AUTHOR ACCEPTED MANUSCRIPT

FINAL PUBLICATION INFORMATION

Do Returns to Education Depend on How and Whom You Ask?

The definitive version of the text was subsequently published in

Economics of Education Review, 60, 2017-10

Published by Elsevier and found at http://dx.doi.org/10.1016/j.econedurev.2017.07.010

THE FINAL PUBLISHED VERSION OF THIS MANUSCRIPT IS AVAILABLE ON THE PUBLISHER'S PLATFORM

This Author Accepted Manuscript is copyrighted by World Bank and published by Elsevier. It is posted here by agreement between them. Changes resulting from the publishing process—such as editing, corrections, structural formatting, and other quality control mechanisms—may not be reflected in this version of the text.

You may download, copy, and distribute this Author Accepted Manuscript for noncommercial purposes. Your license is limited by the following restrictions:

(1) You may use this Author Accepted Manuscript for noncommercial purposes only under a CC BY-NC-ND 3.0 IGO license http://creativecommons.org/licenses/by-nc-nd/3.0/igo.

(2) The integrity of the work and identification of the author, copyright owner, and publisher must be preserved in any copy.

(3) You must attribute this Author Accepted Manuscript in the following format: This is an Author Accepted Manuscript by Serneels, Pieter; Beegle, Kathleen; Dillon, Andrew *Do Returns to Education Depend on How and Whom You Ask?* © World Bank, published in the Economics of Education Review60 2017-10 CC BY-NC-ND 3.0 IGO http://creativecommons.org/licenses/by-nc-nd/3.0/igo http://dx.doi.org/10.1016/j.econedurev.2017.07.010

© 2023 World Bank

Do returns to education depend on how and whom you ask?

Pieter Serneels — University of East Anglia
Kathleen Beegle — The World Bank
Andrew Dillon — Michigan State University

July 2017

Abstract[1]

Returns to education remain an important parameter of interest in economic analysis. A large literature estimates these returns, often carefully addressing issues such as selection into wage employment and endogeneity in terms of completed schooling. There has been much less exploration of whether the estimates of Mincerian returns depend on how information about wage work is collected. Relying on a survey experiment in Tanzania, this paper finds that estimates of the returns to education vary by questionnaire design, but not by whether the information on employment and wages is self-reported or collected by a proxy respondent. The differences derived from questionnaire type are substantial, varying from higher returns of 5 percentage points among the most well educated men to 16 percentage points among the least well educated women. These differences are at magnitudes similar to the bias in ordinary least squares estimation, which receives considerable attention in the literature. The findings demonstrate that survey design matters in the estimation of returns to schooling and that care is needed in comparing across contexts and over time, particularly if the data are generated through different surveys.

Key words: returns to education, survey design, field experiment, development, Tanzania
JEL codes: J24, J31, C83

---

[1] Pieter Serneels (corresponding author) is Associate Professor at the University of East Anglia; his email address is p.serneels@uea.ac.uk. Kathleen Beegle is Lead Economist at the World Bank; her email is kbeegle@worldbank.org. Andrew Dillon is Assistant Professor at Michigan State University; his email is dillona6@msu.edu. The authors would like to thank Joachim De Weerdt, Brian Dillon, Hans Hoogeveen, Andrew Kerr, Mans Söderbom, Helene Bie Lilleør, attendants at the Centre for the Study of African Economies' conference at Oxford University, and two anonymous referees for useful comments. This work has been supported by the World Bank Research Support Budget, the World Bank's Gender Action Plan Trust Fund, and the Norwegian-Dutch World Bank Trust Fund for Mainstreaming Gender. All errors remain the responsibility of the authors.

1. Introduction

Surveys represent a primary approach to data collection in economics and the social sciences. Yet, variation exists in the design and the protocols for survey implementation. These discrepancies may induce nonrandom measurement error, but the potential biases are rarely quantified. This paper investigates whether survey methods matter in estimating the returns to education, a common parameter in economic studies on education, labor, and development. Based on data from a survey field experiment in Tanzania, estimated returns to education vary by questionnaire type (short or more detailed labor questions), but not by whether the respondents were interviewed directly or by proxy (another household member). The effect of survey design features is heterogeneous, that is, they vary by gender and level of education.

The empirical estimation of Mincerian returns to education using earnings functions (Mincer 1958) is a standard approach in economics and frequently involves comparison of returns over time, across countries, and across subgroups within countries (see, among others, Peet, Fink and Fawzi 2015; Psacharopoulos and Patrinos 2004; Shultz 2004; World Bank 2011). Such analysis often relies on estimates from different surveys. Attention has been given to limitations of comparability because of differences in sample coverage in the surveys (selection issues) and the method of analysis (such as nonlinearity). A body of work also analyzes how to produce structurally accurate ("true") estimates of the returns to education, taking endogeneity into account.[2] These estimations provide important inputs to policy debates, especially in low- and middle-income countries in which education attracts tremendous attention as a contributor to economic growth and poverty reduction. However, less consideration has been given to possible discrepancies arising from differences in survey methods.

A growing body of evidence indicates that survey methods matter for the education, labor, and income statistics they generate. This is well illustrated for high income countries, especially the United States (Bound, Brown, and Mathiowetz 2001). Recent evidence confirms this also for developing countries (Bardasi et al. 2011; de Mel, McKenzie, and Woodruff 2009).[3] Survey methods may impact estimates in economic analysis if measurement error varies according to survey design in a systematic way. Nonrandom measurement error in a continuous left-hand side variable in regression analysis would normally not bias ordinary least squares (OLS) point estimates, although it may reduce precision. However, it may lead to bias if the measurement error arises because respondents make systematic or nonrandom errors in

[2] Alternative ways to address this include instrumental variable estimation (such as Angrist and Krueger 1991; Card 2001; Cruz and Moreira 2005), twin studies (Behrman, Rosenzweig, and Taubman 1994, 1996; Bound and Solon 1999; Miller, Mulvey, and Martin 1995) or randomized control trials.

[3] Some work on survey measurement error in high-income countries focuses on how survey replies deviate from true values by using validation studies. For instance, in the case of wages, employee survey replies are compared with employer pay records. Such administrative records are either missing or incomplete in low-income settings.

reporting values. Response by proxy may result in this type of error given that it is potentially prone to strategic guessing. A common example involves husbands who underreport the earnings of their wives (or vice versa), for instance, because they are not aware of the details of their spouse's income activities. At the extensive margin, even whether people work may be underreported.

The study described in this paper has involved the implementation of a field experiment in Tanzania that had variation in two key dimensions of survey design: the level of detail in employment screening questions and the respondents selected, whether self-reporting or proxy reporting. The study investigates whether these differences in survey design yield varying estimates of the returns to education by applying a common econometric approach and identification strategy.

The results indicate that survey methods matter, in particular the use of short versus detailed employment screening questions. While linear OLS estimates suggest that the short questionnaire generates significant differences among men, but not among women, a more advanced analysis that allows for nonlinear returns and accounts for endogeneity finds significantly different results among both men and women. The short questionnaire yields higher returns to education of 5 percentage points among the most highly educated men (secondary or tertiary) and 16 percentage points among the least well educated women (primary). These discrepancies are of a similar or larger magnitude relative to commonly observed biases associated with simple OLS estimation that are the subject of a large literature (Card 1999). The divergence stems from differences in the categorization of people as having wage work caused by the absence of screening questions in the short questionnaire. No differences in estimated schooling returns are observed in proxy responses versus self-responses.

The structure of the paper is as follows. The next section provides background and discusses relevant studies. Section 3 describes the experiment and the estimation strategy. Section 4 provides a description of the data, while section 5 presents the results. Section 6 concludes.

2. Background and literature

Mincer (1974) summarizes how returns to education can be estimated from a simple wage equation, where the dependent variable is the log of wages.[4] Estimation is typically carried out separately by gender because of differences in labor market opportunities among men and women. The focus is on the approach that

[4] Mincer (1974) shows that, if the only cost of attending school an additional year is the opportunity cost of time and if the proportional increase in earnings caused by this additional schooling is constant over the lifetime, then the log of earnings is linearly related to the individuals' years of schooling and the slope reflects the rate of investment in schooling.

studies the effect of years of schooling (S) on the log of wages ($\ln W$), controlling for experience (E) and its squared term (E^2):

$$lnW_i = \beta_0 + \beta_1 S_i + \beta_2 E_i + \beta_3 E_i^2 + \varepsilon_i \tag{1}$$

The coefficient of years of schooling reflects the average returns to education, which represent the change in wages arising from a change in years of schooling.[5] Peet et al. (2015) provide an overview of returns to education estimates worldwide and over time using the Mincerian approach. They conclude that the average returns are 7.6%, with large variations around the mean. This is largely consistent with earlier results of Psacharopolous and Patrinos (2004), who find that returns are in the neighborhood of 10%. Peet at al. (2015) estimate country-specific returns among men that vary from 4.9% in Guatamala (2000) to 13.0% (2011) and, among women, from 5.3% in Ghana (2005–08) to 14.9% in Niger (2011).

