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

Joshua D Merfeld

I wish to thank the Evans School Analysis and Research Group (EPAR) for funding the work that led to this paper. Leigh Anderson, Travis Reynolds, Pierre Biscaye, and participants at the University of Washington's Labor/Development Brownbag series provided comments that greatly improved this paper.

EVANS SCHOOL OF PUBLIC POLICY & GOVERNANCE
UNIVERSITY of WASHINGTON
Evans School Policy Analysis and Research (EPAR)

Outline

Background

- > **Bias in Survey Data**

Data, Summary Statistics, and Methods

- > **Data**
- > **Time Use over Last Week**
- > **Individual Summary Statistics**
- > **Household Summary Statistics**
- > **Methods**

Results

- > **Proxy Respondent Bias**
- > **Effects of Household shocks**

Findings

I examine whether the identity of the survey respondent affects estimates of time allocation within the household.

- **Individuals responding for themselves report higher levels of time use over the previous week than when responding for others.**
- **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.**
- **The identity of the respondent can affect substantive conclusions about the effects of shocks on household time use.**

Background

Survey Data

Survey data are used by empirical researchers in almost all social-science disciplines:

- **Economics**
- **Political Science**
- **Sociology**

Development economists use many surveys: LSMS (this paper), ICRISAT's VDSA, ARIS-REDS, NSS, not to mention all of the data collected by organizations (IFPRI, ILRI, BMGF, etc.) and individual researchers.

Policymakers use conclusions drawn (at least in part) from these data to inform policy.

Bias in Survey Data

Recall bias:

- **People have more difficulties recalling the distant past than the near past (Stone and Shiffman, 2002; Sudman and Bradburn, 1974):**
  - **Consumption surveys - People report lower consumption when using recall than when using diaries (Beegle et al., 2012a; de Weerdt et al., forthcoming)**
  - **Agricultural data - Perhaps not as much recall error as people had assumed? (Beegle et al., 2012b)**

Bias in Survey Data (cont.)

Telescoping and Availability Heuristic:

- **Telescoping - People believe events happened more recently than they actually did (Beckett et al., 2001; Bolger et al., 2003)**
- **Availability Heuristic - People believe easy-to-recall events happen more often than they do (Tversky and Kahneman, 1973), e.g. plane crashes**

Proxy Respondent Bias:

- **Kilic and Moyan (2015) report results from a randomized household survey in which the decision to interview the most knowledgeable household member or a random household member can affect responses**

Difficulties with Collecting Time-Use Data

Simultaneous Activities:

- **Performing multiple activities simultaneously – there is no clear "best method" to allow for simultaneous activities on surveys (Esquivel et al., 2008)**
- **Failure to include secondary activities may lead to undercounting for some activities (Harvey and Taylor 2000)**
  - **Systematic undercounting of hours for some individuals, e.g. women**

Data, Summary Statistics, and Methods

EVANS SCHOOL OF PUBLIC POLICY & GOVERNANCE
UNIVERSITY *of* WASHINGTON
Evans School Policy Analysis and Research (EPAR)

