Policy Research Working Paper 10695

Recording the Time Divide

A Comparative Study of Smartphone- and Recall-Based Approaches to Time Use Measurement

Talip Kilic
Gayatri Koolwal
Wilbert Drazi Vundru
Thomas Daum
Hannes Buchwald
Greg Seymour
Peter Mvula
Alister Munthali
Monice Kachinjika

WORLD BANK GROUP
Development Economics
Development Data Group
February 2024

Policy Research Working Paper 10695

Abstract

Based on a randomized survey experiment in Malawi, this study examines how innovative techniques in time use data collection could sidestep measurement concerns with traditional recall-based time use measurement. The experiment assigns random samples of households, and adult men and women within, to one of two treatment arms on time use measurement: a traditional 24-hour recall time use diary, and a self-administered smartphone-based pictorial time diary, known as the *TimeTracker* app, for real-time data collection. Compared to the recall arm, participation in employment and unpaid domestic and care work is shown to be higher in the smartphone arm for both men and women. The resulting estimates of gender gaps, while continuing to be large, are narrower in the smartphone arm, except for care work where the estimated gender gap increases. The recall treatment leads to substantial underreporting of activities after 6 pm, which otherwise accounts for nearly 30 percent of daily reported time in the smartphone arm. Likewise, the extent of simultaneous activities, particularly among women, is markedly lower in the recall arm. The overall reported time is, however, higher in the recall arm due to the minimum 15-minute duration that was used for recording activities the 24-hour recall diary, while over one-third of activities lasted less than 15 minutes in the smartphone arm. The analysis also shows that using stylized time use questions with a 7-day recall, as opposed to a 24-hour recall diary, results in an even greater overestimation of reported time in employment and unpaid work.

This paper is a product of the Development Data Group, Development Economics. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at tkilic@worldbank.org and gkoolwal@worldbank.org.

Recording the Time Divide:
A Comparative Study of Smartphone- and Recall-Based Approaches to Time Use Measurement

Talip Kilic[a], Gayatri Koolwal[a], Wilbert Drazi Vundru[a], Thomas Daum[b],
Hannes Buchwald[c], Greg Seymour[d], Peter Mvula[e], Alister Munthali[e], Monice Kachinjika[e1]

JEL Codes: C83, J16, J22.

Keywords: Time Use, Employment, Unpaid Work, Recall Bias, Smartphone Time Use Diary, Survey Methodology, Household Surveys, Gender, Malawi, Sub-Saharan Africa.

[1] Corresponding authors: Talip Kilic (tkilic@worldbank.org) and Gayatri Koolwal (gkoolwal@worldbank.org). [a] World Bank; [b] University of Gothenburg; [c] Independent Consultant; [d] International Food Policy Research Institute; [e] Palm Consulting. The Malawi Methodological Experiment on Time Use Measurement was implemented under the World Bank Living Standards Measurement Study (LSMS) program, in collaboration with the Measures for Advancing Gender Equality (MAGNET) initiative and PALM Consulting, the latter of whom implemented the fieldwork. The study benefited from funding from the International Fund for Agricultural Development, the World Bank Knowledge for Change Program, and the World Bank Umbrella Facility for Gender Equality, in partnership with the Bill and Melinda Gates Foundation. The authors are grateful to Andrey Prokhorenko (software programming) and Natalis Lorenz (illustrations) who were instrumental in designing the version of the *Time Tracker* smartphone app that was used for this study; and Diego Zardetto for assistance with the sampling design. We thank Nancy Folbre and Hazel Malapit for valuable comments on the paper, as well as feedback from the participants at the 2023 International Food Policy Research Institute-RISE meeting, the Ninth International Conference of Agricultural Statistics (ICAS), the University of Maryland Conference on Measuring Time Use and Wellbeing, and the World Bank Development Data Group Let's Talk Data Seminar.

1. Background

Representative surveys on individuals' time use are important for understanding economic decision-making among men and women and highlighting broader inequalities within and across households. Effective design, targeting and evaluation of policies/projects related to workforce development, childcare provision, and improvements in transportation services, among others, also require an accurate measurement of men's and women's time use and time use disparities across different sub-populations.

Available time use surveys reveal cross-country gender disparities in time allocations to different activities, including unpaid domestic and care work, which are often missed in labor force surveys and socioeconomic household surveys (Charmes, 2019; Hirway, 2017). While time use data collection has been implemented in varying formats across countries – either as part of stand-alone time use surveys or part of a time use module that is part of a labor force survey or a socioeconomic household survey, time use data collection in low- and middle-income contexts relies either on an interviewer-assisted recall-based time use diary – usually for the last 24 hours, or recall-stylized questions about time allocation to relatively aggregate groups of activities over a given reference period, which can be the last day, the last 7 days or the last 30 days (UN Women, 2021).[2]

These survey design choices have raised concerns regarding potential cognitive errors in recalling time use (Seymour et al., 2020), particularly in contexts with lower literacy and numeracy and where economic activities are less predictable, including in countries where smallholder agriculture is prevalent (Arthi et. al, 2018; Field et al., 2023).[3] Activities conducted in tandem may also be more difficult to recall, depending on how often an activity occurs (regularity) and how distinguishable it is from other activities (saliency) (Menon, 1993; Brown et al., 2007). These errors can occur over shorter reference periods as well (Sudman and Bradburn, 1973).

Although these measurement challenges might have direct implications for our understanding of gender disparities in time use (Floro and Miles, 2003; Budig and Folbre, 2004) and analytical applications that rely on time use data, including the design and evaluation of gender-sensitive policies to improve economic opportunities and well-being for men and women (World Economic Forum, 2021; King et al., 2021), there is scant evidence on how relying on recall-based survey methods may impact accuracy and reliability of time use data. Recall-based diaries can also be time-consuming to implement, and recent work (Field et al., 2023, ILO, 2023; Folbre, 2021) has discussed the role of streamlined time use diaries (by reducing or aggregating the number of activity categories, for example) as an alternative to more intensive data collection.

[2] Within these approaches, however, there is substantial variation across countries and years — including, for example, the choice of activity sets over which to report; the length of the reference period; the reliance on self- versus proxy reporting; the extent of interviewer assistance; and whether time use data collection is designed as a standalone survey or is part of an existing household survey (see UN Women, 2021, for a discussion). Open-ended diaries without pre-defined time intervals have also been used, but rarely (see Chatzitheochari et al.,2018, for a discussion).

[3] Farming follows an inherently irregular schedule that is often determined by factors outside the control of the farmer, such as climate variability, and the specific tasks involved in farming are often very repetitive (e.g., weeding, harvesting, and sorting).

Against this background, we present new findings from a large-scale randomized survey experiment conducted in Malawi to document the relative accuracy of recall-based approaches to time use survey data collection. The experiment assigns random samples of households, and adult men and women within, to one of two treatment arms on time use measurement: a traditional 24-hour recall time use diary; and a self-administered smartphone-based pictorial time diary, known as the *TimeTracker* app. While real-time data collection on time use has been explored in other surveys in higher-income settings (Gershuny et al., 2020), our study is the largest one implemented to date in a low-income country, extending a nascent evidence base regarding the feasibility and utility of smartphone-based time use data collection - previously only available for Zambia, though for a significantly smaller, rural-only sample (Daum et al., 2018, 2019). Beyond the methodological contributions of our work, we reveal new, policy-relevant insights regarding intra-day variation and simultaneity in time use.

The analysis compares men's and women's self-reported time allocation across two survey treatment arms, over an 11-day reporting period, to provide a disaggregated, comparative analysis of real-time and recall-based survey approaches over shorter (24-hour) and longer (7-day) recall periods:

(1) In the first treatment arm, one adult man and adult woman (aged 15-64) in each household received an Android smartphone-based pictorial time use diary app called *TimeTracker*[4] to self-report their time use across a set of activities on a continuous, real-time basis over the period. The app was developed specifically for time use data collection in low-literacy settings (Daum et al., 2018) where respondents do not necessarily have clock-based concepts of time and no prior experience with smartphones.
(2) In the second treatment arm, one adult man and adult woman in each household received a traditional, recall-based 24-hour diary interview on three different days during the reporting period, where interviewers asked respondents to recall their time allocation over the last 24 hours and coded the reported activities into 15-minute intervals based on the respondent's narrative, with up to two simultaneous activities that can be recorded for each 15-minute interval. Reporting in the recall diary arm covered the same activity set as in the smartphone arm.
(3) Finally, at the end of the reporting period, respondents in both treatment arms received a stylized interview-based 7-day recall module covering a subset of activities in the smartphone/recall diary approaches. The 7-day module follows a structure commonly implemented in labor force surveys or socioeconomic household surveys (Seymour et al., 2020), and allows us to investigate how the same groups varied their responses over recall periods of different lengths.

Our findings span three key areas, with important contributions to the literature. First, on participation and time across different activities, we show that the smartphone and recall arms provide very different perspectives, along with an interesting dichotomy in activity sets highlighted across the two arms. Recall, for example, leads to an underreporting of participation in employment and unpaid work, including a lower

[4] *Time Tracke*r was originally developed by the University of Hohenheim and the Institute for Applied Science (University of Media, Stuttgart) (Daum et al., 2018, 2019) and enhanced for this study with World Bank LSMS support. While the app is not yet publicly available, there is a plan is to release it for public use.

share of care work among rural women. On the other hand, we observe higher participation in leisure and self-care activities in the recall arm, including personal care, sleep and food consumption. Importantly, the smartphone arm also reflects greater participation among men in different areas of unpaid work — so that gender gaps in unpaid work in the smartphone sample, while remaining large, are narrowed when compared to the recall arm. Further, we show that the total reported time across most activities is, on average, higher in the recall arm, for both men and men and across urban and rural areas. The main driver of this finding is the set-up of the 24-hour time use diary to record activities in 15-minute intervals, while around a third of activities reported in the smartphone arm in fact last less than 15 minutes, with non-farm employment, unpaid domestic and care work, and self-care activities making up the greatest shares of these tasks. Given the common use of the 15-minute structure of time use diaries (UN Women, 2021; Seymour et al., 2020), our findings highlight how this structure can lead to "rounding" and overestimation of reported time.

Second, the smartphone data collection provides greater insights regarding the diversity and frequency of activities conducted throughout the day and a significantly greater incidence of activities (in particular non-leisure activities) conducted simultaneously with each other vis-à-vis the recall treatment arm. These nuances on time allocation can also help explain some of the drops in participation across activities that we observe in the recall arm. Recall, for example, leads to a substantial drop-off in reporting of active time in the early evening hours (after 6 pm), particularly among women. These hours account for a large share of daily time — about 26 percent of daily time for women and 29 percent of daily time for men in the smartphone arm. We discuss how underreporting of activities in the evening in the recall arm may be due to respondent fatigue - since each recall interview could last up to 30 to 40 minutes – as well as other interview-related factors. We also discuss how activity-linked pop-ups in the smartphone data collection to record the presence of other children, may highlight an important but understudied dimension of passive childcare (Folbre, 2005), over and above respondents' own estimates of childcare.

Third, we show that the stylized recall module with a 7-day reference period leads to even higher estimates of reported daily time compared to the 24-hour recall diary. Moving from real-time data collection to 24-hour recall, and to 7-day recall therefore leads to increasing estimates of time allocation to both employment and unpaid domestic and care work. Recent studies on farm labor (Arthi et al., 2018; Gaddis et al., 2021) have also found that longer recall periods (e.g., over the last 30 days or the past season vis-à-vis the last 7 days) overestimate labor time.

Our findings are consistent over different reporting days, and treatment effects are robust to controlling for a range of individual, household, and community characteristics. We also discuss the successful experience with smartphone-based data collection, including the fact that the drop-offs among the selected participants were limited; that there was high, sustained participation in terms of logged activities and time; and that there was increasing proficiency in the use of the app throughout the study period. These findings can help accelerate the validation of scalable, improved approaches to individual-level data

collection on time use, including efforts to scale up real-time data collection on time use through different approaches.[5]

The paper is structured as follows. Section 2 discusses the study design. Section 3 provides an overview of the *Time Tracker* and metrics on performance in the field. Section 4 compares time allocation across the smartphone and 24-hour recall diary arms, as well as a comparison with time reported in the 7-day recall module administered to both arms. Section 5 presents additional details on intraday variation in reporting and simultaneity of activities across treatment arms. Section 6 concludes with a discussion on future research needs and thoughts on scaling-up smartphone-based time use data collection in large-scale surveys in lower-income settings.

2. Study design

The Malawi Methodological Experiment on Time Use Measurement was implemented from July 2022 to March 2023, across 72 enumeration areas (EAs) that were selected across the nine most populous districts in Southern and Central Malawi, with probability proportional to 2018 Population and Housing Census EA-level household counts. The study period covers part of the dry season, which ends in October, and most of the rainy season, which goes from November to April, in order to capture seasonality in time use as much as possible.

Furthermore, the EA sample was split evenly across Southern and Central Malawi, and by urban/rural status. In each EA, a full household listing was conducted to identify the households that had at least one adult man and one adult woman (working age 15-64). From this universe of households, five households were selected at random for each of the two treatments on time use data collection: (1) an Android-based smartphone diary with the pictorial *Time Tracker* app, self-reported by respondents, and (2) a traditional time use diary with a 24-hour recall, administered by interviewers as part of a face-to-face interview. Within each household, following the administration of the household roster module of the household questionnaire, one adult man and one adult woman were randomly selected as the study participants.

This design yields a target sample of 720 households and 1,440 participants, split evenly by treatment arm and gender. However, due to app performance issues in the first two weeks of data collection (i.e., the first five EAs) that were sorted out swiftly but that led to a partial loss of smartphone data (i.e., for a subset of reported days for the respondents residing in these EAs), the number of sampled households for the smartphone arm was increased by 1 in approximately 25 percent of the remaining EAs to hedge against potential risk of data loss.[6] However, the feared data loss did not end up materializing, thanks to the

[5] This includes other approaches to high-frequency data collection; see McCullough et al. (2023) who implement a high-frequency phone survey on time use in Ethiopia.

[6] There was a total of 19 EAs in which six, as opposed to five, households were selected at random for the smartphone arm to hedge against the potential risk of data loss, following the experience with the first five EAs. Out of the 19 EAs, 11 EAs were visited immediately after the first five EAs, in August-September 2022. Based on the field experience and the incoming data from these 11 EAs in which there was ultimately no data loss, we reverted to sampling five households in each EA for the smartphone arm. Later, in January 2023, we had a single failure in data download from a smartphone on day 5, which we were able to resolve after

updates to the app and the revision to the data retrieval protocol such that the data were successfully downloaded from and subsequently cleared from each smartphone at every check-in. Table 1 provides a breakdown of the final household sample by treatment arm and across the study districts.

Table 1. Distribution of households across districts, by treatment arm

	Smartphone (%)	24-hour recall (%)
Central Malawi		
Nkhotakota	0.03	0.03
Lilongwe	0.13	0.13
Dedza (urban)	0.07	0.07
Ntcheu	0.04	0.05
Lilongwe City (urban)	0.23	0.23
Southern Malawi		
Machinga	0.06	0.06
Zomba	0.06	0.05
Blantyre	0.03	0.03
Thyolo (urban)	0.07	0.07
Mulanje	0.06	0.06
Zomba City (urban)	0.03	0.02
Blantyre City (urban)	0.21	0.21
Number of households	**377**	**358**

Figure 1 presents the structure of the study over the 11-day reporting period, including the visits made by interviewers to households in each treatment arm, and the administration of the 7-day recall module at the end of the survey period (day 11). As part of the study, and to inform a separate line of work in this program (Hocuk et al., in progress) respondents also wore a ActiGraph wGT3X-BT physical activity tracker over the reporting period and had height and weight measurements taken (informed by prior fieldwork in Malawi in 2017 - see Friedman et al., 2023), to inform ongoing research to develop machine learning models to reliably predict time use patterns from physical activity data and to more precisely identify physical activity demands of time use activities, across gender and different age groups.