There is a general awareness that these comparisons suffer from a number of limitations. A first constraint is the difference in data coverage. The estimation sample may contain only formal wage workers and exclude casual and informal wage workers. The data may be obtained from firm surveys that focus on a representative subset of firms rather than workers.[6] Other causes of concern relate to consistency in the method of analysis. Coefficients obtained through the dummy variable approach—see footnote 7—need to be adapted before they can be compared with coefficients obtained through the continuous approach presented here. There is also substantial variation in what is included as the other control variables because some models include occupational variables, leading to lower returns, while others do not. A key concern that has received much attention relates to the estimation method. Coefficients obtained using OLS are biased because they do not take endogeneity of schooling into account. Instrumental variable or control function estimation can potentially address this concern (Blundell, Daerden, and Sianesi 2005). Standard estimates also typically assume linearity and neglect potential heterogeneity in returns across educational levels. Yet, investment in education may depend on whether returns are convex or concave, which is a subject of debate, including in Tanzania. While many of these concerns were long neglected in analysis of developing countries, especially African countries, they are now increasingly being taken into account (Schultz, 2004).[7]

---

[5] The alternative method replaces years of schooling with dummy variables for different levels of education. The returns to education are obtained by dividing the estimated coefficient by the corresponding years of schooling for each level of education over and above the level of education below.

[6] Firm surveys often target a specific sector, for instance, the manufacturing sector, or the private sector (see Söderbom et al. 2006 for an example).

[7] Recent work suggests that the Mincerian model and its underlying assumptions do not necessarily hold in all periods on the US, even when addressing these concerns (see Heckman, Lochner, and Todd 2003). Caution is therefore needed

Little attention has been paid to whether differences in survey methods may affect the estimated returns to education. This is especially relevant for low-income countries, where analysis more often relies on data generated by different surveys. This study focuses on two dimensions of survey design: (1) the questions in questionnaires aimed at identifying the labor classification of individuals and (2) whether the questions are answered by the respondents or by a another person (by proxy).

Several studies have investigated various aspects of questionnaire design, including question style and wording (open versus closed questions, positive versus negative statements, and so on) or the specific place of questions within the survey questionnaire (see Kalton and Schuman 1982 for a review). While the general conclusion is that question wording can have important effects, the direction of these effects is frequently unpredictable. Sustained research efforts to revise employment questions in the United States provide interesting insights. Concerned that irregular, unpaid, and marginal activities may be underreported in the Current Population Survey (CPS) because people may not think of their activity as work, respondents were asked in a debriefing study to categorize different hypothetical situations as “work,” “job,” “business,” and so on. While the majority were able to classify their activities consistent with the CPS definitions, large minorities gave incorrect answers in each vignette. For instance, 38 percent of the respondents categorized nonwork activities as work (Campanelli, Rothgeb, and Martin 1989). A 1991 experiment to evaluate the CPS questionnaire revision used direct screening questions and vignettes for unreported work and found that both the sequence of questions and their wording influenced respondent interpretations of work and affected the employment statistics generated (Martin and Polivka 1995). Specifically and of interest in our setting, the study found that using direct screening questions helped in detecting the underreporting of work related to household businesses and farms, as well as the underreporting of teenage work.

This may be especially relevant in developing countries, where informal work, seasonal employment, and various forms of self-employment (versus wage employment) are common, and, so, labor data are perhaps more complex, making consistent reporting difficult. This may be particularly true in the case of women’s work. Several scholars have expressed concern about the underreporting and undervaluing of women’s work in relying on common survey methods to collect employment data, partly because women who work are less likely to work for a wage rather than some form of self-employment or farm work (Anker 1983; Bardasi et al. 2011; Charmes 1998; Dixon-Mueller and Anker 1988; Mata-Greenwood 2000). In previous research using the survey experiment studied here, the lack of screening questions resulted in lower female

both in the interpretation of the results and in the advice to policy makers. Nonetheless, this is still the dominant approach in the analysis of education and related policies, particularly in low-income countries, and it is thus useful to analyze the sensitivity of the model and the underlying assumptions to survey methods.

employment rates, higher working hours among both men and women who were working, and lower rates of wage work (Bardasi et al. 2011).

Household surveys with a labor content can be broadly categorized into two groups: those making use of multiple detailed questions to identify whether an individual is in the labor force during the reference period (typically the last seven days) and those using simply one broad question to determine this information. To illustrate the importance of this survey design feature in recently implemented surveys, Table 1 provides an overview of national household surveys with labor content in sub-Saharan Africa in 2009–12. Of the 21 surveys, 12 use detailed screening questions, while the remaining 9 use short questions. In practice, the difference between these two survey types can be quite minimal: based on the use of one question ("Did [NAME] work?") versus three ("Did [NAME] work for wages or salary?", "Did [NAME] work in any nonfarm family enterprise or business as either self-employed or family labor?", "Did [NAME] work on the family farm tending livestock?"). So, detailed does not necessarily mean many additional questions, but as few as two more.

Another dimension of survey design that may have implications for parameter estimation is the type of respondent. [8] Household surveys in developing countries often ask a single respondent to answer employment questions about all household members. [9] Proxy respondents may not always provide accurate information, and this may bias the obtained statistics on employment and their distribution (Hussmanns, Mehran, and Verma 1990). One alternative is to interview individual household members above a certain age directly as in the Living Standards Measurement Study surveys and labor force surveys. However, requiring self–reporting from all adults adds logistical and financial burdens in the fieldwork, and in practice survey managers often face a trade-off between information accuracy and the cost to obtain it. An experimental study in the United States investigated the potential bias of proxy versus self-response for health statistics. It found that randomly selected respondents reported fewer health events if they were answering about themselves rather than about other household members (Mathiowetz and Groves 1985). An earlier analysis assessed the implications of survey design for descriptive statistics and found important effects of proxy versus self-reporting on resulting labor statistics in Tanzania, observing that reporting by

[8] Ashenfelter and Krueger (1994) collected proxy and self-reported data on education among identical twins (each reporting their own education and that of the twin). They find that the difference between proxy and self-reported years of schooling is small and not statistically significant.

[9] Response by proxy reflects the common practice of interviewing an informed household member (often the household head or spouse), rather than each household member individually. In practice, proxy respondents are often used if individuals are away from the household or otherwise unavailable in the time allotted to conduct interviews in the enumeration area. This is the general approach followed by standard surveys such as household budget surveys (HBS), household income or consumption expenditure surveys (HICES) and core welfare indicator questionnaires (CWIQ), among others. The exact rules on who is allowed to and who did actually report, are often not carefully documented or reported in the data.

proxy leads to lower employment rates among men than direct reporting, mostly because of underreporting on agricultural activities (Bardasi et al. 2011). This bias is reduced if the proxy respondents are spouses or have some schooling. The consequences of these observed differences in categorization for the estimates of the returns to education or other economic relationships have not been studied using these data.

3. Experimental design and estimation strategy

The survey experiment examined here focused on the two key dimensions discussed above: (1) the screening questions to establish employment status and (2) the type of respondent, namely, self-reported versus proxy response. Households were randomly allocated to one of the four survey assignments based on these two dimensions. The labor module covered all household members ages 10 or above and was used to study the impact of the survey design on the estimates of the returns to education.

On the first of the two dimensions, the longer, detailed module contains three questions to determine employment status, namely, (1) whether the person has worked for someone outside the household (as an employee), (2) whether the person has worked on a household farm, and (3) whether the person has worked in a nonfarm household enterprise. In each case, the response was either yes or no. In the short module, there was only one question to determine employment status, namely, whether the person had done any type of work, which also invited a response of yes or no. The short module reflects the approach followed by more concise surveys used in many low-income countries, such as the core welfare indicator questionnaire (CWIQ) and the welfare monitoring surveys (WMS), as well as other surveys listed in Table 1. In both cases, the questions were asked with respect to events during the seven days previous to the survey interview. The individuals identified as having worked in the last 7 days were then asked a set of questions to gather information on their occupation, employment sector, working hours, and, if wage employed, the employer and wage payments in their main job. The exact question wording is reported in Table A.1 in Annex. [10]

In the second dimension (type of respondent), the method was randomly varied between asking questions directly to the respondent or asking a proxy respondent. The proxy respondent was randomly chosen among household members who were at least 15 years. The selected proxy then each reported on up to two other randomly selected household members ages 10 or older. Because, in actual surveys, proxy respondents are not randomly chosen, but selected on the basis of availability, the experiment did not

[10] The short module differed in one other way from the detailed module in that questions about second and third jobs were dropped from the short module. This paper considers only the labor outcomes in the first job.

exactly mimic the actual conditions that result in proxy responses in household surveys. Random selection of the proxy respondent does, however, allow unbiased estimates for proxy response among a representative sample of potential proxies within the household.[11]

The benchmark against which the proxy and short questionnaire treatments are compared is the self-reported and detailed questionnaire, which is generally considered best practice. Hussmanns et al. (1990), who provide International Labour Organization (ILO) guidelines, and Grosh and Glewwe (2000), who provide detailed guidance for household surveys in developing countries, recommend this approach.

The aim of this paper is to investigate whether point estimates of the returns to education vary depending on survey design. It relies on existing methods of analysis and does not necessarily supply new or more accurate estimates of returns to education for Tanzania. This paper and recent work, such as Peet, et al. (2015) acknowledge the limitations of Mincerian returns to education, while maintaining that Mincerian returns offer parameter estimates useful in education policy.[12] Eq. (2) represents a benchmark specification using OLS to estimate the returns to education.

$$lnW_i = \beta_0' + \beta_1' S_i + \beta_2'(S_i * Short_i) + \beta_3'(S_i * Proxy_i) + \beta_4' Short_i + \beta_5' Proxy_i + \beta_6' X_i + \beta_7' D_i + \varepsilon_i' \quad (2)$$

where the dependent variable is the log of daily wages, constructed as weekly earnings, divided by days worked. S_i is years of schooling. $Short_i$ and $Proxy_i$ are the indicator variables for the two survey assignments: short (versus detailed) and proxy (versus self-reported). X_i refers to age and its squared term, and D_i represents district indicator variables.[13] A test is run to determine whether the coefficient of the interaction term of years of schooling with each of the survey assignments (β_2' or β_3') is significantly different from

---

[11] The design of the survey under study was informed by data of the 2006 Tanzanian Core Welfare Indicator Questionnaire (CWIQ) indicating that the average Tanzanian household includes two to three adults ages at least 15 who could serve as proxies. The sample households in the study survey had an average of 2.7 members ages 15 or older. 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 this setting. 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 or accommodation of response to other's expectations, which would introduce potentially much larger (unobserved) respondent biases. Selecting proxies on the basis of availability, rather than randomly, would have potentially introduced selection bias into the estimates.

