



Evans School Policy Analysis and Research (EPAR)

What Time Is it? Depends Who You Ask:
Respondent Bias in a Ugandan Household Time Use Survey

Joshua Merfeld, C. Leigh Anderson,
Pierre Biscaye and Travis Reynolds

EPAR Technical Report #331 - *Preliminary Draft*

Professor C. Leigh Anderson, Principal Investigator
Professor Travis Reynolds, co-Principal Investigator

June 22, 2016

Abstract

Labor is one of the most productive assets for many rural households in developing countries. Despite the importance of labor—and time use more generally—little research has empirically examined the quality of time-use data in household surveys. Many household surveys rely on respondent recall, the reliability of which may decrease as recall length increases. In addition, respondents often report on time allocation for the entire household, which they may not know or recall as clearly as their own time allocation. Finally, simultaneous activities such as tending children while preparing dinner, may lead to the systematic underestimation of certain activities, particularly those that tend to be performed by women. This paper examines whether the identity of the survey respondent affects estimates of time allocation within the household. Drawing on the Ugandan LSMS-ISA household survey, we find that individuals responding for themselves report higher levels of time use over the previous week than when responding for other household members. Moreover, male respondents tend to underreport time allocation for females over the age of 15 as compared to female respondents, especially time spent on domestic activities. In addition, an analysis of the effects of two economic shocks—having a baby and floods or droughts—suggests that the identity of the respondent can affect substantive conclusions about the effects of shocks on household time use.

Introduction

For many individuals in rural households in developing countries, labor is their most productive asset. As such, many development policies focus on labor, either aiming to provide employment opportunities—the National Rural

Employment Guarantee Scheme in India, for example—or improve labor productivity, often in agriculture (World Bank, 2008). However, in order to appropriately evaluate the efficacy of these interventions, accurate data on labor allocation are required (Carletto et al., 2015). Yet, little research has specifically focused on the quality of these data (Beegle et al., 2012a).

There are reasons to question the accuracy of labor allocation and time-use data more generally. Most surveys rely on some form of recall to collect these data (Beegle et al., 2012a). In addition to time and budget constraints (Beegle et al., 2012a), nationally representative household surveys must also cover all parts of the year due to the seasonal nature of agriculture, further increasing costs (Beegle et al., 2012a; Carletto et al., 2013). Finally, collecting time-use data can be very cognitively demanding on respondents. As such, there is further concern that the quality of survey data—one of the most important sources of data for development researchers (Beegle et al., 2012b)—may be compromised.

This paper investigates one possible source of bias in time-use data. In the Living Standards Measurement Survey (LSMS), a single respondent—generally the “most knowledgeable” household member (Grosh and Glewwe, 2000)—answers most of the questions in a module. The time-use module is no exception. Using three waves of LSMS data from Uganda, we exploit variation in the identity of the survey respondent—across time but within households—to examine how the identity of the respondent affects responses to questions on time allocation. The panel nature of the data aid in identification; we use person fixed effects to control for time-invariant personal characteristics and household-

EPAR uses an innovative student-faculty team model to provide rigorous, applied research and analysis to international development stakeholders. Established in 2008, the EPAR model has since been emulated by other UW schools and programs to further enrich the international development community and enhance student learning.

Please direct comments or questions about this research to Principal Investigators Leigh Anderson and Travis Reynolds at epar.evans.uw@gmail.com.

wave fixed effects to control for time allocation shocks suffered by all individuals within a household in the same year.

Methods

Data are from the World Bank Living Standards Measurement Study (LSMS) Integrated Surveys on Agriculture. Specifically, we use waves two, three, and four of the survey from Uganda, which were collected in 2009/10, 2010/11, and 2011/12, respectively. The panel nature of the data allows us to match households across waves. Of the 3,211 different households in the data, we are able to match 2,345 households across all three waves and 409 households across two waves.

All three waves ask basic questions on time use. Respondents were asked, “*In the last 7 days, how much time in hours did [NAME] spend [doing activity X]?*” In waves three and four, the activities covered were: domestic activities, primary job, secondary job, collecting firewood, fetching water, construction, repairs, food processing, making handicrafts, agriculture, and hunting and fishing. In wave two, domestic activities were not included. A single respondent answered these questions for the entire household.

We employ two methods to explore respondent bias. First, we run regressions using person fixed effects. This controls for time-invariant individual characteristics, such as the relationship of the individual to the household head as well as the general position of the individual within the household. Second, in some regressions we add household/wave fixed effects. In this way, we are also able to sweep out shared variation within the household in the same wave. This helps to control for idiosyncratic shocks at the household level that affect all household members similarly.

In a final set of results, we explore whether any response bias may affect substantive conclusions regarding the effects of household shocks. We look at the effects of two shocks: the presence of a child under the age of two and a drought or flood in the preceding year. We only use household fixed effects in these regressions.

Results

Table 1 presents the first set of findings. When individuals respond for themselves, they tend to report a higher number of total hours accounted for in the previous week than when they respond for others in the household. The mean number of hours accounted for across waves two through four is just 33 hours for prime age adults. Therefore, the coefficient on self represents an effect of more than 10% of the mean. This intuitive finding suggests respondents may not have an accurate idea of the number of hours that other household members spend on

different activities. Table 2 presents results analogous to Table 1, but excluding wave 2, which does not include a question on time spent on domestic activities. The difference in results between Table 1 and Table 2 suggests domestic activities are driving part of these effects.¹ Moreover, the effect is again quite large, as the mean of total time accounted for in waves two and three is approximately 40. Given that domestic activities are often performed by women in developing countries, this finding suggests the identity of the respondent may be important when analyzing data disaggregated by gender.