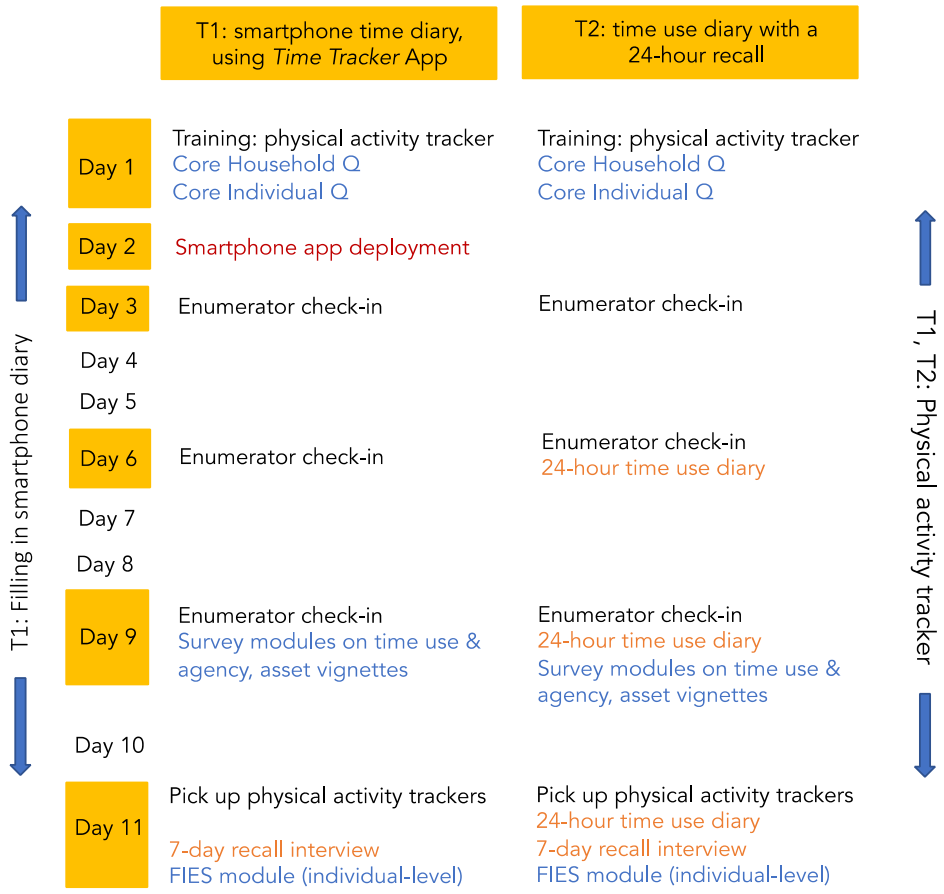
In the smartphone treatment arm, an Android smartphone (Samsung Galaxy A12) was provided to each respondent for the duration of data collection. The smartphones were restricted to only allow respondents to use the *TimeTracker* app on the device.[7] Sixty smartphones were reconditioned and redeployed throughout the fieldwork as survey teams moved from one enumeration area to another, and the teams were equipped with external batteries to recharge the respondents' phones, if needed. As discussed in detail in Section 3, within the *TimeTracker* app, respondents tapped on the images of activities (accompanied by captions in Chichewa) to begin recording time for those specific activities, and tapped the images again to stop recording time when they were done. On the day of deployment (Day 2), interviewers spent time guiding respondents on the use of the app. With the aid of paper illustrations, the interviewers provided an overview of all activities as depicted in *TimeTracker*. Respondents were taught how to navigate

extensive troubleshooting. This incident did, however, stoke fears of data loss, as such we again increased the number of sampled households for the smartphone arm by one household in 8 EAs that were subsequently visited. Overall, 19 EAs were spread across eight out of 9 districts that were covered by the study, and among the 19 EAs, the rural/urban split was 12/7.

[7] Respondents also received a small pouch with a neck strap. A solar charger was also provided to each household in the smartphone treatment arm.

across the activities on the phone and how to press start and stop a given activity and were given time to practice. It was also emphasized that the recording needed to be continuous over the study period. After the initial orientation on day 2, the interviewer check-ins for the smartphone arm were conducted on days 3, 6, 9 and 11 to reinforce the basic tenets of the training provided on Day 2; download and review the time use data from respondents' devices; identify and discuss potential anomalies in the downloaded data; and answer questions/concerns from the respondents.[8]

Figure 1. Overview of field organization and visits to different treatment arms



In the 24-hour recall diary arm, the interviewers visited the respondents three times during the study period (on days 6, 9 and 11 in each EA) to elicit their activities during the previous day, in between the time respondents woke up and when they went to sleep — during which the interviewer used the "time stamps" revealed by the respondent over the course of their conversation to code responses into the diary that was

[8] On the first check-in, the enumerator revisited in depth the entries between the time the enumerator left the household on day 2 and when they returned on day 3. This was done together with each respondent to make sure that he/she understood (a) the icons that are corresponding to each featured activity; (b) the idea of stopping each activity after starting it; and (c) the ability to record simultaneous activities. In this respect, the enumerator/supervisor was able to reinforce the training that was administered on day 2. On the subsequent check-in days, the enumerator and respondent also reviewed the entries to make sure that any anomalies (e.g., activities that were forgotten to be stopped) were addressed using the admin function on the app only accessible to the enumerator – see section 3.2 for a discussion. Any other queries that the respondents had were also addressed.

disaggregated into 15-minute intervals.[9] The recall diary (see Appendix Figure A1) is an adapted version of the time use module that had originally been developed as part of the IFPRI Women's Empowerment in Agriculture Index. A similar module has also been introduced in several country contexts by the International Labor Organization (ILO) as an optional add-on module for labor force surveys (ILO, 2023). The same set of activity categories, including activity wording, were used across treatment arms to allow for one-to-one comparisons in the analysis — covering 56 different categories (see Appendix Table A1) that spanned agriculture, non-farm employment, unpaid domestic and care work, transport, leisure, schooling, personal care, sleeping and eating.

Finally, the 7-day recall module that was administered to both treatment arms on day 11 (see Appendix Figure A2) spanned a subset of 20 key activities that were selected out of the aforementioned 56 activities (i.e., excluding activities pertaining to transport, personal care, eating/drinking, and sleep). The design of the module was informed by similar stylized questions that have been included in labor force surveys and socioeconomic household surveys. The respondents were first asked whether they conducted a given activity in the last 7 days. Then, regarding each activity that they were involved in, respondents were asked to report the number of hours they were engaged in that activity in the last 7 days.

The study also allows for the examination of how each survey approach affects reporting of simultaneity - in terms of participation, combinations of activities, and time spent conducting different activities together. The two treatment arms accounted for simultaneity of activities in different ways. The smartphone app, for example, allowed respondents to select multiple activities they were involved in at the same time, and click off each activity as it was finished. The 24-hour recall diary, on the other hand, first asked respondents what their main activity was for a given time interval, and then whether any secondary activity was conducted simultaneously with it — so that only two activities could be recorded simultaneously for a given 15-minute interval, and that the minimum time activities could be conducted together was 15 minutes. As seen in Appendix Figure A1, the 24-hour recall diary also included a preface to help encourage respondents to think of these secondary activities, including childcare.

As seen in Figure 1, additional survey modules were also implemented for both treatment arms on Day 1 (household and individual questionnaires eliciting a wide range of socioeconomic and demographic attributes) and on Day 9 (on respondents' agency and preferences as well as perceived norms over asset ownership). There was gender matching of respondents and interviewers to help reduce potential unit and item non-response. Table 2 provides means of key individual and household attributes across treatment arms, across education, employment, and ownership of and rights to different assets such as land, financial accounts, and durables.[10] Balance tests across showed that the randomization worked successfully, except for few statistically significant, albeit small, differences across treatment arms.

[9] As a result, while the recall diary organizes the respondent's reported time over the past day into 15-minute intervals, the interviewer does not ask the respondent to report on each 15-minute interval directly. Rather, the interviewer infers this information from the respondent's narrative.

[10] The design of the modules on ownership and rights to land, financial accounts and durables follows recent international guidance on individual-disaggregated asset ownership and rights (United Nations, 2019).

Table 2. Descriptive statistics, by treatment arm

	24-hour recall	Smartphone diary		24-hour recall	Smartphone diary
Demographics			**Employment**		
Household roster: HH head	0.41	0.41	Not employed	0.19*	0.16*
Household roster: spouse of HH head	0.38	0.37	Employed in agriculture	0.33	0.36
Age: 18-24	0.25	0.24	Employed in non-agriculture	0.48	0.49
Age: 25-34	0.26*	0.32*	Other employment characteristics:		
Age: 45-54*	0.12	0.11	Time (mins) to work	32.0**	27.4**
Age: 55+	0.06**	0.04**	Average monthly take-home profits in kwacha (self-employed)	32,799.8	36,813.1
Married (monogamous), patrilineal	0.29	0.33	Average daily earnings in kwacha (wage)	2,350.8*	2,950.5*
Married (monogamous), matrilineal	0.41	0.37	Average daily earnings in kwacha (*ganyu*)	4,421.6	5,095.3
Separated/divorced	0.04	0.04			
Single	0.19	0.19	**Ownership of assets**		
Months resp. is away from HH in last year	0.50	0.49	Owns land, exclusively	0.25	0.23
			Owns land, jointly	0.80	0.78
Education			Has an account: mobile money	0.55	0.58
Highest level of education: primary	0.61	0.59	Has an account: informal savings	0.13	0.15
Highest level of education: secondary or above	0.32	0.34	Has an account: formal fin. institution	0.10	0.09
			Has a mobile phone	0.67	0.68
Household characteristics			Owns a bicycle	0.28	0.28
Number of children <12 years	1.71	1.67	Owns a motorized vehicle	0.05	0.08
Household size	5.04	4.94			
HH has piped water	0.18	0.15	**Number of respondents**	**710**	**753**
HH has electricity	0.25*	0.29*			

Notes:
(1) T-tests of equality of means conducted across treatment groups; ***p<0.01, **p<0.05, *p<0.10.

3. *Time Tracker* Android smartphone app

3.1 Overview

The *Time Tracker* app allows for a continuous, real-time recording of time use that someone carries out during the day. The main screen (Figure 2a) shows respondents easy-to-understand illustrations of typical daily activities, that as mentioned earlier, also correspond one-to-one with activities in the 24-hour recall instrument.[11] Separate versions of the app with images of men and women conducting activities were provided to men and women respondents, respectively, to ensure that certain activities often linked with gender norms (such as household work, driving, and certain forms of employment) were not necessarily under-counted by either women or men. Scrolling activity description in Chichewa were also included at the top of each image to aid in the correct identification of each activity.[12]

[11] The illustrations were customized to different activities, and went through a process of editing and piloting before integrating in the *Time Tracker* app. The app is structured in a way that the illustrations can be replaced easily in order to adapt to different country contexts.

[12] Appendix Table A1 provides the activity list for the smartphone app, along with the pop-ups — as well as the Chichewa captions that accompanied the images.

To begin logging an activity, the respondent taps on an image, initiating a timer, which stops when the respondent taps the image again to stop the activity. The app records the duration of each activity logged. Respondents were also able to select multiple activities at the same time if they were conducting any activities simultaneously. The app also allows for pop-ups (Figure 2b) to collect additional contextual data – in our case, we included pop-up questions to identify what types of food the respondent ate; where and with whom each activity was conducted; and how satisfied the respondent was as a result of performing a particular activity, as a measure of perceived quality of time (Seymour et al., 2020). In particular, as discussed in Section 5, the pop-ups on whom activities were conducted with — including older and younger children — allow a view, albeit imperfect, on potential supervisory or passive care (Folbre, 2021).

The *TimeTracker* data are exported as Excel files from the application to a laptop via Bluetooth or a local Wi-Fi network, and the Excel files are set up with each row including the description of each activity that was recorded by the respondent, along with the columns for the start date, the start time, the activity duration, and the answers to the pop-up questions.

The app also allows for quality checks and editing during the interviewer check-ins. To preserve the real-time aspect of data collection, for example, the app did not allow the respondents to go back and change the activity entries after logging/stopping them. However, as discussed in the next section, only the interviewers could make adjustments in the app if respondents indicated they had erroneously included or omitted entries. These cases were handled at each check-in and auto-labelled by the app as "administrative" and were part of the exported Excel database. The interviewers could also "delete" entries that the respondents erroneously entered – while these entries were kept in the exported database, they are not included in our analyses.

Figure 2. Overview of *Time Tracker* Android smartphone app

(2a) Main screen

Pictured: agricultural activities

Running timer when activity is selected

(2b) Pop-up windows

Pop-up: Food

Pop-up: whom activity conducted with

Pop-up: satisfaction

3.2 Field experience with smartphone-based data collection

The smartphone sample includes a total of 160,880 recorded activities across the respondents, over 9 full days of data collection between deployment (day 2) and pick up (day 11). In what follows, we present specific metrics related to the field experience with smartphone-based data collection. One concern was that respondents might leave activities on too long after pressing them. Ultimately, only 2 percent or fewer entries within an activity category in each day had reported time that was three standard deviations above the daily mean, with no variation over the reporting period. Of all recorded activities, 6.9 percent (or 11,053) were "administrative", i.e., needed adjustment by the interviewers at the time of the check-ins. The comparison of reported time estimates as a result of including versus excluding these administrative entries showed that the administrative adjustments were primarily for recording sleep time.[13]

Figure 3 presents additional metrics on smartphone performance over the reporting period. We exclude the ends of the reporting period (days 2 and 11), since day 2 was when phones were dropped off and considered more of a "training" period on using the smartphone app, and day 11 was when smartphones were picked up. As a result, Figures 3a-3f are informed by a total of 138,609 activities across days 3-10. Some decline is observed in the number of participating respondents over the study period, due mainly to respondent travels out of EAs and their phones being collected early as a result (Figure 3a). However, among reporting respondents, the number of diary entries per day (Figure 3b), as well as average total daily reported time (Figure 3c), were stable over the period (between days 4 and 9 in particular). Notably, the share of activities lasting less than 15 minutes in length was quite substantial for both men and women - ranging around 30 percent for men and 26 percent for women between days 4 and 9, which we will revisit in Section 4. Moreover, over the study period, the respondents improved the accuracy with which they recorded their time in the app - with fewer discontinuities (Figure 3e) as well as entries that required editing by the interviewers (i.e., administrative or deleted entries) (Figure 3f). While the time logged and number of recorded activities were not significantly different across men and women, we observe significant gender differences in Figures 3e and 3f, which can be an important consideration for future surveys and training of different groups of respondents on the app.

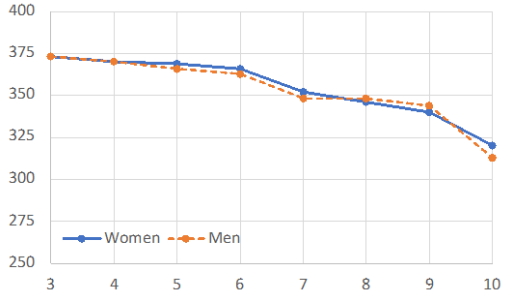
Appendix Table A2 also presents regressions on additional correlates of outcomes in Figures 3e and 3f at the individual-day level, controlling for enumeration area and day of the week fixed effects. The findings reinforce that women were significantly more likely to have entries with time jumps or that required editing, and older participants (55+) as well as participants in non-agricultural employment were also more likely to have administrative or deleted entries. The respondents with secondary or higher education were less likely to have time jumps (as measured in Figure 3e) or needing editing in their smartphone logs. Having a bicycle and electricity in the household were less likely to have entries that needed editing later on.[14]

[13] The results are available upon request. The qualitative evidence from the field was that there were respondents who were disturbed by the phone light or were uncomfortable with the idea of leaving the phone on overnight. As such, they did not log sleep time and left the phone off while they were sleeping.

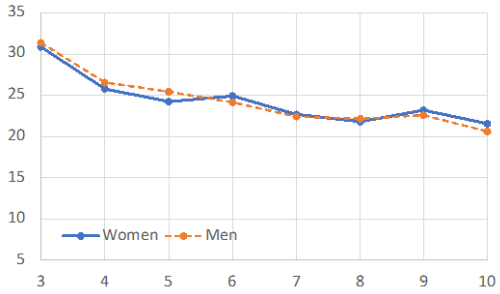
[14] Interestingly, mobile phone ownership did not have a significant effect, but this could also be because many phones owned in the sample may not have been smartphones.

