UNIVERSITÄT
HOHENHEIM

Social and Institutional Change in Agricultural Development
Institute of Agricultural Sciences in the Tropics
(Hans-Ruthenberg-Institute)

The potential of smartphone apps to collect self-recorded data in agricultural households. A study on time-use in Zambia.

Dissertation to obtain the doctoral degree of
Agricultural Sciences (Dr. sc. agr.)

Faculty of Agricultural Sciences
University Hohenheim

submitted by
Thomas Lothar Georg Daum
from Speyer, Germany

2019

UNIVERSITÄT HOHENHEIM

This thesis was accepted as a doctoral dissertation in fulfillment of the requirements for the degree "Doctor of Agricultural Sciences" (Dr. sc. agr.) by the Faculty of Agricultural Sciences at the University of Hohenheim.

Date of thesis submission: 30th April 2019
Date of oral examination: 19th December 2019

Supervisor and Reviewer: Prof. Dr. Regina Birner
Second Examiner: Prof. Dr. Ignace Glorieux
Third Examiner: Prof. Dr. Thomas Berger
Head of the committee: Prof. Dr. Andrea Knierim
Dean of the faculty: Prof. Dr. Ralf T. Vögele

CONTENTS

ACKNOWLEDGEMENTS

This thesis was made possible through many people - above all the farmers who participated in the study and who welcomed me warmly in Zambia. My deepest gratitude goes to Prof. Dr. Regina Birner who, by her curiosity, excitement, confidence and knowledge, supported me throughout the whole thesis and this part of my life. I am equally grateful to Hannes Buchwald who programmed the smartphone application as part of a student project at the University of Media, Stuttgart. His efforts were relentless, his work excellent and he became a good friend. I am thankful to his supervisor Prof. Dr. Ansgar Gerlicher for being open and full of trust towards our ideas. The smartphone app became as easy to use and as beautiful as it is now only thanks Natalis Lorentz who did the illustrations. In Zambia, I am grateful to the Indaba Agricultural Policy Research Institute for helping me to coordinate the field work. I am also grateful to my enumerators and drivers, who helped making this thesis both more thoughtful and enjoyable. In Germany, I am grateful for many colleagues from Hohenheim for their personal and intellectual companionship: Dr. Saurabh Gupta, Dr. Juliet Kariuki, Dr. Lilli Scheiterle, Dr. Athena Birkenberg, Filippo Capezzone, Mary Lubungu, Tilahun Woldie, Adu-Gymafi Poku, Johannes Mössinger, Lutz Heiner-Otto and many others. I am grateful to Linn and Leander Doppler for their continuous and patient administrative support. I am very grateful for the generous and trustful financial support from the Program of Accompanying Research for Agricultural Innovation, which is funded by the German Federal Ministry of Economic Cooperation and Development. PARI above all was an exciting platform to discuss ideas and share experiences. Thank you, Prof. Dr. Joachim von Braun, Heike Baumüller and Oliver K. Kirui. My family and many friends have been part of this experience. I am indebted to them for many things. I owe most to Sabina Cato to whom I dedicate this thesis.

EXECUTIVE SUMMARY

Mobile information and communication technologies (ICTs) have spread across the developing world and are used increasingly by smallholder farmers. While the potential of ICTs, such as smartphone applications, to provide new opportunities for agricultural development is widely acknowledged, the potential to use them as research tools has not been explored. This thesis assesses the potential of smartphone applications for the collection of data from agricultural households in developing countries. Can smartphone applications that use visual tools be used for self-recording of data by the respondents themselves where literacy levels are low? Can such smartphone applications that allow for real-time data recording increase the accuracy of the collected data?

Answering these questions is important as, so far, data from agricultural households are usually collected using surveys, which are prone to recall biases. This is a problem, as researchers, policymakers and development practitioners need reliable data for their work. Poor data can lead to misguided policy recommendations and actions with adverse effects on vulnerable population groups. This can lead to agricultural development trajectories that are socially unequal and unsustainable.

To assess the potential of smartphone apps to collect self-recorded data, a smartphone application called Timetracker was developed as part of this thesis. The Timetracker allows study respondents to record data in real time with the help of illustrations. Recording data in real time reduces recall bias, and using pictures ensures that participants with low literacy can use the application. In its current form, the Timetracker can be used to collect data on time-use and nutrition. Collecting reliable data on time-use and nutrition is key for various strands of research. For example, time-use data are needed to calculate labor productivity and analyze how productivity is affected by new technologies. Time-use data can also help reveal gender-based power relations and asymmetries by pointing out unpaid domestic work. Similarly, nutritional data are crucial for various academic fields and debates. For example, nutritional data are needed to explore the factors determining food and nutrition security, to study how farm diversity affects consumption diversity and to monitor food and nutrition policies and programs.

This study is based on three main chapters, which reflect the main objectives of the whole thesis: 1) to explore and test whether smartphone applications can be used to collect data from rural households in developing countries focusing on time-use and nutrition data, 2) to assess the accuracy of data collected with smartphone applications vis-à-vis recall-based data collection methods, and 3) to use the data to understand the effects of agricultural mechanization on the intrahousehold allocation of time-use within smallholder farming households in Zambia. The first two chapters have a primarily methodological focus. The last chapter is an empirical study.

The second chapter, which addresses the first objective, explores and tests whether smartphone applications can be used as data collection tools in developing countries. For this purpose, the lessons learned from the use of user-oriented smartphone apps in developing countries and from examples where smartphone apps have been used to collect data in developed and developing countries were extracted. Based on these lessons, the Timetracker application was developed, which is presented in detail. The second chapter reflects on the preconditions that need to be fulfilled when using smartphone apps to collect data in developing countries based on the authors' experiences when using the Timetracker in rural Zambia, such as the need to address the

challenge of low literacy levels and social beliefs. The chapter concludes with a discussion on the future potential to use smartphone applications to collect data from smallholder farming systems.

In the third chapter, which addresses the second objective, the accuracy of the collected data is tested. The chapter discusses the potentials and pitfalls of the Timetracker application vis-à-vis the advantages and disadvantages of existing methods to collect time-use data in developing countries. Then, the accuracy of the Timetracker application is tested by comparing the data collected with the application with the data collected using 24-hour recall questionnaires. The results confirm the literature on recall biases, suggesting that using the Timetracker application leads to valid results. Additional methods to validate the collected data are discussed, and limitations and directions for future research are noted.

In the fourth chapter, which addresses the third objective, the collected data are used to explore the effects of agricultural mechanization on farm families. Agricultural mechanization has been rapidly growing in Asian countries and has received growing attention in Africa, but its effects are ambiguous. The chapter investigates the effects of mechanization on the intrahousehold time-use divisions in smallholder farming households in Zambia, paying particular attention to gender, child labor and seasonality. This study was formulated against the background that the adoption of new technologies, policies, and practices can change the intrahousehold time allocation, which may disadvantage vulnerable household members, such as women and children. This study uses compositional data analysis, which accounts for the intrinsic codependence of time-use data, and different regression tools. The results show, for example, that both men and women benefit from agricultural mechanization during land preparation and that a gender differentiation only emerged with mechanization. There is some evidence that the time "saved" is used for off-farm and domestic work. No negative second round effects during weeding and harvesting/processing and no negative effects on children were found. This chapter provides proof of concept that using digital tools can help to collect more reliable socioeconomic data and that compositional data analysis can be used to analyze such data with regard to time use.

Returning to the three objectives formulated above, this thesis showed that 1) picture-based smartphone applications can be used to collect data in rural areas of developing countries with low literacy levels; 2) smartphone applications can help to improve the accuracy of time-use data in developing countries; and 3) having such accurate and detailed data allows to explore the socioeconomic aspects of agricultural development that are otherwise difficult to analyze, such as the effects of agricultural mechanization on the intrahousehold allocation of time-use within smallholder farming households.

This thesis concludes that in addition to improving the accuracy of socioeconomic data collection, smartphone applications may open new research pathways, including through the opportunities provided by real-time data collection and by combining self-recorded data with sensor-recorded data, which may open interesting transdisciplinary research pathways. This thesis suggests that there is a large and still untapped potential for using smartphone applications to collect data on complex agricultural systems in the digital age.

ZUSAMMENFASSUNG

Mobile Informations- und Kommunikationstechnologien (IKT) werden zunehmend auch von Kleinbauern in Entwicklungsländern eingesetzt. Während das Potenzial von IKT, wie Smartphone-Anwendungen, neue Möglichkeiten für die landwirtschaftliche Entwicklung zu bieten, weithin anerkannt ist, wurde deren Potenzial als Forschungsinstrumente bislang kaum erforscht. Diese Dissertation untersucht das Potenzial von Smartphone-Anwendungen zur Erfassung sozioökonomischer Daten von landwirtschaftlichen Haushalten in Entwicklungsländern. Können Smartphone-Anwendungen, die visuelle Elemente verwenden, für die Selbstaufzeichnung von Daten durch Befragte verwendet werden, selbst wenn deren Alphabetisierung gering ist? Können solche Smartphone-Anwendungen, mit denen Daten in Echtzeit erfasst werden, die Genauigkeit der erfassten Daten erhöhen?

Die Beantwortung dieser Fragen ist wichtig, da die Daten von landwirtschaftlichen Haushalten bisher üblicherweise durch Haushaltsbefragungen erhoben werden, die häufig durch Erinnerungsverzerrungen beinflusst sind. Dies ist ein Problem, da Forscher, politische Entscheidungsträger und Entwicklungsakteure verlässliche Daten für ihre Arbeit benötigen. Unzureichende Daten können zu falschen Politikempfehlungen und Politkmaßnahmen führen, die sich negativ auf bestimmte Bevölkerungsgruppen auswirken können. Dies kann zu landwirtschaftlichen Entwicklungspfaden führen, die sozial ungleich und nicht nachhaltig sind.

Um das Potenzial von Smartphone-Apps zur Selbstaufzeichnung von Daten durch Befragte zu bewerten, wurde im Rahmen dieser Dissertation eine Smartphone-Anwendung namens Timetracker entwickelt. Der Timetracker ermöglicht es den Befragten, Daten anhand von Abbildungen in Echtzeit zu erfassen. Das Aufzeichnung von Daten in Echtzeit verringert Erinnerungsverzerrungen und die Verwendung von Bildern stellt sicher, dass Teilnehmer mit geringer Alphabetisierung die Anwendung verwenden können. In seiner jetzigen Form kann der Timetracker verwendet werden, um Daten zu Zeitnutzung und Ernährung zu sammeln.

Zuverlässiger Daten zu Zeitnutzung und Ernährung sind essenziell für verschiedene Forschungsbereiche. Zum Beispiel werden Daten zur Zeitnutzung benötigt, um Arbeitsproduktivität zu berechnen und zu analysieren, wie die Produktivität durch neue Technologien beeinflusst wird. Daten zur Zeitnutzung können auch helfen, geschlechtsspezifische Machtverhältnisse und Asymmetrien aufzuzeigen, indem sie auf unbezahlte häusliche Arbeiten hinweisen. In ähnlicher Weise sind Ernährungsdaten für verschiedene akademische Bereiche von entscheidender Bedeutung. Ernährungsdaten sind beispielsweise erforderlich, um die Faktoren zu untersuchen, die die Ernährungssicherheit bestimmen; um zu untersuchen, wie die Diversität der landwirtschaftlichen Betriebe die Nahrungskonsumvielfalt beeinflusst; und um die Ernährungsstrategien und -programme zu überwachen.

Die Studie basiert auf drei Hauptkapiteln, welche die Hauptfragen der gesamten Dissertation widerspiegeln: 1) zu untersuchen, ob Smartphone-Anwendungen verwendet werden können, um Daten wie Zeitnutzung und Ernährung von ländlichen Haushalten in Entwicklungsländern zu sammeln; 2) zu beurteilen wie genau die mit Smartphone-Anwendungen erfassten Daten im Vergleich zu auf Erinnerung basierenden Datenerhebungsmethoden sind; 3) unter Verwendung der gesammelten Daten zu analysieren, wie sich landwirtschaftliche Mechanisierung auf die Zeitaufteilung innerhalb von kleinbäuerlichen Haushalten in Sambia auswirkt. Die ersten beiden Kapitel sind primär methodisch ausgerichtet. Das letzte Kapitel ist dann eine empirische Studie.

Das zweite Kapitel, das sich mit der ersten Forschungsfrage befasst, untersucht und testet, ob Smartphone-Anwendungen in Entwicklungsländern als Instrumente zur Datenerfassung verwendet werden können. Dazu werden zunächst die Erkenntnisse betrachtet, die aus der Betrachtung von benutzerorientierten Smartphoneanwendung in Entwicklungsländern und anhand von Beispielen, in denen Smartphone-Anwendungen zum Sammeln von Daten in Industrie- und Entwicklungsländern verwendet wurden, gewonnen werden konnten. Basierend auf diesen Erkenntnissen wurde die Timetracker-Anwendung entwickelt, die in diesem Kapital detailliert vorgestellt wird. In dem Kapitel werden die Voraussetzungen beschrieben, die erfüllt sein müssen, wenn Smartphone-Anwendungen zur Erfassung von Daten in Entwicklungsländern verwendet werden sollen. Dies geschieht auf Grundlage der Erfahrungen während der Verwendung des Timetrackers im ländlichen Sambia. Voraussetzungen sind beispielsweise die Notwendigkeit, die Herausforderung niedriger Alphabetisierung und gesellschaftlicher Vorstellungen (*social beliefs*) anzugehen. Das Kapitel schließt mit einer Diskussion über die zukünftigen Potenziale für die Verwendung von Smartphone-Anwendungen zum Sammeln von Daten aus kleinbäuerlichen Landwirtschaftssystemen.

Im dritten Kapitel, das die zweite Forschungsfrage anspricht, wird die Genauigkeit der gesammelten Daten geprüft. In diesem Kapitel werden die Potenziale und Fallstricke der Timetracker-Anwendung im Hinblick auf die Vor- und Nachteile bestehender Methoden zur Erfassung von Zeitnutzungsdaten in Entwicklungsländern erörtert. Anschließend wird die Genauigkeit der Timetracker-Anwendung getestet, indem die mit der Anwendung erfassten Daten mit Daten verglichen werden, die mit 24-Stunden-Recall-Fragebögen erfasst wurden. Die Ergebnisse bestätigen die Literatur zu Erinnerungsverzerrungen, was darauf schließen lässt, dass die Verwendung der Timetracker-Anwendung zu gültigen Ergebnissen führt. Weitere Methoden zur Validierung der gesammelten Daten werden diskutiert und Grenzen und Richtungen für zukünftige Forschung aufgezeigt.

Im vierten Kapitel, das die dritte Forschungsfrage anspricht, werden die gesammelten Daten verwendet, um die Auswirkungen der landwirtschaftlichen Mechanisierung auf landwirtschaftliche Familien zu untersuchen. Die Mechanisierung der Landwirtschaft ist in den asiatischen Ländern rasch gewachsen und hat in Afrika wachsende Aufmerksamkeit erhalten. Ihre Auswirkungen sind jedoch unklar. In diesem Kapitel werden die Auswirkungen der Mechanisierung auf die Zeitnutzung (Aufgabenverteilung) in landwirtschaftlichen Haushalten in Sambia untersucht, wobei Geschlechteraspekte, Kinderarbeit und Saisonalität besonders berücksichtigt werden. Die Studie wurde vor dem Hintergrund formuliert, dass die Einführung neuer Technologien, Politiken und Praktiken die Aufgabenverteilung innerhalb des Haushalts verändern kann, was verletzbare Haushaltsmitglieder wie Frauen und Kinder benachteiligen kann. Die Studie verwendet kompositionell Datenanalyse (*compositional data analysis),* welche die intrinsische Koabhängigkeit (*intrinsic codependence*) von Zeitnutzungsdaten berücksichtigt, sowie verschiedene Regressionsinstrumente. Die Ergebnisse zeigen zum Beispiel, dass sowohl Männer als auch Frauen von der landwirtschaftlichen Mechanisierung während der Landbearbeitung profitieren und dass eine geschlechtsspezifische Arbeitsteilung nur mit zunehmender Mechanisierung auftritt. Es gibt Anzeichen dafür, dass die Zeit, die "gespart" wird, für Arbeit außerhalb der Landwirtschaft und für Hausarbeit verwendet wird. Es wurden keine negativen Zweitrundeneffekte beim Unkraut Jäten und Ernten / Verarbeiten und keine negativen Auswirkungen auf Kinder festgestellt. In diesem Kapitel wird der Nachweis erbracht, dass mit Hilfe digitaler Werkzeuge zuverlässigere sozioökonomische Daten erhoben werden können und dass kompositionell Datenanalyse (*compositional data analysis*), zur Analyse dieser Daten verwendet werden kann.

Um auf die drei oben formulierten Forschungsfragen zurückzukommen, zeigt die Arbeit, dass 1) bildbasierte Smartphone-Anwendungen zum Sammeln von Daten in ländlichen Gebieten von Entwicklungsländern mit niedrigem Alphabetisierungsgrad verwendet werden können, 2) Smartphone-Anwendungen die Genauigkeit von Zeitnutzungsdaten in Entwicklungsländern verbessern können, 3) das mit solch genauen und detaillierten Daten sozioökonomische Aspekte der landwirtschaftlichen Entwicklung untersucht werden können, die ansonsten schwer zu analysieren sind. Das ist zum Beispiel die Auswirkungen der landwirtschaftlichen Mechanisierung auf die Zeitaufteilung innerhalb von kleinbäuerlichen Haushalten.

Die Dissertation kommt zu dem Schluss, dass Smartphone-Anwendungen nicht nur die Genauigkeit der Erfassung sozioökonomischer Daten verbessern, sondern auch neue Forschungspfade eröffnen. Die geschieht vor allem durch die Möglichkeiten der Echtzeit-Datenerfassung und durch die Kombination selbst erfasster Daten mit sensoraufgezeichneten Daten, was interessante transdisziplinäre Forschungsmöglichkeiten aufzeigt. Die Dissertation legt nahe, dass es ein großes und noch nicht ausgeschöpftes Potenzial gibt, Smartphone-Anwendungen zum Sammeln von Daten zu komplexen landwirtschaftlichen Systemen in Entwicklungsländern im digitalen Zeitalter zu nutzen.

LIST OF ACRONYMS AND ABBREVIATIONS

ADP	Animal Draught Power
AIDS	Acquired Immunodeficiency Syndrome
ALR	Additive Log Ratios
Apps	Applications
BMZ	Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (Federal Ministry of Economic Cooperation and Development)
CIMI	Calculator of Inadequate Micronutrient Intake
CoD	Compositional Data
CoDa	Compositional Data Analysis
DF	Degrees of Freedom
DNA	Deoxyribonucleic Acid
G	Gender
GMO	Genetically Modified Organism
GMSA	Groupe Speciale Mobile Association
GPS	Global Positioning System
FAO	Food and Agriculture Organization
Ha	Hectares
HIV	Human Immunodeficiency Virus
HH	Household
IAPRI	Indaba Agricultural Policy Research Institute
ibid	ibīdem, "in the same place"
ICT	Information and Communication Technology
IFPRI	International Food Policy Research Institute
IKT	Informations- und Kommunikationstechnik
ILO	International Labor Organization
ITU	International Telecommunication Union
Kg	Kilograms
Km	Kilometers
L	Liters
M	Manual
MSD	Mean Squared Standard Deviation
NERCIA	New Rice for Africa
P	Power group (mechanization type)
P.a.	Per annum
PARI	Program of Accompanying Research for Agricultural Innovation
PSM	Propensity Score Matching
RALS	Rural Agricultural Livelihood Survey
RCT	Randomized Controlled Trial
SAS	Statistical Analysis Software
SDG	Sustainable Development Goal
SMS	Short Messaging Services

T	Tractor
Tuc	Time use category
US	United States of America

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

1. Introduction

“When you can measure what you are speaking about, and express it in numbers, you know something about it, when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely, in your thoughts advanced to the stage of science.”

William Thomson, often referred to as Lord Kelvin, quote from the lecture on "Electrical Units of Measurement" (1883)

“A man who uses an imaginary map thinking that it is a true one, is likely to be worse off than someone with no map at all”

Ernst Schumacher, quote from the book “Small Is Beautiful: A Study of Economics as if People Mattered” (1973)

Using a phrase of the political economist Charles Lindblom, James C. Scott (1998) argued that pre-modern states were "all thumbs and no fingers” because they lacked enough data from their citizens for the fine tuning of state action. In many cases, this led to misguided policies, sometimes to revolts. Jumping some centuries ahead, Scott argued that modern states have more data but still lack understanding for aspects of life that difficult to measure. A similar story could be told about science. Scientists too collect more and more data but tend to neglect non-measurable aspects of life (Katzner, 2001). And even when measured, things are not always measured correctly, which can be problematic. In the words of Ernst Schumacher quoted above: “a man who uses an imaginary map thinking that it is a true one, is likely to be worse off than someone with no map at all”. This thesis aims to contribute to make what is important measurable and to measure things better.

The focus of this thesis is on improving data collection in the fields of applied agricultural economics, development economics, rural sociology, agronomy and related disciplines who study rural households and farming systems in developing countries, but the findings may be relevant to other disciplines as well. Economists, sociologists and agronomists working with rural households rely on data for their analysis and to derive policy recommendations. Similarly, governments and development practitioners rely on data on rural households. For them, data are needed to prioritize, design, monitor and evaluate development programs and policies. For example, with regard to the 2030 Sustainable Development Goals (SDGs), there is widespread recognition that policy formulation and monitoring progress will require enough and good enough data. Without data, and to restate Charles Lindblom’s formulation from above, governments and development practitioners are "all thumbs and no fingers". They are flying blind. But data collection, especially using surveys is costly. A lack of data therefore constraints particularly low income and middle income countries. Beegle et al. (2016) showed that by 2012 only 27 of 48 African countries had conducted more than two consumption survey since 1990 to measure poverty. With regard to time-use, which is one of the focus areas of this thesis, Rubiano-Matulevich and Kashiwase (2018) showed that 135 countries have never collected data on this. In Africa, since 1960, only 12 detailed, stand-alone, time-use surveys were conducted (Buvinic and King, 2018).

Cleary, it is not enough to just collect data. Data must be of good quality to be useful and not misleading. Data on agricultural households are typically collected using household surveys. However, survey questions are susceptible to recall biases, which can be large. In the 1990s, India changed the recall period of their national surveys on consumption from 30 to 7 days. As a consequence, respondents recalled better how much they consumed. This lead to an "increase" in reported average daily expenditures and a "drop" in the number of poor people by 175 million (Deaton 2013). This is an order of magnitude that makes asking for policy implications needless. Arthi et al. (2018) found equally high recall bias when asking farmers on time-use in rural Tanzania. They asked farmers on how much time they spend on farming either using a post-harvest questionnaire or a weekly survey. When asked on a weekly basis, the average time reported was four times lower. And lastly, Ugandan farmers estimate the value of beans harvested two times higher when using harvesting diaries compared to post-harvest questionnaires (Deininger et al. 2012).

Data on questions that do not rely on recall can be equally erroneous. Carletto et al. (2015) compared self-reported farm sizes with GPS-measured farm sizes. They found that farmers (especially farmers owning little land) tend to overestimate their farm size. According to them, this has two major policy implications: 1) land inequality may be higher than previously assumed; 2) a longstanding debate between agricultural economists on whether small or large farms are more productive may have been based on misleading data. Data have also proven unreliable for improved seeds. Using DNA fingerprinting, various studies have shown that farmers often do not know whether they use improved seeds or not (Floro et al., 2017; Kosmowski et al., 2019). This suggests that studies analyzing the impact of improved germplasm research may be based on unreliable data.

All of the above examples show how unreliable the data underlying much of applied agricultural economics research and related fields can be. Both a lack of data and unreliable data can lead to misguided policy recommendations and actions, which can lead to adverse effects on farmers and more vulnerable household members such as women and children. For example, food and nutrition policies and programs, which are based on wrong or lacking food and nutrition data, may miss their targeted beneficiaries (under- and malnourished household members). Similarly, development strategies that neglect varying effects of new technologies, policies and practices on time-use of men, women, boys and girls in smallholder farming households can fail or have negative consequences. For example, conservation agriculture, which preserves soils but is associated with high weed pressure, has been shown to be not adopted or to lead to time poverty for women (Farnworth et al., 2015). Promoting the new rice variety NERCIA, which allows higher yields but is associated with an increased need for bird scaring, has been shown to prevent children from going to school (Bergman-Lodin et al., 2012).