[12] From an individual labor supply model perspective, Mincerian returns may be considered structural parameters given that they reflect the individual rate of investment in schooling (Mincer 1974). Alternatively, starting from a household model, these returns can be considered reduced form estimates (Card 2001). The former approach may be more relevant for tertiary education, the latter for primary and secondary education.

[13] The results in the subsequent sections are robust to the inclusion or exclusion of the district indicator variables.

zero; rejecting the null provides evidence for the effect either of questionnaire design or of proxy reporting on the estimated returns to schooling.

The literature emphasizes three sources of bias in OLS estimates of returns to education: nonlinearity, endogeneity with respect to schooling, and sample selection, in particular if the focus is on wage workers only.[14] The study assesses whether survey effects are still present when accounting for each of these. To address nonlinearity, a spline function is used that allows returns to vary across levels of education, specifically estimating the following equation:

$$lnW_i = \beta_0'' + \beta_1''S1to7_i + \beta_2''S8to17_i + \beta_3''(S1to7_i * Short_i) + \beta_4''(S1to7_i * Proxy_i) + \beta_5''(S8to17_i * Short_i) + \beta_6''(S8to17_i * Proxy_i) + \beta_7''Short_i + \beta_8''Proxy_i + \beta_9''X_i + \beta_{10}''D_i + \varepsilon'' \quad (3)$$

where *S1to7ᵢ* is the number of years of primary school, and *S8to17ᵢ* is the number of years of postprimary school. The sum of the two values is the total years of school (S_i). Returns are linear if $\beta_1'' = \beta_2''$. A test is then run whether the coefficients of the interaction terms, the years of schooling, and the survey assignments ($\beta_3'', \beta_4'', \beta_5'', \beta_6''$) are significant.[15]

The concern about endogeneity of schooling is often addressed using instrumental variable estimation, while, here, the control function approach is applied as an alternative that is especially attractive in the case of nonlinear returns. A further advantage of control function estimation is that it allows a straightforward inclusion of an interaction term between the endogenous variable and other variables, in this case, the survey assignment indicators.[16] To address endoegneity concerns, Eqs. (2) and (3) are modified by adding a control function term $\widehat{c_i}$ obtained from the first-stage estimation, as follows:

$$S_i = \pi_0 + \pi_1 Z_i + \pi_2 X_{2i} + \pi_3 D_i + c_i \quad (4)$$

[14] For instance, see Blundell, Dearden, and Sienesi (2005); Card (1999); Dickson and Harmon (2011); Glewwe (1996); Heckman, Lochner, and Todd (2003).
[15] No attempt is made to estimate returns to tertiary education specifically because the number of observations is small, particularly among women. However, the results are similar if secondary education is separated from tertiary education. Overall results remain similar when using different cutoff points. For other work using related approaches, see Schady (2003) and Söderbom et al (2006). Note that, because the underlying variable measures the years of schooling within a level of education, the estimated coefficients cannot be interpreted as sheepskin effects (Hungerford and Solon 1987).
[16] In instrumental variable estimation, allowing for interaction between the endogenous variable and the survey assignment dummies requires the three-step procedure presented in Wooldridge (2009), but this is convoluted and less transparent, especially in the presence of two interaction terms.

with the education supply characteristics of the local community as identifying instruments, Z_i, which are correlated with schooling, but not with the unexplained variation in wages ε_i.[17] To isolate the community-specific effects, another community characteristic is also included, namely, distance to an all-weather road; the regression also includes age and age squared (X_{2i}).

Because the detailed and short questionnaires may lead to a different sample of wage workers on which the estimations are carried out, treatment-specific selection is addressed using a correction term. Following the classic Heckman approach, the Inverse Mills ratio is included that is obtained from a first-stage equation that models selection into wage work, λ_i, where this selection term is obtained as follows:

$$P_i = \theta_0 + \theta_1 S1to7_i + \theta_2 S8to17_i + \theta_3 Short_i + \theta_4 Proxy_i + \theta_5 X_{2i} + \theta_6 Z2_i + \theta_7 D_i + v_{2i} \qquad (5)$$

in which P_i indicates whether the individual is a wage worker or not, and the variables contained in the selection equation, but not in the wage equation include marital status and the number of dependents in the household $(Z2_i)$.

While in theory the model is fully identified when using the same variables in the first and second stage, identification then relies entirely on functional form, that is, the nonlinearity of the selection equation. Common practice is followed by including at least one additional identifying variable in the selection equation. While good instruments are often difficult to find, there are valuable candidates, and existing practice is followed.[18] The early literature on labor supply encompasses family formation variables such as marital status and number of children in the selection equation, but not in the wage equation. Later work on the United States and a few other high-income countries shows that marital status can affect earnings among men and, in some cases, among women. Such evidence is, however, absent for developing countries, which offer a very different context. Indeed, marital status is said to affect earnings through two specific channels: specialization and selection (Korenman and Neumark 1998). Regarding the former, the argument is that marriage can increase specialization, especially a husband's, leading to higher productivity and earnings. The selection argument states that more productive workers, who thus also have higher earnings, are more likely to find a partner and become married in the first place. Both of these seem to be largely absent in rural and provincial developing settings, such as the one in the sample here, in which gender roles

[17] Card (2001) discusses other work using variations in the supply of education to identify causal effects.
[18] The best candidate instruments have been investigated in this context, including education reforms in Tanzania. One such reform was implemented in the late 1960s and changed the structure of the education system (Kerr 2011). Another, more recent policy change relates to the introduction of universal primary education (Hoogeveen and Rossi 2011). Both these changes are disqualified as good instruments for the aim here because the former is too old, while the latter is too recent. No other major changes in the organization of education were identified for Tanzania.

are strong, and marriage is relatively early and universal.[19] Although one may not entirely exclude that unobserved characteristics set at a young age drive both marital status and productivity. Evidence on an effect of fertility on male and female earnings is also scarce in developing countries. Piras and Ripani (2005), in a comparison of four countries in Latin America, find little evidence that mothers earn lower wages than women with no children. McCabe and Rosenzweig (1976) consider the possible effects of fertility on wages and argue explicitly that this depends on the compatibility of the specific occupation with child-rearing. The focus in this case is on wage work, which is incompatible with child-rearing, and the number of children is expected to affect the selection into wage work, but not on-the-job productivity.

In contrast, there is strong evidence that family formation, including marriage and fertility, has important effects on time use if markets for household chores and childcare are missing. Being married increases the time devoted to housework, particular among women, while the presence of children, especially small children, increases the time devoted to care among both men and women (World Bank 2011). In this context, family formation affects primarily the probability of being in wage work, that is, working outside the informal sector or performing household chores, and this is confirmed by the data (see section 5). The identifying variables here are marital status and the number of children ages 15 and under in the household. A placebo test that includes these family formation variables in the second-stage equation confirms that these have no effect on wages in this setting. [20]

A challenge with the above approach arises if the education variables determine occupational sorting, as indicated by their statistical significance in the selection equation. This may lead to biased estimation results because the selection correction term is now correlated with the error term in the second stage. To address this, a procedure is followed that is similar to the one applied by Duflo (2001) and originally suggested by Heckman and Hotz (1989). The procedure consists of including the instrumental variables that are used to address the endogeneity of education—Z, that is, the community distance variables used in the control function—in the selection equation and including polynomials of the predicted probabilities of becoming wage work ($\hat{p}$) in the main equation whereby the selection correction term $\widehat{p_i}$ is now obtained from:

$$P_i = \theta_0' + \theta_1' S1to7_i + \theta_2' S8to17_i + \theta_3' Short_i + \theta_4' Proxy_i + \theta_5' X_{2i} + \theta_6' Z2_i + \theta_7' Z_i + \theta_8' D_i + v_{2i} \quad (6)$$

[19] The specialization argument maintains that married men have more time to specialize in professional activities, but it is unclear whether this is the case in rural societies, where unmarried men (and women) often stay with their parents and, hence, do not have to carry out the household chores associated with an independent household. Similarly, marriage is often decided before labor market performance has been revealed, breaking the reverse causality that is a major concern in high-income countries.

[20] All results are similar if younger children (under age 6) and older children (ages 6–15) are considered separately. For reasons of consistency, the same variables are included in the selection equations for men and women.

in which P_i reflects whether the individual is a wage worker or not, and the $Z2$ variables contained in the selection equation, but not in the wage equation include marital status and the number of dependents in the household, while Z reflects the community distance variables. As before, to isolate the community effects, mean distance to an all-weather road is also included.

4. Data and context

The survey experiment was implemented in Tanzania, which has various types of labor market surveys, including Core Welfare Indicator Questionnaires (CWIQ), Labor Force Surveys (LFS) and multipurpose household surveys, such as the household budget survey (HBS). These different data sources have been variably used to estimate returns to education. The experiment that is the subject of this paper is the Survey of Household Welfare and Labor in Tanzania (SHWALITA). The fieldwork was conducted from September 2007 to August 2008 in villages and urban areas in seven districts across Tanzania: one district each in the regions of Dar es Salaam, Dodoma, Manyara, Pwani, and Shinyanga and two districts in the region of Kagera. Households were randomly drawn from the listing of communities and randomly allocated to one of the four survey assignments, that is, the two-by-two survey assignments described above: short or detailed and self-reported or proxy. The sample was 1,344 households, and 336 households were assigned to each of the four survey assignments.[21] Although the sample was not designed to be nationally representative of Tanzania, the districts were selected to capture variations between urban and rural areas and along other socioeconomic dimensions.