Figure 3. Field experience with smartphone-based data collection over days 3-10 [(1)]

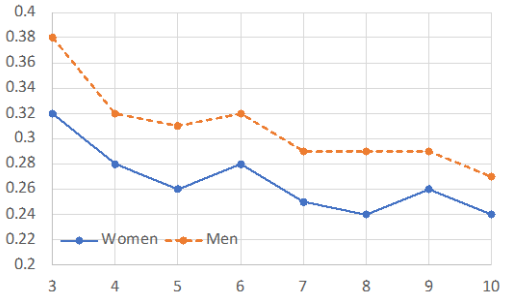
(3a) Number of respondents reporting activities



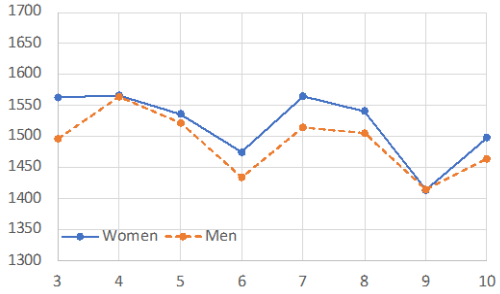
(3b) Number of diary entries per day



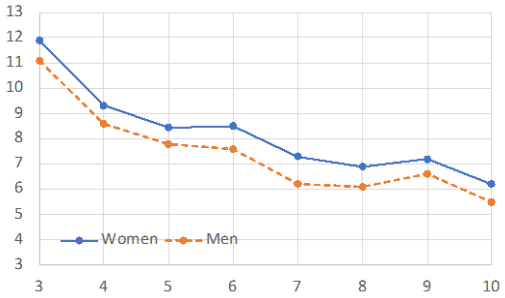
(3c) Share of diary entries that are less than 15 mins long



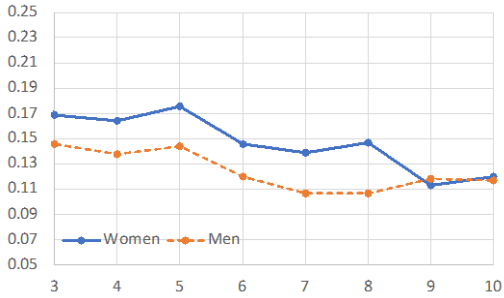
(3d) Average total daily time reported (minutes)



(3e) Number of "jumps" or discontinuities in time (forgetting to log activities) between consecutive activities



(3f) Share of activities with "admin" or "deleted" entries (i.e., respondent needed interviewer's help In adding entries, or removing erroneous ones)



Notes:

(1) For (3e), differences between men and women were statistically significant on Day 3 ($p<0.10$), Days 6 and 8 ($p<0.05$), and Days 7, 9, and 10 ($p<0.01$). For (3f), differences between men and women were statistically significant on Days 4-8 ($p<0.01$). The results from the regressions investigating the additional determinants of (3d) and (3e) are in Appendix Table A2.

4. Daily time allocation across the smartphone, 24-hour recall and 7-day recall approaches

4.1 Time allocation and treatment effects across smartphone and 24-hour recall arms

As shown in Figure 1, the recall treatment was administered at three times during the weeklong survey period for each respondent — day 6 (with the recall diary corresponding to the prior 24 hours, i.e., day 5); day 9 (corresponding to time spent during day 8); and day 11 (corresponding to time spent during day 10). As the smartphone data were collected continually, the analysis in this section presents comparisons of daily time allocation in the smartphone arm matched by day with the recall arm, to account for any potential factors affecting day-to-day changes in time allocation. The results for the aggregated activity categories are presented across (1) agriculture; (2) nonfarm employment; (3) unpaid domestic and care work (obtaining services/resources for the household; cooking and cleaning; and care of children or elderly); (4) transport; (5) schooling/studying; (6) leisure; (7) community activities; (8) personal care (bathing, dressing, care for one's own health issues); (9) sleep; and (10) eating/drinking. The descriptive statistics on the disaggregated activities underlying these categories are presented in the Appendix Tables A3 and A4.

Taking the day 5 comparison as an example,[15] Table 3 shows that the smartphone arm reports significantly higher participation across several activity categories in agriculture, non-agriculture (women), unpaid domestic and care work for the household, schooling/studying, and community activities. This included higher unpaid care work among women in the smartphone arm, but also greater participation in different areas of unpaid work among men. The recall arm, on the other hand, yielded significantly higher participation rates in self-care, such as leisure, personal care, sleeping and eating. We turn to potential reasons for these differences in Section 5, including greater reporting across a range of activities conducted throughout the day in the smartphone arm, as well as multitasked activities. Table 4 also shows that greater participation in employed and unpaid work among men and women in the smartphone arm leads to narrower, albeit still substantial, gender gaps for participation in most of these activities, with a slightly greater gap for care work — 41 percent, compared to 37 percent in the recall arm.

On unconditional time, however, owing likely to the structure of the 24-hour recall module where time devoted to a given activity was recorded for a minimum of 15 minutes, and the fact that smartphone arm respondents recorded large shares of activities for less than 15 minutes (Table 5), the reported daily time in minutes is higher across activity categories in the recall arm, except for transportation. Along with lower participation among men in employed and unpaid work in the recall arm, this leads to higher time-related gender gaps for these activities (Table 4). Kernel density estimates in Figure 4, based on the pooled matched sample of days 5, 8, and 10, also show much greater increases in unconditional time in the recall sample for areas such as leisure, and more so for men. Only unconditional time in transport and some areas of men's unpaid work remain higher in the smartphone arm, indicating that these activities are even further underestimated in the recall arm, given the minimum 15-minute floor for activities in the 24-hour diary.

[15] Appendix Figure A3 shows that there is very little variation when averaging time in activity categories across days 5, 8, and 10, for both men and women.

The gender gaps in time allocation remain quite substantial in the smartphone arm for nonfarm employment, unpaid domestic and care work, transport, and leisure, but are magnified in the recall arm due in large part to the 24-hour diary reporting structure.

Furthermore, Table 5 shows that that the 15-minute interval setup in the 24-hour recall diary is at odds with real-time recorded activity duration across several key areas. For instance, 14-17 percent of smartphone diary entries in agriculture and non-farm employment were recorded for less than 15 minutes. The comparable estimates were 26-29 percent for care work, 39-44 percent for obtaining services and resources for the household, and around 40 percent for eating and drinking. Men were also significantly more likely than women in the smartphone arm to log activities for less than 15 minutes, particularly for non-farm employment, unpaid domestic work, and self-care activities.

Table 3. Participation and time (minutes) in activities, by treatment arm (Day 5 match)

			Unpaid work:										
	Agri-culture	Nonfarm employ-ment	Obtaining services/ resources for HH	Cooking, cleaning in house	Care of children, elderly	TOTAL UNPAID WORK	Transport	School/ studying	Leisure	Comm. Activities	Personal care	Sleep	Eating/ drinking
	(1)	(2)	(3a)	(3b)	(3c)	(3d)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
SHARE PARTICIPATING													
Total													
Recall (obs=696)	0.33***	0.33*	0.49***	0.65	0.27***	0.76***	0.94	0.04***	0.99**	0.04***	0.94***	1.00***	0.99**
Smartphone (obs=734)	0.41***	0.37*	0.63***	0.67	0.34***	0.84***	0.94	0.09***	0.98**	0.14***	0.89***	0.98***	0.96**
Women													
Recall (obs=352)	0.35*	0.23*	0.68	0.98*	0.45***	0.98	0.91	0.03**	0.99	0.05***	0.96***	1.00**	0.99**
Smartphone (obs=365)	0.42*	0.29*	0.72	0.95*	0.55***	0.97	0.91	0.06**	0.97	0.14***	0.90***	0.99**	0.96**
Men													
Recall (obs=344)	0.31***	0.42	0.30***	0.32**	0.08**	0.53***	0.97	0.05***	0.99	0.03***	0.92	1.00***	0.99***
Smartphone (obs=369)	0.41***	0.45	0.53***	0.40**	0.14**	0.71***	0.96	0.11***	0.98	0.14***	0.88	0.97***	0.96***
DAILY TIME, UNCONDITIONAL (MINS)													
Total													
Recall (obs=696)	80.2***	120.2***	39.5	129.7	57.3***	224.3	128.1***	13.9	507.9***	5.3	33.6	536.8	54.8**
Smartphone (obs=734)	53.0***	84.7***	45.9	138.2	38.7***	222.8	162.0***	11.7	374.4***	7.8	34.9	529.3	44.3**
Women													
Recall (obs=352)	74.6***	52.5	50.5	231.5	100.8**	381.6	104.9**	13.0	426.9***	5.5	38.1	549.3	51.6
Smartphone (obs=365)	48.1***	47.8	46.2	240.5	69.6**	356.3	118.7**	6.8	320.1***	6.4	38.3	546.2	48.2
Men													
Recall (obs=344)	85.9***	189.3***	28.4**	25.6*	12.9	65.7**	151.9***	14.9	590.8***	5.0*	28.9	524.0	58.1**
Smartphone (obs=369)	57.8***	121.3***	45.6**	36.6*	8.1	90.4**	204.9***	16.6	428.3***	9.2*	31.6	512.5	47.8**

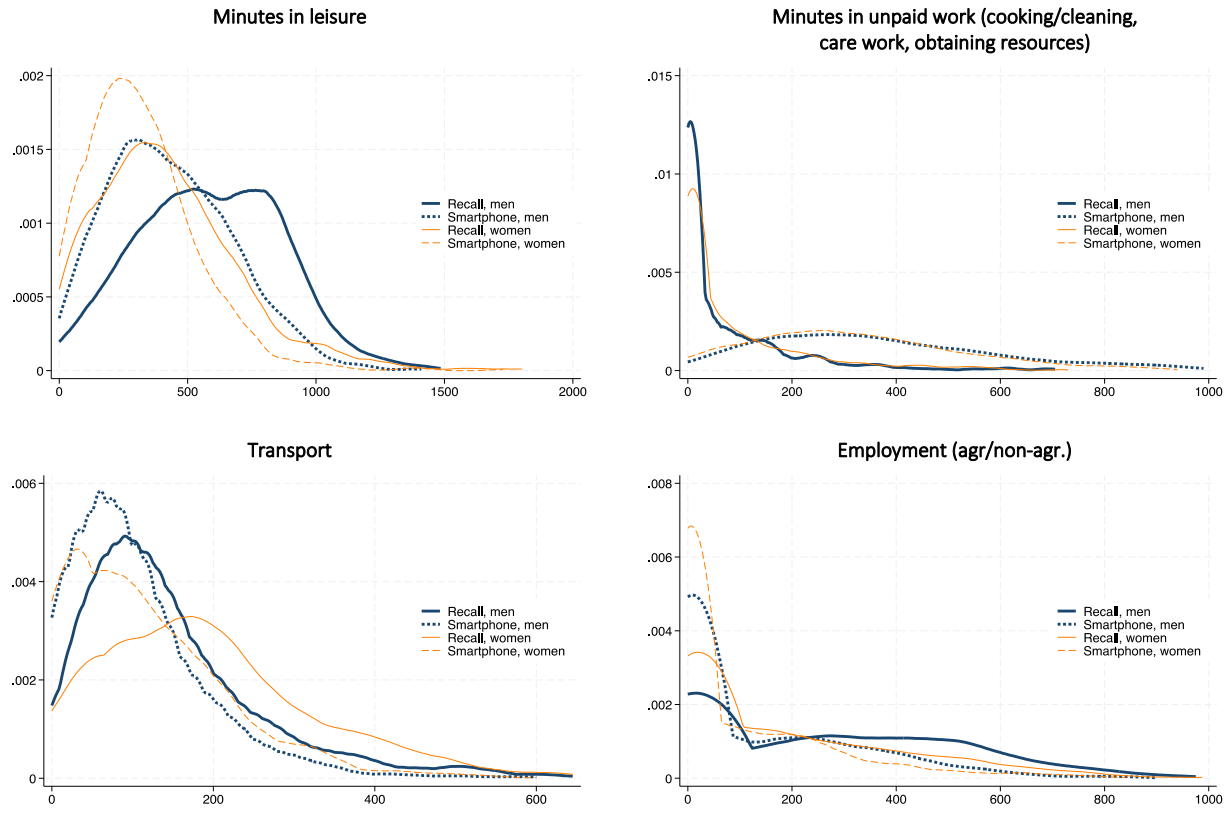
Notes:
(1) Within the total, women's and men's samples, T-tests of equality of means conducted across treatment groups; ***p<0.01, **p<0.05, *p<0.10. **Grey shading =** recall arm yields significantly lower estimates. (2) Average total time for women and men was 1,700 minutes in the recall sample (including main and secondary activities). In the smartphone sample, average total daily time was 1,521 minutes for women, and 1,490 minutes for men.

Table 4. Gender gaps (women – men) in activities, by treatment arm (Day 5 match)

			Unpaid work:										
	Agri-culture	Nonfarm employ-ment	Obtaining services/ resources for HH	Cooking, cleaning in house	Care of children, elderly	TOTAL UNPAID WORK	Transport	School/ studying	Leisure	Com-munity Activities	Personal care	Sleep	Eating/ drinking
	(1)	(2)	(3a)	(3b)	(3c)	(3d)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
SHARE PARTICIPATING													
Gender gap (women-men)													
Recall	0.04	-0.19	0.38	0.66	0.37	0.45	-0.06	-0.02	0	0.02	0.04	0	0
Smartphone	0.01	-0.16	0.19	0.55	0.41	0.26	-0.05	-0.05	-0.01	0	0.02	0.02	0
DAILY TIME, UNCONDITIONAL (MINS)													
Gender gap (women-men)													
Recall	-11.3	-136.8	22.1	205.9	87.9	315.9	-47.0	-1.9	-163.9	0.5	9.2	25.3	-6.5
Smartphone	-9.7	-73.5	0.6	203.9	61.5	265.9	-86.2	-9.8	-108.2	-2.8	6.7	33.7	0.4

Notes:
(1) The gender gap takes the difference in means for women and men in each treatment arm.

Figure 4. Kernel density estimates for daily minutes in key activity categories, by treatment arm and gender



Notes:
(1) For all graphs: p-values from Kolmogorov-Smirnov test of distributional differences (across treatment arms): 0.000
(2) Based on the pooled sample of days 5, 8, and 10.

Table 5. Shares of smartphone diary entries lasting for less than 15 minutes in each aggregate activity category, by gender

	Women	Men
Agriculture	0.17	0.16
Nonfarm employment	0.16*	0.14*
Unpaid work:		
Obtaining services/resources for HH	0.39***	0.44***
Cooking, cleaning in house	0.12***	0.20***
Care of children, elderly	0.26	0.29
Transport	0.41	0.42
School/studying	0.11	0.11
Leisure	0.24***	0.26***
Community activities	0.15	0.15
Personal care	0.30***	0.40***
Eating/drinking	0.38***	0.42***

Notes:
(1) Based on 138,609 activities recorded across Days 3-10 of the survey period (data from deployment on Day 2 were dropped).
(2) T-tests of equality of means conducted across men and women.
***p<0.01, **p<0.05, *p<0.10. Sleep was excluded from the table.

Table 6 presents individual-level, ordinary least squares (OLS) regressions of treatment effects of the recall arm on daily time allocation (unconditional minutes),[16] for the pooled matched sample across days 5, 8, and 10:

$$y_{ijk} = \alpha + \beta Recall + \gamma X_{ij} + \delta_1 A_k + \delta_2 D_{ijk} + \varepsilon_{ijk} \qquad (1)$$

In Equation (1), y_{ijk} represents average daily time in each of the activity categories in Table 3, for individual *i* in household *j* and enumeration area *k*; *Recall* is the dichotomous variable that is equal to 1 if the individual was part of the recall treatment; *X* is a vector of individual- and household-level characteristics, as included in Appendix Table A2; *A* and *D* represents the fixed effects for enumeration areas and days of reporting, respectively; and ε is the error term. The regressions were run separately for the total, men, and women samples, along with an additional disaggregation by urban/rural, in view of geographic differences in employment and transport.

The results in Table 6 reveal similar patterns observed earlier; participation is lower in the recall sample for economic activities in agriculture, unpaid work (especially care work), studying, community activities – but higher in the recall arm for leisure, personal care, sleep, and eating/drinking, as well as some unpaid domestic activities for women, focused mainly on cooking and cleaning. The reported time was lower in the recall arm concerning (a) agricultural activities of both men and women in urban areas and (b) non-farm employment of urban women and rural men. The negative recall treatment effects on participation in unpaid domestic and care work stemmed mainly from rural areas, with some significant negative effects for men in urban areas on care work and obtaining services/resources for the household. Despite this variation in participation, however, treatment effects on time spent in different activities are generally higher in the recall arm, except for time in transport and a subset of unpaid domestic activities for men

[16] Descriptive statistics on conditional minutes are in Appendix Table A5.