There are similar concerns with regard to agricultural mechanization, which has been growing rapidly in Asian countries and has received growing attention in Africa (Diao et al. 2014; Nin-Pratt and McBride 2014; Takeshima, 2017; Wang et al., 2016). While the drivers of mechanization are usually labor limitations, not all household members may benefit from mechanization. For example, mechanized land preparation may allow households to cultivate additional land, which may increase the need for weeding, harvesting/processing or the time spent collecting firewood once forests are cleared, tasks often performed by women and children (Arora, 2015; Blackden and Wodon, 2006; Doss, 2001). Despite such concern, the effects of mechanization on intra-household time allocation have not been examined, notwithstanding some anecdotal evidence.

Against this background and knowledge gaps, this study has three objectives, which correspond with the three papers of this thesis:

1) Explore and test whether smartphone applications can be used to collect data from rural households in developing countries focusing on time-use and nutrition data;
2) Assess the accuracy of data collected with smartphone applications vis-à-vis recall based data collection methods;
3) Use the data to understand the effects of agricultural mechanization on the intra-household allocation of time-use within smallholder farming households in Zambia.

In the subsequent sections of this introductory chapter, the rationale for focusing on time-use (1.11) and nutrition data (1.2) will be explained in detail. In section 1.3, the advantages and disadvantages of existing methods to collect time-use and nutrition data and their suitability to collect data in developing countries will be assessed. This section provides the rationale for exploring the use of smartphone-based data collection methods. In section 1.4, the smartphone application developed for this thesis will be presented. As mentioned above, the data collected were used to analyze social and economic effects of agricultural mechanization on rural households in Zambia. Section 1.5 explains the focus on agricultural mechanization in more detail. Section 1.6 provides background information on the case study country Zambia. Section 1.7 provides an overview of the subsequent structure of the thesis and the three papers that will be presented.

1.1. Significance of time-use data

Time has always mattered for agricultural development. In fact, agricultural development started because the time needed for hunting and gathering rose so high that sedentary farming became an attractive option during the Neolithic Revolution. Much later, time-use changes contributed to explaining why farmers moved from shifting cultivation to annual crop cultivation. Boserup (1965) describes shifting cultivation as an “easy system” with high yields per hour of work. Little time is spent on land clearing and preparation as well as weeding, manuring and the caring for draught animals – mainly because long fallow periods restore soil fertility without the farmers’ contribution. With rising population density, shifting cultivation was replaced by more permanent forms of agriculture, which have higher yields per land but also require more labor, for example, for weeding and applying manure. This later triggered the development and use of labor saving technologies.

The importance of time-use for agricultural development may also be illustrated by the fact that the first time-use surveys worldwide were used to better understand agriculture. In Russia, they were used as early as the late 19th century to better understand the life of peasant farmers. During the 1920s and 1930s, time-use surveys were used for the first time in the United Stated of America and the United Kingdom, mainly to inform the agricultural extension systems (Buvinic and King, 2018; Gershuny, 2011). The examples demonstrate the key role of time for agriculture. In this section, additional examples are presented to highlight the significance of time-use for agricultural economics: a) the role of time-use to understand labor productivity; b) the role of time-use to explore power relations and asymmetries between gender; c) the role of time-use for technology adoption – a vice versa how technology adoption affects time-use, d) the role of time-use for modelling decision making of farmers, e) the role of time-use to obtain a more nuanced understanding of poverty and wellbeing and f) the role of time-use for health outcomes.

Time-use and labor productivity

Many strands of agricultural and development economics research are based on the calculation of agricultural labor productivity. For example, researchers discuss the role of agricultural labor productivity for economic growth at large. Perhaps most famously, Schultz (1953) argued that raising agricultural productivity is a pre-condition for economic growth. In contrast, Lewis (1955) argued that economic growth only happens when labor moves from the agricultural sector, which has a low productivity per worker, to the industrial sector, which has a high productivity per worker. This debate continues to occupy economists (Diao et al, 2017). Other branches of research are

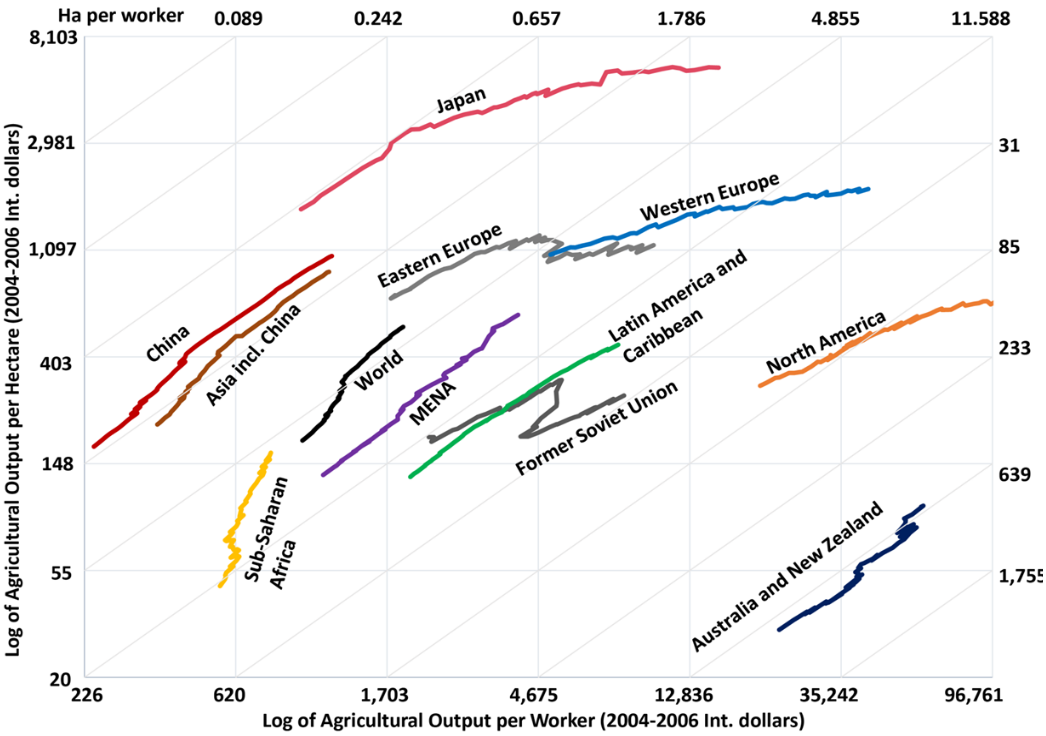


Figure 1. Historical changes in land and labor productivity across the world
Adopted from Fuglie and Rada (2013).

also based on the calculation of agricultural labor productivity. Some use such data to explore why agricultural productivity has grown more in some countries than in others (Gutierrez, 2002).

In this context, the apparently stagnating agricultural labor productivity of African farmers has sparked much concern (see figure 1). This is because, in the absence of alternative sources, agricultural productivity is a key determinant of household income and poverty. Other research branches use agricultural labor productivity data to analyze how differences in labor productivity within and between rural and urban areas influence migration patterns (Lipton, 1980; Goldsmith et al., 2004). In addition, there are various micro-level studies, which have focused, for example, on the effects of HIV/AIDS on labor productivity (Fox et al, 2004) or how different technologies, such as tractors, affect agricultural labor productivity (Adu-Baffour et al., 2018).

In the absence of good enough quality data, much of the above-quoted literature is based on labor productivity *per worker*, which does not take into account how much farmers actually work (McCullough, 2017). This may be a significant omission. An alternative is to measure productivity per days worked. In fact, most household surveys ask farmers to recall how many days they performed certain activities during the previous season. This assumes full work days and that work days are comparable across respondents (for example, by sex and age) – hours worked as well as efforts and skills are ignored (Doss, 2018; McCullough, 2017). More accurate time-use data could help to unveil some of the inaccuracies resulting from coarsely measuring labor input. For example, most lately, McCullough (2017) has shown that the above-mentioned frequently cited productivity gap between agriculture and the non-agricultural sector may be based on misleading evidence. According to her findings, non-agricultural labor is only 1,4 times more productive compared to agricultural labor when looking labor productivity *per hour* (instead of 5 times more productive when looking *per worker*).

Unpaid work, power relations and gender

Besides being needed to calculate agricultural labor productivity, time-use data can be used to explore power relations and gender aspects within households (Bianchi and Milkie, 2010; Blackden and Wodon, 2006; Glorieux et al., 2015). The need to incorporate gender aspects of time-use into policy formulation and development programs is increasingly acknowledged. This is reflected in the Sustainable Development Goals (SDGs) of the United Nations. The SDGs focus on time-use because of the concern that men and women spend their time differently; and women, for example, spend more time on (often unpaid) domestic work and care activities. By showing how much time different household members spend on different activities (such as domestic chores or care), time-use data can make the contribution of women's work more visible as well as reveal power relations and asymmetries.

SDG 5 envisions to "achieve gender equality and empower all women and girls".[1] The related target (5.4) further specifies that this includes the recognition and value of unpaid care and domestic work. The corresponding indicator to track the progress on this target is the "proportion of time spent on unpaid domestic and care work, by sex, age, and location".[2] For policy makers, this can be of relevance as a large share of time spent on unpaid domestic work may hinder women to participate in labor markets and social communities, which may be a barrier to empowerment.

However, 135 countries have yet to collect data on time-use, partly because collecting such country-wide data is very expensive (Rubiano-Matulevich and Kashiwase, 2018). According to Buvinic and King (2018), only 5 % of all nationally representative surveys collected data on unpaid

[1] https://www.un.org/sustainabledevelopment/gender-equality/ (accessed April 20th 2019)
[2] https://www.un.org/sustainabledevelopment/gender-equality/ (accessed April 20th 2019)

domestic work. In its LSMS-ISA surveys, the World Bank includes questions such as “How many hours did you spend yesterday collecting water?”[3] or “How many hours did you spend yesterday collecting firewood (or other fuel materials)?”[4] but mostly does not collect data on domestic and care activities. This data gap constrains efforts to target development programs and policies towards women.

Technology adoption and time-use

A related concern is how new technologies, practices and policies affect the intra household allocation of time-use (Blackden and Wodon, 2006; Doss, 2001; Theis et al., 2018; von Braun and Webb 1989). Ignoring these aspects can lead to promoting technologies, practices and policies, which, for example, increase women’s often already-high labor burden (Blackden and Wodon, 2006). Caution is needed particularly when promoting innovations for smallholder farming households, where men and women have different workloads and duties. This division of labor by gender is based on crops, tasks or both (Doss, 2001). As new policies, practices, and technologies address different crops and tasks, they can affect time-use of men, women, boys and girls differently. Development strategies that overlook these dynamics can fail or have negative effects.

For example, conservation agriculture may not be adopted or may lead to a heavier labor burden for women because of the high requirements for weeding (Giller et al., 2009, Farnworth et al., 2015). Similarly, the new rice variety NERICA has been shown to exacerbate the need for weeding and bird scaring, which can raise women’s workload and prevent children from going to school (Bergman-Lodin et al., 2012). The example of Berman-Lodin highlights the need to pay attention to the time-use patterns of children and adolescents, too, which are particularly but not exclusively relevant when studying child labor. It may thus be a perilous omission that most time-use surveys, in particular in developing countries, focus on adults as collecting time-use data from children and adolescents is challenging.

Agricultural interventions that lead to changes in time-use patterns of parents may determine not only their own wellbeing but also that of their children. For example, there is evidence that interventions that lead to more time spent on agriculture can reduce the time available for domestic work and child care, which is associated with nutritional outcomes, especially of children. However, as noted in a review paper by Johnston et al. (2018) the relationship between changes in parent’s time-use and the nutritional status of children is “complex” and “there is no agreement on the impact” (p. 8). In addition to nutritional outcomes, time-use changes of parents may also be linked to educational outcomes.

---

[3] http://microdata.worldbank.org/index.php/catalog/2936/datafile/F6 (accessed April 20th 2019)
[4] http://siteresources.worldbank.org/INTLSMS/Resources/3358986-1233781970982/5800988-1271185595871/IHS4_Household_Questionnaire_FINAL.PDF (accessed April 20th 2019)

In general, time-use data are key to understanding not only the effects of adopted technologies but also how current time-use patterns affects technology adoption decisions. For example, households may not adopt agricultural technologies that increase income as well as working time, or formulated differently, make them "time poor". This aspect has been explored very little in the literature, potentially because of the lack of good data collection methods on time-use. Lambrecht et al. (2014) argued that farmers weight the benefits of a certain new technology against its labor requirements during adoption decisions (and afterward). Gender plays a large role here again: evidence suggests adopting female-labor-intensive technology can be rejected only when women have higher bargaining power (Fisher et al., 2000).

Farm system modelling and other types of modelling

Time-use data are also relevant when modelling farm household decision-making, which has long tradition in agricultural economics (Hazell and Norton, 1986). Such modelling can be done with single or multiple households (agents), which can then interact with each other and their environment (such as land use and soil). Such models can be used to model the effects of agricultural policies or structural change (Balmann, 1997; Happe et al., 2006), technology diffusion and water allocation (Berger, 2001) and soil fertility dynamics (Schreinemachers et al., 2007), among other things. The core of such models is to model farmers' behavior with the help of simple rules and certain constraints (such as the amount of land available, crop rotations etc.). To be realistic, such models also need time or labor constraints, and thus rely on good time-use data.

Time-use, time poverty and wellbeing

Amartya Sen has shown that poverty has more dimensions than economic wealth (or the lack thereof). Building on his argument, several scholars have suggested to include measures of time-use into measuring poverty (Blackden and Wodon, 2006) or relatedly into the analysis of wellbeing. This field of inquiry contains various normative questions, which involve value judgments. While one could probably reach a scientific consensus that having a higher income is better than a lower one, a consensus cannot be easily reached with regard to care activities, for example. Implicitly, the literature suggests that care activities are work and a reason for "time poverty" (Blackden and Wodon, 2006). But is someone who has more time for his or her children really worse off than one who spends little time with his or her children? Are there not some aspects in child care that are enjoyable, such as teaching a child how to ride a bike or reading a book to them? On the other hand, are there not, perhaps, less enjoyable aspects in child care, such as changing diapers? This raises an important question for this line of research: how useful are time-use data to determine time poverty without knowing the knowledge of the individual contexts l? Looking at the agency over time-use may be pathways forward.

Time-use and health

Time-use patterns are also related to health. For example, Tremblay and Willims (2003) found that time spent on sedentary activities has led to a higher prevalence of obesity among children in Canada. Ng and Pokin (2012), using detailed time-use data, found a decrease of time spent on physically active activities and increase in sedentary time in both developed and developing countries, which can lead to obesity as well as related cardio-metabolic health risks. In addition to changes in diets, changes in time-use patterns therefore seems to be another key driver of an emerging double burden of nutrition in developing countries (Popkin, 2001; Steyn and Mchiza, 2014).

1.2. Significance of food and nutrition data

Across the world, 821 million people do not have access to enough calories and are therefore undernourished (FAO et al., 2018). In addition, close to two billion people lack access to enough micronutrients and are thus malnourished, a phenomenon referred to as "hidden hunger" (IFPRI, 2016). Avoiding both undernourishment and malnourishment is the second goal of the Sustainable Development Goals (SDGs) of the United Nations. Obesity, on the rise in both developed and developing countries with huge health implications, has not been focused on by the SDGs, but it is equally important for the well-being of the world population.[5] With agriculture being the source of food and nutrients, several branches of agricultural economics (but also of development and health economics) are studying aspects related to food and nutrition security, which rely on food and nutrition data: a) the link between farm diversity and consumption diversity; b) food and nutrition policies and programs; c) the economics of bio-fortification; d) intra-household allocation of nutrition; e) causal relations of food and nutrition insecurity determinants; and f) drivers of caloric requirements.

Farm diversity and consumption diversity

The prevalence of both under- and malnutrition has been shown to be particularly high among smallholder farmers (FAO et al., 2018; Pinstrup-Andersen, 2007). Therefore, "agricultural – nutrition linkages" have received much attention as a way to combat under- and malnutrition in the past decades (Dangour et al., 2013; Turner et al., 2013). This is also reflected in the term nutrition-sensitive agriculture. These linkages are well recognized from a food quantity perspective: a high level of farm production raises the overall availability of food, and therefore reduces undernutrition. This has happened during the Asian green revolution when rising farm productions led to higher caloric consumptions (Evenson and Gollin, 2003; Headey and Hoddinott, 2016). In addition to looking at the quantity of food produced, researchers are exploring

[5] Obesity is indirectly referred to in the sub-goal 3.4. Sub-goal 3.4 aims to "reduce by one third premature mortality from non-communicable diseases" (https://sustainabledevelopment.un.org/sdg3, accessed 20th April 2019).

agricultural-nutrition linkages from a food quality perspective, for example, by linking farm diversity with consumption diversity (Carletto et al., 2017; Fanzo, 2017; Jones et al., 2014; Jones, 2017; Koppmair et al., 2017; Sibhatu et al., 2015, Sibhatu and Qaim, 2018). Most of these studies rely on recall data on nutrition to calculate dietary diversity scores. However, although food quantity and types can vary much over the course of a season for subsistence-oriented (and to a lesser extent, market-oriented) poor rural households, most of these studies and their data do not capture seasonality. Ayenew et al. (2018), who distinguish between post-planting and post-harvest season, found an effect of farm diversity on dietary diversity after harvesting but not after planting.

Food and nutrition policies and programs

Given the huge social and economic costs of under- and malnutrition, many states and development organizations use a range of interventions and programs to improve the food and nutrition status of their citizens. This can be through school feeding programs, provision of food by state agencies (such as the Indian Integrated Child Development Service), educational programs, promotion of kitchen gardens, primary healthcare, and through subsidies and taxes, among others. Some of these programs and measures (such as education programs and taxes) can also be used to address the problem of obesity. Designing adequate food and nutrition policies and programs requires a comprehensive picture of the current nutritional status in the target regions. Monitoring and evaluating the effectiveness and efficiency of such policies and programs, again, requires the collection of food and nutrition data. However, data are still often lacking. According to the Malabo Montpellier Panel report "Nourished: How Africa Can Build a Future Free from Hunger and Malnutrition" (2017), African governments "continue to lack the data necessary to effectively combat malnutrition, responses to food crises remain reactive, rather than proactive" (p.11).

Economics of bio-fortification

Food and nutrition policies and programs may also include the support of bio-fortification, an approach motivated by the perceived slow progress rates with the above-mentioned food and nutrition policies and programs. Bio-fortification is plant-breeding (classical or genetic engineering) aiming to enrich staple crops with micronutrients. It has recently gained momentum in the "fight" against malnutrition. This is reflected in the program "Harvest Plus" of the International Policy Research Institute (IFPRI), which aims to bio-fortify staple crops with iron, zinc, and vitamin A. Bio-fortified crops such as vitamin-A-enriched sweet potatoes have been released in several countries. Studies show that such crops can contribute to the reduction of hidden hunger (Bouis and Saltzmann, 2017; Low et al., 2007; White and Broadley, 2009); however, bio-fortification is also criticized. Leaving aside the controversy with genetically engineered bio-fortified crops (such as Golden Rice), critics argues that bio-fortification encourages the simplifications of diets towards fewer and fewer carbohydrate staples (Johns and Eyzaguirre, 2007). It is not the aim of this thesis to evaluate bio-fortification, but clearly, as the food and nutrition policies and programs mentioned above, any studies on the usefulness of bio-fortification require the collection of food and nutrition data.

Intra-household allocation of nutrition

While there is much literature on intra-household resource allocation, nutrition has been neglected in this regard. According to Coates et al. (2018), very few studied have studies the intra-household allocation of food and nutrients: there are only 28 studies providing data from both adults and children and only four of them are from Africa. Harris-Fry (2017) found a similar neglect of this topic in Asia. This can be problematic because food and nutrition policies and programs (see above) may miss their targeted beneficiaries (under- and malnourished individuals) if it is assumed that food and nutrients are equally allocated across the household (Coates et al., 2018). Also, the above-mentioned studies, which discussed the role of farm diversity on consumption diversity, may have missed a point when neglecting such aspects.

Causal relations of food and nutrition insecurity determinants

To understand the drivers of food and nutrition security, various other potential determinants have been explored. Von Braun (2018) summarized some of these determinants, which include: farm-level determinants such as land size and soil quality; access to markets and purchasing power; the price stability of the food system; political economy determinants such as conflicts; structural determinants such as discrimination; but also individual behaviors and social capital, among others. Different researchers have explored some but not all of these aspects. This includes the effects of new farm practices such as improved maize and legume varieties, chemical fertilizers and maize–legume intercropping (all Koppmair et al, 2017), the use of GMOs (Qaim and Kouser, 2013), market access (Koppmair et al., 2017), the spread of supermarkets (Demmler et al., 2018)[6], climate change (Alfania et al., 2019; Tirado et al., 2010), women empowerment (Malapit and Quisumbing, 2015) and conflicts (Dabalen et al., 2014).

Drivers of caloric requirements

One agricultural-nutrition linkage has been forgotten more recently, however. Although the majority of African smallholders relies on hand tools for farming (FAO, 2016), linking ways on how food is produced with nutritional aspects has been neglected. This is despite the fact that heavy physical work is associated with high energy expenditure and the fact that most of these activities need to be performed during the hunger season, when the previous year's harvest is dwindling (Sitko, 2006). In contrast, farming systems that replace requirements for human energy (manual

[6] Interestingly, highly processed food, which is associated with supermarkets, can have effects on both time-use (see previous section) and on nutrition.

labor) with non-human energy (for example, draught animals, machinery, and herbicides) may significantly reduce the daily caloric requirements of farm family members. In the perfect scenario, the reduction of caloric requirements contributes to reduce undernutrition of smallholder farming households, including more vulnerable household members such as women and children. In the extreme, overshooting scenario, this may contribute to the emerging double burden of nutrition in developing countries (Popkin, 2001; Steyn and Mchiza, 2014).

1.3. Data Collection

In this section, the challenges to collect data on time-use and nutrition are outlined. In addition, different methods to collect such data will be discussed. Section 1.3.1 focuses on data collection on time-use. Section 1.3.2 shows aspects related to data collection on nutrition.

1.3.1. Data collection on time-use

The collection of time-use data can be done through several methods, all of which have some advantages and disadvantages. Different methods include household surveys, time-use diaries and direct observations.

Household surveys

Post-harvest surveys are most common in the field of agricultural economics as they are rather inexpensive and therefore allow large sample sizes. However, post-harvest surveys are yielding only very unreliable data. Post-harvest surveys ask questions such as: “how much time did you spent last farming season doing weeding?” It has been shown that such a large recall period leads to poor data. According to a landmark paper on time-use by Juster and Stafford (1991), recall questions “typically prove wide off the mark” (p. 482). Importantly, the quote from Juster and Stafford refers to their work in developed countries. In developing countries data accuracy may be further undermined and bias therefore be large. In Tanzania, for example, Arthi et al. (2018) found that farmers report a farm work time four times lower when asked via a weekly survey instead of a postharvest survey.

Several aspects contribute to the fact that post-harvest surveys often prove “wide off the mark”. The first reason is related to the design of the questionnaires. For example, when asking for the number of days spend weeding (from the example above), the assumption is that different respondents work the same amount of hours per day, which is a strong and often wrong assumption (Doss, 2018; McCullough, 2017). When asking for typical working hours to address this problem, fluctuations of daily working hours are not well captured. Despite such design biases, there are other forms of biases when working with post-harvest surveys. Importantly for researchers in developing countries, respondents may not have a clock- and calendar-based understanding of time-use, which makes any answers arbitrary. But even with knowledge on clock-based time-use, answers may be prone to recall bias. On reason is that most agricultural surveys are to be answered by the "household head", who may not be able to accurately report the work contribution of his or her kin.

In addition, there is evidence that study participants overestimate socially desirable activities and underestimate socially undesirable activities (Hofferth, 1999; Juster and Stafford 1991; Juster et al., 2003). They also underestimate activities are not perceived as work. The influence of societal opinions on work for recalling and reporting time-use can be demonstrated by looking at care activities. For developed countries, Bianchi et al. (2012) and Juster et al. (2003) found that respondents overestimate child care time. This is in sharp contrast to developing countries. In Malawi, care work is not considered as work and "dramatically underreported" (Lentz et al, 2018, p.1).

Besides this social desirability bias, there are forms of bias. Secondary activities are frequently overestimated while sporadic activities may be underestimated (Juster, 2003). In this regard, the seasonality of farming in developing countries may influence the perception of time spent on irregularly performed activities (Arthi et al., 2018). In addition, while respondents in developed countries tend to follow regular and externally structured activities, such as office work, which makes recalling time easier, smallholder farmers in developing countries may have less-structured days, which makes recalling time challenging (Arthi et al., 2018). The role of the intensity of efforts may also play a role for recalling time but has not been studied much (Jodha, 1988). In this regard, physically arduous work may be overestimated. Arthi et al. (2018) speculated that farmers, who know the amount of harvest when using post-harvest questionnaires may overestimate farm labor hours used during good harvests and underestimate hours during bad harvests.