The basic characteristics of the sampled households generally match the nationally representative data from the 2006/07 Household Budget Survey (results not presented here). Household interviews were conducted over a 12-month period, but, because of small samples, the survey assignment effects are not explored across seasons, such as harvest time (with peak labor demand) and the dry season (with low demand). The random assignment of households is confirmed when examining different household characteristics, as reported in Table 2, panel a.

Individuals are classified on the basis of the survey assignment they actually received, which is the result of the initial assignment of their households to one of the four survey assignments, whether the individual is selected to be a proxy respondent or a self-reporting respondent and whether the proxy or self-reporting assignment is realized. In the case of the self-reporting modules, up to two individuals ages over 10 are

[21] Annex Table A.2 provides details of the planned and actual household and individual assignments. The deviation arises from the availability of proxy respondents who needed to be at least 15 years old.

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 the 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. The number of self-reports should therefore be about half the number of proxy reports among households in the proxy assignment. So, in total and by design, there are altogether more self-reports than proxy reports. Because the survey team emphasized the importance of avoiding proxies outside the proxy group treatments, the project was successful at completing self-reports in cases when these were so assigned. In only about 5% of cases was the team unable to interview a person selected to self-report. While there were small deviations from the original design during implementation, the overall realized survey remained close to the planned design (see annex table A.2). Table 2, panel b, reports balance tests across individuals. It shows that allocation across survey assignments is generally well balanced. Mean differences between the proxy and self-reporting assignments exist across age, marital status, and number of children.

The identifying variables for the control function estimation are obtained from Core Welfare Indicator Questionnaire (CWIQ) data collected in the same communities during the same year, but among a separate sample of households. The community mean distance to the closest primary and secondary schools is used as a proxy the local supply of education at the time of the schooling of individuals.

To consider the differences in wages across survey assignments, the plots in Fig. 1 present the kernel density of the log of daily wages for the various subsamples. Fig. 1, panels a and c illustrate the wages obtained from the detailed and short modules among men and women, respectively, with the wage distribution among women mostly to the left of the wage distribution among men. Fig. 1, panels b and d suggest that the gender wage difference is smaller among the proxy and self-reporting groups relative to the detailed and short survey assignments.

The descriptive statistics in Table 3 reveal further differences. While the detailed module yields a lower proportion of labor force participants relative to the short module for both men and women, it generates a higher relative share of wage workers among both sexes. The detailed module also produces lower wages than the short module. Proxy response leads to relatively lower labor force participation and a lower share of wage workers, but produces higher wages compared with self-responses among both sexes. Although the differences in wages between treatments may seem small, they are actually substantial, which is perhaps more clear if they are expressed as monthly wages rather than daily wages. The average daily wage reported in Tanzanian shilling (T Sh) for the detailed versus short survey and for self-reporting versus proxy reporting among men correspond to monthly incomes expressed in U.S. dollars of, respectively, $71,

$83, $65, and $106, while the respective daily wages in Tanzanian shillings among women correspond to monthly wages of $68, $77, $62, and $93.[22] Because the wage equations are estimated for wage workers only, the corresponding estimation sample is much reduced, as is standard in this literature. Wage workers represent 16% and 8% of men and women in the labor force, respectively, while the respective labor force participation rates among men and women are 88% and 84%.

5. Analysis and results

As a benchmark, Eq. (2) is estimated using OLS, which assumes linearity and does not account for endogeneity. The results are reported in Table 4 and suggest that the average returns to education in the sample are 8% among men and 10% among women (columns 1 and 4). These results persist if survey assignment variables are included (columns 2 and 5). Including interaction effects, the results in column 3 indicate that, among men, the short module yields substantially higher returns to education, while the proxy responses do not affect the returns to education. Among women, the interaction effects in either the short or the proxy treatment are insignificant (column 6). [23]

Existing evidence indicates that the returns to education are often nonlinear, particularly in Tanzania (Kerr, 2011; Söderbom, Teal, Wambugu & Kahyarara, 2006). This is confirmed here by estimates of the models that allow for nonlinearity and endogeneity as shown by the plot of a nonparametric kernel regression in Fig. 2, which suggests an S-shaped pattern among men and convexity among women.

A control function term obtained through Eq. (4) is included in the estimation of Eq. (3) to acquire nonlinear returns while also allowing for endogeneity. The results, reported in Table 5 show nonlinearities among women and, possibly, among men. The returns tend to increase with the level of education, but the increase occurs at a declining rate. [24]

Columns 3 and 6 in Table 5 demonstrate that the interaction effect between the short survey assignment and the level of education occurs beyond primary school among men and in primary education among

[22] The respective wages in shillings are T Sh 3951, T Sh 4610, T Sh 3608, and T Sh 5879 among men and T Sh 3798, T Sh 4255, T Sh 3455, and T Sh 5182 among women.
[23] A simple robustness test in which separate regressions are estimated, including each of the two interactions one at a time, gives the same results, as expected given the orthogonality of the survey assignment. An overall alternative approach would be to carry out separate estimation per survey assignment group. This yields the same results. Because of the small sample size, the focus is on pooled results.
[24] The returns are not significantly different across levels of education among men, although this may stem from the small sample size. Overview Table 9 suggests consistently higher returns at greater educational attainment, no matter the estimation method.

women. Additional testing shows that the interaction effects with the short survey assignments are jointly significant among women, but not among men, while the interaction effects with proxy assignments are not jointly significant. The size of the (gender-specific) control function term is substantial, and the inclusion of this term affects the estimation results. As expected, control function estimates of the returns to education are larger than OLS estimates, indicating that the latter are biased because of unobserved characteristics.[25] The last two rows in of Table 5 report an F-test of the joint significance of the interaction term of years of schooling with the respective treatments. They indicate that interaction with the short questionnaire is jointly significant at the 5% level among women, but not among men (p = 0.18). The latter may result from the weaker nonlinearity among men, given that the interaction effect in Table 4 is significant.[26]

Table 6 reports the first-stage estimates for the control function and demonstrates the importance of the identifying variables. As expected, the community mean distance to the nearest secondary school is especially important among men, while the community mean distance to the nearest primary school is of particular relevance among women.[27] To avoid that these variables proxy for community fixed effects, another community characteristic reflecting general isolation is also included, namely, distance to the nearest all-weather road. This variable is then also included in the second-stage regression.

An inspection is carried out to determine whether the instruments perform in the first stage. This is accomplished by using an F-test for joint significance, which is especially relevant in the case of one endogenous regressor (Sanderson and Windmeijer 2016). This yields a test statistic of 7.11 for women and 2.84 for men. Because the test statistics are below 10, this suggests that the instruments are weak (Stock and Yogo 2005). This cannot be rejected by further tests.[28] This is not uncommon in the estimation of returns to education using survey data. In a placebo test whereby one adds the identifying variables to the

[25] The key findings would still be supported even if the OLS estimates were closer to the true effects, as argued in the case of instrumental variable estimation (Boef et al. 2014). There appears to be no corresponding research on this issue for control function estimation. The results also persist if OLS is used for nonlinear estimates (not reported).
[26] As shown in overview Table 9, the pattern and size of the coefficient are confirmed if other estimation methods are used.
[27] Basmann (1960) and Sargan's (1958) tests can reject overidentification among men (p-values of 0.57 and 0.56, respectively), but not among women (p-values of 0.01 in each case). In light of these results, the number of instruments among women was reduced to only the distance to the nearest primary school; the symmetry in the models across gender was thus sacrificed. However, the results are similar throughout if a model is used with both instruments among women.
[28] A conditional likelihood ratio test and a Lagrange multiplier K test are used for men because the number of instruments exceeds the number of endogenous variables. They yield p-values of 0.09 and 0.06, respectively, while, for women, where the model is exactly identified, an Anderson-Rubin test yields a p-value of 0.02.

second-stage regression, but without the control function term, the parameter estimates are not significant.[29]

The next step is to estimate the returns to education with a correction for selection into wage work, that is, adding the selection correction term obtained from estimating Eq. (5). Table 8 presents the results for the selection equation. Two issues stand out. First, education has a significant effect on the probability of being a wage worker for both men and women.[30] This raises an additional challenge and leads to a comparison with an alternative approach (see Section 3). Second, the coefficients of the treatments indicate that using a detailed or short questionnaire has strong effects on who is categorized as a wage worker among both men and women, while the effects are less strong and less robust in the case of proxy versus self-response reporting. The effect of the short questionnaire reflects the key difference of this questionnaire from the detailed module, which includes additional screening questions at the beginning of the questionnaire. Omitting these questions seems to lead to a different categorization of respondents into wage work, and, as a result, the wage equation is estimated on a different sample. The descriptive statistics in Table 3 already indicated that both labor force participation and, especially, the shares of wage workers differ substantially between the detailed and short questionnaires.

The number of children is significant: the p-value is 0.02 for men. Being married has a p-value of 0.11 for women. The instruments are jointly significant at 0.2 for men and 0.06 for women. A placebo test, whereby the family formation variables are included in the wage equation, confirm that these variables have no significant effect on wages. Columns 2 and 4 in Table 8 present the estimates of Eq. (6), which includes both sets of instruments, namely, Z1, the family formation variables, and Z, the community school variables, used in the earlier part of the analysis.

