.8	25.8
Hours collecting firewood last 24 hours (including 0s) *	0.37	0.23	0.40	0.34
Any hours collecting water last 24 hours (%)*	47.3	39.3	48.2	43.5
Hours collecting water last 24 hours (including 0s)	0.34	0.37	0.40	0.37
Month of interview (1=Jan, 12=Dec) *	6.0	5.5	5.9	5.8
N	839	392	842	395

Notes: * indicates statistical difference across at least two pairs at 5%. See NBS (2002) for details on the adult equivalence scales.