To reduce the large recall bias of post-harvest surveys, some surveys have used weekly data collection to reduce recall biases (Arthi et al., 2018). This enhances data accuracy and, if done by phone, is not too expensive. However, time-use researchers regard any recall period beyond two days as too long ago, leading to poor data (Juster et al., 2003). Many time-use researchers, therefore, argue that some form of time-use diaries is the gold standard.

Time-use diaries

Time-use data can also be collected using time-use diaries, where study respondents enter data to 24-hour time grids with 15 to 30-minute slots – either as time goes or at the end of the day. Diaries are considered the most reliable data collection method as they minimize the dependence on recall questions (Chatzitheochari et al. 2017; Juster et al. 2003). However, activities that last shorter than the given intervals may either not be reported or grossly over-reported. Another disadvantage is that time–use diaries rely on text-based questions and must be filled in written form (Chatzitheochari et al., 2017). Therefore, they cannot be used by illiterate study participants, a problem in many developing countries. Enumerators can be used to facilitate data entry. In this case, diaries are filled jointly by the respondents and an enumerator on the subsequent day. Diaries thus become recall based, which can lead to recall bias. Entering data with respondents may lead to a form of enumerator bias, too. There have been experiments with pictorial time diaries (Masuda et al., 2014) but they have been proven cumbersome and coarse as they are based on 30-minute slots.

Direct observations

One way to avoid the problems of recall biases and to address the illiteracy problem is direct observations (Kes and Hema 2006; Paolisso and Hames 2010). However, such observations are costly and are thus typically only used for small samples (Harvey and Taylor 2000; Kes and Hema 2006). While direct observations eliminate recall bias, they are prone to another bias, the so called Hawthorne Effect, which occurs when the observer's presence affects the behavior of the observed (Kes and Hema 2006; Paolisso and Hames 2010).

ICT and smartphone applications

Given the challenges of the above mentioned data collection methods, several research groups collecting data in developed countries have conducted studies on using computer-based or smartphone-based applications for time-use studies. For example, Minnen et al. (2014) developed a modular online time use survey (MOTUS) and researchers from the Netherlands Institute for Social Research developed and pilot-tested an app-based time-use diary (Fernee and Sonck, 2014). While computer-based and smartphone-based time-use diaries make data recording easier, they are facing several limitations. They are based on 15 or 30-minute time slots using a 24-h time grid format, which can make results coarse. In addition, they rely on text elements, and thus, cannot be used when respondents lack literacy.

1.3.2. Data collection on food and nutrition

For the collection of food and nutrition data, researchers can use similar methods as researchers collecting time-use data. This compromises questionnaires working with recall data, nutrition diaries and direct observations. They may also use smartphone applications. In addition, researchers can take anthropometric measurements of respondents and take blood samples. In contrast to time-use data, more variables are of interest with regard to food and nutrition data. This includes expenditures on food, the quantity of food eaten but also the types of food consumed. Thus, different indicators may be of interest, such as money spent on food, months of inadequate household food provisioning, wasting and stunting (reflecting short- and long-term undernutrition), undernutrition-related diseases, food diversity or the level of different vitamins and trace elements in the respondent's blood samples.

Depending on the indicator of interest, researchers can choose different methods from the ones listed above. Some methods are substitutes for collecting the same type of data, others capture correlated types of data, but some are unique in the type of data that they can measure. Fongar et al. (2018) tested various dietary indicators and found that they are positively correlated but find no correlation between such dietary indicators and anthropometric measures, which suggests that dietary indicators are not a good proxy for nutritional status (Jones et al., 2014). This can be because of the influence of sanitation and hygiene as well as the role of unregularly consumed food. Fongar et al. (2018) also found that anthropometric indicators show lower levels of undernutrition than using dietary diversity indicators. In brief, most current assessment methods evaluating nutrient intake are either too inaccurate, too expensive, or too time-consuming.

Household surveys

Some types of food and nutrition data can be collected asking questions with the help of household surveys. For example, a survey can be used to ask about food expenditure as a proxy for food security. In this regard, Brzozowski et al. (2017) and Beegle et al. (2012) warned that the measurement error when recalling food expenditure is substantial. Surveys can also be used to ask about months of inadequate household food provisioning, which are usually asked for a whole year. No studies assessing the accuracy of this type of data could be found, although respondents may over-report months of food shortage to gain access to food and nutrition programs; or they may under report because of shame (Swindale and Bilinsky, 2006a).

To get a more nuanced understanding of what households and people eat, researchers can use different ways to ask for dietary diversity. For example, dietary diversity can be asked through food frequency questionnaire. Food frequency questionnaires ask about the frequency (e.g. daily, weekly) and potentially also about the quantity of food and beverage consumed during the last month or year using checklists. Food frequency questionnaires are inexpensive but associated with high recall and social desirability bias (Shim et al., 2014; Micha et al., 2018).

In addition to food frequency questionnaires, dietary diversity can also be asked for the last 7 days of 24 hours, where the consumption on each day is recorded using recall questions. Here, one can also ask about serving sizes (using standardized sizes) and about cooking methods but both such questions face challenges as well (Micha et al., 2018). Shortening the recall period reduces recall bias but increases the likelihood that day-to-day variation in dietary intakes as well as occasionally consumed food is not captured (Beegle et al., 2012; Koppmair et al., 2017; Micha et al., 2018). Fongar et al. (2018) found that, compared to using a 24–hour recall period, using a 7-day recall period leads to higher measurements of energy consumption and lower measurements of undernutrition.

There is also a concern that 24-hours recall data do not asses “usual” intake accurately, and therefore, for example, does not allow to accurately rank individuals (Micha et al., 2018). Another problem is that, unless repeated throughout the farming season, such surveys do not capture seasonal variation in dietary intakes, both with regard to quantity and diversity. This can be a problem in countries where food availability and food prices vary seasonally. To capture seasonality, one can repeat such surveys throughout the farming season, but this is expensive and it may make respondents alter their diets through self-reflection (Shim et al., 2014). Another challenge is that often up to 150 food items are asked about, which takes time. Such an approach requires highly trained interviewers (Shim et al., 2014; Micha et al., 2018).

To shorten interview times, dietary diversity can also be asked using food groups consumed as proxies. Typically, nine to twelve different food groups (such as vegetables, fruits and meats) are asked for (Swindale and Bilinsky, 2006b). Such dietary diversity scores only roughly determine whether the respondent has eaten certain food groups. Such an approach does not capture the size of servings, which is a constraint, and assumes that farmers are familiar with the concept of food groups (Micha et al., 2018). Both ways of asking about dietary diversity may come with recall and social desirability bias and respondents may also choose to not report actual intake when fatigued by the questionnaire (Shim et al., 2014; Micha et al., 2018). In addition, respondents who do not typically cook their food may be unaware of the ingredients (Shim et al., 2014). Beegle et al. (2012) found that recall questions measure lower levels of consumption than a personal diary, a method that will be discussed in the next section.

Diaries

Given the recall problems associated with using questionnaires, researchers have also used food and nutrition diaries for data collection. Here, respondents self-record food eaten shortly after eating, which reduces recall bias (Shim et al., 2014). For this, respondents need to be well trained. In addition, as filling such diaries can be burdensome, respondents need to be sufficiently motivated to carefully record data. Otherwise, they may not cooperate or not record all of the food eaten. Beegle et al. (2012) found high levels of underreporting for in illiterate and urban households completing household diaries and thus question the superiority of diaries over recall questions. In addition, respondents need to have literacy and numeracy, which can be an obstacle in developing countries (Shim et al., 2014). Researchers have used both household diaries and personal diaries, but the former may suffer from measurement error because a sole respondent may not be able to accurately measure the food intake of the other household members (Beegle et al., 2012).

Direct observations

As with collecting time-use data via direct observations, one can also observe respondent’s food consumption. This has the advantage of not requiring literacy and the fact that one can also observe how food is prepared (Shim et al., 2014). Food can then be recooked in a similar way and the nutrient contents can be analyzed. However, using direct observations to collect food and nutrition data faces similar challenges as with regard to time-use data. In particular, its costly nature and the vulnerability to observer bias.

Smartphone applications

Researchers have also developed smartphone applications that facilitate the recording of dietary intake – either by respondents themselves or with the help of interviewers. Here, respondents and interviewers choose from a pre-coded list of foods and beverages and may also record the servings size. Wald et al. (2017), for example, have developed a tablet-based calculator of inadequate micronutrient intake (CIMI). This app allows recording different typically eaten food items within the specific region, which are linked to a databank with specific nutrient profiles. The app then assesses the levels of energy, protein, and micronutrients absorbed (considering the food composition), which can be compared with recommended levels. CIMI requires a trained interviewer so data are not recorded continuously but for the last 24-hours. CIMI also has challenges recording processed food eaten outside of the homestead.

Other apps allow respondents to record data themselves, but such electronic diaries require both normal and ICT-literacy. There have also been experiments where respondents are allowed to make pictures of food consumed (Kikunaga et al., 2007; Rollo et al., 2011). Such approaches are pointing toward real-time recording of food and nutrition data, reduce respondent's burden and eliminate data entry time and costs, but they are mostly at an experimental stage and open questions remain (Shim et al., 2014). In particular, they are not immune to the above-mentioned challenges of pen-and-paper based diaries (Shim et al., 2014).

Measurements

Some types of food and nutrition data can also be observed. For example, symptoms of vitamin A deficiency can be checked for; however, such a procedure requires highly trained staff (Hendricks and Hussey, 2004). Other types of food and nutrition data can be measured. For example, anthropometric measures of weight, height, head circumference, mid upper arm circumference, and skin-fold thickness can be taken and compared with reference values (Hendricks and Hussey, 2004). Such measures can reveal stunting (an indicator for chronic undernutrition) and wasting (an indicator of recent undernutrition). Anthropometric data are reproducible, taking such measures does not require skilled personnel, and it is relatively inexpensive. However, it also has disadvantages: it can be affected by non-nutritional factors and does not reveal which nutrients exactly are lacking and which types of food have been eaten (Hendricks and Hussey, 2004; Micha et al., 2018).

Biomarkers obtained through urine or blood samples are perhaps the most straightforward and robust method to collect nutrition data. Biomarkers cannot be influenced by recall or social desirability bias and do not require respondent's literacy. However, they also face challenges. First, biomarkers do not allow to assess the intake but rather the status of nutrients, as such they do not reveal what has been eaten and, for some nutrients, also not when they were consumed/stored (Micha et al., 2018). Secondly, collecting biomarker data is expensive;

although, using blood spot samples can significantly reduce such costs. Thirdly, collecting biomarker data through urine and blood samples is invasive, which decreases compliance and participation rate (Micha et al., 2018). Fourthly, biomarkers are not as accurate as often assumed as genetic variability, lifestyle and personal factors (for example, smoking and drinking but also the presence of fever or infections and the use of medicine) as well as dietary factors (for example, nutrient–nutrient interaction) also influence biomarker measures (Hendricks and Hussey, 2004; Micha et al., 2018). Still using biomarkers can be useful. Biomarkers can also be used to validate some of the other methods discussed above.

1.4. Timetracker App

Aiming to build on the advantages of some of the above mentioned data collection methods while overcoming some of their challenges, we developed a smartphone app called Timetracker. The app allows the real-time recording of time-use and nutrition data using visual tools. The real-time recording of data eliminates the recall bias of existing data collection methods. Using visual tools for data entry ensure that illiterate respondents can use the app. With regard to time-use, we have visualized 88 time-use categories (see Figure 2).

Figure 2. Time-use activities of the Timetracker

Figure 3 provides an overview of the Timetracker app. Respondents click on the respective time-use category when they start an activity. When they are done, they click on the activity again and can then start another activity. The app allows to record up to three simultaneous activities (such as hoeing and taking care of a child). Figure 3 also shows two food and nutrition plug-in. These two windows pop up when the activity “eating and drinking” has been terminated. They ask, using four differently filled plates, for how much food was consumed; they also ask, using twelve different food groups suggested by Swindale and Bilinsky (2006b), which types of food where were consumed.

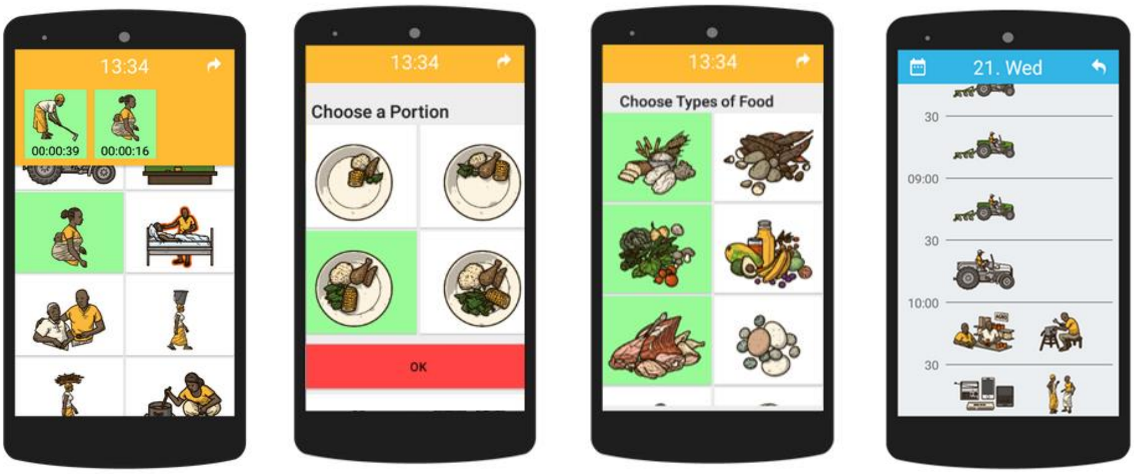


Figure 3. Different screens of the Timetracker

(A) the main interface, (B) a pop up window to enter portion sizes, (C) a pop up window to record food groups eaten, and (D) a control screen showing recorded data.

When using the app, other functions of the smartphone can be blocked. This prevents respondents from being distracted to use the phone for other purposes. It also reduces the temptation to "lose" the smartphones, which were borrowed to the respondents in this case. As a welcomed side effect, blocking other functions of the phones enhances the battery life up to four days. The app has a screen for crosschecking recorded data, which allows enumerators to correct potential mistakes. In this thesis, data recording and submission was done offline but data submission could also be done online. The smartphone application will be described with more details in chapter 2 and chapter 3.

1.5. Agricultural mechanization

Agricultural mechanization growth has been rapid in various Asian countries (Takeshima, 2017; Wang et al., 2016). In addition, based on the recognition that labor, rather than land, limits development for many smallholder farmers, agricultural mechanization has also received growing attention in Africa (Daum and Birner, 2017, Diao et al. 2014; Nin-Pratt and McBride 2014). Agricultural mechanization is, in fact, an umbrella term. Its technologies can be targeted towards different crops and farm operations - ranging from land preparation (e.g., using a tractor and a plough or ripper), on-farm processing (e.g., using a sheller) and other steps along the agricultural value chain.

1.5.1. Agricultural mechanization and time-use

Agricultural mechanization is a labor- (or time-) saving technology. Human labor is replaced by animal or mechanical traction; a technology as such should lead to a rise of agricultural labor productivity. Gollin (2018) reported that farmers in Kenya produce only 1.2 kg maize per hour whereas farmers in Iowa, United States of America, produce 1,470 kg. This does not mean that farmers from Iowa work less - even though they could if they were satisfied with producing the same amount of maize as Kenyan farmers. This shows that the effects of mechanization on time spent on different activities per day are not clear. Do mechanized farmers spend the time saved to cultivate more land, on other farm activities, on off-farm activities or for leisure?

The effects of mechanization on time-use are also unclear when agricultural activities are gendered, where men and women, boy and girls have different workloads and duties, which is the case in many smallholder farming households, (Arora, 2015; Blackden and Wodon, 2006; Doss, 2001). This division of labor by gender can be based on crops, tasks or both. For example, ploughing tends to be done relatively more by men and weeding and processing by women (Alesina, 2011; Raney, 2010; Baanante et al., 1999; Doss, 2001, Foster and Rosenzweig, 1996). As new technologies target different crops and tasks, they may affect men and women, boy and girls differently. Agricultural mechanization can target different farming tasks. The effects of mechanization on time-use thus depend on which tasks and crops are mechanized, on the original allocation of labor between the household members, on how this allocation can be re-negotiated (based on power relations), and on the use of accompanying inputs such as herbicides and hired labor.

Typically, land preparation activities are mechanized first, which may be because land preparation tends to be a main labor bottleneck, but this may, to some extent, also reflect preferences to adopt technologies pertaining to male crops and male activities (Evers and Walters, 2001). With land preparation being mechanized, households may cultivate additional land, which may increase the need for weeding, harvesting, and processing or the time spent collecting firewood once forests are cleared, tasks often performed by women and children (Arora, 2015; Blackden and Wodon, 2006; Doss, 2001). Despite such concerns, the effects of mechanization on intra-household allocation of labor have not been examined by the literature, notwithstanding some anecdotal evidence.

1.5.2. Agricultural mechanization and nutrition

Agricultural mechanization can affect farm household nutrition via different pathways. This thesis will focus on one specific pathway, which could be called energy requirement pathway. The rationale is that the majority of African smallholders use manual labor for farming, which implies (at least seasonally) heavy physical work and high caloric energy requirements. Agricultural mechanization replacing manual labor with non-human energy (for example, draught animals, machinery, and herbicides) may significantly reduce the daily caloric requirements of farm family members. The reduction of caloric requirements can contribute to reducing undernutrition of smallholder farming households, including more vulnerable household members such as women and children. This would be important, given the high prevalence of both under- and malnutrition among smallholder farmers (FAO et al., 2018; Pinstrup-Andersen, 2007). However, the reduction of caloric requirements may also contribute to the emerging double burden of nutrition in developing countries with and associated increased risk of non-communicable diseases (Popkin, 2001; Steyn and Mchiza, 2014).

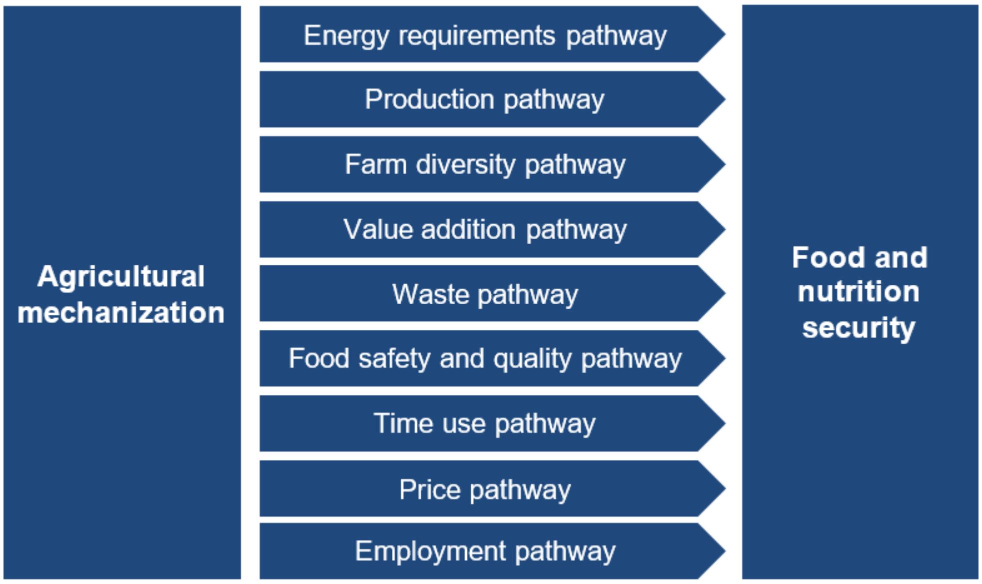


Figure 4. Links between agricultural mechanization and food and nutrition security outcomes

While this thesis focuses on the energy requirements pathway, there are other pathways, some of which have been equally neglected. For example, the well-established production pathway: if agricultural mechanization allows farmers to produce more, this can translate into a higher food consumption and improved diets. But agricultural mechanization could also reduce dietary diversity. For example, when farmers focus on easy-to-mechanize crops such as wheat and maize and do not balance their diets by buying additional food from markets, reduced dietary diversity would be observed, a causal linkage that could be called farm diversity pathway. There can also be a value addition pathway: by peeling, chipping, grating, and drying, farmers can produce cassava chips, starch and flour, which can be sold at higher prices, thus leading to higher incomes and more money for food (Malabo Montpellier Panel, 2018). In addition, there can be a waste pathway: harvesting and processing machines may influence the amount of food waste. Another pathway is the food safety and quality pathway: applying fertilizer and pesticides more precisely can reduce food contamination; drying and cooling, better storage, and transport can help to better preserve food (and nutrients) and reduce contamination with fungi such as aflatoxins and bacteria; enhancing food safety can reduce the risk of diarrhea, which reduces nutrient absorption capacity; and guaranteeing food safety may give farmers access to markets paying higher prices (Malabo Montpellier Panel, 2018). Another pathway is the time-use pathway: agricultural mechanization may alter the time available off-farm work but also activities like cooking. In the latter case, if more time becomes available, more nutritious food may be prepared and consumed rather than processed food (Seymour et al., 2019). There can also be a price pathway if agricultural mechanization reduces costs and thus commodity prices. If only some crops are mechanized and become cheaper, the relative prices for different commodities can change, thus leading to a change in crops grown by farmers and diets of consumers. There may also be other pathways, for example, an employment pathway: if mechanization leads to fewer employment opportunities for agricultural laborers, then this can affect the nutrition status of their households.

1.6. Zambia

The field work for this thesis was done in Zambia. Zambia is considered one of the most sparsely populated countries in Sub-Saharan Africa. Of the whole population, between 60 and 70% depend on agriculture for their livelihoods; they have an average land size of 3.2 ha (Tembo and Sitko; 2013). Most of them do not cultivate all their land because of labor constraints, among other factors. Agriculture is dominated by smallholder farmers, but during the last years, the number of medium-scale farmers (between 5 and 20 ha of land) has grown rapidly. Between 2001 and 2011,

Figure 5. Sub-Sahara-Africa and Zambia

the population of these "emergent farmers" has grown by 62.2%, thus "vastly outstripping the 33.5% growth rate of the total smallholder population" (Sitko and Jayne, 2014, p. 194). Zambia ranks 144 out of 189 countries on the Human Development Index (2018).[7] According to the Global Hunger Index, Zambia has alarming rates of both under- and malnutrition and ranks 115 out of 119 countries.[8]

[7] http://hdr.undp.org/en/2018-update (accessed April 20th 2019)
[8] https://www.globalhungerindex.org/results/ (accessed April 20th 2019)

Figure 6. Provinces of Zambia
Source: https://commons.wikimedia.org/wiki/File:Zambia_provinces_named.png

Data collection took place in the Eastern Province of Zambia. Agriculture is dominated by smallholders (less than 2 ha) and medium-scale farmers (between 5 and 20 ha). Households cultivate on average 2.3 ha (IARPI, 2016), female-headed households cultivate 1.36 ha on average (Tembo and Sitko, 2013). Farmers usually own more land than this, which they cannot cultivate due to labor shortages. The main crops grown are the staple crop maize, groundnuts (as subsistence or cash crops) as well as cash crops such as cotton, sunflower, and tobacco. Farming is characterized by a short rainy season and an extensive dry season. Mechanization levels are low: 1 % of the farmers use own or hired mechanical traction and 57 % use animal traction – mainly for land preparation (IARPI, 2016). The access to inputs such as herbicides and improved seeds is limited, but fertilizers, which are subsidized, are commonly used (IARPI, 2016). Accordingly, land- and labor-productivity is low, resulting in low household incomes. 90 % of the rural population live on less than US$ 1.25 day (IARPI, 2016).

1.7. Research topics and overview of the thesis

The thesis consists of three papers, which correspond to the three objectives outlined above. The first paper (second chapter) discusses the potentials to use smartphone apps as a new method to collect more accurate data on smallholder farming systems. The chapter first reflects on the need to collect accurate data and to what extent existing data collection methods fulfill this need. It then discusses whether smartphone apps can be used to collect better data. For this, the chapter first condenses lessons learned from the use of user-oriented smartphone apps (apps which provide services to farmers or the rural population). It also reflects on the existing examples where apps have been used to collect data from respondents both in the developed and developing world. The chapter then presents the Timetracker app developed for this thesis. It extensively reflects on preconditions that need to be addressed when using smartphone apps to collect data from smallholder farmer in developing countries. Finally, the chapter discusses the advantages and disadvantages of using such apps as well as future research avenues for the use of such apps.