Table 7 presents the second-stage estimation results. Columns 1 and 3 report the estimates from the classic Heckman approach and relying on the first-stage results presented in Table 8 columns 1 and 3. Columns 2 and 4 report the estimates from the alternative Heckman-Hotz approach that relies on the first-stage presented in Table 8, columns 2 and 4. Both sets of results confirm that the returns to education among

---

[29] In the case of instrumental variable estimation, standard errors may be too small if weak instruments are used, and similar concerns may arise in control function estimation, although it appears there is no research establishing this. To verify the robustness of the results to weak instruments, an instrumental variable estimation is carried out of the model in Table 4. For the linear case, this yields point estimates similar to the control function estimation, but with different standard errors. However, the focus is on the linear coefficient of years of schooling because of the limitations of instrumental variable estimation in dealing with nonlinearity and interaction effects with endogenous variables. Confidence intervals are estimated that are robust to weak instruments, leading to estimates of 0.20 [0.006; 1.306] for men and 0.24 [0.016; 1.459], for women. These estimates are close to the equivalent estimates in Table 5, although the latter are for the nonlinear case, and suggest that the results are robust to weak instruments.

[30] To maintain consistency between first- and second-stage equations, nonlinearity is allowed. Descriptive statistics and kernel density regression also suggest nonlinear education effects on the selection into wage work (not reported).

men are higher among those beyond primary education if the short questionnaire is used, while the returns among women are higher for primary-educated if the short module is used. As before, an F-test for joint significance of the interaction terms between years of schooling and the short survey assignment indicates significance among women at the 5% level, but not among men. A similar test on the joint significance of interaction with the proxy survey assignment has low significance.

Table 9 summarizes the estimates on the returns to education across estimation methods, focusing on the differences between the detailed questionnaire and the short questionnaire because the difference with the proxy assignment is not significant. Despite the small sample sizes, the results are consistent across estimation methods.[31] Using the detailed questionnaire with self-response as the best practice reference, the results indicate that the short questionnaire systematically overestimates the returns among men who have attained beyond primary schooling and women who have attained primary education. The pattern is consistent across different estimation methods. The preferred models, which also account for selection, confirm this bias. There, the estimated returns among men who have secondary or higher education and among women who have primary education are 5 and 16 percentage points higher, respectively, if the short module is used. These differences arise because fewer men and women at these education levels—who represent the majority of men and women wage workers, respectively—are categorized as wage workers in the short questionnaire. In light of the existing debate concerning the convexity of returns to education in developing countries and, especially, in Tanzania, (Söderbom et al. 2006) these results suggest that the convexity may have been overestimated in studies using data obtained from short questionnaires.

These results make the general point that questionnaire design can have both significant and substantial effects on the estimation of structural or reduced form parameters. That these effects are found despite the small sample size of the data provides strong evidence that survey design matters, particularly in estimating the returns to education among men and women across different levels of schooling. The findings highlight the need for caution in comparing estimates obtained from data generated through diverse survey methods. At a practical level, they also underline the importance of consistency and best practice in survey design.[32]

---

[31] As an additional check on robustness, the results are compared with those obtained from median regressions. They are similar. The small sample size prohibits estimation for other quantiles. Lee (2007) shows that quantile estimation can follow the usual control function approach to account for endogeneity for linear estimation if models are additive in observed covariates and unobservables. Arellano and Bonhomme (2017) present a similar argument for including a selection correction term when using control function estimation.

[32] In principle, it would be interesting to compare biases in estimates between rural and urban areas. However, because the returns to education are estimated on the sample of wage workers only, sample size is already limited, and splitting this sample further into rural and urban subsamples results in smaller sample sizes, especially among women. For example, if one re-estimates Table 4 for rural and urban areas, the interaction effects among men are similar in magnitude in urban and rural areas, although they lose significance because of higher variance and smaller

6. Conclusion

This paper investigates whether measurement error because of variation in survey design matters in estimating returns to education using a survey experiment. Households in Tanzania were randomly assigned variations in survey design: a commonly used short module versus a detailed labor module and self-response compared with response by proxy. The experiment shows that, among both men and women, the estimated returns to education differ according to the survey instrument (short or detailed screening questions for employment), but not according to the type of respondent. The short questionnaire leads to biased estimates of the returns to education relative to the detailed questionnaire. The biases are substantial and significant, resulting in higher estimates in the short questionnaire, typically ranging from 5 percentage points higher among men educated beyond primary, school to 16 percentage points higher among primary-educated women. These results are robust to accounting for nonlinearity in education, the endogeneity of education, and selection into wage work by making use of commonly applied estimation and identification methods. The results are consistent with suggestive evidence from qualitative research, including respondent debriefing studies in the United States showing that screening questions can have important effects on the labor statistics they generate.

These observed differences are similar in magnitude to the estimation bias related to endogeneity that is the subject of considerable attention in the literature on returns to education. Estimation bias because of survey design is also similar in magnitude to the differences in estimated returns between sexes and across levels of education and public, formal private, and informal private sectors observed in the region. This therefore deserves attention.[33]

This paper does not aspire to obtain more accurate estimates of the returns to education in Tanzania. The data were not collected for this purpose, and the sample is also small and not nationally representative. Nonetheless, it is useful to consider the existing results on Tanzania particularly because they rely on data from different surveys. Nerman and Owens (2010), using the 2001 and 2007 waves of the nationally representative household budget surveys, estimate returns to education to explain the demand for education and report OLS estimates of 0.3% to 16.3%, depending on the subgroup, and not controlling for endogeneity. Using nationally representative cross-sectional data from the 2001 and 2006 Integrated

sample size. Among women, the point estimate is higher in rural areas relative to urban areas, but the small sample sizes make clear conclusions difficult.

[33] A review of key overview papers on the region indicates differences of between 2 and 18 percentage points across these dimensions (see Kuepie, Nordman, and Roubaud 2009; Schultz 2004; Teal and Baptist 2014).

Labor Force Surveys (IFLS), Kerr (2011) estimates returns at between 8% and 13% by using OLS. Peet et al. (2015) estimate the returns to education in 2004–10 at 11.0% and varying between 9.2% among men and 14.6% among women. Results that allow for nonlinearities suggest that returns are strongly convex, but, if endogeneity is also addressed by exploiting a change in the education system in the mid-1960s, the returns are concave and higher at the lower levels of education, which the authors argue reflects an ability bias.[34] These results also shed light on earlier findings of Söderbom et al. (2006), who, using data for employees in the manufacturing sector in Tanzania in 1993, 1994, 1999, and 2001, also find a convex earnings function after taking endogeneity into account. Their estimates exceed the OLS estimates, which may be a consequence of self-selection on ability into the manufacturing sector. The estimates here also align well with those on other African countries. Schultz (2004) reports wage gains of 5–20% for each year of schooling in five African countries, while Peet et al. (2015) find overall returns between 3% and 12% in nine African countries.[35]

The results described in this paper demonstrate that measurement error related to survey methods biases estimation of parameters, such as Mincerian returns to education, and indicate that care is needed in the comparison of these returns both across countries and within countries if surveys are designed differently.

7. References

Angrist, Joshua D., and Alan B. Krueger. 1991. "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics* 106 (4): 979–1014.

Anker, Richard Bruce. 1983. "Female Labour Force Participation in Developing Countries: A Critique of Current Definitions and Data Collection Methods." *International Labour Review* 122 (6): 709–24.

Arellano, Manuel and Stéphane Bonhomme. 2017. "Sample Selection in Quantile Regression: A Survey." CEMFI Working Paper 1702 (January), Centro de Estudios Monetarios y Financieros, Madrid.

Ashenfelter, Orley, and Alan B. Krueger. 1994. "Estimates of the Economic Returns to Schooling from a New Sample of Twins." *American Economic Review* 84 (5): 1157–73.

Bardasi, Elena, Kathleen Beegle, Andrew Dillon, and Pieter Serneels. 2011. "Do Labor Statistics Depend on How and to Whom the Questions Are Asked? Results from a Survey Experiment in Tanzania." *World Bank Economic Review* 25 (3): 418–47.

Basmann, R. L. 1960. "On Finite Sample Distributions of Generalized Classical Linear Identifiability Test Statistics." *Journal of the American Statistical Association* 55 (292): 650–59.

Behrman, Jere R., Mark Rosenzweig, and Paul Taubman. 1994. "Endowments and the Allocation of

[34] While the instrument reflects an exogenous policy change that extended primary school from four to seven years in the mid-1960s, the extent to which this is applicable to the entire sample of analysis is unclear. Given that the policy change took place 40 years ago, it had an immediate impact on those who were ages 10–13 in the mid-1960s or 50–53 in the mid-2000s. This is likely to be a minority in the IFLS sample, in which the average age was 38.

[35] Schultz (2004) reports estimates on Burkina Faso, Côte d'Ivoire, Ghana, Kenya, and Nigeria; Peet, Fink, and Fawzi (2015) report estimates on Côte d'Ivoire, Ethiopia, Ghana, Malawi. Niger, Nigeria, South Africa, Tanzania, and Uganda.

Schooling in the Family and in the Marriage Market: The Twins Experiment." *Journal of Political Economy* 102 (6): 1134–74.

Behrman, Jere R., Mark Rosenzweig, and Paul Taubman .1996. "College Choice and Wages: Estimates Using Data on Female Twins." *Review of Economics and Statistics* 73 (4): 672–85.

Blundell, Richard, Lorraine Dearden, and Barbara Sianesi. 2005. "Evaluating the Effect of Education on Earnings: Models, Methods, and Results from the National Child Development Survey." *Journal of Royal Statistical Society Series A* 168 (3): 473–512.

Boef, Anna G. C., Judith van Paassen, M. Sesmu Arbous, Arnoa Middelkoop, Jan P. Vandenbroucke, Saskiaa le Cessie, and Olaf M. Dekkers. 2014. "Physicians Preference-Based Instrumental Variable Analysis: Is It Valid and Useful in a Moderate-Sized Study?" *Epidemiology* 25 (6): 923–27.