(particularly for cooking and cleaning in rural areas). One takeaway is that where there is interest in areas around employment, unpaid work, transport and other non-leisure activities, the recall arm provides greater estimates of reported time, but among a smaller, less representative subset of participating individuals.

Table 6. Treatment effects of 24-hour recall diary on participation and daily time allocation

			Unpaid work:										
	Agri-culture	Nonfarm employ-ment	Obtaining services/ resources for HH	Cooking, cleaning in house	Care of children, elderly	TOTAL UNPAID WORK	Transport	School/ studying	Leisure	Com-munity Activities	Personal care	Sleep	Eating/ drinking
	(1)	(2)	(3a)	(3b)	(3c)	(3d)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PARTICIPATION (Y=1 N=0)													
All areas													
Total (obs=4,108)	-0.06***	-0.01	-0.08***	-0.03*	-0.04**	-0.02	0.01	-0.03***	0.02***	-0.07***	0.05***	0.04***	0.04***
Women (obs=2,049)	-0.06***	-0.03	0.00	0.03***	-0.04*	0.06***	0.01	-0.03***	0.02***	-0.05***	0.06***	0.02***	0.03***
Men (obs=2,059)	-0.06***	0.00	-0.16***	-0.09***	-0.04**	-0.11***	0.01	-0.02	0.02***	-0.09***	0.05***	0.06***	0.05***
Rural													
Total (obs=2,081)	-0.04	-0.03	-0.13***	-0.08***	-0.07***	-0.06***	0.00	-0.03**	0.01*	-0.09***	0.05***	0.04***	0.04***
Women (obs=1,042)	-0.04	-0.00	-0.06*	0.03*	-0.09**	0.05***	0.01	-0.04***	0.01	-0.06***	0.08***	0.03***	0.03***
Men (obs=1,039)	-0.03	-0.06**	-0.21***	-0.19***	-0.04**	-0.17***	-0.01	-0.02	0.01	-0.11***	0.03	0.05***	0.04***
Urban													
Total (obs=2,027)	-0.08***	-0.00	-0.03	0.04*	-0.02	0.01	0.01	-0.03*	0.03***	-0.05***	0.06***	0.04***	0.05***
Women (obs=1,007)	-0.07***	-0.07**	0.06*	0.03**	-0.01	0.07***	0.01	-0.03	0.03***	-0.04**	0.04**	0.02***	0.04***
Men (obs=1,020)	-0.10***	0.05	-0.12***	0.01	-0.04*	-0.06*	0.01	-0.02	0.03**	-0.06***	0.07***	0.06***	0.06***
DAILY TIME, UNCONDITIONAL (MINS)													
All areas													
Total (obs=4,108)	26.2***	28.8***	0.1	-5.3	18.4***	23.8***	-30.8***	6.1**	141.2***	-2.3*	-2.6	11.3**	6.6***
Women (obs=2,049)	18.7***	8.8	6.3**	-0.1	32.1***	53.5***	-11.4**	3.8	108.2***	-0.3	-3.0	10.8*	3.4*
Men (obs=2,059)	34.8***	49.4***	-5.6	-13.1***	4.5*	-9.6	-50.9***	8.9*	170.1***	-3.9*	-2.7	14.0*	10.1***
Rural													
Total (obs=2,081)	55.5***	14.4*	1.3	-8.7	23.9***	29.5***	-34.5***	3.1	106.6***	-2.7	3.0	19.7***	11.0***
Women (obs=1,042)	38.7***	4.7	1.9	5.0	43.3***	66.3***	-10.8	-1.2	73.3***	-2.7	4.5	13.5	9.6***
Men (obs=1,039)	76.7***	22.9*	2.4	-17.5***	6.7*	-2.0	-59.1***	6.5	130.4***	-2.2	2.7	28.1***	12.1***
Urban													
Total (obs=2,027)	-0.7	39.7***	-4.1	-2.1	8.7	11.2	-26.9***	10.1*	179.0***	-1.4	-8.0*	2.6	2.4
Women (obs=1,007)	0.3	8.7	9.8**	-7.8	15.1	29.9	-11.3*	8.2	151.4***	2.2	-9.2	9.7	-2.8
Men (obs=1,020)	-3.4	75.4***	-16.7*	-7.7	1.8	-18.5*	-41.7***	11.3	207.5***	-4.9	-6.5**	-1.7	6.7*

Notes:
(1) OLS regressions controlling for EA and day of the week Fes; the standard errors are clustered at the household-level. ***p<0.01, **p<0.05, *p<0.10. **Grey shading =** recall arm yields significantly lower estimates.
(2) Based on a pooled individual-level sample for days 5, 8 and 10 (each individual appeared three times in the sample).
(3) Descriptive statistics on conditional minutes are in Appendix Table A5.

4.2 Comparisons with time use data based on stylized questions with a 7-day recall

In this section, we compare average daily minutes based on stylized questions on time use with a 7-day recall (obtained by dividing the total reported weekly hours by 7) with the average daily minutes under the 24-hour recall and smartphone treatments (computed across days 5, 8 and 10). The comparisons are provided for the activity categories that were common across the three survey approaches. Table 7 shows that the reliance on stylized time use questions with a 7-day recall leads to higher average daily time across

activity categories, including in agriculture, nonfarm employment, unpaid domestic work (particularly among women), and care work for both men and women. A key exception is leisure, which is underreported by several hours per day relative to the smartphone and 24-hour recall arms.

Table 7. Average daily unconditional minutes across activity categories according to 7-day recall, 24-hour recall, and smartphone diary

	Agri-culture	Nonfarm employ-ment	Unpaid work: obtaining services/ resources for HH	Unpaid work: cooking and cleaning in house	Unpaid work: Care of children, elderly	Community activities	School/ studying	Leisure
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Total								
Smartphone arm:								
7-day recall (obs=738)	68.3***	167.8***	49.7**	160.0***	88.3***	6.6	14.7	111.5***
Smartphone diary (obs=735)	52.6***	86.3***	41.9**	133.4***	38.2***	8.4	11.0	366.4***
24-hour recall arm:								
7-day recall (obs=702)	71.9	164.0***	44.1	141.5*	87.0***	6.5	19.1	102.4***
24-hour diary (obs=703)	78.1	115.9***	40.9	128.0*	56.4***	6.3	17.4	506.9***
Women								
Smartphone arm:								
7-day recall (obs=357)	58.8	107.4***	76.9***	273.4***	139.5***	6.8	11.6	117.9***
Smartphone diary (obs=366)	50.3	52.0***	43.1***	230.4***	68.9***	7.5	8.7	314.2***
24-hour recall arm:								
7-day recall (obs=349)	57.1	98.0***	67.3***	247.1*	133.6***	7.4	12.7	99.2***
24-hour diary (obs=353)	68.7	59.1***	47.9***	231.0*	100.3***	6.5	11.4	430.1***
Men								
Smartphone arm:								
7-day recall (obs=365)	79.2***	226.6***	23.9***	48.8**	37.0***	6.7	17.9	105.4***
Smartphone diary (obs=369)	54.9***	120.4***	40.9***	37.2**	7.7***	9.3	13.3	418.2***
24-hour recall arm:								
7-day recall (obs=349)	87.5	229.2***	21.4***	35.7***	39.0***	5.7	25.8	106.1***
24-hour diary (obs=350)	87.6	173.2***	33.7***	24.1***	12.1***	6.1	23.4	584.3***

Notes:
(1) Asterisks are indicated on the smartphone diary and 24-hour recall estimates if they were significantly different, respectively, from the 7-day module. ***p<0.01, **p<0.05, *p<0.10
(2) In the smartphone and recall treatment arms, the average daily minutes were calculated across days 5, 8 and 10, given the focus on these days in the rest of the paper. Computing the average daily minutes estimates for the smartphone arm across all 9 days of reporting yields estimates that are comparable to those computed across days 5, 8 and 10. The 7-day recall estimates are generated by dividing the total reported weekly hours by 7.
(3) Activities (1)-(8) are the common aggregate activity categories across treatment arms; transport was not included in the 7-day recall.

In turn, we estimate the adapted versions of Equation 1 separately for the recall I and smartphone (S) samples, this time by using the 7-day recall data as the comparison (Table 8):

$$y_{ijk}^{R} = \tilde{\alpha} + \tilde{\beta} Recall + \tilde{\gamma} X_{ij}^{R} + \tilde{\delta} A_{k}^{R} + \epsilon_{ijk}^{R} \quad (2a)$$

$$y_{ijk}^{S} = \hat{\alpha} + \hat{\beta} Smartphone + \hat{\gamma} X_{ij}^{S} + \hat{\delta} A_{k}^{S} + \epsilon_{ijk}^{S} \quad (2b)$$

The regressions are run at the individual-level and the dependent variables are the average daily minutes across eight aggregate activity categories that overlapped with the 7-day recall module. The average daily minutes for the 24-hour recall and smartphone treatment arms are computed across days 5, 8 and 10 – though, the estimates for the effect of smartphone treatment are qualitatively similar if the average daily

time is computed across all 9 days of reporting.[17] Table 8 reports the 7-day recall treatment effects.[18] The results, in conjunction with the 24-hour recall and smartphone comparison in Section 4.1, suggest that at higher aggregations of recall time – moving from the last 24 hours to 7 days – the estimated time for key activity areas around employment and unpaid work go up. This is extremely important for understanding individuals' productivity, for example, as well as opportunity costs for re-allocating their time across other areas. The substantial underestimation of leisure time in the 7-day recall module vis-à-vis the smartphone and 24-hour recall approaches, also indicates that the 7-day recall module may be insufficient for understanding the true distribution of daily time across activities.

Table 8. Overall treatment effects of smartphone and 24-hour recall approaches compared to 7-day recall

	Effect of smartphone and recall arms on average daily minutes spent in:							
	Agri-culture	Nonfarm employ-ment	Unpaid work: obtaining services/ resources for HH	Unpaid work: cooking and cleaning in house	Unpaid work: Care of children, elderly	Community activities	School/ studying	Leisure
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Total								
Smartphone arm	-16.1***	-81.9***	-8.2**	-26.5***	-49.5***	1.6	-3.6***	255.6***
24-hour recall arm	6.1	-49.1***	-3.4	-13.8***	-30.1***	-0.1	-2.0	404.4***
Women								
Smartphone arm	-7.8*	-55.9***	-33.9***	-42.0***	-70.1***	0.8	-2.6*	196.6***
24-hour recall arm	11.8**	-40.8***	-19.4***	-16.4**	-32.7***	-0.8	-1.3	331.5***
Men								
Smartphone arm	-24.5***	-107.0***	17.0***	-11.4**	-29.4***	2.6	-4.6**	313.3***
24-hour recall arm	0.3	-57.8***	12.6***	-11.6***	-27.4***	0.5	-2.3	477.7***

Notes:
(1) OLS regressions controlling for EA and day of the week FEs; the standard errors are clustered at the household-level. Regressions control for the same individual and household-level characteristics as in eq. (1). Asterisks are indicated on the smartphone diary and 24-hour recall estimates if they were significantly different, respectively, from the 7-day module. ***p<0.01, **p<0.05, *p<0.10

5. Factors underlying reporting differences across recall and smartphone arms: Intraday variation in time allocation and simultaneity

Apart from average treatment effects, the data collected across both treatment arms in the experimental study allow for a more disaggregated view of time allocation that has been less well-examined in time use surveys. This includes how men and women manage their time within the day — as well as how intensively they use their time, as measured by different activities being conducted simultaneously. As we discuss in this section, some of these nuances can help explain some of the findings we observed in Table 6. For example, despite the 24-hour recall arm's 15-minute floor for time recording, we find that the smartphone arm reveals greater time across leisure and non-leisure activities during potentially productive times of the day. We also find a much higher incidence of activities conducted simultaneously in the smartphone arm, and across a wider range of activities than the recall arm, including non-leisure activities. In so doing, we

[17] The results are available upon request.
[18] For treatment effects, all days between days 3-10 were used in the smartphone sample, and days 5, 8 and 10 were used in the recall arm.

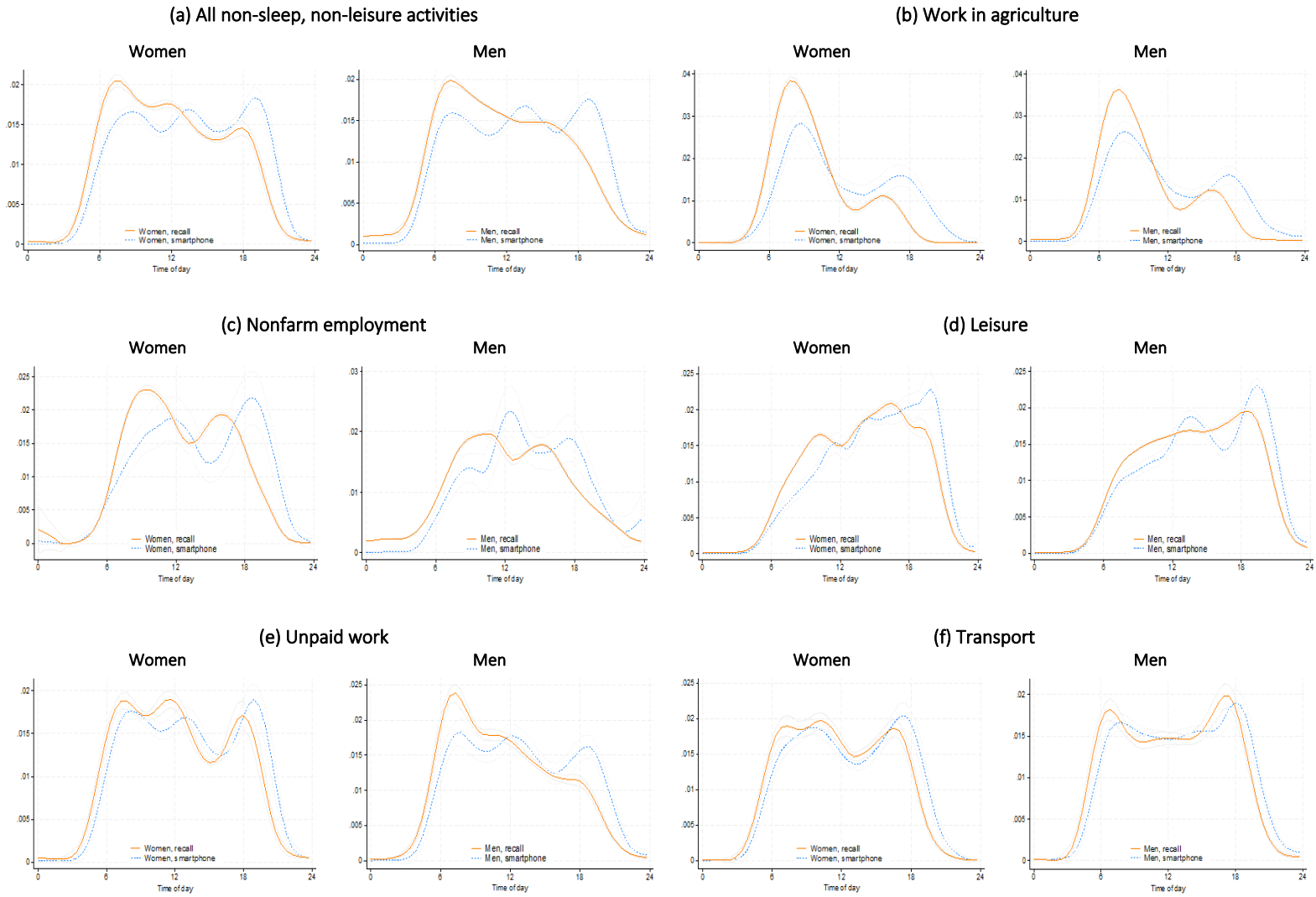
discuss how real-time, smartphone data collection can reveal greater insights on individuals' opportunity costs of time, along with a wider view of different activities they are involved in.