The second paper (third chapter) specifically focuses on using apps for time-use studies. The advantages and disadvantages of existing methods to collect time-use data are discussed, particularly with a view on collecting data in developing countries. The chapter then compares the accuracy of two methods to collect time-use data: a) conventional 24-hour recall questionnaires and b) the developed smartphone app for this thesis, the Timetracker app. The results suggest that the app leads to valid results. Further methods to validate the data collected with apps as well as other limitations are discussed. Directions for future research are shown.

The third paper (fourth chapter) is an empirical study investigating the effects of agricultural mechanization on the intra-household division of time-use in smallholder farming households in Zambia, paying special attention to seasonality, gender and age. The study uses compositional data analysis, which accounts for the intrinsic codependence of time-use data, as well as different regression tools. The results both confirm and question existing literature on time-use, agriculture, and gender. In general, the study finds that agriculture to be less gendered than often assumed. With regard to land preparation, a gender differentiation only emerges with mechanization. As mechanized land preparation may lead to area expansion and a higher labor demand during weeding and harvesting, the chapter tests for potential negative second round effects during subsequent farming steps; but finds no negative effects. Time saved due to mechanization is channeled to multiple different activities, which makes detection of significant differences difficult. There is some evidence of increased time spend on off-farm work (by women) and domestic work (by men). The results provide a proof-of-concept that using picture-based smartphone apps can help to collect data on difficult-to-measure but potentially highly relevant research areas such as how new technologies, practices, and policies affect the allocation of time-use within households.

In the last chapter, the methodological and empirical contributions of the study are discussed. Limitations of the Timetracker application are shown and potential ways to address these limitations are described. The chapter concludes with suggestions for future research.

1.8. References

Adu-Baffour, F., Daum, T., and Birner, R. 2019. Can small farms benefit from big companies' initiatives to promote mechanization in Africa? A case study from Zambia. Food Policy, 84, 133-145.

Arora, D. 2015. Gender Differences in Time-Poverty in Rural Mozambique. Review of Social Economy, 73(2), 196-221.

Arthi, V., Beegle, K., de Weerdt, J., Palacios-López, A. 2018. Not your average job: Measuring farm labor in Tanzania. Journal of Development Economics, 130, 160–72.

Alesina, A. F., Giuliano, P., and Nunn, N. 2011. On the origins of gender roles: Women and the plough (No. 17098). National Bureau of Economic Research.

Alfania, F., Andrew, D., Fisker P., Casco, M. 2019. Vulnerability to stunting in the West African Sahel. Food Policy, 83, 39-47.

Ayenew, H. Y., Biadigilign, S., Schickramm, L., Abate-Kassa, G., and Sauer, J. 2018. Production diversification, dietary diversity and consumption seasonality: panel data evidence from Nigeria. BMC Public Health, 18(1), 988.
Baanante, C., Thompson, T. P., & Acheampong, K. 1999. Labour contributions of women to crop production activities in three regions of West Africa: an analysis of farm-survey data. Institute of African Studies Research Review, 15(1), 80-100.
Balmann, A., 1997. Farm-based Modelling of Regional Structural Change: A Cellular Automata Approach. European Review of Agricultural Economics, 24, 85-108.
Beegle, K., Christiaensen, L., Dabalen, A., and Gaddis, I. 2016. Poverty in a rising Africa. The World Bank. Washington, D.C.
Beegle, K., De Weerdt, J., Friedman, J., and Gibson, J. 2012. Methods of household consumption measurement through surveys: Experimental results from Tanzania. Journal of Development Economics, 98(1), 3-18.
Berger, T., 2001. Agent-based models applied to agriculture: a simulation tool for technology diffusion, resource use changes and policy analysis. Agricultural Economics, 25(2/3), 245-260.
Bergman Lodin, J., Paulson, S., and Mugenyi, M. S. 2012. New seeds, gender norms and labor dynamics in Hoima District, Uganda. Journal of Eastern African Studies, 6(3), 405-422.
Bianchi, S. M., Sayer, L. C., Milkie, M. A., Robinson, J. P. 2012. Housework: Who did, does or will do it, and how much does it matter? Social Forces, 91, 55–63.
Bianchi, S. M. and Milkie M. A. 2010. Work and family research in the first decade of the 21st century. Journal of Marriage and Family, 72, 705-725.
Blackden, M., Wodon, Q., 2006. Gender, Time-use, and Poverty in Sub-Saharan Africa. World Bank Working Paper No. 73. The World Bank. Washington, D.C.
Boserup, E. 1965. The Condition of Agricultural Growth. Allen and Unwin. London.
Bouis, H. E., and Saltzman, A. 2017. Improving nutrition through biofortification: a review of evidence from HarvestPlus, 2003 through 2016. Global Food Security, 12, 49-58.
Brzozowski, M., Crossley, T. F., and Winter, J. K. 2017. A comparison of recall and diary food expenditure data. Food Policy, 72, 53-61.
Buvinic, B., King, E. 2018. Invisible No More? A Methodology and Policy Review of How Time-use Surveys Measure Unpaid Work. Data2X.
Carletto, C., Corral, P., and Guelfi, A. 2017. Agricultural commercialization and nutrition revisited: Empirical evidence from three African countries. Food Policy, 67, 106-118.
Carletto, C., Gourlay, S., and Winters, P. 2015. From guesstimates to GPStimates: Land area measurement and implications for agricultural analysis. Journal of African Economies, 24(5), 593-628.
Chatzitheochari, S., Fisher, K., Gilbert, E., Calderwood, L., Huskinson, T., Cleary, A., Gershuny, J. 2017. Using new technologies for time diary data collection: Instrument design and data quality findings from a mixed-mode pilot survey. Social Indicators Research 137, 379–390.
Coates, J., Patenaude, B., Rogers, P., Roba, A., Woldetensay, Y., Tilahun, F., Spielman, K. 2018. Intra-household nutrient inequity in rural Ethiopia. Food Policy, 81, 82-94.
Dabalen, A. L., and Paul, S. 2014. Effect of conflict on dietary diversity: Evidence from Côte d'Ivoire. World Development, 58, 143-158.
Dangour, A.D., Hawkesworth, S., Shankar, B., Watson, L., Srinivasan, C.S., Morgan, E.H., Haddad, L., Waage, J. 2013. Can nutrition be promoted through agriculture-led food price policies? A systematic review. BMJ Open, 3(6), e002937.
Daum, T., and Birner, R. 2017. The neglected governance challenges of agricultural mechanisation in Africa insights from Ghana. Food Security, 9(5), 959-979.
Deaton, A. 2013. The great escape: health, wealth, and the origins of inequality. Princeton University Press. Princeton.
Deininger, K., Carletto, C., Savastano, S., Muwonge, J. 2012. Can diaries help in improving agricultural production statistics? Evidence from Uganda. Journal of Development Economics, 98(1), 42–50.
Demmler, K. M., Ecker, O., and Qaim, M. 2018. Supermarket shopping and nutritional outcomes: A panel data analysis for urban Kenya. World Development, 102, 292-303.
Diao, X., McMillan, M., and Wangwe, S. 2017. Agricultural Labour Productivity and Industrialisation: Lessons for Africa. Journal of African Economies, 27(1), 28-65.
Diao, X., Cossar, F., Houssou, N., and Kolavalli, S. 2014. Mechanization in Ghana: Emerging demand, and the search for alternative supply models. Food Policy, 48, 168-181.
Doss, C. R. 2018. Women and agricultural productivity: Reframing the Issues. Development Policy Review, 36(1), 35-50.
Doss, C. R. 2001. Designing agricultural technology for African women farmers: Lessons from 25 years of experience. World Development, 29(12), 2075–2092.
Evenson, R. E., and Gollin, D. 2003. Assessing the impact of the Green Revolution, 1960 to 2000. Science, 300(5620), 758-762.

Evers, B., and Walters, B. 2001. The Model of a Gender-Segregated Low-Income Economy Reconsidered: Evidence from Uganda. Review of Development Economics, 5(1), 76-88.
Fanzo, J. C. 2017. Decisive decisions on production compared with market strategies to improve diets in rural Africa. Journal of Nutrition, 147(1), 1–2.
FAO, IFAD, UNICEF, WFP and WHO. 2018. The State of Food Security and Nutrition in the World 2018. Building climate resilience for food security and nutrition. Food and Agriculture Organization. Rome.
FAO. Agricultural mechanization: a key input for sub-Saharan Africa smallholders. Integrated Crop Management Vol. 23. Food and Agriculture Organization. Rome.
Farnworth, C. R., Baudron, F., Andersson, J. A., Misiko, M., Badstue, L., and Stirling, C. M. 2016. Gender and conservation agriculture in East and Southern Africa: towards a research agenda. International Journal of Agricultural Sustainability, 14(2), 142-165.
Fernee, H., Sonck, N. 2014. Measuring smarter: time-use data collected by smartphones. Electronic International Journal of Time-use Research, 11(1), 94–96.
Fisher, M., R. Warner, and W. Masters. 2000. Gender and agricultural change: Crop-livestock integration in Senegal. Society and Natural Resources, 13(3), 203–222.
Floro IV, V. O., Labarta, R. A., Becerra López-Lavalle, L. A., Martinez, J. M., and Ovalle, T. M. 2018. Household Determinants of the Adoption of Improved Cassava Varieties using DNA Fingerprinting to Identify Varieties in Farmer Fields: A Case Study in Colombia. Journal of Agricultural Economics, 69(2), 518-536.
Fongar, A., Gödecke, T., Aseta, A., and Qaim, M. 2018. How well do different dietary and nutrition assessment tools match? Insights from rural Kenya. Public Health Nutrition, 22(3), 391-403.
Foster, A. D., & Rosenzweig, M. R. 1996. Comparative advantage, information and the allocation of workers to tasks: Evidence from an agricultural labour market. The Review of Economic Studies, 63(3), 347-374.
Fox, M. P., Rosen, S., MacLeod, W. B., Wasunna, M., Bii, M., Foglia, G., and Simon, J. L. 2004. The impact of HIV/AIDS on labour productivity in Kenya. Tropical Medicine and International Health, 9(3), 318-324.
Fuglie, K., and Rada, N. 2013. Resources, Policies, and Agricultural Productivity in Sub-Saharan Africa. USDA-ERS Economic Research Report 145.
Gershuny, J. 2011. Time-use surveys and the measurement of national well-being. Centre for Time-use Research, Department of Sociology, University of Oxford. Office of National Statistics. Swansea.
Giller, K. E., Witter, E., Corbeels, M., and Tittonell, P. 2009. Conservation agriculture and smallholder farming in Africa: the heretics' view. Field Crops Research, 114(1), 23-34.
Glorieux, I., Minnen, J., Tienoven, T. P. V., Deyaert, J., and Mészáros, E. 2015. Evolutions in time-use and division of labour of men and women. Revue Interventions économiques. Papers in Political Economy, (53).
Goldsmith, P. D., Gunjal, K., and Ndarishikanye, B. 2004. Rural–urban migration and agricultural productivity: the case of Senegal. Agricultural Economics, 31(1), 33-45.
Gollin, D. 2018. Farm Size and Productivity. FAO, IFAD, ISPC/CGIAR, World Bank Expert Consultation: Focusing Agricultural and Rural Development Research and Investment on Achieving SDGs 1 and 2: https://ispc.cgiar.org/sites/default/files/files/events/Joint%20Initiative%202018/Gollin.pdf 11th January 2018. Rome.
Gutierrez, L. 2002. Why is agricultural labour productivity higher in some countries than others? Agricultural Economics Review, 3(1), 58-72.
Happe, K., Kellermann, K., Balmann, A. 2006. Agent-based analysis of agricultural policies: an Illustration of the Agricultural Policy Simulator AgriPoliS, its adaptation and behavior. Ecology and Society, 11(1), 49.
Harris-Fry, H., Shrestha, N., Costello, A., and Saville, N. M. 2017. Determinants of intra-household food allocation between adults in South Asia–a systematic review. International Journal for Equity in Health, 16(1), 107.
Harvey, A. S., Taylor, T. 2000. Designing household survey questionnaires for developing countries: Lessons from 15 years of the Living Standards Measurement Survey. The World Bank. Washington, DC.
Hazell, P., Norton, R. 1986. Mathematical programming for economic analysis in agriculture. Macmillan. New York.
Headey, D. D., and Hoddinott, J. 2016. Agriculture, nutrition and the green revolution in Bangladesh. Agricultural Systems, 149, 122-131.
Hendricks, M. K., and Hussey, G. 2004. The field assessment of nutrition. In Gershwin, M. E., Nestel, P., and Keen, C. L. (Eds.). Handbook of Nutrition and Immunity (pp. 19-47). Springer. Heidelberg.
Hofferth, S. 1999. Family reading to young children: Social desirability and cultural biases in reporting. Workshop on measurement and research on time-use. National Research Council. Washington, DC.
IAPRI (Indaba Agricultural Policy Research Institute). 2016. Rural Agricultural Livelihoods Survey: 2015 Survey Report. Lusaka.

IFPRI (International Food Policy Research Institute). 2016. Global Nutrition Report 2016: From Promise to Impact: Ending Malnutrition by 2030. Washington, D.C.
Jodha, N. S. 1988. Poverty debate in India: A minority view. Economic and Political Weekly, 23, 2421–2428.
Johns, T., and Eyzaguirre, P. B. 2007. Biofortification, biodiversity and diet: A search for complementary applications against poverty and malnutrition. Food Policy, 32(1), 1-24.
Johnston, D., Stevano, S., Malapit, H. J., Hull, E., and Kadiyala, S. 2018. 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.
Jones, A. D. 2017. On-farm crop species richness is associated with household diet diversity and quality in subsistence- and market-oriented farming households in Malawi. Journal of Nutrition, 147(1), 86–96.
Jones, A. D., Ickes, S. B., Smith, L. E., Mbuya, M. N., Chasekwa, B., Heidkamp, R. A., Menon, P., Zongrone, A. A., Stoltzfus, R. J. 2014. World Health Organization infant and young child feeding indicators and their associations with child anthropometry: a synthesis of recent findings. Maternal and Child Nutrition, 10(1), 1-17.
Juster, F.T., Ono, H., Stafford, F.P. 2003. An assessment of alternative measures of time-use. Sociological Methodology. 33 (1), 19–54.
Juster, F. T., Stafford, F. P. 1991. The allocation of time: Empirical findings, behavioural models, and problems of measurement. Journal of Economic Literature, 29, 471–522
Katzner, D. W. 2012. Unmeasured information and the methodology of social scientific inquiry. Springer Science and Business Media. Heidelberg.
Kikunaga, S., Tin, T., Ishibashi, G., Wang, D. H., and Kira, S. 2007. The application of a handheld personal digital assistant with camera and mobile phone card (Wellnavi) to the general population in a dietary survey. Journal of Nutritional Science and Vitaminology, 53(2), 109-116.
Kosmowski, F., Aragaw, A., Kilian, A., Ambel, A., Ilukor, J., Yigezu, B., and Stevenson, J. 2019. Varietal identification in household surveys: results from three household-based methods against the benchmark of DNA fingerprinting in southern Ethiopia. Experimental Agriculture, 53(3), 371-385.
Koppmair, S., Kassie, M., and Qaim, M. 2017. Farm production, market access and dietary diversity in Malawi. Public Health Nutrition, 20(2), 325-335.
Kes, A., Hema, S. 2006. In Gender and time poverty in sub-Saharan Africa. World Bank Working Paper No. 73. The World Bank. Washington, DC.
Lambrecht, I., B. Vanlauwe, R. Merckx, and M. Maertens. 2014. Understanding the process of agricultural technology adoption: Mineral fertilizer in eastern DR Congo. World Development, 59, 132–146.
Lentz, E., Bezner Kerr, R., Patel, R., Dakishoni, L., and Lupafya, E. 2018. The Invisible Hand that Rocks the Cradle: On the Limits of Time-use Surveys. Development and Change, 50(2), 301-328.
Lewis, W. A. 1955. The Theory of Economic Growth. George Allen and Unwin. London.
Lipton, M. 1980. Migration from rural areas of poor countries: the impact on rural productivity and income distribution. World Development, 8(1), 1-24.
Low, J. W., Arimond, M., Osman, N., Cunguara, B., Zano, F., and Tschirley, D. 2007. A food-based approach introducing orange-fleshed sweet potatoes increased vitamin A intake and serum retinol concentrations in young children in rural Mozambique. The Journal of Nutrition, 137(5), 1320-1327.
Masuda, Y. J., Fortmann, L., Gugerty, M. K., Smith-Nilson, M., Cook, J. 2014. Pictorial approaches for measuring time-use in rural Ethiopia. Social Indicators Research, 115, 467–82.
Malabo Montpellier Panel. 2018. Mechanized: Transforming Africa's agriculture value chains. International Food Policy Research Institute (IFPRI) and Malabo Montpellier Panel. Dakar.
Malabo Montpellier Panel. 2017. Nourished: How Africa Can Build a Future Free from Hunger and Malnutrition. International Food Policy Research Institute (IFPRI) and Malabo Montpellier Panel. Dakar.
Malapit, H. J. L., and Quisumbing, A. R. 2015. What dimensions of women's empowerment in agriculture matter for nutrition in Ghana? Food Policy, 52, 54-63.
McCullough, E. B. 2017. Labor productivity and employment gaps in Sub-Saharan Africa. Food Policy, 67, 133-152.
Micha, R., Coates, J., Leclercq, C., Charrondiere, U. R., and Mozaffarian, D. 2018. Global Dietary Surveillance: Data Gaps and Challenges. Food and Nutrition Bulletin, 39(2), 175-205.
Ng, S. W., and Popkin, B. M. 2012. Time-use and physical activity: a shift away from movement across the globe. Obesity Reviews, 13(8), 659-680.
Nin-Pratt, A., and McBride, L. 2014. Agricultural intensification in Ghana: Evaluating the optimist's case for a Green Revolution. Food Policy, 48, 153-167.
Paolisso, M., Hames, R. 2010. Time diary versus instantaneous sampling: A comparison of two behavioral research methods. Field Methods, 22, 357–77.
Pinstrup-Andersen P. 2007. Agricultural research and policy for better health and nutrition in developing countries: A food systems approach. Agricultural Economics, 37, 187–198.

Popkin, B. M. 2001. The nutrition transition and obesity in the developing world. The Journal of Nutrition, 131(3), 871-873.
Qaim, M., and Kouser, S. 2013. Genetically modified crops and food security. PloS one, 8(6), e64879.
Raney, T., Anriquez, G., Croppenstedt, A., Doss, C., Gerosa, S., Lowder, S., Matuschke, I., and Skoet, J. (2010). The Role of Women in Agriculture. Agricultural Development Economics Working Papers ESA No 10-02. Food and Agriculture Organization. Rome.
Rollo, M. E., Ash, S., Lyons-Wall, P., and Russell, A. 2011. Trial of a mobile phone method for recording dietary intake in adults with type 2 diabetes: evaluation and implications for future applications. Journal of Telemedicine and Telecare, 17(6), 318-323.
Rubiano-Matulevich, E. and Kashiwase, H. 2018. Why time-use data matters for gender equality—and why it's hard to find. The Data Blog. From https://blogs.worldbank.org/opendata/why-time-use-data-matters-gender-equality-and-why-it-s-hard-find (retrieved March 24 2019). World Bank. Washington, DC.
Schreinemachers, P., Berger, T., Aune, J.B. 2007. Simulating soil fertility and poverty dynamics in Uganda: A bio-economic multi-agent systems approach. Ecological Economics, 64(2), 387-401.
Schultz, T. W. 1953. The Economic Organisation of Agriculture. McGraw-Hill. New York.
Scott, J. C. 1998. Seeing like a state: How certain schemes to improve the human condition have failed. Yale University Press.
Seymour, G., Masuda, Y. J., Williams, J., and Schneider, K. 2019. Household and child nutrition outcomes among the time and income poor in rural Bangladesh. Global Food Security, 20, 82-92.
Sibhatu, K. T., and Qaim, M. 2018. Review: Meta-analysis of the association between production diversity, diets, and nutrition in smallholder farm households. Food Policy, 77, 1-18.
Sibhatu, K. T., Krishna, V. V., and Qaim, M. 2015. Production diversity and dietary diversity in smallholder farm households. Proceedings of the National Academy of Sciences, 112(34), 10657-10662.
Sitko, N. J., and Jayne, T. S. 2014. Structural transformation or elite land capture? The growth of "emergent" farmers in Zambia. Food Policy, 48, 194–202.
Shim, J. S., Oh, K., and Kim, H. C. 2014. Dietary assessment methods in epidemiologic studies. Epidemiology and Health, 36, e2014009.
Steyn, N. P., and Mchiza, Z. J. 2014. Obesity and the nutrition transition in Sub-Saharan Africa. Annals of the New York Academy of Sciences, 1311(1), 88-101.
Swindale, A., Bilinsky, P. 2006a. Household Dietary Diversity Score (HDDS) for Measurement of Household Food Access: Indicator Guide. Food and Nutrition Technical Assistance Project, Academy for Educational Development. Washington, DC.
Swindale, A., and Bilinsky, P. 2006b. Development of a universally applicable household food insecurity measurement tool: process, current status, and outstanding issues. The Journal of Nutrition, 136(5), 1449-1452.
Takeshima, H. 2017. Overview of the evolution of agricultural mechanization in Nepal: A focus on tractors and combine harvesters. IFPRI Discussion Paper 1662. International Food Policy Research Institute. Washington, DC.
Tembo, S., and Sitko, N. 2013. Technical Compendium: Descriptive Agricultural Statistics and Analysis for Zambia. Lusaka: Working Paper 76, Indaba Agricultural Research Institute (IAPRI). Lusaka.
Theis, S., Lefore, N., Meinzen-Dick, R., and Bryan, E. 2018. What happens after technology adoption? Gendered aspects of small-scale irrigation technologies in Ethiopia, Ghana, and Tanzania. Agriculture and Human Values, 25(3), 671-684.
Tirado, M. C., Cohen, M. J., Aberman, N., Meerman, J., and Thompson, B. 2010. Addressing the challenges of climate change and biofuel production for food and nutrition security. Food Research International, 43(7), 1729-1744.
Tremblay, M. S., and Willms, J. D. 2003. Is the Canadian childhood obesity epidemic related to physical inactivity? International Journal of Obesity, 27(9), 1100.
Turner, R., Hawkes, C., Waage, J., Ferguson, E., Haseen, F., Homans, H., Hussein, J., Johnston, D., Marais, D., McNeill, G., Shankar, B. 2013. Agriculture for improved nutrition: the current research landscape. Food Nutrition Bulletin, 34, 369–377.
von Braun, J. 2018. Innovations to Overcome the Increasingly Complex Problems of Hunger. ZEF Working Paper 167. Center for Development Research. University of Bonn. Bonn.
Von Braun, J., and Webb, P. J. 1989. The impact of new crop technology on the agricultural division of labor in a West African setting. Economic Development and Cultural Change, 37(3), 513-534.
Wald, J., Asare, E., Nakua, E., Lambert, C., Biesalski, H., Gola, U., Nohr, D. 2017. Dietary assessment using the CIMI approach: A case study from three districts of the Ashanti region in Ghana. Presentation at the 3rd International Conference on Global Food Security, Cape Town, South Africa.
Wang, X., Yamauchi, F., and Huang, J. 2016. Rising wages, mechanization, and the substitution between capital and labor: evidence from small scale farm system in China. Agricultural Economics, 47(3), 309-317.

White, P. J., and Broadley, M. R. 2009. Biofortification of crops with seven mineral elements often lacking in human diets–iron, zinc, copper, calcium, magnesium, selenium and iodine. New Phytologist, 182(1), 49-84.

2. Smartphone apps as a new method to collect data on smallholder farming systems in the digital age: A case study from Zambia

***Daum, T.**, Buchwald, H., Gerlicher, A., and 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. https://doi.org/10.1016/j.compag.2018.08.017*

Abstract

Across the developing world, the spread of mobile- and smartphones has led to a surge in mobile services for rural populations. While the potentials of mobile services to provide development opportunities for smallholder farmers are widely acknowledged, the potentials to use smartphone applications to collect data on smallholder farming systems are little explored. Yet, researchers studying farming systems need good quality data. So far, data on smallholder farming systems is typically collected using household surveys. Survey questions are prone to recall biases, however, which can be substantial. This paper assesses whether smartphone can be used to collect data in real time and thus increase the accuracy of socioeconomic and agronomic data collection. In this paper, we present a smartphone application that was developed for this purpose. We use the application to analyze the effects of agricultural mechanization on intra-household time-use and nutrition in rural Zambia. While the early, descriptive results shed interesting light on the effects of mechanization, the contribution of this study is primarily methodological. The study highlights the potentials of using smartphone applications to collect socioeconomic and agronomic data on smallholder-farming systems, potentially in real time. It also suggests ways to combine data recorded by respondents with built-in sensors of smartphones and external sensors and thus shows fascinating new pathways for researchers in the digital age.