Bound, John, Charles Brown, and Nancy Mathiowetz. 2001. "Measurement Error in Survey Data." In *Handbook of Econometrics*, vol. 5, edited by James J. Heckman and Edward Leamer, 3705–3843. Handbooks in Economics Series 2. Amsterdam: North-Holland, Elsevier Science.

Bound, John, and Gary Solon. 1999. "Double Trouble: On the Value of Twins-Based Estimation of the Return to Schooling." *Economics of Education Review* 18 (2): 169–82.

Campanelli, Pamela C., Jennifer M. Rothgeb, and Elizabeth A. Martin. 1989. "The Role of Respondent Comprehension and Interviewer Knowledge in CPS Labor Force Classification." In *Proceedings of the Survey Research Methods Section*, 425–30. Alexandria, VA: American Statistical Association.

Card, David. 1999. "The Causal Effect of Education on Earnings." In *Handbook of Labor Economics*, vol. 3A, edited by Orley C. Ashenfelter and David Card, 1801–63. Handbooks in Economics Series 5. Amsterdam: North-Holland, Elsevier Science.

Card, David. 2001. "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems." *Econometrica* 69 (5): 1127–60.

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

Cruz, Luiz M., and Marcelo J. Moreira. 2005. "On the Validity of Econometric Techniques with Weak Instruments: Inference on Returns to Education Using Compulsory School Attendance Laws." *Journal of Human Resources* 40 (2): 393–410.

de Mel, Suresh, David J. McKenzie, and Christopher Woodruff. 2009. "Measuring Microenterprise Profits: Must We Ask How the Sausage Is Made?" *Journal of Development Economics* 88 (1): 19–31.

Dickson, Matt, and Colm Harmon. 2011. "Economic Returns to Education: What We Know, What We Don't Know, and Where We Are Going, Some Brief Pointers." *Economics of Education Review* 30 (6): 1118–22.

Dixon-Mueller, Ruth, and Richard Bruce Anker. 1988. "Assessing Women's Economic Contributions to Development." Background Papers for Training in Population, Human Resources, and Development Planning Series 6, International Labour Office, Geneva.

Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review* 91 (4): 795–813.

Glewwe, Paul. 1996. "The Relevance of Standard Estimates of Rates of Return to Schooling for Educational Policy: A Critical Assessment." *Journal of Development Economics* 51 (2): 267–90.

Grosh, Margaret E., and Paul Glewwe, eds. 2000. *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Development Study*. 3 vols. Washington, DC: World Bank.

Heckman, James J., and V. Joseph Hotz. 1989. "Choosing among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training." *Journal of the American Statistical Association* 84 (408): 862–74.

Heckman, James J., Lance J. Lochner, and Petra E. Todd. 2003. "Fifty Years of Mincer Earnings Regressions."

IZA Discussion Paper 775, Institute for the Study of Labor, Bonn, Germany.

Hoogeveen, Johannes, and Mariacristina Rossi. 2011. "Saving Goat and Cabbages? Enrollment and Grade Achievement after the Introduction of Free Primary Education in Tanzania." Paper presented at the Centre for the Study of African Economies's 25th Anniversary Conference, "Economic Development in Africa," St. Catherine's College, Oxford University, Oxford, March 20–22.

Hungerford, Thomas, and Gary Solon. 1987. "Sheepskin Effects in the Returns to Education." *Review of Economics and Statistics* 69 (1): 175–77.

Hussmanns, Ralf, Farhad Mehran, and Vijay Verma. 1990. *Surveys of Economically Active Population, Employment, Unemployment and Underemployment: An ILO Manual on Concepts and Methods*. Geneva: International Labour Office.

Kalton, Graham, and Howard Schuman. 1982. "The Effect of the Question on Survey Responses: A Review." *Journal of the Royal Statistical Society* 145 (1): 42–57.

Kerr, Andrew. 2011. "Estimating the Returns to Education in Tanzania." Chapter 2, PhD dissertation, Economics Department, Oxford University, Oxford, UK.

Korenman, Sanders D., and David Neumark. 1998. "Marriage, Motherhood, and Wages." In *Women in the Labor Market*, vol. 2, edited by Marianne A. Ferber, 192–214. International Library of Critical Writings in Economics 90. Northampton, MA: Edward Elgar Publishing.

Kuepie, Mathias, Christophe J. Nordman, and François Roubaud. 2009. "Education and Earnings in Urban West Africa." *Journal of Comparative Economics* 37 (3): 491–515.

Lee, Sokbae. 2007. "Endogeneity in Quantile Regression Models: A Control Function Approach." *Journal of Econometrics* 141 (2): 1131–58.

Martin, Elizabeth A., and Anne E. Polivka. 1995. "Diagnostics for Redesigning Survey Questionnaires: Measuring Work in the Current Population Survey." *Public Opinion Quarterly* 59 (4): 547–67.

Mata-Greenwood, Adriana. 2000. "Incorporating Gender Issues in Labour Statistics." STAT Working Paper, Bureau of Statistics, International Labour Office, Geneva.

Mathiowetz, Nancy A., and Robert M. Groves. 1985. "The Effects of Respondent Rules on Health Survey Reports." *American Journal of Public Health* 75 (6): 639–44.

McCabe, James L., and Mark R. Rosenzweig, 1976. "Female Labor-Force Participation, Occupational Choice, and Fertility in Developing Countries." *Journal of Development Economics* 3 (2): 141–60.

Miller, Paul W., Charles Mulvey, and Nick Martin. 1995. "What Do Twins Studies Reveal about the Economic Returns to Education? A Comparison of Australian and U.S. Findings." *American Economic Review* 85 (3): 586–99.

Mincer, Jacob. 1958. "Investment in Human Capital and Personal Income Distribution." *Journal of Political Economy* 66 (4): 281–302.

Mincer, Jacob. 1974. *Schooling, Experience, and Earnings*. Vol. 2 of *Human Behavior and Social Institutions*. Cambridge, MA: National Bureau of Economic Research; New York: Columbia University Press.

Nerman, Måns, and Trudy Owens. 2010. "Determinants of Demand for Education in Tanzania: Costs, Returns, and Preferences." Working Papers in Economics 472, University of Gothenburg, Göteborg, Sweden.

Peet, Evan D., Günther Fink, and Wafaie W. Fawzi. 2015. "Returns to Education in Developing Countries: Evidence from the Living Standards and Measurement Study Surveys." *Economics of Education Review* 49 (December): 69–90.

Piras, Claudia, and Laura Ripani. 2005. "The Effects of Motherhood on Wages and Labor Force Participation: Evidence from Bolivia, Brazil, Ecuador, and Peru." Sustainable Development Department Technical Paper WID-109, Inter-American Development Bank, Washington, DC.

Psacharopoulos, George, and Harry Anthony Patrinos. 2004. “Returns to Investment in Education: A Further Update.” *Education Economics* 12 (2): 111–34.

Sanderson, Eleanor, and Frank Windmeijer. 2016. "A Weak Instrument F-Test in Linear IV Models with Multiple Endogenous Variables.” *Journal of Econometrics* 190 (2): 212–21.

Sargan, J. D. 1958. “The Estimation of Economic Relationships Using Instrumental Variables.” *Econometrica* 26 (3): 393–415.

Schady, Norbert R. 2003. “Convexity and Sheepskin Effects in the Human Capital Earnings Function: Recent Evidence for Filipino Men.” *Oxford Bulletin of Economics and Statistics* 65 (2): 171–96.

Schultz, T. Paul. 2004. “Evidence of Returns to Schooling in Africa from Household Surveys: Monitoring and Restructuring the Market for Education.” *Journal of African Economies* 13 (Suppl.2): ii95–ii148.

Söderbom, Måns, Francis Teal, Anthony Wambugu, and Godius Kahyarara. 2006. “The Dynamics of Returns to Education in Kenyan and Tanzanian Manufacturing.” *Oxford Bulletin of Economic and Statistics* 68 (3): 261–88.

Stock, James H., and Motohiro Yogo. 2005. “Testing for Weak Instruments in Linear IV Regression.” In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, edited by Donald W. K. Andrews and James H. Stock, 80–108. New York: Cambridge University Press.

Teal, Francis, and Simon Baptist. 2014. “Technology and Productivity in African Manufacturing Firms.” *World Development* 64: 713–25.

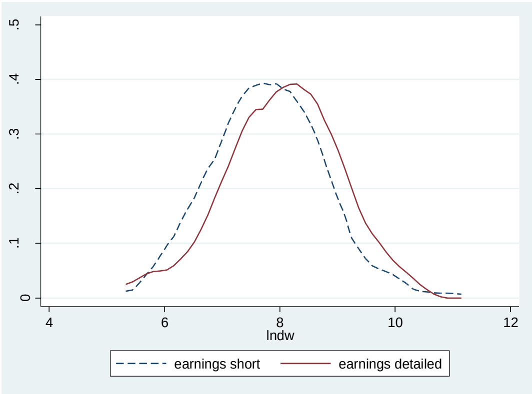
Wooldridge, Jeffrey M. 2009. “On Estimating Firm-Level Production Functions Using Proxy Variables to Control for Unobservables.” *Economics Letters* 104 (3): 112–14.