5.1 Intraday variation in time allocation

Across treatment arms, when looking at the intraday distribution of time across key activity categories (presented in Figure 5 by local polynomial smooth plots of the share of activities reported throughout the day),[19] we observe large differences in time allocation across such areas as agriculture, nonfarm employment, leisure, and unpaid work. While non-sleep, non-leisure activity reporting is much higher early in the day for the recall arm, there is a substantial drop-off in reporting for nearly all non-sleep activities by mid-day or early evening around 6 pm. The smartphone arm, on the other hand, reveals much more evenly distributed reporting, with drop-offs towards sleep only much later in the day (around 9 pm). The share of daily time in activities after 6pm accounts for about 26 percent for women and 29 percent for men the smartphone sample—compared to 17 and 19 percent for women and men in the recall sample, respectively—leading to a substantial loss of information when moving to recall-based methods. Furthermore, Table 9 highlights how time in activities such as unpaid work and non-farm employment begins to fall significantly in the recall arm vis-à-vis the smartphone arm after 12 pm, along with agriculture and leisure after 6 pm. Substantial amounts of daily time are spent by smartphone respondents in unpaid work and non-farm employment after 6 pm (17 and 33 minutes in nonfarm employment for women and men, respectively, and 83 and 17 minutes, respectively, in unpaid work). Although the recall arm tends to report higher time in leisure throughout the day, this trend flips after 6 pm as well. Men and women in the smartphone arm also spend nearly twice as much time in transport-related activities after 6 pm, compared to the recall arm. These trends are likely a reason for underreporting of participation across key activities in the recall vis-à-vis smartphone arm (Table 6).

[19] The results are based on the pooled Day 5, 8 and 10 samples. 95 percent confidence intervals are included as well.

Figure 5. Local polynomial smooth plots: share of activities reported at different times of the day



Notes: (1) Local polynomial smooth plots, with 95% confidence intervals. Pooling Days 5, 8, 10 for both treatment arms

Table 9. Minutes spent within the day across key activity categories, by treatment arm

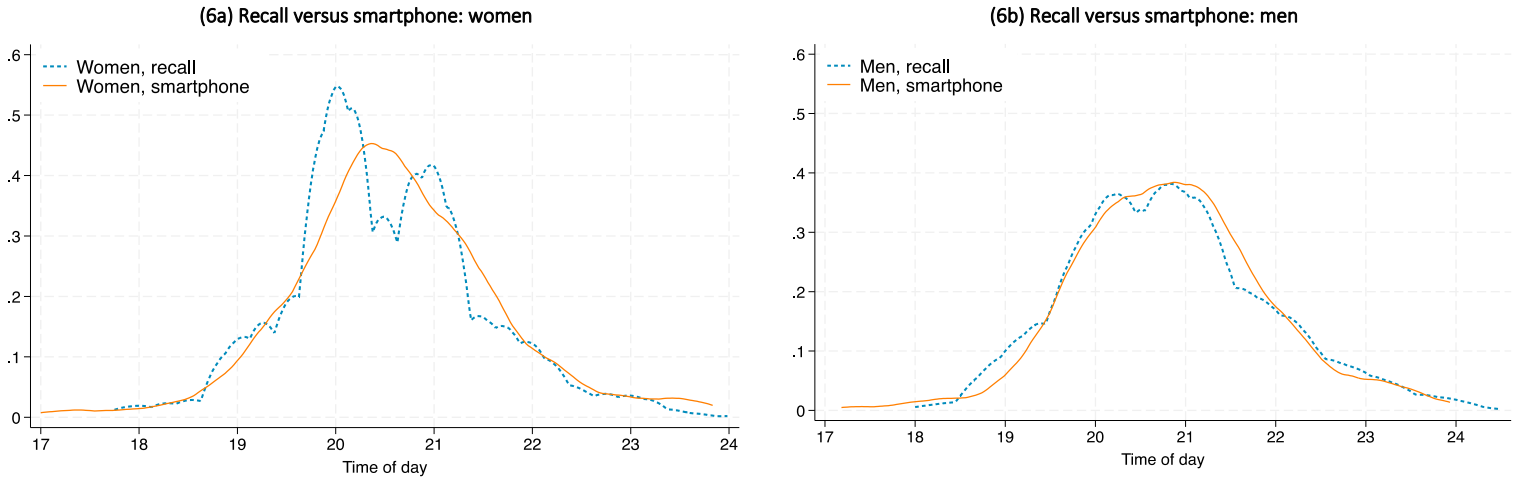
	Total		Women		Men	
	Recall	Smartphone	Recall	Smartphone	Recall	Smartphone
Morning (12am-11:59am)						
Agriculture	59.6***	35.2***	53.1***	35.3***	66.2***	35.1***
Non-agriculture	54.9***	18.1***	28.4***	11.1***	81.7***	24.9***
Leisure	158.3***	75.4***	135.2***	61.7***	181.7***	89.0***
Unpaid work	109.2***	82.6***	177.4***	134.4***	40.0***	31.4***
Transport	62.8	65.7	55.2	52.7	70.4**	78.5**
Eating/drinking	8.4***	6.8***	8.3***	6.6***	8.4**	7.0**
Afternoon/evening/night (12pm-11:59pm)						
Agriculture	18.9	17.9	15.7	16.1	22.3	19.7
Non-agriculture	58.2***	70.1***	29.5**	41.1**	87.2	98.7
Leisure	306.2**	294.1**	267.8	256.6	345.2	331.1
Unpaid work	103.7***	133.3***	176.6***	210.3***	29.8***	57.1***
Transport	64.7***	95.1***	48.7***	64.5***	81.0***	125.3***
Eating/drinking	44.4***	39.9***	42.7	40.6	46.1***	39.2***
Evening/night (6pm-11:59pm)						
Agriculture	1.1***	3.2***	0.7***	2.2***	1.5**	4.1**
Non-agriculture	14.3***	25.1***	6.0***	16.9***	22.7**	33.3**
Leisure	106.8***	133.5***	86.5***	106.3***	127.4***	160.5***
Unpaid work	31.9***	49.6***	56.0***	82.6***	7.5***	17.0***
Transport	13.7***	34.5***	7.0***	20.5***	20.6***	48.4***
Eating/drinking	21.7	21.1	20.7	21.4	22.7*	20.8*

Notes:
(1) T-tests of equality of means conducted across treatment groups; ***p<0.01, **p<0.05, *p<0.10.
(2) Based on a pooled individual-level sample for days 5, 8 and 10 (each individual appeared three times in the sample).
(3) Grey shading = recall arm yields significantly lower estimates.

One possibility for these differences is respondent fatigue in the recall arm, given the structure and relatively long duration of the interview. Respondents might be more actively reporting time in different productive activities at the start of the interview, and by the end of the narrative (i.e., for hours later in the day) they lose interest or track of their time. Figure 6 gives some indication that this may be the case, revealing a bimodal sleep distribution for recall respondents that is particularly pronounced for women. A significantly greater share of recall respondents, particularly women, report earlier sleep times that would conclude the diary interview — 64 percent of women in the recall arm reported sleep times between 5 pm-8:59 pm, versus 54 percent in the smartphone arm. For men, these shares were 50 and 45 percent, respectively. This supports, to some extent, that fatigue may be a key factor. Other factors may be at play as well, however, including potential difficulties among those reporting earlier sleep times to pinpoint in the interview exactly when they went to sleep.[20]

[20] Also see Kaplan et al., 2020, for a discussion on recalling sleep times using the American Time Use Survey.

Figure 6. Kernel density graphs: time of sleep for men and women



Notes:
(1) The estimates pool the Day 5, 8 and 10 samples for each treatment arm. K-S p-values for differences across treatment arms in (6a) and (6b): 0.000.

Therefore, the smartphone arm provides greater variation in intraday reporting and could potentially provide an easier way to avoid issues around understanding time in the day – given how early people are reporting sleep time. If time use statistics have typically been analyzed on a day-level aggregation, however, a natural question arises as to why we should potentially care about intraday time allocation. In addition to our understanding of how men and women organize their time, the distinctions we observed across treatment arms of reporting during morning versus evening hours also have important implications for difficult-to-answer questions around infrastructural interventions affecting time use. This includes, for example, the role of household electrification in supporting time allocation in different productive areas during evening and nighttime hours (see, for example, Agénor and Agénor, 2023; Ribeiro et al., 2021; van de Walle, et. al., 2017). Taking electricity and access to energy sources more broadly, high-quality intraday data on time allocation can help us better understand the role of energy infrastructure on individuals' economic and other intrahousehold decision making. As an example, Appendix Table A6 presents OLS regressions showing that for women in the smartphone arm in particular, electricity access (as opposed to no access, or household lighting through battery-operated torches) are significantly associated with time allocation in nonfarm employment and leisure during the evening/nighttime, but with no effects for women in the recall sample.[21] The choice of survey approach to time use data collection can therefore provide very different perspectives in understanding how time allocation is linked to energy infrastructure and transport.[22]

[21] The regressions control for the same right hand side variables as in Equation (1), and examine effects separately by treatment arm, men and women, as well as employed men and women — given the findings in Figure 5 and Table 9 regarding agriculture and nonfarm employment.
[22] In separate estimations, we also found fewer effects of electricity or motorized transport on time allocation in the morning hours across both treatment arms, supporting the relevance of the intraday distinction.

5.2 Simultaneity of activities

Accurately measuring multitasking — when two or more activities are conducted simultaneously — is critical for measuring total work burden and gender differences in time allocation (Craig and Brown, 2016). A longstanding literature has pointed to the relevance of data on multitasking for improving our estimates of women's participation and time in unpaid work — based on traditional time use surveys showing that women regularly conduct different unpaid domestic and care activities together, as well as alongside other activities in employment and leisure (Floro, 1995; Budig and Folbre, 2004; Irani and Vemireddy, 2021). Accurate quantification of these overlapping time constraints can in turn inform the effective design of care policies and measuring progress towards the Sustainable Development Goal to recognize and value unpaid work across countries. Moreover, accurate quantification of the incidence and time spent in simultaneously conducted activities can help us answer open questions around the links between multitasking and a range of outcomes related to health, leisure, productivity, and stress (Bittman and Wajcman, 2000; Floro and Miles, 2003; Manhart, 2004; Craig and Brown, 2016, 2017; Zaiceva-Razzolini, 2022).

In this section, we show that the mode of data collection also matters in assessing the role of multitasking in individuals' time allocation, and with implications for understanding of women's unpaid work burden. As discussed earlier, the smartphone diary allowed respondents to select multiple activities they were involved in at a given time, allowing for a continuous examination of how different activities were layered with each other throughout the day. The recall diary, alternatively, recorded (when applicable) a single secondary activity that respondents were involved in alongside the main activity for a given 15-minute interval. Appendix Figure A1 shows that the recall diary preface read out to respondents suggested secondary activities that might include collecting water or cooking, alongside caring for children. This yields different approaches to capturing simultaneous activities — while the smartphone diary did not distinguish between main and secondary activities (every activity was selected on its own, without any limits on the number of simultaneous activities that can be recorded), the recall diary did distinguish between the two, and encouraged respondents to consider a specific secondary activity that might have typically been done alongside the main one. As a result, close to 10 percent of respondents in the smartphone treatment arm often reported three or more activities (although typically not more than three) within a 15-minute or shorter duration, compared to the maximum ceiling of two activities in the recall diary.

The analysis shows that nearly all (around 95 percent, for both treatment arms) of women's daily multitasked non-leisure time was from conducting different unpaid domestic and care activities together. However, if we consider leisure activities (see Appendix Table A1 for the breakdown of leisure activities), the share of women's multitasked time in unpaid work was greater in the smartphone arm (38 percent, compared to 23 percent in the recall arm). As discussed below, this stems from greater reporting in the recall arm of multitasked time in leisure; the smartphone arm, on the other hand, revealed higher participation in multitasked employment, unpaid work, and transport activity.

To provide a richer understanding of different types of activities that were conducted together in both treatment arms, Tables 10 and 11 present heat map charts for the share of women and men, respectively,

conducting different activities together for the pooled matched samples of days 5, 8 and 10.[23] The tables are broken out by aggregate activity category, excluding sleep. Since the data in each treatment arm were collected by disaggregated activities within these larger activity categories, the tables include cases where different activities within a given aggregate category were conducted with each other. A key finding is that the smartphone arm, for both men and women, provides a much more diverse picture of different activities being conducted together than the recall arm. The latter primarily revealed leisure being conducted with different activities (including other leisure activities). In the smartphone arm, for example, 17 percent of women and 13 percent of men conducted transport activities while also obtaining services/resources for the household, compared to zero in the recall arm. In addition, greater shares of men and women in the smartphone arm conducted unpaid activities with agriculture and nonfarm employment, as well as transport with employment. While the smartphone arm also captured a large share of respondents conducting leisure with other activities — and the extent to which leisure is "shared" across other activities is also important to understand — the smartphone approach still reveals a much richer view of different combinations of activities individuals are involved in.

[23] Sleep was excluded from the table since it was rarely reported together with other activities.

Table 10. Average share of respondents conducting distinct activities together within a day (pairwise distribution), women

Smartphone sample

		Agri-culture	Nonfarm employ-Ment	Obtaining services/ resources for HH	Cooking, cleaning in house	Care of children, elderly	Transport	School/ studying	Leisure	Comm. Activities	Personal care	Eating/ drinking
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Agriculture	(1)	0.03										
Nonfarm employment	(2)	0.01	0.01									
Obtaining services/ resources for HH	(3)	0.01	0.03	0.04								
Cooking, cleaning in house	(4)	0.04	0.05	0.14	0.02							
Care of children, elderly	(5)	0.03	0.03	0.07	0.17	0.01						
Transport	(6)	0.05	0.05	0.17	0.11	0.11	0.14					
School/studying	(7)	0.00	0.00	0.00	0.00	0.00	0.01	0.00				
Leisure	(8)	0.09	0.08	0.11	0.24	0.16	0.24	0.03	0.36			
Community Activities	(9)	0.00	0.00	0.00	0.01	0.01	0.01	0.00	0.02	0.00		
Personal care	(10)	0.01	0.01	0.01	0.07	0.04	0.05	0.00	0.11	0.01	0.02	
Eating/drinking	(11)	0.02	0.04	0.02	0.18	0.09	0.04	0.00	0.28	0.00	0.07	0.01

Recall sample

		Agri-culture	Nonfarm employ-Ment	Obtaining services/ resources for HH	Cooking, cleaning in house	Care of children, elderly	Transport	School/ studying	Leisure	Comm. Activities	Personal care	Eating/ drinking
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Agriculture	(1)	0.00										
Nonfarm employment	(2)	0.00	0.00									
Obtaining services/ resources for HH	(3)	0.00	0.00	0.00								
Cooking, cleaning in house	(4)	0.01	0.02	0.02	0.00							
Care of children, elderly	(5)	0.03	0.04	0.08	0.21	0.00						
Transport	(6)	0.00	0.00	0.00	0.00	0.17	0.00					
School/studying	(7)	0.00	0.00	0.00	0.00	0.00	0.00	0.00				
Leisure	(8)	0.10	0.11	0.15	0.47	0.23	0.34	0.01	0.23			
Community Activities	(9)	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00		
Personal care	(10)	0.00	0.00	0.00	0.01	0.04	0.00	0.00	0.06	0.00	0.00	
Eating/drinking	(11)	0.00	0.01	0.00	0.01	0.16	0.00	0.00	0.45	0.00	0.00	0.00

Notes:
(1) Pooled sample of individuals across days 5, 8 and 10.