Key words

Smartphone apps; Methodology; ICTs; Data collection; Time-use; Nutrition

2.1. Introduction

Across the developing world, ownership of mobile phones is rapidly increasing (ITU, 2016). This has led to a surge of mobile tools, which help smallholder farmers to access agricultural, health, educational, and financial services (Baumüller, 2012). The more recent rise of smartphones now creates possibilities to also use applications that are based on visual tools (which make them usable for users with no or low literacy levels) and that may work with data obtained from in-built sensors. There is a consensus, that such applications offer new potentials for smallholders to improve their agricultural production systems. For example, smallholder farmers in the Senegal may use a cloud-based decision-support tool to apply fertilizer more accurately (Saito et al., 2015). And Argentinian farmers may use a smartphone app to better time fungicide applications (Carmona et al., 2017). These examples show the potentials that apps offer for farmers. However, they may also provide new opportunities for researchers to collect data in complex smallholder farming systems. So far, these potentials have been little explored.

For socioeconomic studies, data on smallholder farming systems is typically collected through household surveys. In agronomic studies, there has been an increasing emphasis on on-farm research, partly motivated by donor priorities (De Roo et al., 2017). In on-farm research, there is typically a need to collect not only agronomic data from the field experiments, but also to collect data from the participating farmers through interviews (Vugt et al., 2017). As smallholders do not usually keep any records, researchers collecting socioeconomic or agronomic data from smallholders rely on recall questions. There is strong evidence, that the error that may be introduced through recall bias can be substantial (Arthi et al., 2017; Deininger et al., 2012). Increasing the frequency of data collection is, however, associated with major costs. The use of smartphone apps can considerably reduce these costs and increase the accuracy of data collection, if the farmer himself or herself enters data in real time. In this paper, we present a smartphone app that was developed for this purpose. We demonstrate our experience using this app when analyzing the effects of agricultural mechanization on intra-household time-use and nutrition in smallholder farming systems in Zambia.

We chose to focus on mechanization because there has been a renewed interest to promote agricultural mechanization in Africa (Daum and Birner, 2017). The intra-household effects of mechanization are, however, ambiguous. Positive effects may include increased income through the expansion of the area that a household cultivates. However, the expansion of the land area cultivated by a household may, increase the burden of labor for activities that are not yet mechanized, such as weeding and harvesting, which are often carried out by women and children (Blackden and Wodon, 2006). The changes in (female) time-use may in turn alter the nutrition status of the household members, including the children (Johnston et al., 2015). Analyzing such intra-household effects requires the collection of data on time-use and nutrition from different household members over an entire cropping season.

Intra-household data is difficult or expensive to collect using conventional methods. Household surveys are prone to large recall biases (Arthi et al., 2017; Juster et al., 2003). Time-use diaries where respondents fill out a 24-hours' time grid that is divided to 15 or 30 minutes slots with pre-coded activities, are an alternative, but they are burdensome and difficult to use if respondents cannot read or write and lack a “modern” or clock-based concept of time (Kes and Hema, 2006). Direct observations eliminate recall biases and address the problem of illiteracy and the lack of “modern” concepts of time. However they are expensive and the presence of the researchers may affect the behavior of the observed (Kes and Hema, 2006). The collection of nutrition data faces similar challenges. As a consequence, there is a lack of reliable intra-household data on time-use and nutrition in smallholder farming systems, which makes it difficult for governments and development practitioners to prioritize and design development programs and policies as well as to measure their effects (Johnston et al., 2015).

Against this background, we developed smartphone application called “Time-Tracker”. We successfully applied this easy-to-use application in 62 households in the Eastern Province of Zambia. The households were selected to represent different levels of mechanization. The application allowed respondents to record their daily activities and nutritional intake in real-time to avoid recall bias. The application is picture-based so that all user groups can to participate. Altogether, 2790 days of data were collected. While the results shed interesting light on the intra-household effects of mechanization, the contribution of this study is primarily methodological and seeks to highlight the potentials of using smartphone applications as well as sensors to study socioeconomic and agronomic aspects of smallholder farming systems.

2.2. Methodological Considerations

During the last decade, the number of mobile phones has increased rapidly in the developing world (ITU, 2016). In Sub-Saharan Africa, the rate of unique mobile subscribers was 43% in 2016, as compared to 66% worldwide (GMSA, 2017). The rate of adults who own a mobile phone was higher than this: for example, 83% and 82% of adults owned a phone in Kenya and Ghana, respectively in 2014 (PEW, 2015). A quarter of phones used are smartphones (GMSA, 2017). In Zambia, even in rural areas 54% of households have at least one person owning a mobile phone (IARPI, 2016). The growth of mobile phone ownership has led to a surge of mobile services that aim to address the challenges faced by rural populations (Baumüller, 2012; Aker et al., 2016). For example, Saito et al. (2015) developed a cloud-based decision-support tool for field-specific fertilizer recommendation in Senegal. And Bueno-Delgado et al. (2017) developed a tool that helps farmers to minimize fertilizer costs.

While there is widespread consensus about the development opportunities that these mobile services offer, some challenges remain. For example, the text-based nature of most mobile services is a barrier for low- and illiterate users (Aker et al., 2016). The recent rise of smartphones allows developing of visual apps that may also help to overcome the challenges faced by low- and non-literate users. While the rise of smartphones provides new opportunities both with regard to the use of visuals and the use of its built-in sensors, there are still few app-based mobile services (Baumüller, 2016). One of them is “Hello Tractor” – a geo-tracking service that helps farmers to access tractor services in Nigeria, however it still contains text-elements.

While there are several examples of user-oriented mobile services, as shown above, the potentials to use mobile phones as research tools are little explored, especially in the developing world. One notable exception is a pilot study by the Makerere University in Uganda to monitor the spread of pests with camera phones (Quinn et al., 2011). Another exception is an SMS-tool to evaluate the impact of farmer training developed by Technoserve in Tanzania (Baumüller, 2012). One of the first pilot studies to actually use a smartphone application as a research tool in a developing country was done by a research group in rural Bangladesh (Bell et al., 2016). Their application allows participants to answer questions related to social data on a weekly basis. While making a strong case for the use of smartphone apps for research purposes in rural areas, the developed application is still mostly text-based. In brief, there are promising but few attempts to use mobile phones, and especially smartphone apps, as research tools, more importantly, most of the existing apps are text-based thereby excluding low and non-literate users.

In contrast to developing countries, there are already several examples of studies using smartphone apps as research tools in developed countries. Most of them are related to health questions such as "MyHeart Counts" (see https://med.stanford.edu/myheartcounts.html). There are also attempts to use smartphone apps for time-use studies. For example, the Netherlands Institute for Social Research developed an app-based time-use diary to test its viability as a research tool (Fernee and Sonck, 2014). However, this is implicitly based on a cumbersome and imprecise 24-hour time grid format and relies on text-based questions. While this attempt shows the potentials of using apps for time-use research, it is difficult to use in developing countries where large parts of the population are illiterate or lack a "modern" or clock-based concept of time.

Given the lack of adopted data collection methods for developing countries, there are few studies analyzing time-use within households and even fewer focusing on rural areas in developing countries (Blackden and Wodon, 2006; Johnston, 2015). Existing studies frequently rely on survey-questions but are prone to high recall and social desirability errors (Chatzitheochari et al., 2017; Juster and Stafford, 1985). An interesting attempt to find a method for time-use research in developing countries is the use of pictorial-diary-sets, which Masuda et al. (2014) pilot-tested with ten households in Ethiopia. The set consisted of a booklet, activity stickers and a timer, which beeps at 30-minute intervals. Whenever the timer beeps, respondents would place a sticker in the booklet, which reflected with their activity at that time. While having advantages compared to other data collection methods, this procedure is still cumbersome, does not capture simultaneous activities and is inaccurate given the use of 30 minutes intervals.

2.3. Materials and Methods

To make use of the fast spread of ICTs and to overcome the challenges of conventional data collection methods, we developed an open source, Android-based smartphone application called "Time-Tracker" that allows respondents to record data themselves (see 2.3.1).[9] We applied the app with 62 households in the Eastern Province of Zambia, which represent different levels of farm mechanization (see 2.3.2).

[9] The code can be accessed via https://github.com/HannesBuchwald/TimeTracker.

2.3.1. Data Collection Technique

The “Time-Tracker” is intuitive and fast to use. The data-entry is picture-based to ensure usability by those who are illiterate and/or have never used smartphones before and to ensure that respondents do not develop an “entry fatigue”. The main screen (see Figure 7) shows buttons with drawings of 88 typical daily activities, ranging from farming to household chores to social activities. The participants press a bottom with the picture of the respective activity (e.g., weeding by hand) when he or she starts that activity, which starts recording (by triggering a time stamp). The button is pressed again when the activity is terminated, which ends recording (again triggering a time stamp). This real-time recording reduces recall biases. To capture simultaneity, the participants can select up to three activities at the same time (e.g., weeding by hand and caring for a baby). The participants can also indicate whether they do piecework, meaning work as agricultural laborers on the field of others (by pressing the respective activity three times). The currently recorded activities are always displayed on top of the screen (normally highlighted in green but highlighted in blue when the respective activity is done in piecework).

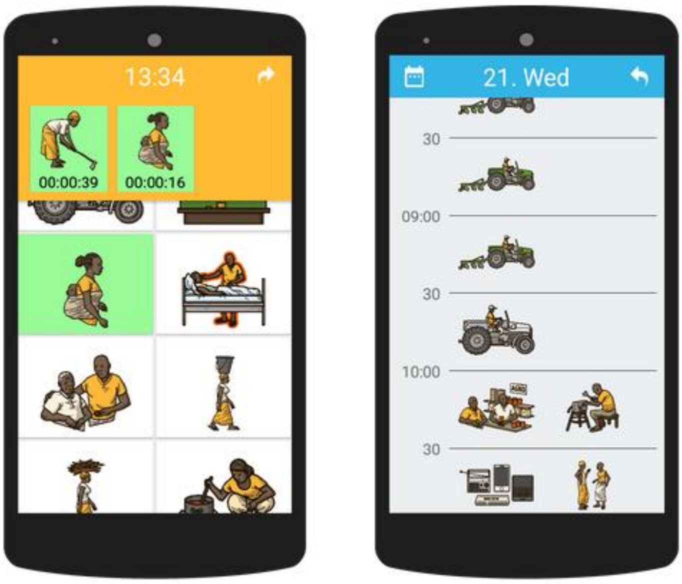


Figure 7. Main screen of application (left) and second screen (right)

A “plug-in” has been designed if the selected activity is “eating” (Figure 8). In this case, two new windows open when the activity is terminated. The first window shows four differently filled plates. Here, the respondents can record the quantities of food consumed. The second window shows twelve different food groups as suggested by Swindale and Bilinsky (2006), which allows to calculate food diversity scores.

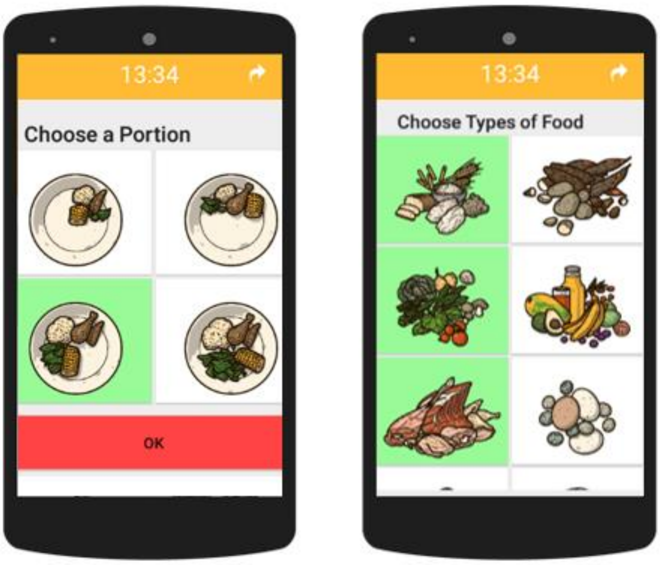


Figure 8. Nutrition “Plug-In”

Respondents and researchers can flip to a second screen to review recorded activities (shown in a 24-hour time grid with 30 minute intervals).[10] This screen allows the researchers to crosscheck the data that was recorded with the respondent and to thus identify potential gaps or wrongly entered data. The researchers can correct potential mistakes by activating a hidden button, which is not shown to the respondents. The daily activities are stored on the smartphone every 15 minutes. The data can quickly be transferred from the application to a laptop (PC) via a local Wi-Fi network.

2.3.2. Data Collection Site and Sampling Procedure

Smallholders dominate agriculture in the Eastern Province of Zambia; where the average size of land cultivated is 2.3 ha (IARPI, 2016). Farmers grow mainly maize but also cotton, sunflower, groundnuts and tobacco. Farming is characterized by a short rainy season and an extensive dry season. 1 % of the households use own or hired mechanical power and 57 % use animal traction - at least for some crop husbandry operations (IARPI, 2016). The access to inputs such as herbicides and improved seeds is limited, only fertilizers are commonly used (IARPI, 2016). Accordingly, land and labor-productivity is low, resulting in a low household income - 90 % of the rural population live on less than 1.25 US$/day (IARPI, 2016).

We used two-stage-random-sampling to select respondents based on the population of the nationally representative Zambian Rural Agricultural Livelihood Survey (RALS). First, we randomly selected four survey clusters in the Eastern Province where at least five households were non-mechanized, five households used animals and one household was mechanized in 2014/2015. Second, in each cluster we randomly selected five non-mechanized households, five animal power-using households and five to six mechanized households from the RALS-

[10] The time-use data is aggregated only for graphical purposes to avoid extensive scrolling.

population under the condition that households have one adult man, one adult woman and at least one child. Additional households were randomly selected based on lists from the District Agriculture and Cooperatives Offices when the RALS-population was exhausted according to this criterion.

In total, we selected 62 households. 20 of them used only manual labor, 20 used draft-animals and 22 used tractors services for land preparation that was provided by tractor owners. In each household, the head of the household, one spouse and one child (alternating between boys and girls) were trained to use the app. As time-use varies along the season, the respondent used the application over a period of three days at five points of the 2016/2017 farming season. This includes times around land preparation, fertilizer application/planting, weeding, harvesting and processing. To avoid selection biases, respondents were lent smartphones and waterproof pouches to carry them. In total we used 50 smartphones; 45 to 48 respondents used the application at a time. One smartphone costs around 90 US-Dollars. The phones were blocked so that only the application could be used to avoid the temptation of keeping the phone or of using it for purposes other than data recording. Restricted use increased the battery-life to up to five days. A research assistant visited the participating households once a day to check if the smartphones worked properly and if the data was entered correctly. When the battery-life was below 50 %, respondents were given power-banks to charge their smartphones.

2.4. Results

In this section we reflect on our experiences of introducing the app (2.4.1.) and report some illustrative findings (2.4.).

2.4.1. Preconditions for using the smartphone app

This section describes the steps that were used to meet the challenges of using a smartphone app as a research tool in rural areas of low-income countries, which are characterized by high levels of illiteracy and limited prior exposure to smartphones.

2.4.1.1. Need for pre-testing

We found it essential to pre-test the application in the selected study area, in our case rural areas of Zambia, to ensure that the users do not have problems implementing the app. During the pre-test, two main aspects were crucial: the design of the application and the illustrations used. To be able to adjust the design, it was highly valuable that the app-programmer participated in the pre-testing in Zambia. This allowed us to make changes directly in the field and to immediately test them with our respondents, the members of smallholder families. We ensured that male and female family members of different age groups participated in the pre-test. As an example of the adjustment identified by the pre-test, we changed the clicking-style from short to long press after we found that the respondents accidently clicked activities. We also changed the size of the illustrations after we realized that some of the respondents could not identify small illustrations because they had debility of sight without access to glasses (a common problem in rural areas of low-income countries). In addition, we experimented with grouping the 88 activities into sub-folders (having an overarching picture for them) to allow for quicker navigation through the activities. This turned out to be not easy to understand for most of the farmers so that we decided to show all activities at one screen but to still group them thematically.[11] We also cross-checked all illustrations with smallholder farmers from the selected study site to ensure that they were unambiguous and easy to understand taking the local pictorial language into account. Several illustrations had to be adjusted as a consequence. The cross checking was also crucial to ensure that no activities were overlooked. Some new illustrations had to be designed, for example, charcoal making. Pre-testing also allowed us to illustrate the food group items according to local diets. In our case, the illustrator did not participate directly in the pre-testing. However, to facilitate the process, we found it would be useful to have the illustrator on site for the pre-testing as well.

2.4.1.2. Need for considering social context

When introducing smartphone apps in low income-countries where smartphones are not yet widely used, it is important to take the role of village authorities, social dynamics and beliefs (e.g. in Satanism) within the village into account. As it is good practice in such research, we first approached the village authorities and explained the purpose of the research project. Particular emphasis was placed on explaining why a smartphone app was used as a research tool and why the participating households had to be selected randomly. The research team also explained that the participating households received a small gift (e.g., caps) for their participation, but the smartphones were only made available for the period of data collection and could not be used for other purposes. Without a careful explanation of these conditions, suspicions and tensions may easily arise in the village community that certain households received special favors. It was also important to explain the conditions to the selected household members, so that they were motivated to participate and to use the app carefully. None of the randomly selected respondents in our study declined to participate.

2.4.1.3. Need for training

We found it was crucial to provide sufficient training to the research participants. We trained a maximum of three persons in one household at a time for approximately 60 minutes. We found that the respondents had no difficulties to understand the logic and function of the app. Some respondents were unfamiliar with the use of a touch screen, so we allowed them to practice the scrolling through and touching of activities first. We then went through all of the illustrations to ensure that they are clearly understood - which was usually the case. The nutrition “plug-in”

[11] Importantly, this has not led to respondents’ fatigue (e.g., clicking only at activities at the top screen). Our data, shows not skewness towards activities at the beginning.

required more time since we had to carefully explain the 12 different food groups. We practiced the use of the application together with the respondents using some explanatory stories, such as the following: Mary goes hoeing, then she takes the bicycle to buy groceries, then she cooks while taking care of a baby, then she eats *nshima* (a common dish in Zambia made of maize flour), vegetables and eggs. They respondents then used the "Time-Tracker" App during one afternoon to practice before the actual recording started.

Due to the careful training, all selected respondents were able to use the app. The age of our respondents ranged from 6 to 90 years and we found the application to be user-friendly to all age groups. Only two respondents who had problems with their eyesight mentioned that even after adjusting the size of the illustrations, they were difficult to read, but they were still able to enter data. We found that respondents, who are illiterate as well as those without any experience with smartphones, or even any type of phones, had no problems with using the app. This indicates that the use of a smartphone app does not lead to selection biases if appropriate training is provided. Respondents reported enjoying using the app, handled the smartphones with care and recorded their time-use with much discipline. For example, in the first round of data collection, only 0.6 %, of the data had to be entered or changed with the help of the research assistants who visited the households once a day, because respondents forgot to enter data or clicked the wrong activity. One of the smartphones got lost during data collection.

2.4.2. Illustrative Results

To provide proof for the concept of using a smartphone app such as "Time-Tracker" in rural areas of low-income countries, we report descriptive results on the time spent by household members and on dietary intake during the weeding season. The difference reported across mechanization categories are not causal effects as we do not yet correct for confounding factors. Out of the five data collection points, the weeding season was selected because it is one major labor bottlenecks of farming and because it allows us to pre-test our hypothesis that mechanization increases the work burden for women during weeding.

Table 1 provides an overview of how respondents spend their time during the weeding season. On average, respondents entered 20.29 activities each day, lasting on average 52.32 minutes (excluding sleeping). For brevity, Table 1 shows only activities that are done for longer than 15 minutes per day on average. Households using hand tools spend significantly more time on farming than households using animal and mechanical traction. Members from households that use animal and mechanical traction also spend more time on mobility and transportation activities such as walking and social life activities. Across farm-power-categories, we found that, expectedly, females spend significantly more time on household chores and care and that males spend more time on social life activities, expanded farming activities and mobility and transport-related activities. Children spend almost the same amount of time on farming activities as do adults.

Activity categories	P1: Manual			P2: Animal			P3: Mechanical			P-Significance	G-Significance	PG-Significance
	G1: Male (60)	G2: Female (60)	G3: Child (54)	G1: Male (60)	G2: Female (60)	G3: Child (60)	G1: Male (66)	G2: Female (66)	G3: Child (63)			
Farming	296	281	262	230	235	253	206	202	187	*0.000*	*0.834*	*0.717*
Land Preparation	14	25	13	35	3	35	49	9	31	*0.331*	*0.033*	*0.051*
Planting	10	12	8	4	20	11	10	14	13	*0.878*	*0.258*	*0.795*
Fertilization	22	9	10	40	39	25	25	14	10	*0.014*	*0.208*	*0.934*
Weeding	234	224	219	147	158	172	120	151	120	*0.000*	*0.792*	*0.673*
Extended Farming [1]	41	11	24	66	7	12	88	30	23	*0.062*	*0.000*	*0.348*
Mobility/ Transport	119	108	122	221	102	132	224	109	158	*0.000*	*0.000*	*0.002*
Care	2	32	48	4	102	17	4	18	9	*0.002*	*0.000*	*0.000*
Household Chores	19	192	126	8	238	112	17	177	119	*0.410*	*0.000*	*0.045*
Personal Care	836	778	789	767	730	782	776	786	817	*0.029*	*0.104*	*0.196*
Social Life	170	83	144	187	130	227	207	164	212	*0.002*	*0.000*	*0.572*
Total	1493	1492	1515	1508	1544	1544	1532	1510	1528	*0.141*	*0.529*	*0.648*

Table 1. Analysis of variance of time-use

By sex and age (G) and farm-power-category (P). In Minutes/Day. In total, 543 Observations. Number of observations of each group (e.g. P1G1) in brackets. [1]Marketing, Animal Husbandry, Hunting, Fishing, Gathering, Charcoal Making, Maintaining and Repairing, Farm Administration, Vegetable Garden. The categories of community work, meetings, education and construction are not shown here. P values for comparative means based on gender and age (G), farm-power-category (P) and both (PG) are reported in italics. P values below 0.1 and 0.05, 0.01 indicate that mean differ significantly at the 10% and 5% and 1% level, respectively.

As shown in Table 1, households using hand tools spend significantly more time on weeding than households that use animal and mechanical traction. This is surprising since households who are able to use animal or mechanical traction for land preparation tend to expand their farm size, which results in a higher demand for weeding. However, these households also use animals for weeding and/or they use knapsack sprayers to apply herbicides (Table 2). Also, mechanization may suppress weed growth. Thus, we do not see an increase of (female) time spent on weeding due to mechanization. In contrast, the time spent on weeding by females' declines nearly by half from 224 to 151 minutes per day. The time men spend on weeding declines much more sharply, however, especially when we look at manual weeding.

Weeding	P1: Manual			P2: Animal			P3: Mechanical			P-Significance	G-Significance	PG-Significance
	G1: Male (60)	G2: Female (60)	G3: Child (54)	G1: Male (60)	G2: Female (60)	G3: Child (60)	G1: Male (66)	G2: Female (66)	G3: Child (63)			
Manual	175	209	199	104	135	131	57	144	94	*0.000*	*0.003*	*0.483*
Manual Piecework	46	10	7	0	13	9	0	0	0	*0.001*	*0.191*	*0.002*
Animal	4	5	12	37	10	32	45	7	22	*0.010*	*0.005*	*0.145*
Knapsack	7	0	0	5	0	0	16	0	3	*0.242*	*0.002*	*0.627*
Total	234	224	219	147	158	172	120	151	120	*0.000*	*0.792*	*0.673*

Table 2. Analysis of variance of weeding time-use

By sex and age (G) and farm-power-category (P). In Minutes/Day. Only activities that are done for more than 15 min per day on average are shown. P values for comparative means based on gender and age (G), farm-power-category (P) and both (PG) are reported in italics. P values below 0.1 and 0.05, 0.01 indicate that mean differ significantly at the 10% and 5% and 1% level, respectively.

Table 3 shows data from the nutrition “plug-in”. The table indicates that households that use manual labor consume lower amounts of food (as calculated by the sum of the daily portions) than households using animal or mechanical traction. This is a potential problem since they are likely to have higher energy requirements because they spend more time on heavy physical work (Table 1). Interpreting the findings, one has to take into account that the size of the portion that the respondent indicated on the smartphone app is based on his or her subjective assessment. This may explain why male respondents tend to record lower daily portions per day than females and children. Table 3 also shows the daily food diversity scores, an indicator that was less dependent on a subjective assessment than food quantity. Households that use animals or tractors have significantly higher daily food diversity scores than non-mechanized households. Households that use animals or tractors also consume more meat products than non-mechanized households. This effect was statistically significant and the magnitude was considerable. The same applies to the consumption of fats and oils. Interestingly, members of non-mechanized households consumed more vegetables, potentially wild mushrooms and green leafy field vegetables, which grow during the rainy season and are basically available for free.