Data – Uganda LSMS

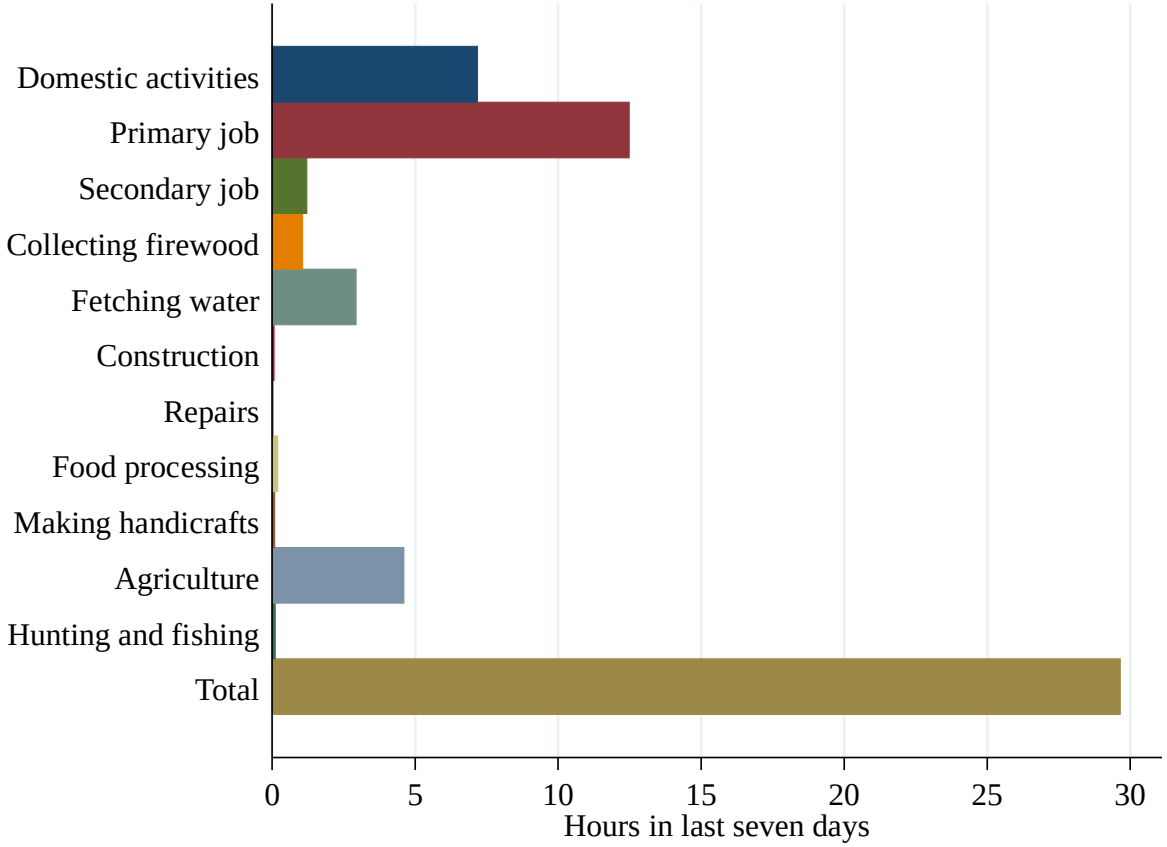
- Living Standards Measurement Survey (LSMS). Implemented by participating countries' statistical agencies and coordinated by the World Bank.
- Uganda:
  - Uganda Bureau of Statistics
  - Three waves in this paper: 2009/10, 2010/11, 2011/12
  - I am able to match 2,345 households across all three waves and an additional 409 households across two waves, while 357 households appear in only one wave.
  - Wave two does not include domestic activities question.