Table 3 presents a final set of results from a pseudo-policy simulation to examine whether these differences in responses affect substantive conclusions regarding the effects of household shocks. The table reports on the effects of two household shocks—the birth of a baby and droughts or floods—and shows that the estimated effect of the shock on household time allocation depends on the gender of the respondent. The coefficients can be interpreted as the effect of the shock on the specified household members in affected households.² Male respondents report large differences in the effect of these shocks for male and female household members, whereas female respondents report almost no differences.

Conclusions

In this paper, we find that individuals report higher time allocation for themselves than for other household members. We also find that male respondents underreport—relative to female respondents—the time female household members spent in the previous week, particularly in domestic activities. Moreover, these biases can affect substantive conclusions regarding the effects of household-level shocks.

Overall, these findings suggest that the identity of the respondent is an important consideration for survey design. In particular, the gender of the survey respondent may affect responses regarding certain activities within the household. If the identity of the respondent affects estimates of household shocks, empirical researchers might consider using respondent characteristics as control variables in regression models. This is of course not guaranteed to remove all bias. However, the addition of these characteristics as controls may help researchers identify when the identity of the respondent is affecting results.

Table 1. Survey Respondent and Reported Hours

	Person Fixed Effects			Person and Household Fixed Effects	
	(1)	(2)	(3)	(4)	(5)
Respondent is self	4.495***	4.356***	4.168***	3.314***	3.293***
Respondent is male		0.007	-0.044		
Age		1.004*	1.306*		0.961
Age x age		-0.017*	-0.017*		-0.010
Household Size			-2.188***		
Household size x household size			0.097***		
Observations	19,222	19,061	19,061	19,222	19,222

Standard errors are in parentheses. Standard errors are clustered at the household level. Columns 1-3 include person fixed effects, while columns 4 and 5 include both person fixed effects and household/wave fixed effects. Regressions include waves 2, 3, and 4. Wave 2 does not ask about domestic hours.

* p<0.10, ** p<0.05, *** p<0.01

Table 2. Survey Respondent and Reported Hours - Excluding Wave 2

	Person Fixed Effects			Person and Household Fixed Effects	
	(1)	(2)	(3)	(4)	(5)
Respondent is self	7.998***	8.036***	7.969***	5.803***	5.760***
Respondent is male		-1.231	-1.265		
Age		2.210*	2.361*		0.725
Age x age		-0.015	-0.014		0.001
Household Size			-1.445		
Household size x household size			0.051		
Observations	19,222	19,061	19,061	19,222	19,222

Standard errors are in parentheses. Standard errors are clustered at the household level. Columns 1-3 include person fixed effects, while columns 4 and 5 include both person fixed effects and household/wave fixed effects. Regressions include only waves 3 and 4.

* p<0.10, ** p<0.05, *** p<0.01

Table 3. Effect of Baby/Agricultural Shocks on Gender by Respondent (Linear Combinations)

	Baby in Household		Drought or Flood	
	(1)	(2)	(3)	(4)
Female Respondent	2.421		-5.366***	
Male Respondent	-0.064		-7.322***	
Female respondent, male household member		2.262		-5.458***
Female respondent, female household member		2.232		-5.540***
Male respondent, male household member		-1.920		-8.789***
Male respondent, male household member		4.467*		-3.784*

Standard errors are in parentheses. Standard errors are clustered at the household level. Each column pertains to one regression, and all four regressions include household fixed effects. Only waves three and four are included in these models. Columns 1 and 2 look at the effects of having a baby-defined as a child less than two years of age-in the household, while columns 3 and 4 look at the effects of an agricultural shock-a drought or flood. In columns 1 and 3, we do not separate out the effect of the shocks on male and female time allocation. In columns 2 and 4, we allow the effect of the shocks to vary by gender of the individual. In all four specifications, we include an interaction term between the effect of the shock and the gender of the respondent. The coefficients are linear combinations of the different interaction terms and thus show the overall effect of the shock on different household members, depending on the gender of the respondent. Also included as control variables in all regressions are age, age squared, household size, number of children, whether the respondent is self, region/wave fixed effects, and a set of dummy variables for relationship to household head.

* p<0.10, ** p<0.05, *** p<0.01

Endotes

¹ The full paper also contains a table (Table A1 in the appendix) looking explicitly at domestic activities using the last two waves. The results confirm that domestic activities are indeed responsible for at least part of the effect.

² In other words, the coefficient for “Female respondent” represents the estimated effect of the shock on households with a female respondent. Similarly, the coefficient for “Male respondent, female member” represents the estimated effect of the shock on female household members in households with a male respondent.

References

- Beegle, K., Carletto, C., & Himelein, K. (2012a). Reliability of recall in agricultural data. *Journal of Development Economics*, 98(1):34-41.
- Beegle, K., De Weerd, J., Friedman, J., & Gibson, J. (2012b). Methods of household consumption measurement through surveys: Experimental results from Tanzania. *Journal of Development Economics*, 98(1):3-18.
- Carletto, C., Jolliffe, D., & Banerjee, R. (2015). From tragedy to renaissance: improving agricultural data for better policies. *The Journal of Development Studies*, 51(2):133-148.
- Carletto, C., Zezza, A., & Banerjee, R. (2013). Towards better measurement of household food security: Harmonizing indicators and the role of household surveys. *Global Food Security*, 2(1):30-40.
- Grosh, M. & Glewwe, P. (2000). *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Measurement Study, Volume 2*. Washington, DC: World Bank.
- World Bank (2008). World development report: Agriculture for development.