Nutrition "Plug-In"	P1: Manual			P2: Animal			P3: Mechanical			P-Significance	G-Significance	PG-Significance
	G1: Male (60)	G2: Female (60)	G3: Child (54)	G1: Male (60)	G2: Female (60)	G3: Child (60)	G1: Male (66)	G2: Female (66)	G3: Child (63)			
Cereals	**1.9**	**1.9**	**2.1**	**1.8**	**2.0**	**2.0**	**1.7**	**1.7**	**1.8**	*0.011*	*0.152*	*0.860*
Vegetables	**1.5**	**1.5**	**1.7**	**1.4**	**1.5**	**1.5**	**1.4**	**1.3**	**1.4**	*0.018*	*0.223*	*0.715*
Roots/Tubers	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.1	*0.000*	*0.850*	*0.903*
Fruits	0.2	0.2	0.3	0.3	0.3	0.3	0.4	0.4	**0.4**	*0.071*	*0.602*	*0.879*
Meats	0.2	0.2	0.2	**0.5**	**0.5**	**0.5**	**0.5**	**0.4**	**0.5**	*0.000*	*0.828*	*0.930*
Eggs	0.2	0.3	**0.4**	0.3	0.4	0.3	0.3	0.4	0.3	*0.666*	*0.611*	*0.450*
Fish/Seafood	0.2	0.2	0.2	0.3	0.3	0.2	0.3	0.3	0.3	*0.191*	*0.757*	*0.947*
Pulses/ Legumes/ Nuts	0.1	0.1	0.1	0.3	0.3	0.2	0.2	0.2	0.3	*0.002*	*0.925*	*0.719*
Milk products	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	*0.273*	*0.719*	*0.897*
Oils/Fats	0.2	0.3	0.3	**0.6**	**0.5**	**0.5**	**0.6**	**0.6**	**0.6**	*0.000*	*0.915*	*0.989*
Sugar/Honey	0.1	0.1	0.0	0.1	0.1	0.1	0.2	0.1	0.1	*0.126*	*0.380*	*0.967*
Condiments	0.0	0.1	0.0	0.1	0.0	0.1	0.2	0.1	0.2	*0.000*	*0.471*	*0.717*
Daily DD score	**3.0**	**3.2**	**3.2**	**4.1**	**4.0**	**4.0**	**4.4**	**4.1**	**4.4**	*0.000*	*0.801*	*0.772*
Daily portion sum	**1.0**	**1.1**	**1.2**	**1.3**	**1.4**	**1.5**	**1.3**	**1.3**	**1.4**	*0.000*	*0.069*	*0.866*

Table 3. Analysis of variance of nutrition "plug-in"

During weeding time by sex and age (G) and farm-power-category (P). A value of 2 means that respondents eat the respective item twice a day. A value of 0.5 means that respondents eat the respective item every second day. P values for comparative means based on gender and age (G), farm-power-category (P) and both (PG) are reported in italics. P values below 0.1 and 0.05, 0.01 indicate that mean differ significantly at the 10% and 5% and 1% level, respectively. Bold numbers indicate values of above 0.4 to improve readability.

2.5. Discussion

This discussion of the results is divided into three sections. The first section discusses the potentials and challenges of the “Time-Tracker” app. The second section discusses the potential of expanding the application to study additional aspects of smallholder farming systems. In section three, we discuss the potentials of using smartphone apps in rural areas of low-income countries more generally.

The analysis of the data collected by using the “Time-Tracker” app shows that this approach made it possible to gain highly valuable insights on the intra-household implications of agricultural mechanization. It would have been difficult, if not impossible, to collect this data using conventional data collection methods. The recording of data in real time nearly eliminates recall biases, which are inherent to other data collection methods – especially, household surveys. The use of visual tools instead of text ensures that even illiterate people can use the application, which has been a major drawback of time-use diaries. Also, the recording of data in real time makes the data more precise than data collected in the 30 minutes intervals typically used by time-use diaries. The “Time-Tracker” also addresses the problem of time-use diaries that respondents tend to under-report activities that are much shorter than the 30 minutes intervals (Chatzitheochari et al., 2017; Kelly et al., 2015). Using the app, respondents can record any lengths of activity, even activities lasting less than a minute. However, respondents may not record very short activities as this may be not feasible or inconvenient in some situation (for example, when a baby starts crying and need immediate help).

The entering of data by respondents themselves reduces the costs of data collection, which is a major drawback of using direct observations to collect time-use data. Also, compared to direct observations, respondents are less likely to change their behavior while using the application because they are not directly observed by a researcher (Kelly et al., 2015). The number of participants (in this case, 186) that can be reached is higher than the typical sample size of direct observations. However, it was lower than the sample size of a typical household survey. However, this study was only a proof-of-concept study.

The cost for the development of the app is a fixed cost, which can be spread over the total number of applications. This number can be quite high since the “Time-Tracker” is a highly flexible tool that can be used in different contexts with limited adjustments. Depending on the focus of the study, the activity sets can be adapted. In our study, we used only one illustration representing all livestock related activities (herding, feeding, rearing etc.). For a study focusing on livestock keepers, one could split this activity into several different ones. Likewise, the illustration-sets could be adapted for other professions and for urban settings.

A feature that makes the “Time-Tracker” valuable for many different applications is the “plug-in” opportunities. In our case, we used the app to “plug in” a module on nutrition. Other “plug-in” modules could be developed, as well. For example, respondents could provide details on fertilizer application or pesticide application when they are doing this activity. Likewise, they could record information on their use of agricultural extension services or financial services, and this information could be passed on to service providers, which may use this information to improve their services.

The app could also be further developed to offer participants the opportunity to take photos that can subsequently be analyzed by the researchers. The mobile tool developed by Quinn et al. (2011) showed that camera phones could be used to study the spread of pests. The "Time-Tracker" may also be linked to fitness-trackers that measure physical activities. This would make it possible to estimate more accurately the food energy requirements caused by physical labor. The app could also be linked with motion, environmental and position sensors that are already integrated into newer smartphones. The use of position sensors may be interesting for studies that focus on land use, studies with pastoralists, or studies looking at migration patterns. The app may also be used to validate agricultural plot-sizes, an important and often bad-stated variable if agricultural productivity indicators are to be assessed (Carletto et al., 2013). In general, the use of sensor-equipped devices provides additional opportunities for transdisciplinary studies with a socioeconomic *and* a natural science focus (plant production). While it could be problematic to overload the application, the use of sensors makes it possible to collect data "en passant" and without burdening the respondents.

Cameras, fitness-trackers and position sensors (GPS) may also be used to ensure data quality. For example, using fitness-trackers researchers could validate whether the time-use activities recorded by the respondents fit with the physical activity levels recorded with the help of the fitness-trackers. Similarly, researchers could validate, for example, whether respondents are really on the field when recording a farm activity by using geo-tracking. Fitness-trackers and position sensors could be used ex-post during data cleaning. However, they could also be used while respondents record data. In this case, in-built-reminders may be used, for example, when respondents do not record a new activity for a long time or when the recorded activities do not seem to match physical activity ratios and geographical positions. However, even without such extra devices, data quality was good. In Daum et al. (forthcoming), we compare the data recorded with the app with data collected with the help of 24-hours recall questions and find a high data quality.

While offering a range of new research opportunities, the potential challenges of using smartphone apps as a research tool need to be considered, as well. In particular, it is important to pay attention to the ethical implications of this approach. Data on time-use, physical activities, nutrition and location are personal data, which may have sensitive implications. In addition to standard good practices in research, such as ensuring informed consent and anonymity, special measures are required to ensure the safety of the collected data. To avoid that data is improperly accessed we therefore stored the data on the device itself and transmitted the collected data without using the Internet. Still, it would be more rigorous to encrypt the data.

While using mobile services (and to a lesser extent smartphone apps) is increasingly common among rural populations in developing countries, the use of mobile phones, especially smartphones, as research tools was thus far rarely considered as viable. This study shows that smartphone apps such as the “Time-Tracker” can indeed by used as research tools in rural areas of developing countries. While certain preconditions must be fulfilled, e.g., the addressing the challenges of low and non-literate users, this study suggests that smartphone apps open a new world for research – both to improve existing data collection methods and to create entirely new research approaches. As outlined above for the research in smallholder farming systems, smartphone apps provide not only the potentials to use visual aids but also to use cameras and built-in-sensors. Most of the areas for new research discussed above focused on smallholder farming systems but smartphone apps can equally be used to study other aspects of rural areas, for example related to health, governance and education. This suggests that there is a big and still untapped potential to use smartphone apps for research in rural areas of low-income countries. As smartphone penetration levels are low in rural areas of developing countries, today’s researchers may still need to lend smartphones to respondents to avoid selection bias. However, smartphone ownership is raising rapidly even in the remotest corners of the world. This would allow for crowd sourcing and the realization of citizen science.

2.6. Conclusions

The study provides proof of concept that smartphone apps, if appropriately designed, can serve as a reliable, affordable and participatory tool for data collection in complex smallholder farming systems. The experience with the “Time-Tracker” indicates that such smartphone apps can be well used in rural areas of low-income countries, where illiteracy levels are high and previous exposure to smartphones is low. This study also proves that a smartphone app can make it possible to collect data that is normally difficult or costly to obtain using other methods. While we focused on the collection of time-use and nutrition data, there is much wider untapped potential. In our case, we used an application focused on time-use to “plug in” a module on nutrition. Similarly, other “plug-in” modules could be developed for further transdisciplinary uses of such smartphone apps. Smartphone application may also be linked with motion, environmental and position sensors of smartphones. Overall, the “Time-Tracker” provides a promising example of new approaches for research on agricultural systems in the digital age. In our study, the smartphones were provided to the respondents as research tools for the period of data collection, since smartphone ownership in rural areas of developing countries is still low. However, this may change rapidly. With increasing ownership of smartphones, apps such as “Time-Tracker” can also be used for approaches of “crowd sourcing” and “citizen science” that include the rural poor in low-income countries.

2.7. References

Aker, J.C., Ghosh, I., Burrell, J., 2016. The promise (and pitfalls) of ICT for agriculture initiatives. Agric. Econo. 47 (S1), 35–48.

Arthi, V., Beegle, K., De Weerdt, J., Palacios-López, A., 2018. Not your average job: measuring farm labor in Tanzania. J. Dev. Econ. 130, 160–172.

Baumüller, H., 2016. Agricultural service delivery through mobile phones: local innovation and technological opportunities in Kenya. In: Gatzweiler, F.W., von Braun, J. (Eds.), Technological and Institutional Innovations for Marginalized Smallholders in Agricultural Development. Springer, pp. 143–159.

Baumüller, H., 2012. Facilitating agricultural technology adoption among the poor: the role of service delivery through mobile phones. ZEF Working Paper Series No. 93.

Bell, A., Ward, P., Killilea, M., Tamal, M., 2016. Real-time social data collection in rural Bangladesh via a ‘Microtasks for Micropayments’ platform on android smartphones. PLoS ONE 11 (11).

Blackden, M., Wodon, Q., 2006. Gender, Time-use, and Poverty in Sub-Saharan Africa. World Bank Working Paper No. 73.

Bueno-Delgado, M.V., Molina-Martínez, J.M., Correoso-Campillo, R., Pavón-Mariño, P., 2016. Ecofert: An Android application for the optimization of fertilizer cost in fertigation. Comput. Electron. Agric. 121, 32–42.

Carletto, C., Savastano, S., Zezza, A., 2013. Fact or artifact: The impact of measurement errors on the farm size–productivity relationship. J. Dev. Econ. 103, 254–261.

Carmona, M.A., Sautua, F.J., Pérez-Hernández, O., Mandolesi, J.I., 2018. AgroDecisor EFC: first Android™ app decision support tool for timing fungicide applications for management of late-season soybean diseases. Comput. Electron. Agric. 144, 310–313.

Chatzitheochari, S., Fisher, K., Gilbert, E., Calderwood, L., Huskinson, T., Cleary, A., Gershuny, J., 2018. Using new technologies for time diary data collection: instrument design and data quality findings from a mixed-mode pilot survey. Soc. Indic. Res. 137 (1), 379–390.

Daum, T., Buchwald, H., Gerlicher, A., Birner, R., 2019. Times have changed. Using a pictorial smartphone app to collect time-use data in rural Zambia. Field Methods 31 (1) (forthcoming).

Daum, T., Birner, R., 2017. The neglected governance challenges of agricultural mechanisation in Africa. Insights from Ghana. Food Secur. 9, 959.

Deininger, K., Carletto, C., Savastano, S., Muwonge, J., 2012. Can diaries help in improving agricultural production statistics? Evidence from Uganda. J. Dev. Econ. 98 (1), 42–50.

De Roo, N., Anderson, J., Krupnik, T., 2017. On-farm trials for development impact? The organization of research and the scaling of agricultural technologies. Exp. Agric. 1–22.

Fernee, H., Sonck, N., 2014. Measuring smarter: time-use data collected by smartphones. Electr. Int. J. Time-use Res. 11 (1), 94–96.

GSMA, 2017. The Mobile Economy. Sub-Saharan Africa 2017. Groupe Speciale. Mobile Association, Zürich.

IAPRI (Indaba Agricultural Policy Research Institute), 2016. Rural Agricultural Livelihoods Survey: 2015 Survey Report.

ITU, 2016. The World in 2016: ICT Facts and Figures. International Telecommunication Union, Geneva.

Juster, F.T., Ono, H., Stafford, F.P., 2003. An assessment of alternative measures of time-use. Sociol. Methodol. 33 (1), 19–54.

Juster, F.T., Stafford, F.P., 1985. Time, Goods, and Well-Being. Survey Research Center, Institute for Social Research, University of Michigan.

Johnston, D., Stevano, S., Malapit, H.J.L., Hull, E., Kadiyala, S., 2015. Agriculture, gendered time-use, and nutritional outcomes: a systematic review. IFPRI Discussion Paper 01456.

Kes, A., Hema, S., 2006. Gender and time poverty in sub-Saharan Africa. In Gender, time-use, and poverty in sub-Saharan Africa. World Bank Working Paper No. 73, pp. 13–38.

Kelly, P., Thomas, E., Doherty, A., Harms, T., Burke, Ó., Gershuny, J., Foster, C., 2015. Developing a method to test the validity of 24 hour time-use diaries using wearable cameras: a feasibility pilot. PloS One 10 (12).

Masuda, Y.J., Fortmann, L., Gugerty, M.K., Smith-Nilson, M., Cook, J., 2014. Pictorial approaches for measuring time-use in rural Ethiopia. Soc. Indic. Res. 115 (1), 467–482.

PEW, 2015. Cell Phones in Africa: Communication Lifeline. Press Release Pew Research Center, Washington, DC.

Quinn, J., Leyton-Brown, K, Mwebaze, E., 2011. Modelling and monitoring crop disease in developing countries. In: Presented at the Twenty-Fifth AAAI Conference on Artificial Intelligence, San Francisco, 7–11 Aug 2011.

Saito, K., Diack, S., Dieng, I., N’Diaye, M.K., 2015. On-farm testing of a nutrient management decision-support tool for rice in the Senegal River valley. Comput. Electron. Agric. 116, 36–44.

Swindale, A., Bilinsky, P., 2006. Household Dietary Diversity Score (HDDS) for Measurement of Household Food Access: Indicator Guide. Food and Nutrition Technical Assistance Project, Academy for Educational Development, Washington, DC.
Van Vugt, D., Franke, A.C., Giller, K.E., 2017. Participatory research to close the soybean yield gap on smallholder farms in Malawi. Exp. Agric. 53 (3), 396–415.

3. Times Have Changed Using a Pictorial Smartphone App to Collect Time–Use Data in Rural Zambia

***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, https://doi.org/10.1177/1525822X18797303*

Abstract

One challenge of collecting socioeconomic data, such as data on time-use, is recall biases. While time-use researchers have continuously developed new methods to make data collection more accurate and easy, these methods are difficult to use in developing countries, where study participants may have low literacy levels and no clock-based concepts of time. To contribute to the closing of this research gap, we developed a picture-based smartphone app called Time-Tracker that allows data recording in real time to avoid recall biases. We pilot tested the app in rural Zambia, collecting 2,790 data days. In this article, we compare the data recorded with the app to data collected with 24-hours recall questions. The results confirm the literature on recall biases, suggesting that using the app leads to valid results. We conclude that smartphone apps using visual tools provide new opportunities for researchers collecting socioeconomic data in developing countries.

3.1. Introduction

One major challenge of collecting data through surveys is recall biases, and there is strong evidence that recall biases can be substantial with regard to time-use (Chatzitheochari et al. 2017; Juster and Stafford 1991; Juster et al. 2003; Kan and Pudney 2008). In developing countries, time–use research is particularly challenging because study participants may have high rates of illiteracy and populations may lack clock-based concepts of time (Harvey and Taylor 2000; Kes and Hema 2006). In light of these challenges, time–use research in developing countries has often been neglected, which makes it difficult for governments and development practitioners to prioritize and design development programs and policies and to measure their effects. In contrast, time–use researchers working in developed countries have long focused on developing methods to collect more accurate time–use data. Recognizing that recall periods beyond two days lead to unreliable data, time–use researchers have considered daily time–use diaries to be the gold standard (Chatzitheochari et al. 2017; Juster et al. 2003). However, such time–use diaries are burdensome to fill and the administrative and processing costs are high (Chatzitheochari et al. 2017). This has led to calls for better time–use methods, potentially using electronic devices (Chatzitheochari et al. 2017; Minnen et al. 2014; Paolisso and Hames 2010; Seymour et al. 2017).

Following these calls, several research groups have developed app-based time–use diaries (Chatzitheochari et al. 2017; Fernee and Sonck 2014; Runyan et al. 2013). While these efforts have shown the potential of using smartphone apps for time–use research, these text-based apps are not applicable for research in developing countries. One noteworthy exception that aims to address the challenges of time–use researchers in developing countries is Masuda et al. (2014) who test piloted a pictorial diary set in Ethiopia. This set contains a book with a grid, activity stickers, and a clock, which beeps every 30 minutes. When the clock beeps, participants place a sticker in the book that reflects their then current activity. While being accessible for people with low or no literacy and without clock-based concepts of time, this method still seems cumbersome, and it does not allow for capturing simultaneous activities, which may, given the of 30-minute interval, consequently provide inaccurate data.

To sum up, time–use researchers in developing countries lack suitable data collection methods and, therefore, reliable data. To address the lack of suitable methods, we developed a picture-based smartphone application called Time-Tracker that allows study participants to record time-use in real time to avoid recall biases. The app can be combined with pop-up questions, a feature that we used to ask questions on quantity and quality of food consumed.

We used the app in Eastern Province of Zambia to collect approximately 2,790 data days on the time-use of farm families throughout an entire farming season. This was done as part of a larger study assessing the impacts of agricultural mechanization on intrahousehold labor allocation. This study was done because agricultural mechanization has received renewed attention in many developing countries (Daum and Birner 2017). However, the effects of mechanization on labor are ambiguous. The mechanization of activities that tend to be done by men, such as land preparation, may lead to land expansion and thereby a higher labor burden for non-mechanized activities that tend to be done by women, such as weeding.

We compared the data collected with the app, which we used as a benchmark, with answers from 24-hour recall questions. This comparison allowed us to explore how and why recall biases differ for different activities and for different household members. In brief, this article aims to contribute to the development of much needed methods to collect more accurate time–use data in developing countries.

3.2. Methodological Considerations

Data collection methods time–use researchers can adopt, including their advantages and disadvantages, are summarized in Table 4. They are further discussed in subsequent sections with a specific focus on their suitability for developing countries. Table 4 also depicts the expected advantages and disadvantages of using smartphone apps for collecting time–use data, such as the one introduced in this article.

Criteria	Surveys (Seasonal)	Surveys (Weekly)	Diaries (Daily/Weekly)	Observations (Real Time)	Apps (Real Time)
Recall bias	High	Medium/high	Low	No	Low
Social desirability bias	Yes	Yes	Yes	Yes	Low
Costs	Low	Medium/high[a]	Low	High	Medium
Respondent burden	Low	Medium	High	Medium/high	Low/medium

[a] Depending on whether questions are asked by phone or face-to-face (Arthi et al. 2018).

Table 4. Advantages and disadvantages of methods to collect time–use data

3.2.1. Weekly and Seasonal Surveys

Most time–use studies rely on household surveys using recall questions such as: How much time did you spend last week/last farming season doing X? Using household surveys is inexpensive and allows for large sample sizes. However, the answers to survey questions "typically prove wide off the mark" (Juster and Stafford 1991: 482). Several aspects contribute to this, some of which are general problems and some of which are more pronounced in developing countries.

In general, study participants overestimate activities that are socially desirable and underestimate activities that are nondesirable and activities that they or the society do/does not perceive as work (Hofferth 1999; Juster and Stafford 1991; Juster et al. 2003). This is one reason why the length of activities may be reported differently by men and women. For example, Bianchi et al. (2012) found that men overestimate their contribution to household work by 70% in the United States. Study participants frequently overestimate secondary activities such as childcare (Juster et al. 2003). There is no clear agreement about whether sporadic activities are underestimated (Juster et al. 2003) or not (Menon 1993). The role of the intensity of different activities has not been studied much but may play role as well (Jodha 1988). There is a consensus that regular and externally structured activities, such as office work, have low biases (Juster et al. 2003).

Some of the challenges mentioned are more severe in developing countries, with regard to agriculture, for the following reasons. First, study participants may lack a clock-based concept of time. Second, compared to people from developed countries, people from developing countries tend to have less-structured days, which makes recalling time-use more difficult (Arthi et al. 2018). Third, the seasonality of farming may have effects on the perception of time spent on activities that are performed highly irregularly (Arthi et al. 2018). In view of these challenges, it is problematic that rural livelihood surveys frequently use postharvest recall questions that cover the entire farming season.

Arthi et al. (2018) found that Tanzanian farmers report a work time that is four times higher when asked via a postharvest instead of a weekly survey, which suggests that the long-standing debate on whether small or large farms are more efficient and thus whether agricultural development should be based on small or large farms (Collier and Dercon 2014; Larson et al. 2014) may be based on unreliable data. One reason for the overestimation of farm labor may be that postharvest recall questions force study participants to make “cognitively taxing calculations which result in labour inferences that appear to be based on recent rather than representative experiences” (Arthi et al. 2018:19). They also speculate that, in view of the harvest produced, labor can be overstated during good harvests and understated during bad harvests. Another problem is that agricultural surveys are often designed to be answered by the “household head” who may underestimate the work contribution of his or her kin.

Using weekly recall questions does substantially reduce the recall biases associated with household surveys (Arthi et al. 2018). However, they still do not meet the implicit standard of time–use researchers who argue that recall period beyond two days leads to unreliable data. Also, weekly data collection may be associated with high costs unless study participants are contacted by phone, which may lead to excluding study participants without phones.

Usually, comparing recall data is difficult because not only are different types of activities recalled as having lasted different times they are also recalled differently by different genders, social groups, and people from different countries with different familiarities with clock-based concepts of time. This makes intrahousehold comparisons of time-use very difficult. It also makes comparisons between different social groups challenging, for example, farmers and pastoralists, or Germans and Ghanaians.

3.2.2. Time–Use Diaries

Studies may also use time–use diaries in which study participants fill out 24-hour time grids that are divided into 15- or 30-minute slots either freely or using pre-coded activities. Time–use diaries are considered the most reliable and accurate data collection method, as they are less prone to recall problems as well as social desirability bias (Chatzitheochari et al. 2017; Juster et al. 2003; Paolisso and Hames 2010). It has been argued that time–use diaries are “the only valid measurement of time-use, and less expensive substitutes are substantially lower quality and have systematic biases” (Juster and Stafford 1991:482).

However, time–use diaries involve text-based questions and are burdensome to complete (Chatzitheochari et al. 2017). Therefore, diaries are not a viable option for developing countries unless they are filled with the help of interviewers, which may lead to biases. An exception is the above-described pictorial time diaries used by Masuda et al. (2014). While the use of pictures allows low-literacy and illiterate participants to use these diaries, they are still cumbersome to fill. In addition, they are based on 30-minute slots, which may affect study participants to underreport activities that are regularly performed throughout the day but are shorter than 30 minutes each time they are performed (Chatzitheochari et al. 2017; Kelly et al. 2015).