UGANDA

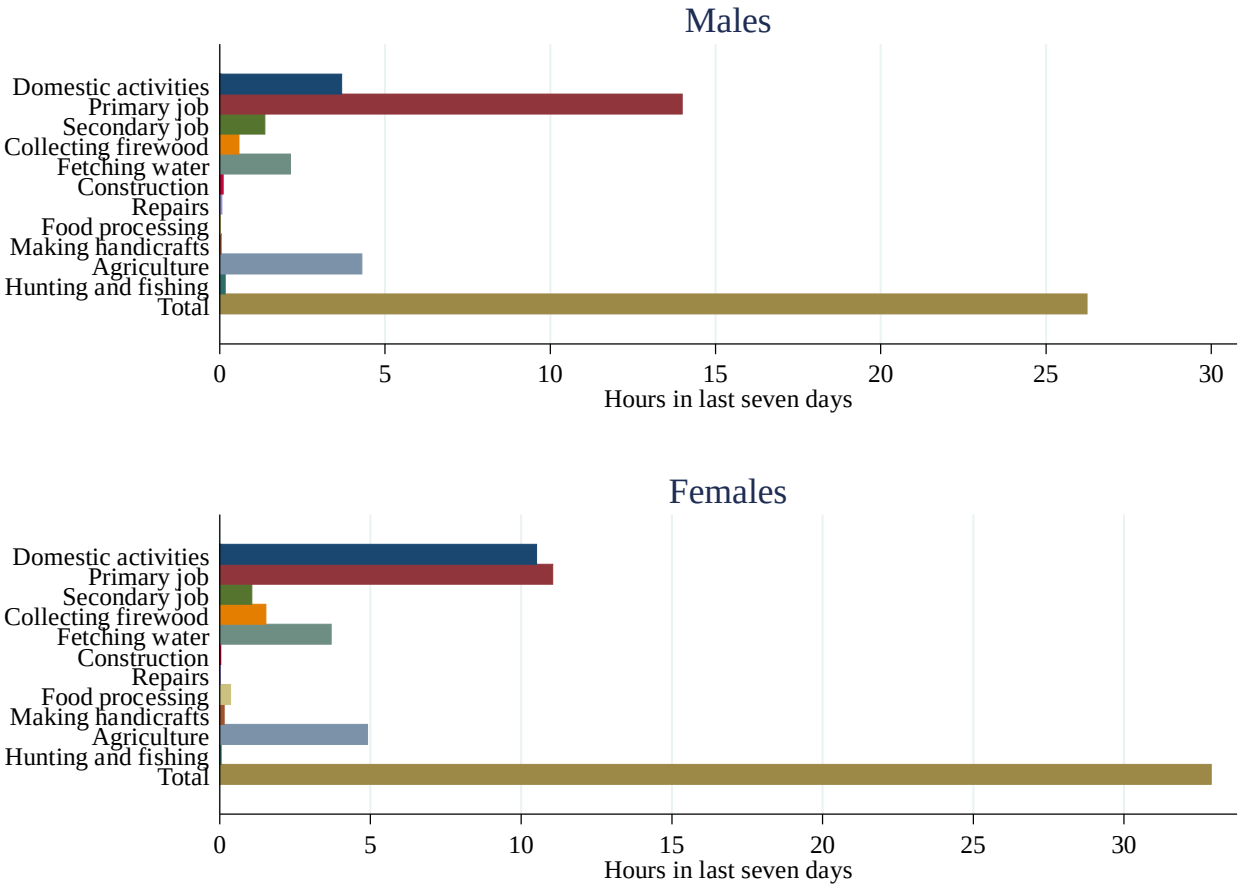
- **World Bank statistics (2015):**
  - **Population: 39.03 million**
  - **GNI per capita: USD 1,780 (PPP)**
  - **GDP growth: 5.0%**
  - **Inflation: 5.2%**

Google Maps

Average hours spent over last seven days



Average hours by gender



Individual Summary Statistics

	Wave 2 mean/sd	Wave 3 mean/sd	Wave 4 mean/sd
Male	0.494	0.487	0.489
	(0.500)	(0.500)	(0.500)
Age	23.999	24.267	24.284
	(17.542)	(17.967)	(18.037)
Adult (age>= 15)	0.606	0.601	0.606
	(0.489)	(0.490)	(0.489)
Child (age< 15)	0.394	0.399	0.394
	(0.489)	(0.490)	(0.489)
Elderly (age>= 60)	0.055	0.059	0.060
	(0.228)	(0.236)	(0.237)
Respondent Is Self	0.327	0.290	0.279
	(0.469)	(0.454)	(0.448)
Observations	13802	12533	13527

Household Summary Statistics

	Wave 2 mean/sd	Wave 3 mean/sd	Wave 4 mean/sd
Respondent is male	0.427	0.391	0.387
	(0.495)	(0.488)	(0.487)
Adults in household (age>= 15)	3.444	3.974	4.511
	(2.062)	(2.432)	(2.872)
Children in household (age< 15)	2.940	3.115	3.041
	(2.139)	(2.214)	(2.210)
Newborn (< 1 year old) in household - yes/no	0.178	0.173	0.169
	(0.382)	(0.378)	(0.375)
Baby (< 2 years old) in household - yes/no	0.343	0.329	0.322
	(0.475)	(0.470)	(0.467)
Agricultural shock (drought or flood) in last year in household - yes/no	0.474	0.293	0.237
	(0.499)	(0.455)	(0.425)
Observations	2889	2618	2803

Methods

- To explore the effect of gender and proxy respondent bias on survey responses, I use person fixed effects.
  - I also include wave dummies, the age of the individual, and household size.
- To further identify proxy respondent bias, I add household-wave fixed effects.
- In final estimation, I explore the effects of household shocks.
  - I look at the effects of a baby in the household and the effects of a flood/drought on individual time allocation.