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

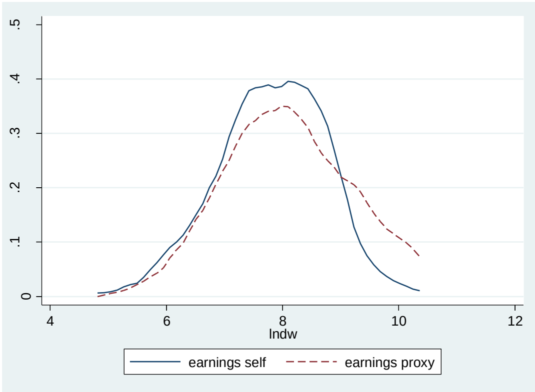
8. Figures and Tables

Figure 1. Kernel density of earnings by survey assignment

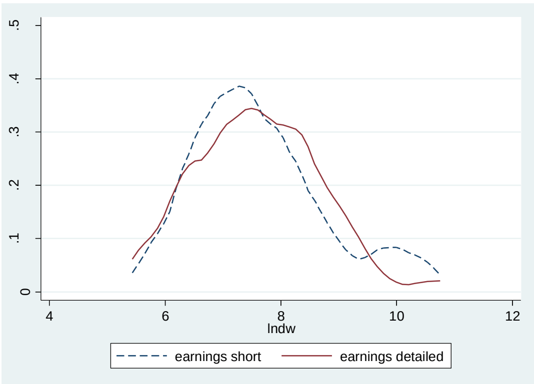
a. Men, detailed versus short module



b. Men, self-reporting versus proxy module



c. Women, detailed versus short module



d. Women, self-reporting versus proxy module

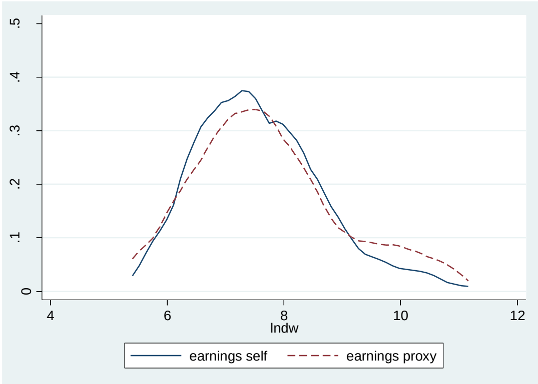
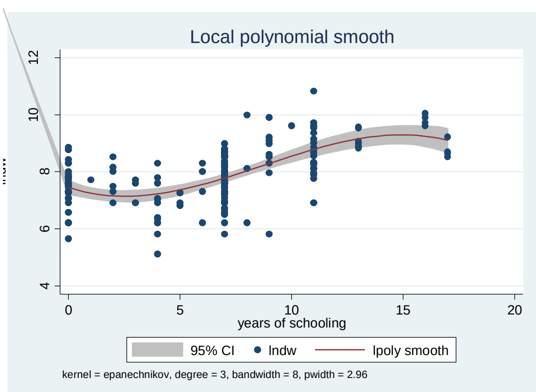


Figure 2. Kernel Regression Plot, Wage Equation

a. Men



b. Women

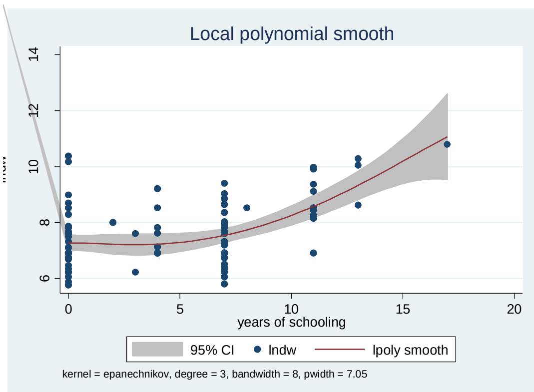


Table 1. Overview of Types of Recent Surveys, Sub-Saharan Africa, 2009-12

Country	Survey name	Abbreviation	Year	Type of Survey
Botswana	Botswana Core Welfare Indicators Survey	BCWIS	2009	detailed
Cameroon	Second Survey on Employment and the Informal Sector	EEIS2	2010	detailed
Malawi	Integrated Household Survey	IHS	2010	detailed
Rwanda	Enquête Intégrale sur les Conditions de Vie des Ménages	EICV	2010	detailed
Uganda	Uganda National Panel Survey	UNPS	2010	detailed
Zambia	Living Conditions Monitoring Survey	LCMS	2010	detailed
Niger	Living Standards Survey	LSS	2011	detailed
Sierra leone	Sierra Leone Integrated Household Survey	SLIHS	2011	detailed
Tanzania	Household Budget Survey	HBS	2011	detailed
South Africa	General Household Survey	GHS	2011	detailed
Mauritius	Continuous Multi Purpose Household Survey	CMPHS	2012	detailed
Nigeria	General Household Survey	GHS_2	2012	detailed
Swaziland	Household Income and Expenditure Survey	HIES	2009	short
The Gambia	Integrated Household Survey	IHS	2010	short
Lesotho	Household Budget Survey	HBS	2010	short
Madagascar	Enquête Intégrale sur les Conditions de Vie des Ménages	EICVM	2010	short
Sao Tome & Principe	Inquérito aos orçamentos familiars	IOF	2010	short
Senegal	Enquête de Suivi de la Pauvreté au Sénégal	ESPS	2011	short
Togo	Core Welfare Indicators Questionnaire	QUIBB	2011	short
Ethiopia	National Labour Force Survey	UEUS	2012	short
Ghana	Living Standards Survey	LSS	2012	short

Table 2. Balance Tests

a. Household characteristics, by survey assignment of household

Household characteristics	*Households by survey assignment*				*F-test of equality of coefficients across groups* [a]
	Detailed Self	*Detailed Proxy*	*Short Self*	*Short Proxy*	
Head: years of schooling	4.70	4.83	4.68	4.63	0.92
Head: age	47.7	45.8	45.8	46.48	0.40
Head: female	0.19	0.19	0.20	0.22	0.84
Head: married	0.75	0.71	0.74	0.72	0.63
Household size	5.28	4.87	4.99	5.38	0.07
Number of members ages 15+	2.66	2.53	2.56	2.71	0.30
Number of households	336	336	336	336	

b. Individual characteristics, by survey assignment of household

	Household by survey assignment				*F-test for equality of coefficients across groups* [a]	
Household characteristics	*Detailed Self*	*Detailed proxy*	*Short self*	*Short Proxy*	*across all assignments*	*between detailed and short*
Years of schooling	4.6	4.3	4.5	4.3	0.26	0.77
Age	33.9	28.9	34.4	29.4	0.00	0.43
Male	0.54	0.53	0.50	0.53	0.21	0.09
Married	0.55	0.48	0.57	0.45	0.00	0.86
Number of children under age 6	1.1	1.1	1.0	1.2	0.05	0.86
Number of household members ages 65+	0.2	0.2	0.3	0.3	0.68	0.31
Mean community distance to primary school	2.3	2.3	2.3	2.3	0.99	0.87
Mean community distance to secondary school	7.5	7.7	7.5	7.7	0.89	0.96
Mean community distance to all-weather road	3.5	3.7	3.6	3.7	0.86	0.94
Number of observations[b]	942	530	937	536		

a. The F-test tests the equality of coefficients across the groups in a regression of each of the household and individual characteristics on group indicators with clustered household standard errors.
b. The number of observations reflect the actual individual assignments. This corresponds to the data reported in Table A.2. Panel 4, final column.

Table 3. Descriptive Statistics

a. Individual sample, by survey assignment

	Men				Women			
	Detailed, self	Detailed, proxy	Short, self	Short, proxy	Detailed, self	Detailed, proxy	Short, self	Short, proxy
All observations	437	250	472	251	505	280	465	285
Labor force participation (%)	88%	78%	92%	86%	81%	75%	88%	88%
Labor force (number of observations)	386	196	436	215	407	210	411	251
Wage workers (as % of labor force)	22%	15%	14%	11%	12%	9%	6%	4%
Daily wages, T Sh	2968	6638	4482	4931	2637	6853	5233	2007
	(2398)	(9928)	(4462)	(5260)	(3993)	(10045)	(9842)	(1646)
Wage workers (number of observ.)	82	30	60	24	50	19	23	10
LFP estimation sample[a] (Table 8)	1404				1525			
Wage worker estimation sample (Table 4-7)	192				99			

b. Wage workers

Characteristics, assignment	*men*	*women*
Daily wages, T Sh	4321	4066
Years of schooling	6.60	4.89
0 years of schooling, %	16%	37%
1-7 years of schooling, %	59%	44%
8-11 years of schooling, %	17%	15%
12-17 years of schooling, %	8%	4%
Age	33.82	34.09
Married	64%	57%
Number of children below age 6	0.88	1.08
Detailed –self	29%	42%
Detailed proxy	18%	.15%
Short self	34%	31%
Short proxy	19%	12%
Total number	192	99

Note: Standard deviations in parentheses.
a. Because of missing variables, 6 men and 10 women were excluded from the labor force participation analysis.4 men and 3 women were excluded from the wage worker estimation for the same reason.