3.2.3. Direct Observations

Direct observations can eliminate recall biases and address the illiteracy problem (Kes and Hema 2006; Paolisso and Hames 2010); however, direct observations are expensive and thus reduce potential sample sizes (Harvey and Taylor 2000; Kes and Hema 2006). Assuming that a trained research assistant costs US$30/day (working maximum of eight hours/day), the cost of observing 2,790 days (the number of days recorded in this study) would have been US$167,000 (not including the costs for organization, cars, and accommodation). In addition, the observer's presence may affect the behavior of the observed, the so-called Hawthorne Effect (Kes and Hema 2006; Paolisso and Hames 2010).

3.3. Method

Incorporating strengths of some of the abovementioned time–use methods while overcoming some of their challenges, we developed an open-source smartphone app called Time-Tracker (the code can be accessed through https://github.com/HannesBuchwald/TimeTracker), which allows study participants to record data themselves. To ensure that study participants can easily record and capture the social context of the study area, the app was developed in close collaboration with farmers. To guarantee that populations with low or no literacy can participate, the app works only with pictures; data recording was designed to be as simple as possible to allow study participants with no experience with mobile phones or smartphones to effectively and easily use the app and to make sure that study participants do not develop "entry fatigue," losing the motivation to carefully record data. Figure 9 shows the main screen of the app, which displays pictures of 88 typical daily activities that we selected and designed together with the local population.

Figure 9. The data entry (left) and data control (right) screen
The screen on the right allows study participants to see the recorded data. The researchers can activate a hidden button and correct potential mistakes.

To start recording an activity, study participants click on the respective picture (e.g., hand hoeing): up to three activities can be recorded at the same time to capture simultaneity (e.g., hand hoeing and caring for a baby). To stop recording an activity, the participant clicks on the respective picture again. Activities are thus recorded in real time, which reduces recall biases. Activities participants are performing at that moment are displayed on top of the screen.

The study participants can also indicate whether the work was done on their own field or on the field of others as agricultural laborers (with a triple click on the respective drawing). However, some study participants had difficulties with this mechanism. The research team, therefore, frequently cross-checked with the participants whether work was done on the field of others.

In addition to time–use data, study participants can record food intake. To make this possible, we designed a plug-in that opens when the activity “eating and drinking” was terminated. Study participants are then shown four differently filled plates to record the quantity of food consumed. Afterward, they are shown different food groups (e.g., cereals, roots, and vegetables), which allows them to indicate the diversity of the food they ate.

Before study participants used the app, we introduced them to it. First, we practiced the use of touch screens; then, we clarified questions about the pictures. Moreover, the study participants practiced using the app with the help of explanatory stories. For example, using their local language, they were told to record the following story using the app: “Christian goes to the field to hoe while listening to the radio; one the way back, he uses the bicycle of a friend. After reaching home, he eats vegetables and nshima” (a maize dish in Zambia).

The participants were given a smartphone with the app for the data-recording period. We lent them a smartphone to avoid selection biases, which would have occurred if we had only selected study participants who already owned a smartphone. We used 50 smartphones costing US$90 each, which allowed us to work with 45–48 respondents at a time. In total, the costs of developing and administering the app were around US$40,000 (including the smartphone costs). Thus, each collected datum cost US$13.4.

The smartphones were configured, so that only the time–use app could be used. Blocking other smartphone functions was done to ensure that smartphone use did not alter daily routines of the participants. The blocking had two positive side effects. First, it may have reduced any temptation to “lose” the smartphone. Second, the blocking enabled only the app to run, which extended the battery life up to five days. When the battery level was below 50%, we distributed power banks.

To ensure that the smartphones worked properly, the research team made daily random checks. This approach also allowed the team to double-check whether study participants recorded the activities that they actually performed and to help participants manually enter activities that they may have forgotten. These corrections were possible through a second screen of the smartphone app, which is shown in Figure 9. Data recording and submission were done off-line, and after the recording phase, the research team uploaded the data from the smartphone to a laptop using a local Wi-Fi network. The participants received small gifts such as caps and fabrics as appreciation for their participation in the study.

3.4. Study Site and Sampling

The study was conducted as part of a larger research project that aimed at assessing the effects of agricultural mechanization in Eastern Province of Zambia. For this, we had used a two-stage sampling procedure to select 62 farm families with different mechanization levels based on the population of the nationally representative Rural Agricultural Livelihood Survey. The households were located across four different communities, all of which have been dominated by smallholder farmers. On average, farmers cultivate 2.3 hectares—mainly maize but also cotton, sunflower, groundnuts, and tobacco. Farming is characterized by a short rainy season and an extensive dry season. Most of the farming activities are done manually (1% of the households use, own, or hired mechanical power for land preparation and 57% use animal traction). Land and labor productivity is low, and 90% of the rural population live on less than US$1.25/day (all above from Indaba Agricultural Policy Research Institute 2016). Table 5 shows selected sample characteristics. In each of the selected households, the head of the household, one spouse, and one child (alternating between boys and girls) were trained to use the smartphone app. After the training, data were collected at five different stages during the 2016–2017 season: land preparation, planting, weeding, harvesting, and processing. At each point, households recorded their time-use for three days.

Variable	Mean
Sample size	62
Household size (members)	7.1
Age male (years)	47.4
Age female (years)	39
Age app using child (years)	15.6
Education level	8.9
Land size cultivated (hectares)	5.2
Farm income	US$1,532

Table 5. Sample Characteristics
Education levels range from 1=first grade to 20=master degree

This study focuses on the last two stages, harvesting and processing. During these stages, study participants were also asked 24-hour recall questions after the last day of using the smartphone app. Twenty-four-hour recall questions made it possible to compare data recorded with the app to data elicited from recall questions. The recall questions asked were about major time–use categories such as farming or household chores (giving clear examples of different time–use categories). To reduce potential “adding up” problems, the day to be recalled was split into five time–use intervals.

3.5. Results

All sampled household members were able to use the app, which indicates that the use of a well-designed and explained smartphone app for data collection does not lead to selection bias. Figure 10 shows the average age of study participants. The wide range of study participants (from seven to 92 years) suggests that all age-groups can use the Time-Tracker. Also, participants with low or no literacy were able to use the app as well as study participants who had no prior exposure to smartphones.

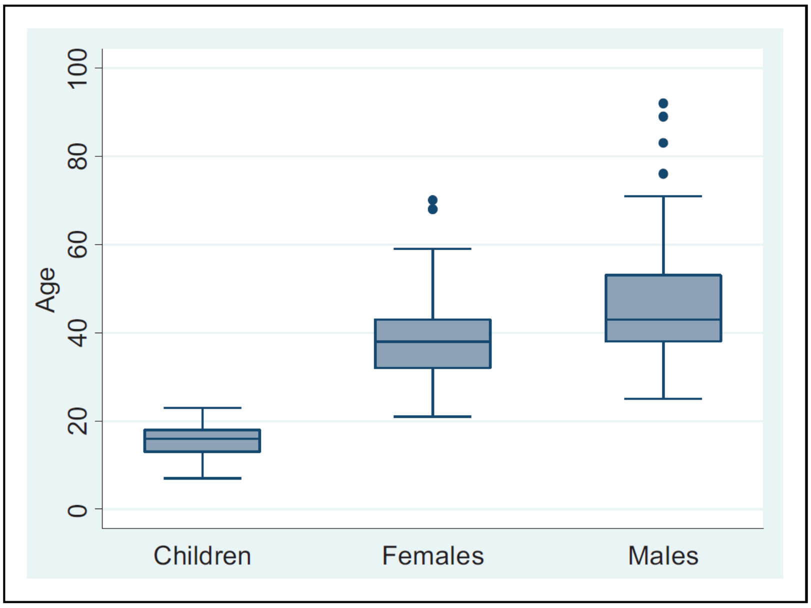


Figure 10. Age distribution of participants

Our experiences suggest that study participants enjoyed working with the app, handled the smartphones with care, and took data recording very seriously. The data collected with the smartphone app appear to be of high quality and this can be seen from Table 6, which indicates that only a small percentage of the data had to be entered or corrected by the research team who supervised data collection.

Data Collection Rounds	Data Entered/Corrected by Research Team because Participants Forgot Entering or Clicked Wrong Activity (%)
Land preparation	0.6
Planting	0.8
Weeding	0.7
Harvesting	2.9
Processing	2.3

Table 6. Percentage of data entered/changed by enumerators

Table 7 shows an example of a typical data day, showing both the time-use and food intake. During the entire data collection phase, only one smartphone disappeared and one accidently cracked, but still worked.

						Food				
Activity	ID	Date	Start	End	Piecework	Quantity	Cereals	Vegetables	...	Diversity Score
Sleeping	1	May 06, 2017	00:00:00	05:34:28	No					
Personal Hygiene		(Saturday)	05:34:51	05:46:48	No					
Walking (unloaded)			05:47:07	06:02:07	No					
Harvesting (manual)			06:02:27	12:01:59	No					
Walking (unloaded)			12:02:08	12:28:22	No					
...			...	...	...					
Personal hygiene			19:34:30	19:47:53	No					
Eating + drinking			19:47:59	20:08:57	No	1	1	1	1	3
Chatting			20:09:46	20:31:38	No					
Sleeping			20:31:50	00:00:00	No					

Table 7. Example of a data day

Table 8 compares durations of different time–use categories recorded by participants using the app and as they are recalled the next day. The differences indicate the magnitude of recall biases of 24-hour recall questions. Table 8 shows that some activities are recalled as lasting significantly longer (overestimated) compared to the time recorded with the smartphone app. For example, farming activities are recalled as having lasted between 29% and 47% longer. Of activities perceived as lasting significantly shorter (underestimated), social life activities, which were recalled as lasting 45–54% less than the app-recorded time, had the biggest difference.

Minutes per activity (Daily)	Round I			Round II		
	Real Time (App)	24-hour Recall	Difference (%)	Real Time (App)	24-hour Recall	Difference (%)
Farming	200.3 (15.0)	258.0 (16.5)	29***	149.9 (14.9)	220.9 (17.9)	47***
Expanded farming	26.5 (6.8)	30.5 (9.4)	15	47.3 (9.3)	21.0 (6.3)	−56**
Off-farm work	11.2 (6.6)	18.2 (7.9)	63	11.5 (5.8)	12.4 (6.0)	8
Community work	2.2 (2.0)	3.0 (3.0)	36	4.2 (2.7)	3.7 (3.0)	−12
Meeting	7.5 (5.1)	19.8 (5.6)	164	4.2 (2.2)	8.3 (2.6)	98
Mobility	146.8 (12.6)	101.4 (9.9)	−31***	121.6 (11.4)	89.8 (7.2)	−26**
Education	25.9 (8.8)	28.4 (7.9)	10	36.0 (10.1)	34.5 (9.5)	−4
Care for others	24.2 (6.1)	118.5 (14.7)	390***	26.1 (7.8)	103.7 (12.7)	297***
Household chores	125.8 (12.4)	95.8 (9.8)	−24*	124.5 (11.3)	107.2 (10.5)	−14
Construction	5.5 (3.8)	9.1 (3.3)	65	6.1 (3.9)	16.9 (5.7)	177
Sleeping	710.3 (11.2)	608.0 (11.5)	−14***	718.5 (11.7)	603.7 (13.7)	−16***
Personal care	77.7 (4.3)	62.3 (3.4)	−20***	75.1 (4.4)	65.9 (5.2)	−12
Social life	172.2 (15.7)	77.9 (9.2)	−55***	193.3 (16.2)	105.2 (9.2)	−46***

Table 8. Comparison of time-use recorded with app and 24-hour recall questions
In two data collection rounds (harvesting and processing). The values of standard deviations are given in brackets. Asterisks indicate significantly differing mean values: ***p < .01, **p < .05, *p < .1.

Table 9 combines rounds I and II but splits the study participants into different categories (males, females, and children). It shows that activities such as farming, household chores, and social activities are gendered. Table 9 suggests that different respondent groups recall different activities with differing accuracies. For example, construction seems to be well recalled by females and children but recalled as lasting much longer than app-recorded time by males. Males did recall their contribution to household chores rather accurately, whereas, surprisingly, women did not.

	Round I and Round II								
	Males			Females			Children		
Minutes per Activity (Daily)	Real Time (App)	24-hour Recall	Difference (%)	Real Time (App)	24-hour Recall	Difference (%)	Real Time (App)	24-hour Recall	Difference (%)
Farming	184.2 (19.5)	247.7 (21.4)	34**	213.1 (17.6)	270 (20.4)	27**	117.4 (17.1)	191.8 (20.8)	63***
Expanded farming	77.7 (13.5)	65.6 (14.9)	−16	12.6 (5.4)	2.7 (1.4)	−79*	20.8 (8.7)	9.3 (7.1)	−55
Off-farm work	21.9 (10.0)	34.3 (13.0)	57	4.9 (4.8)	10.3 (6.0)	110	7.4 (7.4)	0 (0)	−100
Community work	6.4 (4.0)	5.5 (4.4)	−14	2.7 (2.7)	4.2 (4.0)	56	0.1 (0.1)	0 (0)	−100
Meeting	14.9 (8.0)	22.0 (7.5)	48	1.3 (1.0)	9.2 (3.5)	608**	1.2 (1.1)	11.1 (4.3)	825**
Mobility	197.3 (18.3)	112.5 (10.4)	−43***	89.7 (10.5)	97.3 (11.7)	8	118.2 (12.1)	74.1 (8.5)	−37***
Education	0 (0)	0 (0)	0	2.8 (2.3)	3.6 (3.3)	29	101.6 (20.6)	102.2 (18.5)	1
Care for others	6.2 (3.3)	42.6 (8.5)	−587***	51.1 (11.3)	180.4 (18.2)	253***	14.2 (7.0)	102.2 (18.7)	620***
Household chores	9.9 (2.9)	12.1 (2.7)	22	233.3 (13.0)	172.2 (12.5)	−26***	120.8 (13.9)	114.4 (12.6)	−5
Construction	10.2 (6.6)	36.4 (9.3)	257***	1.9 (1.9)	1.1 (0.6)	−42	5.6 (4.8)	1.3 (0.9)	−77
Sleeping	732 (13.5)	657.8 (11.7)	−10***	702.1 (12.4)	604 (13.4)	−14***	709.8 (16.6)	549.7 (19.7)	−23***
Personal care	69.5 (3.9)	76 (7.5)	9	77.8 (3.8)	60.2 (3.5)	−23***	82.4 (8.0)	55.3 (4.1)	−33***
Social life	205.1 (22.6)	109.7 (12.0)	−47***	140.9 (14.1)	73.3 (9.1)	−48***	209.6 (21.5)	93.8 (13.1)	−55***
Sample size	99	99		109	109		87	87	

Table 9. Comparison of time-use recorded with app and 24-hour recall questions by different respondents
In two data collection rounds (harvesting and processing). The values of standard deviations are given in brackets. Asterisks indicate significantly differing mean values: ***p < .01, **p < .05, *p < .1.

We also analyzed the time–use difference, treating the difference between recall estimates and app-recorded data as a measurement error in recall data. These errors are normally distributed for all time–use categories. We tested whether the measurement error in recall data is significantly correlated with age and education but found no evidence (see Figure 11).

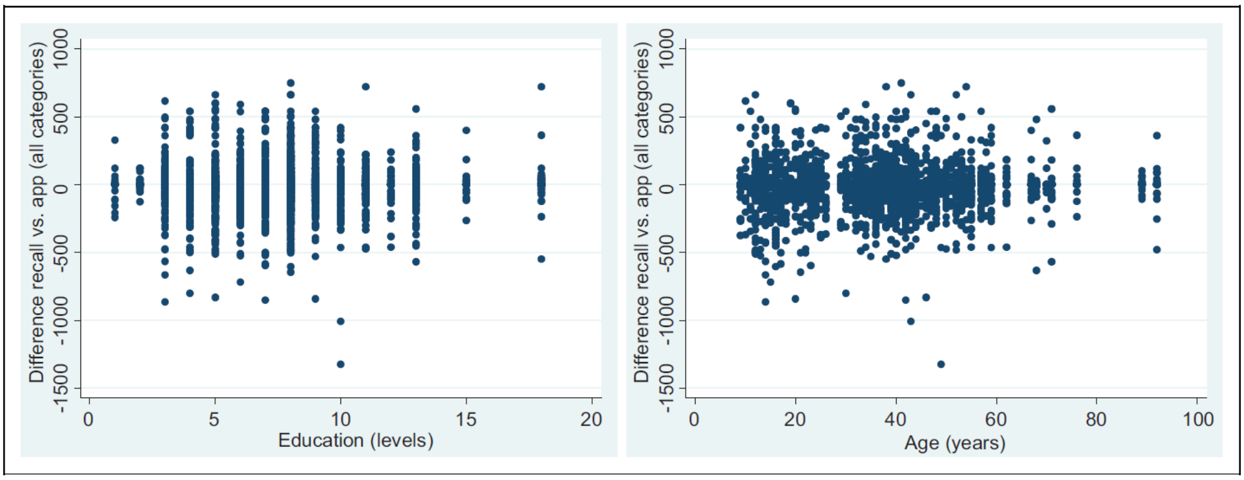


Figure 11. Recall error by educational level and age

In appendix 1 and the online supplementary of this article, we indicate the accuracy of recall data both graphically and by reporting the mean squared standard deviation (MSD) around the 1:1 line for different demographic characteristics (male, female, and children). As suggested by Gauch et al. (2003), we separated the MSD into three categories to obtain a deeper understanding into recall accuracy: squared bias, showing translation; nonunity slope, showing rotation; and lack of correlation, showing scatter. We also graphically indicated whether the magnitude of errors vary systematically with the size of recall estimates. The results suggest that the longer an activity's duration was recalled, the more likely participants overestimated the activity's duration, which was the case for all time–use categories (not only the ones shown here). This observation seems more pronounced with regard to farming and household chores (a and b) than social life (c).

3.6. Discussion

3.6.1. Comparative Advantage over Existing Methods

The Time-Tracker combines advantages of existing time–use data collection methods and overcomes their respective disadvantages. The app allows study participants to record time-use in real time, which reduces the recall biases associated with household surveys. Using pictures, the app overcomes the text hurdle of existing paper and electronic time–use diaries. Also, the smartphone app has overcome some drawbacks of the pictorial diary sets of Masuda et al. (2014) described above: Instead of recording activities based on a 30-minute time slot, which inevitably introduces inaccuracies, especially for tasks that last less than 30 minutes, data recording with the smartphone app is straightforward and can be done in real time. We found that the study participants recorded data with great care and accuracy. We also found that the use of a smartphone app did not lead to selection bias, as all members of all sampled households were able to use the app. This suggests that well-designed smartphone apps may be used by time–use researchers to collect data in developing countries.

3.6.2. The Role of Recall Biases

To validate the data collected with the Time-Tracker, we compared the data recorded using the app with data collected through widely used 24-hour recall questions. We find that regular and uniformly structured activities such as going to school are recalled as lasting the same time as recorded with the smartphone app. This also validates that the data recorded with the app are accurate. Confirming existing literature, we find that the length of socially desirable (and arduous) activities, such as farming, is overstated. This confirms Arthi et al. (2018), who found that reported farm labor decreased by a factor of four when the recall period was reduced from postharvest data collection to weekly data collection. Our results (moving from 24-hour recall questions to real-time data) suggest that actual working time may even be lower. This implies that the labor productivity of farm households may be much higher than commonly estimated based on recall studies.

We found that, similar to farming time, the duration of care activities was reported to be longer than the duration from recall questions. This could partly be explained by findings in the literature, which suggest that durations of secondary activities, such as care taking, are often overestimated (Juster et al. 2003). However, study participants may not have recorded all care activities, especially when they lasted only short periods of time. Confirming Hofferth (1999) and Juster et al. (2003), we find that recalled durations of social life activities are significantly shorter than durations recorded with the app. The reason to understate the duration of such activities may be that study participants perceive these activities to be less desirable by society. Interestingly, male household members were not found to overestimate their contribution to household chores, a contrast with the findings of Bianchi et al. (2012) from the United States. Potentially, the difference is due to the fact that household chores are not seen as desirable for males in Zambia.

3.6.3. Limitations and Directions for Future Research

It is important to note that the time–use difference reported may be underestimated because the study participants, who had used the Time-Tracker the previous days, were already sensitized about time-use during recall questions. To avoid the inherent difficulty of comparing two data collection methods at the same time, one would need to conduct a randomized control trial as done by Arthi et al. (2018).

By using the Time-Tracker at various times of the farming season, we were able to capture seasonality. However, data recording only at specific (even if well-chosen) points of time does not make it possible to extrapolate the data over the entire farming season. For example, a participant may weed her fields 120 minutes/day for 15 days (a total of 1,800 minutes), whereas another may weed her fields 90 minutes/day for 30 days (a total of 2,700 minutes). Looking at specific data points (such as one day), one may wrongly conclude that the second participant works less. Data collection using a postharvest questionnaire would capture the difference in days weeded, but since it is based on recall questions, this information may be inaccurate. This implies a need for further research to find ways to collect data over extended periods of time while ensuring that study participants do not develop a fatigue. One solution may be gamification of the app, making app usage more attractive by using game-design elements, or allowing study participants to collect airtime credits. Furthermore, research can aim to find ways to better extrapolate data for the entire farming season.

There are additional ideas on how to further improve such smartphone applications. So far, similar to most time–use research, the Time-Tracker does not capture the intensity of efforts. This shortcoming could be addressed by combining the app with fitness trackers. Also, the Time-Tracker does not indicate whether activities are perceived as enjoyable; absence of this function can be improved by asking participants whether they enjoyed the activity with the help of pop-up windows (similar to Fernee and Sonck 2014). Thinking some steps ahead, data collected with smartphone apps may be validated using cameras and built-in sensors.

3.6.4. Recommendations

While we recommend using smartphone apps for time–use researchers in developing countries, several aspects need to be kept in mind. First, the app design as well as the selection and drawing of the illustrations need to be done in close collaboration with study participants. Second, when introducing the app, it is important to consider the role of village authorities, social dynamics, and beliefs. For example, it is key to explain how participating households are selected to avoid social tensions. Third, it is important to have sufficient training on how to handle smartphones and the app.

3.7. Conclusion

While time–use researchers have been continuously developing more accurate and user-friendly methods to collect time–use data, these efforts have been largely restricted to the developed world. This skewness has resulted in a lack of methods that could be adopted to collect time–use data in developing countries where study participants may have low literacy levels and no clock-based concept of time. In this article, we presented a picture-based smartphone app that allows researchers to collect time–use data with high precision in developing countries. The results suggest that well-tailored smartphone apps that use visual tools provide new and much needed pathways for time–use researchers working in developing countries.

3.8. References

Arthi, V., Beegle, K., de Weerdt, J., Palacios-López, A. 2018. Not your average job: Measuring farm labor in Tanzania. Journal of Development Economics 130:160–72.

Bianchi, S. M., Sayer, L. C., Milkie, M. A., Robinson, J. P. 2012. Housework: Who did, does or will do it, and how much does it matter? Social Forces 91:55–63.

Chatzitheochari, S., Fisher, K., Gilbert, E., Calderwood, L., Huskinson, T., Cleary, A., Gershuny, J. 2017. Using new technologies for time diary data collection: Instrument design and data quality findings from a mixed-mode pilot survey. Social Indicators Research 137:379–90.

Collier, P., Dercon, S. 2014. African agriculture in 50 years: Smallholders in a rapidly changing world? World Development 63:92–101.

Daum, T., Birner, R. 2017. The neglected governance challenges of agricultural mechanisation in Africa—Insights from Ghana. Food Security 9:959–79.

Fernee, H., Sonck, N. 2014. Measuring smarter: Time-use data collected by smartphones. Electronic International Journal of Time-Use Research 11:94–96.

Gauch, H. G., Hwang, J. T., Fick, G. W.. 2003. Model evaluation by comparison of model-based predictions and measured values. Agronomy Journal 95:1442–46.

Harvey, A. S., Taylor, T. 2000. Designing household survey questionnaires for developing countries: Lessons from 15 years of the Living Standards Measurement Survey. Washington, DC: The World Bank.

Hofferth, S. 1999. Family reading to young children: Social desirability and cultural biases in reporting. National Research Council Workshop on measurement and research on time-use. Washington, DC: National Research Council.

Indaba Agricultural Policy Research Institute. 2016. Rural Agricultural Livelihoods Survey 2015 Survey Report. https://images.agri-profocus.nl/upload/post/RALS_2015_Survey_Report__Finalv-_edited1456827249.pdf (accessed January 8, 2018).

Jodha, N. S. 1988. Poverty debate in India: A minority view. Economic and Political Weekly 23:2421–28.

Juster, F. T., Ono, H., Stafford, F. P. 2003. An assessment of alternative measures of time-use. Sociological Methodology 33:19–54.

Juster, F. T., Stafford, F. P. 1991. The allocation of time: Empirical findings, behavioural models, and problems of measurement. Journal of Economic Literature 29:471–522.