Table 11. Average share of respondents conducting distinct activities together within a day (pairwise distribution), men

Smartphone sample

		Agri-Culture	Nonfarm employ-ment	Obtaining services/ resources for HH	Cooking, cleaning in house	Care of children, elderly	Transport	School/ studying	Leisure	Comm. Activities	Personal care	Eating/ drinking
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Agriculture	(1)	0.03										
Nonfarm employment	(2)	0.01	0.01									
Obtaining services/ resources for HH	(3)	0.02	0.04	0.02								
Cooking, cleaning in house	(4)	0.02	0.01	0.02	0.00							
Care of children, elderly	(5)	0.00	0.00	0.01	0.01	0.00						
Transport	(6)	0.06	0.10	0.13	0.04	0.02	0.18					
School/studying	(7)	0.00	0.00	0.00	0.00	0.00	0.01	0.01				
Leisure	(8)	0.10	0.12	0.11	0.10	0.04	0.41	0.03	0.50			
Community Activities	(9)	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.03	0.00		
Personal care	(10)	0.01	0.02	0.01	0.02	0.01	0.08	0.00	0.15	0.01	0.01	
Eating/drinking	(11)	0.01	0.07	0.02	0.04	0.01	0.11	0.00	0.39	0.01	0.06	0.01

Recall sample

		Agri-culture	Nonfarm employ-ment	Obtaining services/ resources for HH	Cooking, cleaning in house	Care of children, elderly	Transport	School/ Studying	Leisure	Comm. Activities	Personal care	Eating/ drinking
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Agriculture	(1)	0.00										
Nonfarm employment	(2)	0.00	0.00									
Obtaining services/ resources for HH	(3)	0.00	0.01	0.00								
Cooking, cleaning in house	(4)	0.00	0.00	0.00	0.00							
Care of children, elderly	(5)	0.00	0.00	0.00	0.00	0.00						
Transport	(6)	0.01	0.03	0.02	0.00	0.01	0.00					
School/studying	(7)	0.00	0.00	0.00	0.00	0.00	0.00	0.00				
Leisure	(8)	0.11	0.22	0.10	0.11	0.04	0.43	0.01	0.42			
Community Activities	(9)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00		
Personal care	(10)	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.03	0.00	0.00	
Eating/drinking	(11)	0.00	0.03	0.00	0.00	0.01	0.01	0.00	0.72	0.00	0.01	0.00

Notes:
(1) Pooled sample of individuals across days 5, 8 and 10.

Table 12 also provides descriptive statistics on both the share of individuals and average daily shared time (in minutes) in conducting activities simultaneously, across different activity combinations — and, with a focus on how unpaid work, transport and employment were conducted with other activities. With a few exceptions, the share of individuals conducting activities simultaneously across men and women is significantly higher in the smartphone arm. Looking strictly at non-leisure activities, 74 percent of women and 60 percent of men in the smartphone arm conducted non-leisure activities together, compared to 37 and 13 percent of women and men in the recall sample (these shares for the recall arm were 84 and 85 percent, respectively, for all activities). This reinforces the role of leisure in simultaneous activity reporting in the recall diary approach. Within the smartphone arm, 31 percent of women and 17 percent of men also conducted unpaid work with transport. These shares were 10 and 15 percent, respectively, for transport with either agricultural or nonfarm employment. The shares in the recall arm, on the other hand, were much lower. Unpaid work and agricultural or nonfarm employment were conducted together among 15 percent of women and 7 percent of men in the smartphone sample, compared to 9 and 1 percent of women and men, respectively, in the recall arm.

Only for women's care work is participation significantly higher in the recall arm, regardless of whether the other activity was in leisure or non-leisure. This may be due to the preface of the 24-hour diary, specifically suggesting care work as a potential secondary activity that the respondents could include (see Appendix Figure A1). For women, cooking/cleaning with any other activity was significantly higher in the recall arm, but much of this is due to leisure being the other activity, since the trend flips in Table 12 (B2). On reported time, however, since activities in the recall sample could be conducted together for a minimum of 15 minutes, the recall arm often reveals significantly higher time conducted together in simultaneous activities – though, the magnitude of the differences is quite small given the substantially lower participation rates, and, in the case of men, often not significant. An understudied area, however, that the smartphone data provides some initial insights on, is the prevalence of passive care, where children are present with adults but not directly engaged in their activities (Folbre et al., 2005). Through the *TimeTracker* pop-up data on whom activities were conducted with (Table 13), we find that about 16 percent of women's daily time, or about 144 minutes daily, is spent with young children (aged 0-9 years) in activities that are not actively reported as care or sleep (and thus over and above the direct time reported earlier in childcare). Most of these activities are in other areas of unpaid work, transport, or leisure, and while performing these activities, women spend roughly twice as much time with young children compared to men. Gender differences in the extent of time devoted to activities in the presence of older children (aged 10-17 years), however, are narrower, reflecting important differences in the way we need to consider passive care in time use data collection. Although the *TimeTracker* pop-ups provide an imperfect view of passive care, the data shed some light on a significantly less-understood dimension of childcare that may not be well-captured by simply asking individuals whether they were involved in childcare provision.

Table 14 also presents treatment effects, using the same structure as Equation (1), and looking at the recall arm's effect on whether (1) activities were conducted together under different combinations, and (2) shared minutes across those activities. The estimates are presented for the total, women, and men samples, as well as by all areas/urban/rural. The results underscore the significantly lower incidence simultaneous activities in the recall arm, even when including leisure activities in column (A1). This holds

except for men in urban areas. The magnitude of underreporting widens across non-leisure activities, as well as carrying out unpaid work and agricultural/nonfarm employment with non-leisure activities. On shared time across non-leisure activities, apart from the large positive effects of 24-hour recall treatment on conducting unpaid work with non-leisure among women, most of the recall treatment effects on shared minutes lose significance, and the recall treatment has a significant negative effect on shared minutes among rural men's non-leisure activities (column B2). Overall, the smartphone arm provides a much more granular view of the extent to which activities are conducted together, across leisure and non-leisure activities.

Table 12. Descriptive statistics on simultaneity of activities by treatment arm

	Share of individuals conducting activities simultaneously (*s*)				Average daily shared time in conducting different activities simultaneously (mins)			
	Women		Men		Women		Men	
	Smart-phone	Recall	Smart-phone	Recall	Smart-phone	Recall	Smart-phone	Recal
(A) Any activities together:								
Across any activity categories	0.89***	0.84***	0.86***	0.85***	125.0***	267.3***	109.6***	272.3***
Across non-leisure activities	0.74***	0.37***	0.60***	0.13***	52.4**	61.7**	15.7*	11.9*
(B1 Unpaid work with any activity	0.68	0.69	0.38***	0.25***	79.3***	150.8***	13.1***	25.8***
Specific areas of unpaid work with any activity:								
Obtaining services/resources for HH	0.38***	0.24***	0.25***	0.13***	10.8***	14.1***	6.6***	12.9***
Cooking and cleaning	0.56***	0.62***	0.17***	0.12***	47.3***	79.9***	4.4***	7.8***
Care of children and/or elderly	0.31***	0.36***	0.08***	0.05***	42.7***	82.1***	2.3***	5.7***
(B2) Unpaid work w/any non-leisure activity	0.61***	0.35***	0.29***	0.05***	48.0***	60.7***	4.2	3.2
Specific areas of unpaid work with any non-leisure activity:								
Obtaining services/resources for HH	0.35***	0.11***	0.20***	0.02***	8.8***	5.4***	2.1	1.7
Cooking and cleaning	0.48***	0.26***	0.12***	0.02***	30.5***	23.9***	1.7***	0.8***
Care of children and/or elderly	0.27**	0.30**	0.05***	0.02***	32.6***	56.6***	0.8	1.3
(B3) Unpaid work w/any transport activity	0.31***	0.17***	0.17***	0.02***	8.7***	14.3***	1.7	1.5
Specific areas of unpaid work with any transport activity:								
Obtaining services/resources for HH	0.17***	0.01***	0.13***	0.02***	2.1***	0.07***	1.1	0.8
Cooking and cleaning	0.11***	0.002***	0.04***	0.002***	1.4***	0.03***	0.3***	0.03***
Care of children and/or elderly	0.11***	0.17***	0.02**	0.01**	5.6***	14.2***	0.3	0.6
(C) Agricultural or non-agr. work with:								
Any activity	0.34***	0.26***	0.35	0.37	20.8***	52.4***	26.3***	100.6***
Any non-leisure activity	0.26***	0.10***	0.26***	0.07***	10.2*	13.2*	7.1	8.8
Any unpaid work	0.15***	0.09***	0.07***	0.01***	7.7***	12.3***	1.2	0.9
Any transport activity	0.10***	0.002***	0.15***	0.03***	0.8***	0.04***	2.7	4.7

Notes:
(1) T-tests of equality of means conducted across treatment arms, for men and women separately; ***p<0.01, **p<0.05, *p<0.10.
(2) Pooled sample of individuals across days 5, 8 and 10.

Table 13. Smartphone data: average daily time (mins), and share of daily time, spent in passive care across non-care and non-sleep activities

	Pop-up data: activity conducted with:					
	Any children (aged 0-17)		Children aged 0-9		Children aged 10-17	
	Women	Men	Women	Men	Women	Men
Average daily minutes:						
Agriculture	8.8	7.0	3.1	1.8	6.2	5.5
Nonfarm employment	5.5***	16.3***	3.2	5.4	3.0***	13.0***
Unpaid work:						
Obtaining services/resources for HH	5.9***	2.4***	3.9***	0.8***	2.3	1.9
Cooking, cleaning in house	42.1***	5.1***	29.6***	1.6***	16.9***	3.7***
Transport	21.5	19.1	14.8***	5.4***	8.1***	14.7***
School/studying	3.8	5.8	0.02	0.5	3.7	5.3
Leisure	100.9	107.3	67.1***	48.2***	57.4**	73.3**
Community activities	1.7	2.7	0.6	0.2	1.2	2.7
Personal care	6.1**	1.2**	4.8**	0.6**	3.2	0.6
Eating/drinking	22.2***	13.9***	16.4***	7.5***	10.3***	7.6***
All non-care and non-sleep activities	**218.4*** **	**180.8*** **	**143.5*** **	**72.0*** **	**112.3***	**128.4***
Share of daily time (within activity categories; among those involved):						
Agriculture	0.17**	0.12**	0.07***	0.03***	0.10	0.09
Nonfarm employment	0.13	0.11	0.08***	0.03***	0.05*	0.09*
Unpaid work:						
Obtaining services/resources for HH	0.13**	0.09**	0.09***	0.04***	0.04	0.06
Cooking, cleaning in house	0.18**	0.14**	0.13***	0.06***	0.07**	0.10**
Transport	0.19***	0.10***	0.14***	0.03***	0.07	0.08
School/studying	0.27	0.25	0.08	0.04	0.19	0.21
Leisure	0.30***	0.25***	0.21***	0.12***	0.15	0.17
Community activities	0.20	0.17	0.09**	0.02**	0.12	0.16
Personal care	0.07***	0.04***	0.05***	0.02***	0.02	0.03
Eating/drinking	0.50***	0.31***	0.36***	0.16***	0.22***	0.18***
All non-care and non-sleep activities	**0.24*** **	**0.19*** **	**0.16*** **	**0.07*** **	**0.12***	**0.13***

Notes:
(1) Based on 1,026 and 1,035 observations for women and men respondents, respectively, for Days 5, 8 and 10 of the survey period.
(2) T-tests of equality of means conducted across men and women. ***p<0.01, **p<0.05, *p<0.10
(3) After each recorded activity in the *Time Tracker* app, respondents were asked whom the activity was conducted with.

Table 14. Treatment effects: effect of recall arm on average daily participation and shared time (minutes) in conducting different activities simultaneously

	Effect of recall arm:							
	(A) Across any activity categories (including leisure):		(B) Across non-leisure activities:		(C) Unpaid work with any non-leisure activity:		(D) Work in agriculture/non-farm emp. With any non-leisure activity:	
	Conducted activities together (Y=1 N=0)	Shared minutes	Conducted activities together (Y=1 N=0)	Shared minutes	Conducted activities together (Y=1 N=0)	Shared minutes	Conducted activities together (Y=1 N=0)	Shared minutes
	(A1)	(A2)	(B1)	(B2)	(C1)	(C2)	(D1)	(D2)
All areas								
Total	-0.03*	158.17***	-0.42***	3.14	-0.25***	6.45*	-0.17***	2.74
Women	-0.05***	144.92***	-0.36***	10.62	-0.25***	13.58**	-0.15***	4.46*
Men	-0.01	168.95***	-0.47***	-4.77	-0.24***	-1.03	-0.20***	0.76
Rural areas								
Total	-0.09***	136.39***	-0.46***	3.65	-0.27***	10.17**	-0.20***	1.87
Women	-0.11***	131.90***	-0.42***	19.53**	-0.29***	23.27***	-0.17***	8.10**
Men	-0.08***	141.65***	-0.51***	-9.86***	-0.25***	-1.06	-0.23***	-2.80
Urban areas								
Total	0.04***	176.12***	-0.39***	0.01	-0.23***	0.21	-0.16***	3.36
Women	0.02	153.99***	-0.32***	-1.33	-0.24***	0.56	-0.14***	0.80
Men	0.05**	194.34***	-0.46***	0.73	-0.25***	-1.00	-0.19***	5.06

Notes:
(1) OLS regressions controlling for EA and day of the week Fes; the standard errors are clustered at the household-level. ***p<0.01, **p<0.05, *p<0.10
(2) Based on a pooled individual-level sample for days 5, 8 and 10 (each individual appeared three times in the sample).
(3) Regressions control for the same individual and household-level characteristics as in eq. (1).

6. Discussion

This study reports on a large-scale survey experiment that was implemented for the first time in a low-income country to provide a comparative assessment of real-time, smartphone-based self-administered time use diaries versus recall-based, interviewer-assisted approaches to time use measurement. The experimental study design, repeated measurements of smartphone- and recall-based data, and the scope of the smartphone app functionality unlock several empirical insights regarding the practical field experience with smartphone-based data collection and the quality of reporting across smartphone and recall-based approaches. The analysis shows that compared to real-time time use data collection using a pictorial smartphone app, the reliance on a recall-based 24-hour diary leads to underestimation of participation in employment and unpaid work; mischaracterizes intra-day variation in time use with policy-relevant implications; fails to capture the true extent of multitasking, particularly among women; and yet, results in overestimation of time devoted to a wide range of activities. The latter is due to the minimum 15 minute-duration that was used for recording activities the 24-hour recall diary – while over one-third of activities lasted less than 15 minutes in the smartphone arm. The design of the 24-hour recall diary also results in "rounding" of estimated time, which is not a concern in the use of the open-ended smartphone time use diary. The analysis also demonstrates that compared to the smartphone time use diary, the reliance on stylized time use questions with a 7-day recall (as opposed to a 24-hour recall diary) results in an even greater overestimation of time devoted to employment and unpaid work. Ultimately, recall-based time use measurement methods provide an incomplete and misleading representation of men's and women's time use outcomes in this study population.

Looking forward, survey practitioners can consider using the existing *TimeTracker* app and training materials for deployment in alternative contexts. The app illustrations and translations can be updated with ease, and the fixed costs of smartphone procurement can be managed by procuring a limited pool of smartphones that can redeployed throughout the fieldwork. The app pop-ups can also be tailored to collect additional policy-relevant contextual information regarding reported activities, including passive childcare provision, as demonstrated in this study. At the same time, there are key issues that need to be addressed prior to proposing concrete recommendations for scaling up real-time time use data collection in low- and middle-income countries. First, it would be essential to replicate our study in alternative contexts, including with additional recall variants that have emerged in recent literature (Field et al., 2023; McCullough et al., 2023), to verify the external validity of our findings and strengthen the evidence base regarding the utility and requirements of improved approaches to time use measurement. Second, future research should provide guidance on whether the duration of smartphone data collection can be reduced, for example from nine to three full days of reporting, to save on costs without losing the analytical utility of the resulting data. Finally, survey practitioners could benefit from guidance on whether real-time, smartphone-based data collection can be limited to a sub-sample with the idea of relying on model-based within-survey imputation of time use measures – again as a cost saving mechanism while improving the quality of time use data collected in large-scale surveys.

References

Agénor, P-R, Agénor, M. 2023. "Access to infrastructure and women's time allocation: Implications for growth and gender equality." *Journal of Macroeconomics* 75.

Arthi, V., Beegle, K., De Weerdt, J., and Palacios-López, A. 2018. "Not Your Average Job: Measuring Farm Labor in Tanzania." *Journal of Development Economics* 130: 160-172.

Bittman, M., and Wajcman, J. (2000). The rush hour: The character of leisure time and gender equity. *Social Forces*, 79(1): 165–189.

Blackden, C. M., Wodon, Q. 2006. *Gender, Time Use, and Poverty in Sub-Saharan Africa*. World Bank.

Brown, N. R., Williams, R. L., Barker, E. T., Galambos, N. L. 2007. "Estimating Frequencies of Emotions and Actions: A Web-Based Diary Study." *Applied Cognitive Psychology*.

Budig, M.J., Folbre, N. 2004. "Activity, proximity or responsibility? – Measuring parental childcare time." In: Folbre, N. and M. Bittman (eds.), *Family time – The social organization of care*, Routledge, New York, 51-68.

Charmes, J. 2019. "The Unpaid Care Work and the Labor Market: An Analysis of Time Use Data Based on the Latest World Compilation of Time-Use Surveys." International Labor Organization, Geneva.