Results

Results I

Survey Respondent and Reported Hours – All Waves

	Person Fixed Effects			Person and HH/Wave Fixed Effects	
	(1)	(2)	(3)	(4)	(5)
Respondent is self	4.495***	4.356***	4.168***	3.314***	3.293***
	(0.612)	(0.618)	(0.617)	(0.594)	(0.594)
Respondent is male		0.007	-0.044		
		(0.599)	(0.600)		
Age		1.004*	1.306*		0.961
		(0.498)	(0.511)		(0.653)
Age x age		-0.017*	-0.017*		-0.010
		(0.007)	(0.007)		(0.008)
Household size			-2.188***		
			(0.489)		
Household size x household size			0.097***		
			(0.024)		
Observations	19222	19061	19061	19222	19222

Results II

Survey Respondent and Reported Hours – Waves 3 and 4

	Person Fixed Effects			Person and HH/Wave Fixed Effects	
	(1)	(2)	(3)	(4)	(5)
Respondent is self	7.998***	8.036***	7.969***	5.803***	5.760***
	(0.951)	(0.962)	(0.961)	(0.839)	(0.839)
Respondent is male		-1.231	-1.265		
		(0.913)	(0.914)		
Age		2.210*	2.361*		0.725
		(0.933)	(0.946)		(0.942)
Age x age		-0.015	-0.014		0.001
		(0.014)	(0.014)		(0.013)
Household size			-1.445		
			(0.796)		
Household size x household size			0.051		
			(0.038)		
Observations	14184	14071	14071	14184	14184

Results III

Effect of Shocks by Gender of Respondent – Wave 3 and 4

	Baby		Ag. Shock	
	(1)	(2)	(3)	(4)
	Total time accounted for	Total time accounted for	Total time accounted for	Total time accounted for
Female respondent	2.421		-5.366***	
	(1.554)		(1.337)	
Male respondent	-0.064		-7.322***	
	(1.725)		(1.474)	
Female respondent, male member		2.262		-5.458***
		(1.870)		(1.674)
Female respondent, female member		2.232		-5.540***
		(1.653)		(1.482)
Male respondent, male member		-1.920		-8.789***
		(1.802)		(1.528)
Male respondent, female member		4.467*		-3.784*
		(2.279)		(2.096)

Conclusions

- Individuals report higher time allocation for themselves.
- Male respondents underreport individual time allocation.
  - This is especially evident for domestic hours.

- Surveys should be designed with respondent in mind.
- Empirical researchers should consider including interviewer characteristics as controls.

Evans School Policy Analysis & Research Group (EPAR)

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

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 C. Leigh Anderson and Travis Reynolds at epar.evans.uw@gmail.com.

EVANS SCHOOL OF PUBLIC POLICY & GOVERNANCE
UNIVERSITY of WASHINGTON
Evans School Policy Analysis and Research (EPAR)

Survey Respondent and Domestic Hours

	(1) Domestic activities b/se	(2) Domestic activities b/se	(3) Domestic activities b/se	(4) Domestic activities b/se
Respondent is self	3.034***	3.034***	3.021***	3.044***
	(0.321)	(0.329)	(0.329)	(0.327)
Respondent is male		-0.706**	-0.731**	-0.769**
		(0.301)	(0.301)	(0.300)
Age		0.315	0.413	0.360
		(0.329)	(0.335)	(0.352)
Age x age		-0.001	-0.001	-0.001
		(0.005)	(0.005)	(0.005)
Household size			-0.161	-0.243
			(0.269)	(0.275)
Household size x household size			-0.010	-0.008
			(0.013)	(0.013)
Region/Wave Fixed Effects	No	No	No	Yes
Observations	14184	14071	14071	14071