Table 4. Returns to Education: Survey Assignments and Interaction Effects in OLS

	Men			*Women*		
	(1) $ln(w)$	*(2)* $ln(w)$	*(3)* $ln(w)$	*(4)* $ln(w)$	*(5)* $ln(w)$	*(6)* $ln(w)$
Years of schooling	0.08***	0.08***	0.04**	0.10***	0.10***	0.08**
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.04)
Years of schooling X short			**0.06****			**0.05**
			(0.03)			**(0.05)**
Years of schooling X proxy			**0.04**			**0.00**
			(0.03)			**(0.06)**
Short		0.06	-0.35		-0.06	-0.29
		(0.11)	(0.22)		(0.22)	(0.33)
Proxy		0.22	-0.09		0.25	0.23
		(0.14)	(0.28)		(0.19)	(0.37)
D : District dummies	yes	yes	yes	yes	yes	yes
Observations	192	192	192	99	99	99
R-squared	0.44	0.45	0.47	0.38	0.39	0.39

Note: All regressions include control variables: age, age squared, and a constant.
Standard errors, clustered at the household level, in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 5. Returns to Education, Allowing for Nonlinearity and Endogeneity

	Men			*Women*		
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
	lnw	*lnw*	*lnw*	*lnw*	*lnw*	*lnw*
Years of schooling 1 to 7	0.13	0.14	0.13	0.21*	0.22**	0.17
	(0.101)	(0.101)	(0.103)	(0.109)	(0.109)	(0.108)
Years of schooling 8 to 17	0.18*	0.19*	0.17*	0.31***	0.33***	0.34***
	(0.097)	(0.098)	(0.099)	(0.113)	(0.113)	(0.115)
Years of schooling 1to 7 X short			0.06			**0.15****
			(0.043)			**(0.064)**
Years of schooling 8 to 17 X short			**0.05***			-0.01
			(0.028)			(0.045)
Years of schooling 1 to7 X proxy			-0.01			0.02
			(0.055)			(0.078)
Years of schooling 8 to 17 X proxy			0.03			-0.02
			(0.034)			(0.052)
Survey assignment variables Short, Proxy	no	yes	yes	no	yes	yes
$\hat{c}$: control function term men / women	yes	yes	yes	yes	yes	yes
D : District dummies	yes	yes	yes	yes	yes	yes
Observations	192	192	192	99	99	99
R-squared	0.49	0.50	0.52	0.45	0.46	0.52
F-statistic for joint test years of schooling X short			1.69			3.65**
F-statistic for joint test years of schooling X proxy			0.94			0.37

Note: All regressions include control variables: age, age squared, community mean distance to nearest all season road, and a constant.
Standard errors, clustered at the household level, in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 6. First Stage to Obtaining the Control Function Term

	Years of Schooling	
	Men	*Women*
Community mean distance to primary school	-0.21	-0.62***
	(0.203)	(0.234)
Community mean distance to secondary school	-0.17**	
	(0.078)	
D : District dummies	yes	yes
Observations	192	99
R-squared	0.32	0.38

Note: All regressions include control variables: age, age squared, community mean distance to nearest all season road and a constant.
Standard errors, clustered at the household level, in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 7. Returns to Education, Allowing for Nonlinearity, Endogeneity and Selection Correction

	Men		*Women*	
	(1) *Heckman* *lnw*	*(2)* *Heckman - Hotz* *lnw*	*(3)* *Heckman* *lnw*	*(4)* *Heckman - Hotz* *lnw*
Years of schooling 1 to 7	0.12	0.13	0.17	0.17
	(0.103)	(0.103)	(0.106)	(0.110)
Years of schooling 8 to 17	0.15	0.16	0.50***	0.29**
	(0.103)	(0.103)	(0.160)	(0.126)
Years of schooling 1 to 7 X short	0.06	0.06	**0.14****	**0.16****
	(0.042)	(0.043)	**(0.060)**	**(0.064)**
Years of schooling 8 to 17 X short	**0.05***	**0.05***	-0.01	0.03*
	(0.027)	**(0.029)**	(0.049)	(0.018)
Years of schooling 1 to 7 X proxy	-0.01	-0.01	0.03	0.02
	(0.054)	(0.055)	(0.079)	(0.082)
Years of schooling 8 to 17 X proxy	0.02	0.03	-0.01	-0.01
	(0.033)	(0.034)	(0.056)	(0.051)
Survey assignment variables Short, Proxy	yes	yes	yes	yes
$\hat{c}$: control function term men / women	yes	yes	yes	yes
λ : Mills term men / women	yes	no	yes	no
$h(\hat{p})$: Predicted probability wage worker men / women	no	yes	no	yes
D : District dummies	yes	yes	yes	yes
Observations	192	192	99	99
R-squared	0.51	0.51	0.54	0.52
F-statistic for joint test years of schooling X short	1.73	1.64	3.47**	2.83*
F-statistic for joint test years of schooling X proxy	0.88	0.98	0.33	0.21

Note: All regressions include control variables: age, age squared, mean distance to nearest all season road, and a constant. Standard errors, clustered at the household level, in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 8: First Stage Selection Equation, Wage Worker

	Men		*Women*	
	(1)	*(2)*	*(3)*	*(4)*
	Heckman	*Heckman- Hotz*	*Heckman*	*Heckman-Hotz*
Years of schooling 1 to 7	0.01	0.01	-0.02	-0.02
	(0.019)	(0.019)	(0.020)	(0.020)
Years of schooling 8 to 17	**0.04*****	**0.03****	**0.07*****	**0.07*****
	(0.014)	**(0.014)**	**(0.020)**	**(0.020)**
Short	**-0.23****	**-0.23****	**-0.38*****	**-0.38*****
	(0.090)	**(0.091)**	**(0.112)**	**(0.112)**
Proxy	-0.15	-0.14	-0.12	-0.12
	(0.097)	(0.097)	(0.115)	(0.116)
Married	-0.11	-0.11	-0.20	-0.20
	(0.138)	(0.138)	(0.125)	(0.125)
Number of children	-0.06**	-0.06**	-0.04	-0.04
	(0.026)	(0.026)	(0.030)	(0.030)
Community mean distance to primary school		-0.02		-0.00
		(0.026)		(0.025)
Community mean distance to secondary school		0.00		
		(0.010)		
Community mean distance to nearest all-season road	no	yes	no	yes
D : District dummies	yes	yes	yes	yes
Observations	1,404	1,404	1,525	1,525

Note: All regressions include control variables: age, age squared, and a constant.
Standard errors, clustered at the household level, in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table 9: Returns to Education Estimates, Overview

	Men						*Women*					
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*	*(10)*	*(11)*	*(12)*
	OLS	*Linear estimates after controlling for endogeneity and selection*	*Non-linear returns*	*Non-linear returns after controlling for endogeneity*	*Non-linear returns after controlling for endogeneity and selection Heckman*	*Non-linear returns after controlling for endogeneity and selection Heckman - Hotz*	*OLS*	*Linear estimates after controlling for endogeneity and selection*	*Non-linear returns*	*Non-linear returns after controlling for endogeneity*	*Non-linear returns after controlling for endogeneity and selection Heckman*	*Non-linear returns after controlling for endogeneity and selection Heckman-Hotz*
	Table 4 column 3	*Table A.3 column 1*	*Not reported*	*Table 5 column 3*	*Table 7 column 1*	*Table 7 column 2*	*Table 4 column 6*	*Table A.3 column 2*	*Not reported*	*Table 5 column 6*	*Table 7 column 3*	*Table 7 column 4*
Detailed self	0.04**	0.14	0.004 0.05**	0.13 0.17*	0.12 0.15	0.13 0.16	0.08**	0.21*	-0.01 0.16***	0.17 0.34***	0.17 0.50***	0.17 0.29**
Short	0.10***	0.20**	0.06 0.05*	0.19* 0.22**	0.18 0.20*	0.19* 0.21**	0.13***	0.05	0.13** 0.14***	0.22** 0.33***	0.21** 0.49***	0.33** 0.32**
Difference short - detailed self	0.06**	0.06**	0.06*	0.06 0.05*	0.06 0.05*	0.06 0.05*	0.05	0.26**	0.14** -0.02	0.15** -0.01	0.14** -0.01	0.16** 0.03*

Standard errors, clustered at the household level, in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Annex

Table A.1. Screening Questions to Establish Employment Status, Short and Detailed Questionnaire

Short questionnaire	Detailed questionnaire
1. Did [NAME] do any type of work in the last seven days? Even if for 1 hour. YES...1 (»3) NO.....2 (2. question repeated for the past 12 months)	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 (go to 3) NO.....2 (2. question repeated for the past 12 months) 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 (go to 5) NO.....2 (4. question repeated for the past 12 months) 5. During the past 7 days, has [NAME] worked on his/her own account or in a business enterprise belonging to he/she or someone in your household, for example, as a trader, shop-keeper, barber, dressmaker, carpenter or taxi driver? YES...1 (go to 7) NO.....2 (6. question repeated for the past 12 months)

Table A.2.: Planned and Actual Survey Assignments

	Household survey assignment				
	Detailed self-reported	*Detailed proxy response*	*Short self-reported*	*Short proxy response*	*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 age 15+ [a]					
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 the assumption about household composition [a b]					
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
Numbers of observations for different groups					
Detailed	638	834	0	0	1472
Short	0	0	636	837	1473
Self	606	336	601	336	1879
Proxy	32	498	35	501	1066

a Assuming that each household has at least 2 members ages 10+ to be randomly selected for self-reporting
b Assuming that each household has one member age 15+ and an average of 2.5 householdmembers ages 10+. Thus, there are 1.5 *336 other members to be reported on by proxy.

Table A.3. Returns to Education, After Controlling for Endogeneity but Not Nonlinearity

	Men *(1)* *lnw*	*Women* *(2)* *lnw*
Years of schooling	0.14	0.21*
	(0.098)	(0.114)
Years of schooling X short	0.06**	0.05
	(0.028)	(0.041)
Years of schooling X proxy	0.04	-0.01
	(0.033)	(0.051)
Survey assignment variables Short, Proxy	yes	yes
$\hat{c}$: control function term men / women	yes	yes
$h(\hat{p})$: Predicted probability wage worker men / women	yes	yes
D : District dummies	yes	yes
Observations	192	99
R-squared	0.49	0.48

Note: All regressions include control variables age, age squared, mean distance to nearest all season road, and a constant
Standard errors, clustered at the household level, in parentheses; *** p<0.01, ** p<0.05, * p<0.1.