Chatzitheochari, S., Fisher, K., Gilbert, E., Calderwood,L., Huskinson, T., Cleary, A., and J. Gershuny. 2018. "Using New Technologies for Time Diary Data Collection: Instrument Design and Data Quality Findings from a Mixed-Mode Pilot Survey." *Social Indicators Research* 137(1): 379–390.

Craig, L. and Brown, J. E. 2017. "Feeling Rushed: Gendered Time Quality, Work Hours, Nonstandard Work Schedules, and Spousal Crossover." *Journal of Marriage and Family* 79(1): 225-242.

Craig, L. and Brown, J. E. 2016. "The Multitasking Parent: Time Penalties, Dimensions and Gender Differences." in *The Economics of Multitasking*. Editors: Kalenkoski CM, Foster G. 1, 33-59. Palgrave Macmillan, UK.

Daum, T., Buchwald, H., Gerlicher, A., and Birner, R. 2019. "Times Have Changed: Using a Pictorial Smartphone App to Collect Time–Use Data in Rural Zambia." *Field Methods* 31(1): 3-22.

Daum, T., Buchwald, H., Gerlicher, A., & Birner, R. (2018). "Smartphone apps as a new method to collect data on smallholder farming systems in the digital age: A case study from Zambia." *Computers and electronics in agriculture*, 153: 144-150.

Field, E., Pande, R., Rigol, N., Schaner, S., Stacy, E., Moore, C. T. 2023. "Measuring Time Use in Rural India: Design and Validation of a Low-Cost Survey Module." *Journal of Development Economics* 164.

Floro, M. S. 1995. "Women's well-being, poverty, and work intensity." *Feminist Economics* 3.

Floro, M. S. and Miles, M. 2003. "Time Use, Work and Overlapping activities: Evidence from Australia." *Cambridge Journal of Economics* 27: 881–904.

Folbre, N. 2021. "Quantifying Care: Design and Harmonization Issues in Time-Use Surveys." UN Women. Available at:
https://data.unwomen.org/sites/default/files/documents/Publications/Quantifying%20Care.pdf

Folbre, N., Yoon, J., Finnoff, K., and Fuligni, A. S. 2005. "By what measure? Family time devoted to children in the United States." *Demography* 42(2):373-90.

Frazis, H., and Stewart, J. 2012. "How to Think about Time-Use Data: What Inferences Can We Make About Long- and Short-Run Time Use from Time Diaries?" *Annals of Economics and Statistics*, Number 105/106, January/June.

Friedman, J., Gaddis, I., Kilic, T., Martuscelli, A., Palacios-Lopez, A., Zezza, A. 2023. "The Distribution of Effort: Physical Activity, Gender Roles, and Bargaining Power in an Agrarian Setting." *World Bank Economic Review* 37(1): 93-111.

Gaddis, I., Oseni, G., Palacios-Lopez, A., Pieters, J. 2021. "Measuring Farm Labor: Experimental Evidence from Ghana." *World Bank Economic Review* 35(3).

Gershuny, J., Harms, T., Doherty, A., Thomas, E., Milton, K., Kelly, P., Foster, C. 2020. "Testing Self-Report Time-Use Diaries against Objective Instruments in Real Time." *Sociological Methodology* 50(1): 318-349.

Hirway, I. 2017. *Mainstreaming Unpaid Work, Time Use Data in Developing Policies*. Oxford University Press, New Delhi.

International Labor Organization (ILO), 2023. *Measurement of unpaid domestic and care work.* Room Document 7, 21st International Conference of Labor Statisticians, Geneva, 11-20 October 2023. Available at:
https://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/meetingdocument/wcms_896583.pdf

ILO, 2018. *Survey methods to improve measurement of paid and unpaid work: Country practices in time use measurement*. Room Document No. 18, 20th International Conference of Labour Statisticians, Geneva, 10-19 October 2018.

Irani, L. and Vemireddy, V. 2020. "Getting the measurement right! quantifying time poverty and multitasking from childcare among mothers with children across different age groups in rural north India." *Asian Population Studies*.

Johnston, D., Stevano, S., Malapit, H. J., Hull, E., and Kadiyala, S. 2018. "Review: Time Use as an Explanation for the Agri-Nutrition Disconnect: Evidence from Rural Areas in Low and Middle-Income Countries." *Food Policy* 76: 8-18.

Kaplan, R. L., Kopp, B., Phipps, P. 2019. "Contrasting Stylized Questions of Sleep with Diary Measures from the American Time Use Survey." In Beatty, P., Collins, D., Kaye, L., Padilla, J. L., Willis, G., Wilmot, A. (eds). *Advances in Questionnaire Design, Development, Evaluation and Testing*. Wiley.

King, E., Randolph, H. L., Floro, M. S., Suh, J. 2021. "Demographic, Health, and Economic Transitions and the Future Care Burden." *World Development* 140.

Malapit, H., Quisumbing, A., Meinzen-Dick, R., Seymour, G., Martinez, Elena M., Heckert, J., Rubin, D., Vaz, A., and Yount, K. M. 2019. "Development of the Project-Level Women's Empowerment in Agriculture Index (pro-WEAI)." *World Development* 122 (October): 675–92.

Manhart, K. (2004). The limits of multitasking. *Scientific American Mind*, 14(5): 62–67

McCullough, E., McGavock, T. and Assefa, T. W. 2023. "Calling for Time: Examining Bias in Time Use Measurement using High-frequency Phone Surveys." Presentation at the Ninth International Conference of Agricultural Statistics (ICAS), Washington, DC.

Menon, G. 1993. "The Effects of Accessibility of Information in Memory on Judgments of Behavioral Frequencies." *Journal of Consumer Research* 20(3): 431-440.

Ribeiro, F.G., Souza, A. P., Carraro, A. 2021. "Rural electrification and agricultural family time allocation decisions." *Applied Economics* 53(16).

Seymour, G., Malapit, H., Quisumbing, A. 2020. "Measuring Time Use in Developing Country Agriculture: Evidence from Bangladesh and Uganda." *Feminist Economics* 3: 169-199.

Sudman, S., Bradburn, N. M. 1973. "Effects of Time and Memory Factors on Response in Surveys." *Journal of the American Statistical Association* 68(344): 805–15.

UN Women. 2021. *Measuring Time Use: An Assessment of Issues and Challenges in Conducting Time-Use Surveys with Special Emphasis on Developing Countries Methodological Inconsistencies, Harmonization Strategies, And Revised Designs*. Available at: https://data.unwomen.org/sites/default/files/documents/Publications/Measuring%20time%20use.pdf

United Nations. 2019. *Guidelines for Producing Statistics on Asset Ownership from a Gender Perspective.* New York. Available at: https://unstats.un.org/edge/publications/docs/Guidelines_final.pdf

Van de Walle, D., Ravallion, M., Mendiratta, V., Koolwal, G. 2017. "Long-Term Gains from Rural Electrification in India." *The World Bank Economic Review* 31(2): 385–411.

Watanabe, H., Chikaraishi, M., Maruyama, T. 2021. "How Different are Daily Fluctuations and Weekly Rhythms in Time-Use Behavior Across Urban Settings? A Case in Two Japanese Cities." *Travel Behaviour and Society* 22: 146-154.

World Economic Forum. 2021. *Global Gender Gap Report*. Available at: https://www3.weforum.org/docs/WEF_GGGR_2021.pdf

Zaiceva-Razzolini, A. "Multitasking." GLO Discussion Paper No. 1173, Global Labor Organization (GLO), Essen.

Appendix

Appendix Figure A1. 24-Hour Recall Time Use Diary
Adapted from the IFPRI Women's Empowerment in Agriculture Index Time Use Module

MODULE: TIME ALLOCATION

PLEASE RECORD A LOG OF THE ACTIVITIES FOR THE INDIVIDUAL IN THE LAST COMPLETE 24 HOURS (STARTING YESTERDAY MORNING AT 4 AM, FINISHING 3:59 AM OF THE CURRENT DAY). THE TIME INTERVALS ARE MARKED IN 15 MIN INTERVALS. MARK ONE PRIMARY ACTIVITY FOR EACH TIME PERIOD BY ENTERING THE CORRESPONDING ACTIVITY CODE IN THE BOX. A SECONDARY ACTIVITY (OPTIONAL) CAN BE ENTERED IN CASE OF SIMULTANEOUS ACTIVITIES.

THIS FORM MUST BE ADMINISTERED TO THE **RESPONDENT HIMSELF/HERSELF**.

Now I'd like to ask you about how you spent your time during the past 24 hours. We'll begin from yesterday morning, and continue through to this morning. This will be a detailed accounting. I'm interested in everything you did (i.e. resting, eating, personal care, work inside and outside the home, caring for children, cooking, shopping, socializing, etc.), even if it didn't take you much time. I'm also interested in how much time you spent caring for children, as well as other activities you were involved in while you did some other activity (e.g., collecting water while carrying a child or cooking while watching after a sleeping child).

	Night								Morning				Day																			
	4:00				5:00				6:00				7:00				8:00				9:00				10:00				11:00			
1a. Primary Activity (WRITE ACTIVITY CODE)																																
2a. Secondary Activity (WRITE ACTIVITY CODE)																																

	Day								Evening				Night																			
	16:00				17:00				18:00				19:00				20:00				21:00				22:00				23:00			
3a. Primary Activity (WRITE ACTIVITY CODE)																																
4a. Secondary Activity (WRITE ACTIVITY CODE)																																

24:00				1:00				2:00				3:00			

5. All things considered, how satisfied are you with your life as a whole these days? Using this card on which 1 means you are "completely dissatisfied" and 10 means you are "completely satisfied" where would you put your satisfaction with your life as a whole?

1	2	3	4	5
Completely dissatisfied				Completely satisfied

(Code one number from scale)

6. How have you felt about the activities you were engaged in the last day?

HAPPY.....1
NEITHER HAPPY NOR UNHAPPY..2
UNHAPPY.....3

ACTIVITY CODES

A.....Sleeping

B.....Eating/drinking

C1.....Personal: bathing, dressing

D1.....Sick/not active, at home
D2.....Sick/not active, in healthcare center

Schooling
E1..... Studying outside of school
E2.....At school

Agriculture
F1.....Fishing or tending to livestock
F2.....Crops: land clearing and preparation, manual
F3.....Crops: clearing land, with animal
F4..... Crops: planting, manual
F5.....Crops: manual watering
F6.....Crops: pumping water, using pump
F7.....Crops: weeding, manual
F8.....Crops: fertilizer/pesticide application, manual
F9.....Crops: harvesting, manual
F10.....Crops: drying and storage
F11.....Crops: processing, manual
F12.....Crops: processing, mechanical

Other market activity for income
G1.....Service or office-related occupations
G2.....Manual or *ganyu* labor, or manufacturing
G3.....Selling goods in the market
G4.....Gathering foodstuffs, hunting

Unpaid work for one's own household
H1.....Shopping
H2.....Cooking, cleaning in house
H3.....Collecting fuel/firewood
H4.....Collecting water
H5.....Stitching/knitting; weaving mats, baskets
H6.....Building and repairs
H7.....Maintenance of vehicles, electronics
H8.....Household finances
H9.....Care of elderly and sick adults
H10.....Care of children
H11.....Visit bank or office for services
H12.....Discussion with survey team

Transport
I1.....Walking, **not hauling** heavy load
I2.....Walking, **hauling** heavy load
I3..... Running
I4.....Bicycle, **not hauling** heavy load
I5.....Bicycle, **hauling** heavy load
I6.....Motorbike, **not hauling** heavy load
I7.....Motorbike, **hauling** heavy load
I8.....Truck/tractor/ox cart/bus, **not hauling** heavy load
I9.....Truck/tractor/ox cart/bus, **hauling** heavy load
I10..... Bicycle, as a passenger
I11..... Motorcycle, as a passenger

Leisure
J1.....Reading for leisure
J2.....Using devices (mobile phone, computer, TV)
J3.....Resting/doing nothing
J4.....Chatting with others inside/outside household
J5.....Celebrations
J6.....Religious activities
J7.....Exercise/playing games

K1.....Public works projects
K2..... Local community meetings

L.....Other (specify)

Appendix Figure A2. 7-Day Stylized Recall Module

CODE	ACTIVITY	Did you conduct this activity over the last 7 days? YES....1 NO.....2 >> NEXT CODE	**ENUMERATOR - AFTER ASKING Q2 FOR ALL ACTIVITIES, ASK THE FOLLOWING FOR THOSE WHERE Q2=1:** During the last 7 days, that is from [DAY] up to yesterday, how many hours did you work on these activities?
101	Livestock or fishing activities		
102	Crop-related agriculture on a household farm		
103	Run or do any kind of non-agricultural or non-fishing household business, big or small		
104	Help in any of the household's non-agricultural or non-fishing household businesses		
105	Casual or ganyu labor		
106	Work for a wage, salary, commission, or any payment in kind, excluding ganyu		
107	Engage in an unpaid apprenticeship for anyone that is not a member of the household		
301	Cleaning the house, washing clothes, cooking or shopping for the household		
302	Paying household bills, or visiting the bank for financial services		
303	Repairs or maintenance work him/her-self (e.g. fixing broken appliances or fixtures, painting walls, repairing vehicles), or any time doing construction work him/her-self to renovate, extend or build the household's dwelling		
304	Fetching water from natural or public sources for use by the household		
305	Collecting firewood or other natural products for use as fuel by the household		
306	Providing care, help or assistance to household members aged 18 years or older because of a disability, illness, or problems related to old age **ENUMERATOR: READ: For example: Administering medication, feeding, helping them with bathing and personal hygiene, etc.**		
307	Looking after children aged 17 years or younger living in this household **ENUMERATOR: READ: For example: Bathing playing with children, taking children to school, sports or other activities, instructing, tutoring or helping children with homework, advising or talking with teens about their problems, etc**		
401	Participating in local public works projects (involved in construction of local public infrastructure, for example), or participating in local community or government meetings		
402	Studying or in school/classes for his/her own education		
403	Reading, watching TV or mobile devices for one's own entertainment, socializing, participating in religious occasions, exercise/playing sports		

Appendix Table A1. Activity list and pop-ups in the *Time Tracker* Smartphone app, along with Chichewa translations included as captions to images

(1) Activities

	Type of activity	Chichewa Translation
1. Market work: agriculture *(Ntchito za ulimi)*	*Agricultural*	*Za ulimi*
	Fishing or tending to livestock	*kuwedza nsomba/kuweta ziweto*
	Crops: land clearing and preparation , manual	*Kulima ndi manja*
	Crops: clearing land, with animal	*Kulima ndi ziweto*
	Crops: planting, manual	*Kubyzala ndi manja*
	Crops: manual watering	*Kuthilira ndi vava*
	Crops: pumping water, using pump	*kuthilira ndi pompi*
	Crops: weeding, manual	*Kupalira ndi manja*
	Crops: fertilizer application/pesticide, manual	*Kuthira fetereza/mankhwala ndi manja*
	Crops: harvesting, manual	*Kukolora ndi manja*
	Crops: drying and storage	*Kuyanika ndi kusunga*
	Crops: processing, manual	*kukonza ndi manja*
	Crops: processing, mechanical	*kukonza ndi makina*
2. Market work: other activities *(Ntchito zina)*	*Other activities*	*Ntchito Zina*
	Service or office-related occupations	*Ntchito zamaofesi ndi zina*
	Manual/ganyu labor or manufacturing	*Ntchito zamanja/Ganyu, kapena ntchito yopanga zinthu zina*
	Selling goods in market	*Kugulitsa zinthu (katundu) ku msika*
	Gathering foodstuffs, hunting	*Kusonkhanitsa zakudya, kusaka*
3. Unpaid work (for one's own household) *(Ntchito zapakhomo)*	Shopping	*Kugula zinthu*
	Cooking, cleaning in house	*Ntchito zapakhomo*
	Collecting fuel/firewood	*Kukatola nkhuni*
	Collecting water	*kutunga madzi*
	Stitching/knitting, weaving baskets and mats	*Kusoka/kuluka Malaya, kuluka dengu, basiketi, mkeka ndi kusoka mphasa*
	Building and repairs	*Kukonza zipangizo zoononeka zapanyumba*
	Maintenance of vehicles, electronics	*Kukonza galimoto, zamagetsi*
	Household finances	*chuma cha pakhomo*
	Care of elderly and sick adults	*Kusamalira achikulire ndi anthu akulu oti akudwala*
	Care of children	*Kusamalira ana*
	Visit bank or office for services	*kukapeza thandizo ku bank kapena ku office*
	Discussion with survey team	*Kukambirana ndi akafukufuku*
4. Transport	Walking, not hauling heavy load	*Kuyenda, osanyamula katundu olemetsa*
	Walking, hauling heavy load	*Kuyenda utanyamula katundu olemetsa*
	Running	*Kuthamanga*
	Bicycle, not hauling heavy load	*Njinga yakapalasa, yosanyamula katundu olemetsa*
	Bicycle, hauling heavy load	*Njinga yakapalasa, yonyamula katundu olemetsa*
	Motorbike, not hauling heavy load	*Njinga ya moto, yosanyamula katundu olemetsa*
	Motorbike, hauling heavy load	*Njinga ya moto, yonyamula katundu olemetsa*
	Truck/tractor/ox cart/bus, not hauling heavy load	*Galimoto/ thirakitala/ngolo/ Bus, yosanyamula katundu olemetsa*
	Truck/tractor/ox cart/bus, hauling heavy load	*Galimoto/ thirakitala/ngolo/ Bus, yonyamula katundu olemetsa*
	Bicycle, as passenger	*Njinga yakapalasa, yokwezedwa*
	Motorcycle, as passenger	*Njinga ya moto, yokwezedwa*
5. Schooling *(Kuphunzira)*	Studying outside of school	*Kuphunzira kunja kwa sukulu*
	At school	*Kuphunzira ku sukulu*
6. Leisure *(Zisangalaro)*	Reading for leisure	*Kuwerenga*
	Using devices	*kugwiritsa ntchito zipangizo zamakono*
	Resting/doing nothing	*Kupumula*
	Chatting with others inside/outside the household	*Kucheza ndi ena pakhomo/kunja kwa khomo*
	Celebrations	*Zisangalalo*
	Religious activities	*Zamapemphero*
	Exercise/playing games	*Masewera olimbitsa thupi*
7. Community meetings/activities *(Misonkhano / Ntchito zachituko)*	Public works projects	*Kugwira nawo ntchito zakudera*