Individual Stats – Relationship to Head

	Wave 2 mean/sd	Wave 3 mean/sd	Wave 4 mean/sd
Relationship to Head			
Head	0.201	0.208	0.207
	(0.400)	(0.406)	(0.405)
Spouse of head	0.133	0.137	0.138
	(0.340)	(0.344)	(0.345)
Son/daughter of head	0.452	0.448	0.446
	(0.498)	(0.497)	(0.497)
Grandson/daughter of head	0.094	0.095	0.099
	(0.292)	(0.293)	(0.299)
Parent of head/spouse	0.005	0.005	0.006
	(0.072)	(0.073)	(0.076)
Sibling of head/spouse	0.022	0.018	0.017
	(0.147)	(0.133)	(0.128)
Nephew/niece of head	0.039	0.038	0.034
	(0.195)	(0.191)	(0.181)
Other relatives	0.036	0.035	0.040
	(0.187)	(0.184)	(0.196)
Servant	0.010	0.007	0.006
	(0.099)	(0.081)	(0.079)
Non-relative	0.006	0.006	0.005
	(0.079)	(0.076)	(0.073)
Observations	13802	12533	13527

Individual Stats – Time Use

	Wave 2 mean/sd	Wave 3 mean/sd	Wave 4 mean/sd
Time Use in Previous Week			
Domestic activities		6.962	7.382
		(9.417)	(9.719)
Primary job	13.274	12.242	12.710
	(19.767)	(18.262)	(19.134)
Secondary job	1.686	1.393	1.043
	(6.771)	(5.725)	(4.931)
Collecting firewood	1.171	0.989	1.141
	(2.990)	(1.911)	(2.381)
Fetching water	2.751	2.981	2.898
	(5.108)	(4.752)	(4.722)
Constructing dwelling, buildings, roads, etc.	0.100	0.080	0.054
	(1.707)	(1.263)	(1.073)
Repairing dwelling, buildings, roads, etc.	0.055	0.052	0.033
	(1.026)	(0.758)	(0.629)
Milling and food processing for household	0.102	0.143	0.244
	(1.232)	(1.046)	(1.415)
Making handicrafts for household	0.145	0.096	0.091
	(1.782)	(1.637)	(1.363)
Agriculture	4.751	4.283	4.908
	(9.869)	(9.164)	(10.073)
Hunting and fishing	0.106	0.088	0.121
	(1.962)	(1.745)	(1.910)
Total time accounted for	24.142	29.013	30.237
	(28.320)	(28.842)	(29.715)
Observations	13802	12533	13527

Effect of Baby/Ag Shock

	(1) Total time accounted for b/se	(2) Total time accounted for b/se	(3) Total time accounted for b/se	(4) Total time accounted for b/se
Male respondent (yes=1)	-1.878	-2.302*	-2.228*	-2.520*
	(1.018)	(1.044)	(0.980)	(1.002)
Respondent is self	7.769***	8.654***	7.730***	8.304***
	(0.723)	(0.813)	(0.719)	(0.782)
Panel A - Female Respondent				
Family has baby (¡2 years)	2.421	2.262		
	(1.554)	(1.870)		
Baby times female		-0.030		
		(1.591)		
Agricultural shock			-5.366***	-5.458**
			(1.337)	(1.674)
Ag shock times female				-0.082
				(1.661)
Panel B - Male Respondent				
Baby	-2.485	-4.182*		
	(1.530)	(1.868)		
Baby times female		6.417**		
		(2.374)		
Ag shock			-1.956	-3.331
			(1.574)	(1.931)
Ag shock times female				5.087*
				(2.375)
Observations	9414	9414	9414	9414

Estimating Equations

One:

$$y_{iht} = a_i + \beta S_{iht} + \gamma M_{ht} + X_{iht} + \psi_t + \varepsilon_{iht}$$

Two:

$$y_{iht} = a_i + \phi_{ht} + \beta S_{iht} + X_{iht} + \varepsilon_{iht}$$

Three:

$$y_{iht} = a_h + \beta Z_{ht} + \gamma (Z_{ht} \times M_{ht}) + X_{iht} + \varepsilon_{iht}$$