	Local community meetings	*kukhala nayo pa misonkhano yakudera*
8. Personal care/health (*Ukhondo)*	Personal care (bathing, dressing, etc) Sick/not active: at home Sick/not active: in healthcare center	*Za ukhondo(kusamba/kuvala etc)* *Kudwalira pakhomo* *Kudwalira ku chipatala*
9. Sleep *(Kugona)*	Sleep	*Kugona*
10. Eating *(Kudya)*	Eating/drinking	*kudya/Kumwa*

(2) Pop-ups

	Pop-up	**Chichewa Translation**
1. Pop-up: who was present (appeared after each activity was ended; can select multiple) *(Munali ndi ndani)*		
	Man/men aged 18 and older	*Mamuna wa zaka 18 kapena kuposela apo*
	Woman/women aged 18 and older	*Mkazi wa zaka 18 kapena kuposela apo*
	Children below age 10	*Ana ochepela zaka 10*
	Children age 10-17	*Ana amene ali ndi zaka 10-17*
	By myself	*Pa ndekha*
2. Pop-up: food groups, if food was selected (can select multiple *(Magulu a zakudya (mutha kusankha zingapo)*		
	Cereals	*Gulu la tirigu*
	Roots/tubers	*Mitsitsi*
	Vegetables	*Zamasamba*
	Fruits	*Zipatso*
	Meats	*Nyama*
	Eggs	*Mazira*
	Fish/seafood	*Nsomba*
	Pulses/legumes/nuts	*Za gulu la nyemba*
	Milk products	*Mkaka*
	Oils/fats	*Mafuta*
	Sugar/honey	*Shuga/Uchi*
	Tea/coffee, spices	*Tiyi/khofi, Zokometsera zakudya*
3. Pop-up: how do you feel (appeared after each activity was ended) *(Mukuganiza bwanji)*		
	Happy	*Okondwa*
	Neutral	*Pakatikati*
	Unhappy	*Osakondwa*

Appendix Table A2. Individual-day level regressions:
Correlates of (Figure 4d) Number of "jumps" or discontinuities in time (forgetting to log activities) between consecutive activities; and (Figure 4e) Share of activities with "admin" or "deleted" (i.e., non-author) entries [(2)]

	Number of time jumps	Admin or deleted entries in smartphone log (Y=1 N=0)
	(1)	(2)
Female (Y=1 N=0)	0.63**	0.10***
HH Head	-0.24	-0.04
Age: 18-24	-0.34	-0.04
Age: 25-34	-0.42	-0.03
Age: 45-54	-0.27	0.05
Age: 55+	-1.17	0.11**
Marital status: married	0.27	0.02
Religion: not Christian	0.01	0.04
Secondary or higher education	-0.63**	-0.10***
Main occupation: agriculture	0.47	0.04
Main occupation: non-agriculture	0.29	0.07**
Has rights to sell/bequeath land	-0.14	-0.01
Has a bicycle	0.42	-0.06**
Has a motorized vehicle	0.04	-0.00
Has a mobile phone	-0.04	0.02
Has a financial account	-0.04	-0.03
HH: Number of children 12 and under	0.21	0.00
HH: permanent construction	-0.38	0.01
Main source of light in HH: electricity	0.10	-0.10**
Main source of light in HH: batteries/torch	0.37	-0.00
Observations	5,644	5,646
R-squared	0.235	0.131

Notes:
(1) Regressions control for day of the week, and enumeration-area (EA) fixed effects; standard errors are clustered at the household level. ***p<0.01, **p<0.05, *p<0.10.

Appendix Table A3. Participation in the last day across disaggregated activity categories (Day 5)

	24-hour recall arm	Smartphone diary arm
Agriculture		
Fishing or tending to livestock	0.03**	0.06**
Crops: land clearing and preparation, manual	0.10**	0.13**
Crops: land clearing and preparation, w/animal	0.01	0.02
Crops: planting, manual	0.03	0.04
Crops: manual watering	0.06	0.08
Crops: pumping water, using pump	0.00*	0.01*
Crops: weeding, manual	0.08**	0.11**
Crops: fertilizer/pesticide application, manual	0.01**	0.03**
Crops: harvesting, manual	0.06	0.04
Crops: drying and storage	0.01	0.02
Crops: processing, manual	0.03	0.04
Crops: processing, mechanical	0.03	0.04
Nonfarm employment		
Service or office-related work	0.06	0.07
Manual or ganyu labor, or manufacturing	0.14	0.15
Selling goods in the market	0.11**	0.17**
Gathering foodstuffs, hunting	0.04	0.05
Unpaid work		
Shopping	0.23***	0.33***
Cooking, cleaning in house	0.65	0.67
Collecting fuel/firewood	0.05***	0.10***
Collecting water	0.27	0.30
Stitching/knitting; weaving mats, baskets	0.00***	0.04***
Building and repairs	0.04	0.06
Maintenance of vehicles, electronics	0.01***	0.05***
Household finances	0.00***	0.05***
Care of elderly and sick adults	0.04	0.05
Care of children	0.24***	0.32***
Visit bank or office for services	0.01**	0.03**
Transport		
Walking, not hauling heavy load	0.85*	0.88*
Walking, hauling heavy load	0.31	0.29
Running	0.00***	0.12***
Bicycle, not hauling heavy load	0.05***	0.12***
Bicycle, hauling heavy load	0.04	0.06
Motorbike, not hauling heavy load	0.02***	0.05***
Motorbike, hauling heavy load	0.01	0.02
Truck/tractor/ox cart/bus, not heavy load	0.04***	0.06***
Truck/tractor/ox cart/bus, hauling heavy load	0.02**	0.04**
Bicycle, as a passenger	0.02***	0.08***
Motorcycle, as a passenger	0.01***	0.05***
Schooling		
Studying outside of school	0.03**	0.05**
At school	0.03**	0.05**
Leisure		
Reading for leisure	0.01***	0.15***
Using devices (mobile phone, computer, TV)	0.29***	0.45***
Resting/doing nothing	0.50***	0.82***
Chatting with others inside/outside household	0.97***	0.81***
Celebrations	0.05***	0.10***
Religious activities	0.19***	0.27***
Exercise/playing games	0.08	0.09
Community activities		
Public works projects	0.00**	0.02**
Local community meetings	0.04***	0.12***
Personal care	0.93***	0.88***
Sick/not active, at home	0.01***	0.07***
Sick/not active, in healthcare center	0.01***	0.04***
Sleep	1.00***	0.98***
Eating/drinking	0.99***	0.96***
Number of respondents	**696**	**735**

Notes:
(1) T-tests of equality of means conducted across treatment groups; ***p<0.01, **p<0.05, *p<0.10.

Appendix Table A4. Time (mins) in the last day across disaggregated activity categories (Day 5)

	24-hour recall arm	Smartphone diary arm
Agriculture		
Fishing or tending to livestock	2.22	1.63
Crops: land clearing and preparation, manual	25.43**	17.17**
Crops: land clearing and preparation, w/animal	0.24	0.15
Crops: planting, manual	7.22*	2.93*
Crops: manual watering	8.69**	4.72**
Crops: pumping water, using pump	0.04	0.29
Crops: weeding, manual	19.18	17.81
Crops: fertilizer/pesticide application, manual	1.36	1.37
Crops: harvesting, manual	9.18**	2.29**
Crops: drying and storage	0.84	0.41
Crops: processing, manual	4.38	2.82
Crops: processing, mechanical	1.40	1.40
Nonfarm employment		
Service or office-related work	27.07	23.05
Manual or ganyu labor, or manufacturing	51.08**	31.66**
Selling goods in the market	40.11	29.28
Gathering foodstuffs, hunting	1.90**	0.72**
Unpaid work		
Shopping	9.40	9.24
Cooking, cleaning in house	129.70	138.16
Collecting fuel/firewood	4.12	5.33
Collecting water	13.92	13.99
Stitching/knitting; weaving mats, baskets	1.06*	3.52*
Building and repairs	9.48	8.12
Maintenance of vehicles, electronics	0.26**	4.83**
Household finances	0.04***	0.50***
Care of elderly and sick adults	6.14**	1.74**
Care of children	51.21**	36.99**
Visit bank or office for services	1.27	0.34
Transport		
Walking, not hauling heavy load	86.14***	115.93***
Walking, hauling heavy load	19.38**	13.00**
Running	0.04***	1.95***
Bicycle, not hauling heavy load	4.98**	9.44**
Bicycle, hauling heavy load	7.87*	3.95*
Motorbike, not hauling heavy load	1.36	2.10
Motorbike, hauling heavy load	2.20	1.34
Truck/tractor/ox cart/bus, not heavy load	3.28	4.52
Truck/tractor/ox cart/bus, hauling heavy load	1.19**	4.78**
Bicycle, as a passenger	1.31	2.30
Motorcycle, as a passenger	0.39**	2.71**
Schooling		
Studying outside of school	3.66	2.01
At school	10.28	9.73
Leisure		
Reading for leisure	0.63***	7.59***
Using devices (mobile phone, computer, TV)	55.58	59.40
Resting/doing nothing	46.53***	102.06***
Chatting with others inside/outside household	360.04***	170.49***
Celebrations	5.95	8.42
Religious activities	27.65	22.11
Exercise/playing games	11.51**	4.37**
Community activities		
Public works projects	0.69	0.73
Local community meetings	4.59	7.08
Personal care	27.76	26.96
Sick/not active, at home	2.69	7.22
Sick/not active, in healthcare center	3.10	0.76
Sleep	536.81	529.28
Eating/drinking	54.78**	47.98**
Number of respondents	**696**	**735**

Notes:
(1) T-tests of equality of means conducted across treatment groups; ***p<0.01, **p<0.05, *p<0.10.

Appendix Table A5. Conditional time across activities (Day 5 match, smartphone and recall arms)

	Agri-culture	Nonfarm employ-ment	Unpaid work: Obtaining resources for HH	Cooking, cleaning in house	Care of children, elderly	TOTAL UNPAID WORK	Transport	School/ studying	Leisure	Comm. activities	Personal care	Sleep	Eating/ drinking
	(1)	(2)	(3a)	(3b)	(3c)	(3d)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
DAILY TIME, CONDITIONAL (MINS)													
Total													
Recall (obs=696)	243.7***	368.4***	80.7	198.4	213.4***	297.0**	136.6***	334.7***	513.0***	131.3***	35.8	536.8	55.3**
Smartphone (obs=734)	127.7***	228.1***	73.3	205.2	113.4***	265.4**	173.1***	132.8***	383.8***	55.7***	39.2	541.8	49.9**
Women													
Recall (obs=352)	213.5***	225.5***	74.3*	235.5*	224.5***	389.3	115.4**	382.5***	431.8***	114.7**	39.7	549.3	52.2
Smartphone (obs=365)	115.7***	165.0***	63.8*	252.3*	127.3***	367.3	130.5**	108.4***	329.1***	46.0**	42.5	553.8	49.9
Men													
Recall (obs=344)	278.6***	449.2***	95.7	80.6	153.1***	124.3	156.9***	300.9***	596.0***	156.8**	31.6	524.0	58.4**
Smartphone (obs=369)	139.5***	268.1***	86.3	92.6	59.0***	127.2	213.0***	146.1***	437.8***	65.3**	35.8	529.7	49.8**

Notes:

(1) Within the total, women's and men's samples, T-tests of equality of means conducted across treatment groups; ***p<0.01, **p<0.05, *p<0.10.

(2) Grey shading = recall arm yields significantly lower estimates.

(3) Recall estimates include time spent on main + secondary activities (see Appendix Table A1 for the 24-hour recall time use module).

(4) Average total time for women and men was 1700 minutes in the recall sample (including main and secondary activities). In the smartphone sample, average total daily time was 1521 minutes for women, and 1490 minutes for men.

Appendix Figure A3. Average daily time (minutes) across main activity categories, by check in day, treatment arm, and men/women

Appendix Table A6. OLS regressions: energy/fuel-related correlates of minutes spent in evening/nighttime (6pm onwards)

	Minutes spent in evening/nighttime (6pm onwards) in:							
	Smartphone arm				Recall arm			
	Agriculture (1)	Nonfarm employment (2)	Leisure (3)	Unpaid work (4)	Agriculture (1)	Nonfarm employment (2)	Leisure (3)	Unpaid work (4)
Employed women (obs = 803 for smartphone, obs=755 for recall)								
Main source of light: electricity	1.29	49.54**	49.79	-34.76	-0.67	5.74	2.82	10.66
Main source of light: battery-powered devices	0.94	20.74	3.12	-18.24	-0.48	-0.49	-11.23	-3.76
Respondent owns motorized vehicle	-1.53	76.72*	40.96	20.34	1.72	4.85	-16.47	15.58
All women (obs = 1,023 for smartphone, obs=1,037 for recall)								
Main source of light: electricity	-0.21	27.46*	41.71*	-0.29	-0.42	3.24	9.47	8.00
Main source of light: battery-powered devices	0.88	15.08	2.74	-1.98	-0.17	-0.62	-7.48	-4.19
Respondent owns motorized vehicle	-1.27	68.74**	46.45*	16.10	1.07	3.78	-6.55	10.53
Employed men (obs = 952 for smartphone, obs=910 for recall)								
Main source of light: electricity	-4.02	6.04	28.45	-12.98	2.65*	7.58	25.54**	-2.05
Main source of light: battery-powered devices	-3.11	-1.46	7.28	-9.34	2.74	3.26	-0.10	0.16
Respondent owns motorized vehicle	-8.87	39.81**	-24.16	-8.04**	-0.61	-10.42	21.54	12.04*
All men (obs =1,035 for smartphone, obs=1,023 for recall)								
Main source of light: electricity	-4.62	-0.72	18.99	-6.34	2.66*	7.46	25.25**	-1.83
Main source of light: battery-powered devices	-4.44	-4.34	3.82	-6.37	2.76	3.25	-0.70	0.46
Respondent owns motorized vehicle	-6.28	36.50**	-5.72	-6.92**	-0.66	-11.38	22.90	12.01*

Notes:

(1) OLS regressions controlling for EA and day of the week FEs; the standard errors are clustered at the household-level. ***p<0.01, **p<0.05, *p<0.10

(2) Based on a pooled individual-level sample for days 5, 8 and 10 (each individual appeared three times in the sample).

(3) Regressions also control for the same individual and household-level socioeconomic characteristics as in eq. (1).