Kan, M. Y., Pudney, S. 2008. Measurement error in stylized and diary data on time-use. Sociological Methodology 38:101–32.

Kelly, P., Thomas, E., Doherty, A., Harms, T., Burke, Ó. Gershuny, J., Foster, C. 2015. Developing a method to test the validity of 24 hour time-use diaries using wearable cameras: A feasibility pilot. PLoS One 10:e0142198.

Kes, A., Hema, S. 2006. In Gender and time poverty in sub-Saharan Africa. World Bank Working Paper No. 73, Washington, DC, pp.13–38.

Larson, D. F., Otsuka, K., Matsumoto, T., Kilic, T. 2014. Should African rural development strategies depend on smallholder farms? An exploration of the inverse-productivity hypothesis. Agricultural Economics 45:355–67.

Masuda, Y. J., Fortmann, L., Gugerty, M. K., Smith-Nilson, M., Cook, J. 2014. Pictorial approaches for measuring time-use in rural Ethiopia. Social Indicators Research 115:467–82.

Menon, G. 1993. The effects of accessibility of information in memory on judgments of behavioural frequencies. Journal of Consumer Research 20:431–40.

Minnen, J., Glorieux, I., van Tienoven, T., Daniels, S., Weenas, D., Deyaert, J., Van den Bogaert, S., Rymenants, S.. 2014. Modular online time-use survey (MOTUS)—Translating an existing method in the 21st century. Electronic International Journal of Time-Use Research 11:73–93.

Paolisso, M., Hames, R. 2010. Time diary versus instantaneous sampling: A comparison of two behavioral research methods. Field Methods 22:357–77.

Runyan, J. D., Steenbergh, T. A., Bainbridge, C., Daugherty, D. A., Oke, L., Fry, B. N. 2013. A smartphone ecological momentary assessment/intervention “app” for collecting real-time data and promoting self-awareness. PLoS One 8:e71325.

Seymour, G., Malapit, H. J., Quisumbing, A. 2017. Measuring time-use in development settings. Policy Research Working Paper No. 8147, World Bank, Washington, DC.

4. Of trackers and tractors. Using a smartphone app and compositional data analysis to explore the link between mechanization and intra-household allocation of time in Zambia.

***Daum, T.**, Capezzone F., Birner, R. 2019. Of trackers and tractors. Using a smartphone app and compositional data analysis to explore the link between mechanization and intra-household allocation of time in Zambia. Under review with Agricultural Economics.*

Abstract

Digital tools may help to study socioeconomic aspects of agricultural development that are difficult to measure such as the effects of new technologies, policies and practices on the intra-household allocation of time. As new technologies, policies and practices may target different crops and tasks, they can affect time-use of men, women, boys and girls differently. Development strategies that overlook such effects can fail or have negative consequences for women and children. In this paper, the effects of agricultural mechanization on time-use in smallholder farming households in Zambia were investigated. For this, a novel data collection method was used: a pictorial smartphone application that allows real-time recording of time-use to eliminate recall bias. Existing studies analysing intra-household allocation of resources often focus on adult males and females. This study paid particular attention to boys and girls. The study also addressed seasonal variations. For data analysis, compositional data analysis was used, which yields higher accuracy than univariate analysis by accounting for the co-dependence and sum constraint of time-use data. The study found that women benefit relatively more from mechanization during land preparation, which leads to gender differentiation; for households using manual labor, such differentiation was not found. There was some evidence that the time "saved" is used for off-farm and domestic work. No negative second-round effects during weeding and harvesting/processing and no negative effects on children were found.

4.1. Introduction

During the last years, various researchers have used digital tools to enhance data accuracy in the field of applied agricultural economics. For example, researchers have used GPS devices to measure plot sizes (Carletto et al., 2015a), fitness-trackers to capture energy expenditure (Zanello et al., 2017) and satellites to assess yields (Lobell et al., 2018). However, for socioeconomic data, researchers still largely rely on household surveys, which are prone to recall bias (Arthi et al., 2018; Bell et al., 2019; Fraval et al., 2018). The lack of reliable data collection methods for socioeconomic aspects that are difficult to recall has led to data suffering from poor quality and the neglect of potentially highly relevant research areas (Carletto et al., 2015b). One such research area is the effects of technology adoption (such as tractors or herbicides) or the exposure to new policies on the time-use within smallholder farming households in developing countries.

The need to monitor intra-household time-use effects when promoting technologies and designing policies is widely acknowledged (Blackden and Wodon, 2006; Bryceson, 2019; Doss, 2013; Theis et al., 2018). Since smallholder farming is often associated with a gender-division of labor, which can be based on crops, tasks or both, new technologies and policies can affect adult men and women as well as boys and girls differently (Blackden and Wodon, 2006; Doss et al., 2001; Quisumbing et al., 1995). Development strategies that overlook these dynamics can fail or have negative consequences for vulnerable household members. For example, promoting conservation agriculture may lead to more labor for women because of the increased weeding required (Farnworth et al., 2015); and the exacerbated need for bird scaring associated with the new rice variety NERICA has been shown to prevent children from going to school (Bergman-Lodin et al., 2012). There are concerns that such time-use changes have negative consequences on nutrition and child care (Johnston et al., 2018). In this paper, the time-use effects of agricultural mechanization are explored. Mechanization is unfolding rapidly in various Asian countries (Takeshima, 2017; Wang et al., 2016) and has received growing attention in Africa (Daum and Birner, 2017; Benin, 2015; Diao et al. 2014).[12] Notwithstanding some anecdotal evidence, the effects of mechanization on intra-household time allocation have not been examined.

While the need to carefully monitor time-use effects of new technologies and policies is widely acknowledged, studying such effects empirically has been hampered by a lack of suitable data collection methods. Post-harvest questionnaires and 24-hour recall questions are prone to recall bias; time-use diaries require literacy and a familiarity with clock-based concepts of time; direct observations are expensive and associated with observer bias (Arthi et al., 2018; Daum et al., 2019). In this study, therefore, a smartphone application called Timetracker is used, which is based on visual tools. It allows real-time recording of time-use to reduce recall bias. The Timetracker was used to collect 2790 days of time-use data in Zambia during different seasons to capture seasonality. The Timetracker has the advantage of allowing data recording by children. This is a unique contribution, as existing studies analyzing intra-household allocation of resources - and time-use, which are far fewer in number and often qualitative in nature - focus mainly on adults (Doss, 2013). This is despite that 60% of all child labor is in agriculture, affecting around 100 million girls and boys (ILO, 2019).

Researchers studying gender differences in time-use have often studied time spent on different activities in isolation of each other (Arora, 2015). This can be misleading since total time-use always sums up to 24 hours. Also, time-use is intrinsically collinear and codependent: an increase

[12] Agricultural mechanization is an umbrella term and its technologies can be targeted towards different crops and farm operations.

in time spent on one activity reduces the time available for other activities (Chastin et al. 2015; Gupta et al., 2018). Standard statistical techniques fail to account for this and result in spurious correlations (Pearson, 1897). To address these challenges, compositional data analysis is used in this paper (Atchinson, 1986; Bacon-Shone, 2011). Compositional data analysis has been used in different disciplines such as soil science, biology, geochemistry and medicine (Bacon-Shone, 2011) but has not been applied within the agricultural economics field.

In addition to finding a reliable data collection method and dealing with the structure of time-use data, the third challenge when studying the effects of mechanization on time-use is establishing causality. Ideally, a randomized control trial would take place. In the context of agricultural mechanization, which is adopted ad-hoc, a randomized control approach is difficult to implement. An alternative would be to use propensity score matching (PSM). In this study, the data collection method was novel and aiming for a sample size large enough for PSM was considered risky. Hence, cross-sectional data was used to compare time-use across differently mechanized households. The study uses multiple linear regression models to account for covariates and builds on economic theory but the use of cross-sectional data remains a limitation. Thus, the study concentrates on the first two challenges related to time-use data, data collection and analysis, and should be understood as a proof-of-concept case study. The paper has three major objectives: 1) providing a proof-of-concept that using digital tools can help to collect more reliable socioeconomic data; 2) introducing compositional data analysis to agricultural economics; 3) exploring how time-use differs by levels of mechanization, paying particular attention to gender, child labor and seasonality.

In section 4.2, the potential effects of agricultural mechanization on the intra-household allocation of time-use are discussed and four research hypotheses are derived. In section 4.3, study site and sampling procedure are described and the Timetracker is presented. In addition, the section discusses how the challenges of time-use data can be addressed using compositional data analysis. In section 4.4, the hypotheses are answered. Section 4.5 discusses and concludes.

4.2. Background and research hypotheses

In many African smallholder farming households, men and women have different workloads and duties (Arora, 2015; Blackden and Wodon, 2006; Quisumbing et al., 1995). For example, ploughing tends to be done relatively more by men and weeding and processing by women (Alesina, 2011; Baanante et al., 1999); however, such gender roles can clearly vary across space and time (Lambrecht et al., 2017) and have also been questioned (Palacios-Lopez et al., 2017). Little is known about the different roles of children in farm households. As new technologies target different crops and tasks, they may affect men and women, boy and girls differently. As there is evidence that households favor technologies that can be directed to male crops and activities (Evers, 2001), there are various examples where the introduction of a new technology increased women's burden (Agarwal, 1985; Bergman-Lodin et al., 2012; Kumar, 1994; Doss, 2001; Quisumbing et al., 1995). Bergman-Lodin et al. (2012) also found negative effects on children; however, most studies do not focus on children. An increase in women's time-use may likely come from where women do not have the bargaining power to reject more labor-intense technologies or to demand a re-allocation of activities (Fisher et al., 2000).

The time-use effects of mechanization depend on which tasks or crops are mechanized, how accompanying inputs such as herbicides and hired labor are used, what the original labor allocation was, and how this allocation can be re-negotiated. For understanding how time-use allocation may be re-negotiated, both unitary and bargaining models have been proposed (Alderman et al., 1995; Doss, 2013) but this is not the focus of this paper. Typically land preparation is mechanized first because land preparation tends to be a labor bottleneck. However, it may also reflect preferences to adopt technologies which can be directed to male-focused activities (Evers and Walters, 2001). With land preparation being mechanized, household may cultivate additional land, which may increase the need for weeding, harvesting/processing or the time spend collecting firewood once forests are cleared, which are tasks often performed by women and children (Arora, 2015; Blackden and Wodon, 2006; Doss, 2001). This was observed in India on anecdotal basis by Mukhopadhyay (1984) who found that the mechanization of ploughing (which was a male activity) led to more land cultivated and a higher workload for women since they were then "dealing with bigger crops over a larger acreage without mechanization of any of the operations they control" (p.58). A similar phenomenon may be observable in Zambia since there is evidence on mechanization increasing the area under cultivation (Adu-Baffour et al., 2018).

Based on the theoretical framework sketched above, four research hypotheses can be derived, which will be tested in section 4.

H1: Land preparation is predominantly a male activity
H2: Mechanized land preparation benefits males relatively more than females
H3: The time "saved" by mechanized land preparation is used differently by gender.
H4: On mechanized farms, females spend more time on weeding and harvesting/processing compared to females on non-mechanized farms

4.3. Study site, data collection method and sampling

4.3.1. Study Site

The study was conducted in the Eastern Province, which is one of Zambia's most important smallholder agricultural regions. The average size of land cultivated is 2.3 hectares - mainly maize, cotton, sunflower, groundnuts, and tobacco are grown (IAPRI, 2016). Farming is rain-fed and constrained by an extensive dry season. The emergence of medium-scale farmers as observed by Jayne et al. (2016) has led to more farmers owning tractors and providing services to neighboring farmers but the access to mechanizations remains low: 1% of the households use own or hired tractors for land preparation and 57% of the farmers use own or hired animal traction on a least one plot (IAPRI, 2016).

4.3.2. Data collection methods and sampling

As outlined above, time-use is difficult to measure, especially in developing countries. To address this challenge, a smartphone application called Timetracker was used, which is based on visual tools and allows real-time recording of 88 time-use categories (Figure 12; Daum et al., 2018 and Daum et al., 2019). The app allows to record up to three activities at a time but the focus here is on primary activities.

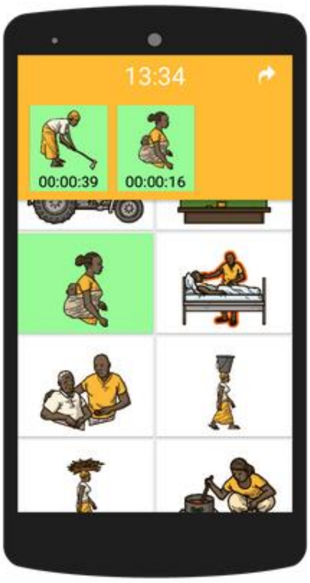


Figure 12. The Timetracker

The Timetracker was used to collect data from 62 households: 20 used manual labor, 20 used animal power and 22 used mechanical traction for land preparation, which are henceforward abbreviated with “manual”, “animal” and “tractor” households. Based on the nationally representative Rural Agricultural Livelihood Survey, households were selected using a two-stage random sampling procedure. First, four communities were sampled based on the criteria that more than five households used manual labor, more than five households used animals, and more than five used tractors for land preparation. Second, five manual-, five animal-, and five to six tractor-households were randomly selected, who in each household had at least one adult male, one adult female, and one child. If not enough households could be identified based on these criteria, missing households were randomly added from lists of the District Agriculture and Cooperatives Offices. In each household, household head, spouse and the oldest child used the Timetracker for three days at five different points of the 2016/2017 cropping season. This resulted in 2790 data days. Since the smartphone app was used in rotation in four different communities, data was collected on 60 different days. This paper focuses on land preparation, weeding and harvesting/processing season. At the end of the season, a household survey was conducted. Table 1 provides descriptive statistics about the selected households.

Variable	Manual-HHs	Animal-HHs	Tractor-HHs	Difference
Household characteristics				
Household size	6.6 (0.3)	7.8 (0.5)	6.7 (0.4)	
Gender head male (%)	95% (0.1)	100% (0)	95% (0.4)	
Age (years)	49.7 (3.8)	45.1 (2.5)	47.3 (2.9)	
Education level head (0-18)	6.8 (0.7)	8.5 (0.8)	10.5 (0.9)	***
Agronomic Characteristics				
Land cultivated (ha)	2.3 (0.2)	4.8 (0.9)	8.4 (1.3)	***
Land owned (ha)	2.5 (0.4)	5.9 (1.5)	19.8 (6.6)	***
Crop diversity	3.1 (0.2)	3.7 (0.2)	3.5 (0.2)	
Frequency of animal draught weeding	0.32 (0.1)	0.69 (0.12)	0.51 (0.1)	**
Maize yield (tons/ha)	1.9 (0.4)	2.6 (0.4)	3.6 (0.4)	***
Fertilizer per ha cultivated (kg)	110.5 (30.4)	190.3 (33.2)	216 (43.6)	
Pesticide per ha cultivated (l)	1.5 (1.0)	8.8 (3.3)	5.4 (2.5)	
Tropical livestock unit[1]	0.8 (0.2)	7.4 (1.8)	6.4 (1.7)	***
Hired labor (hours per cultivated ha)				
Land preparation	4.1 (2.8)	7.0 (5.6)	3.9 (2.1)	
Weeding	5.3 (5.3)	14.2 (10.9)	20.7 (10.0)	
Harvesting	8.8 (8.8)	7.6 (5.6)	16.7 (7.4)	
Child labor (hours per cultivated ha)				
Land preparation	38.3 (18.9)	29.1 (7.9)	11.8 (6.9)	
Weeding	45.4 (15.2)	59.9 (23,9)	16.6 (7.6)	
Harvesting	43.9 (15.8)	42.7 (12.4)	12.2 (4.6)	*
Socioeconomic Characteristics				
Log income	7.8 (0.36)	9 (0.28)	10.3 (0.21)	***
Share off-farm income	35% (13.2)	17% (7.1)	33% (6.6)	
Month with food shortage	2.4 (0.4)	1.5 (0.4)	1.2 (0.4)	*
Distance to nearest market (km)	6.7 (1.2)	6.6 (1.6)	4.4 (1.5)	
Extension contacts (p.a.)	1.9 (0.3)	2.6 (0.5)	2.2 (0.5)	
Access finance	10% (0.1)	10 % (0.1)	23% (0.1)	
Sample size	20	20	22	

Table 10. Sample characteristics
Standard errors in brackets. Differences of means are indicated with *, **, and ***, which denote differences at the 10 %, 5 %, and 1 % level. 1The following weights were used: cattle=0.7, sheep=0.1, goats=0.1, pigs=0.2, chicken=0.01.

In addition, six focus groups discussions were conducted (three with men and three with women). Visual tools were used to facilitate discussion. For example, respondents were asked to judge activities according to the perceived work toil and enjoyableness. For this, a large sheet of paper with two crossing axes indicating work toil (from hard work to no work) and enjoyableness (from enjoyable to not enjoyable) was used. Respondents were given stickers representing different activities. The stickers were placed within the framework once consensus was reached.

4.3.3. Statistical Analysis

The data collected always has positive numbers and sums up to fixed sum of 1440 minutes (24 hours) per day. Time spent on different activities is correlated and co-dependent: an increase in time spent on one activity reduces the time available for other activities. Such a data structure is known as compositional data (CoD) and requires special attention due to two features: sum constraint and correlation (Atchinson, 1986; Bacon-Shone, 2011). A simple series of univariate analyses, where each time-use category is analyzed separately is incapable to account for these features. A multivariate analysis, where all categories are analyzed simultaneously, can account for correlation but not for the sum constraint. The latter constraint can be addressed by fitting multivariate models to log-transformed ratios of the categories of a composition, so called log-ratios (lr), which are assumed to be logit-normal distributed (Bacon-Shone, 2011). Such methodology has been coined compositional data analysis (CoDA). CoDA yields higher accuracy than univariate analysis (e.g. Chastin et al. 2015; Gupta et al. 2018).

In this study, the values of single categories underwent an additive lr-transformation (alr), where each category is divided by a reference category and the resulting ratios were transformed by taking the natural logarithm (Bacon-Shone, 2011). A set of K=9 categories was constructed from the raw data, which resulted in k=K-1, i.e. k=8 lr. Table 11 shows the aggregated categories. The category 'personal care' was used as common reference category for log-ratio transformation. Atchison (1986) showed that conclusions about relations of compositions are independent of which category is chosen as reference.

	Group	Sub activities
1	Crop farming	
1.1	Land Preparation	Land clearing, hoeing, plowing, harrowing, dibbling, potholing, ripping, ridging and raking (all with different power sources)
1.2	Weeding	Weeding by hand or using draught animals, knapsack sprayers, boom sprayers, and pest and disease control
1.3	Harvesting/processing	Harvesting, bundling, drying, storing, bagging, shelling, grinding, pounding, milling, winnowing (all with different power sources)
2	Crop farming (others)	Planting, applying fertilizer, applying manure, guarding of crops, watering as well as the activities that are not specifying the respective season (for example weeding and harvesting/processing activities during land preparation season)
3	Rural livelihood activities	Beverage preparation, marketing, animal husbandry, hunting, fishing, gathering food and grasses, charcoal making, maintaining/repairing, farm administration, vegetable garden, construction (household and community), meeting, cooking (community)
4	Off-farm and seasonal labor	Off-farm activities and the above mentioned farm activities as hired labor
5	Transportation	Walking, motorbike, bicycle, animal cart, car/van, bus, tractor (all of which can be loaded or unloaded)
6	Education	
7	Domestic	Care of children, sick and old, fetching water, collecting firewood, cooking (household), cleaning, washing pots and clothes, buying groceries
8	Leisure	Resting, media, religion, chatting, sports, dancing, making music
9	Personal care	Sleeping, being sick, eating, drinking, personal hygiene

Table 11. Aggregation of time-use activities to overall groups

A complication was that some activities were not done by every participant, resulting in zero values where a log-ratio transformation could not have been defined. Commonly, zeros in CoDA are subdivided into 'structural' or 'essential' zeros, where the category is truly empty or 'rounded' zeros, where the number is below a detection limit (Martin-Fernandez et al. 2003). Empty time-use categories could represent structural zeros as an activity may not have been performed by a participant. Martin-Fernandez et al. (2003) suggest analyzing the data separately for subjects performing and not performing a certain activity. However, the data recordings of the subjects' daily activities during three subsequent days are too short to conclude that subjects would come from different populations, one with a certain activity, the other one without. Moreover, it seems reasonable to consider zero values as being under a detection limit if e.g. an activity is performed for periods seemingly too short to be worth recording. Hence, multiplicative replacement - a method recommended for rounded zeros - was used and zeros were replaced by the small amount of one minute (ibid.).

A multivariate model was used to study the dependence of alr of time-use on mechanization and gender. As the sampling was stratified by communities with three different members of each household sampled, the multivariate model for analysis was extended to account for the possible correlations of observations within communities and households. The following multivariate linear mixed model was fitted to the alr-transformed data:

$$\begin{pmatrix} \log\left(\frac{y_{1ijml}}{y_{9ijml}}\right) \\ \vdots \\ \log\left(\frac{y_{8ijml}}{y_{9ijml}}\right) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_8 \end{pmatrix} + \begin{pmatrix} c_{1i} \\ \vdots \\ c_{8i} \end{pmatrix} + \begin{pmatrix} \tau_{1j} \\ \vdots \\ \tau_{8j} \end{pmatrix} + \begin{pmatrix} \gamma_{1l} \\ \vdots \\ \gamma_{8l} \end{pmatrix} + \begin{pmatrix} (\tau\gamma)_{1jl} \\ \vdots \\ (\tau\gamma)_{8jl} \end{pmatrix} + \begin{pmatrix} h_{1ijm} \\ \vdots \\ h_{8ijm} \end{pmatrix} + \begin{pmatrix} e_{1ijml} \\ \vdots \\ e_{8ijml} \end{pmatrix}, \qquad (1)$$

where the response vector contains the log-transformed ratios of time of 8 time-use categories divided by the reference category ‘personal care’. Each time-use category of each household member with gender *l* of each household *m* with mechanization level *j* from community *i* underwent this transformation. μ_1 to μ_8 are the fixed effects of time-use categories (tuc) 1 to 8, c_{1i} to c_{8i} are the random tuc-specific effects of the *i*-th community, τ_{1j} to τ_{8j} are the tuc-specific fix effects of the *j*-th mechanization type with the levels ‘manual, ‘animal’ and ‘tractor’. γ_{1l} toγ_{1l} are the tuc-specific fix effects of the *l*-th gender with levels: ‘female adult’, ‘male adult’, ‘girl’ and ‘boy’. $(\tau\gamma)_{1jl}$ to $(\tau\gamma)_{8jl}$ are the tuc-specific interaction terms of gender and mechanization type. h_{1ijm} to h_{8ijm} are the tuc-specific random household effects and e_{1ijml} to e_{8ijml} are the residual error terms. Time-use-category-specific random effects for community, household and residual error were assumed to have a multivariate normal distribution with mean zero and tuc-specific variances. Individual covariance parameters were estimated for all pairs of tucs, resulting in the following variance-covariance structure for communities:

$$\begin{pmatrix} c_{1i} \\ \vdots \\ c_{8i} \end{pmatrix} \sim MVN \left[\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{c1}^2 & \cdots & \sigma_{c1,8}^2 \\ \vdots & \ddots & \vdots \\ \sigma_{c8,1}^2 & \cdots & \sigma_{c8}^2 \end{pmatrix} \right],$$

for households

$$\begin{pmatrix} h_{1ijm} \\ \vdots \\ h_{8ijm} \end{pmatrix} \sim MVN \left[\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{h1}^2 & \cdots & \sigma_{h1,8}^2 \\ \vdots & \ddots & \vdots \\ \sigma_{h8,1}^2 & \cdots & \sigma_{h8}^2 \end{pmatrix} \right],$$

and for residual errors:

$$\begin{pmatrix} e_{1ijml} \\ \vdots \\ e_{8ijml} \end{pmatrix} \sim MVN\left[\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{e1}^2 & \cdots & \sigma_{e1,8}^2 \\ \vdots & \ddots & \vdots \\ \sigma_{e8,1}^2 & \cdots & \sigma_{e8}^2 \end{pmatrix}\right],$$

resulting in a total of 108 variance-covariance parameters to estimate. Model (1) was fitted to the data of the three seasons separately.

Model parameters were estimated using the HPMIXED procedure of SAS (Version 9.4). Variance components were estimated by the method of restricted maximum likelihood (REML) and subsequently transferred to the MIXED procedure, which was used for inferences on fixed effects. Model assumptions of normal distribution of residuals and homogeneity of variance were graphically assessed. The presence of (multiplicatively replaced) zeros necessarily led to slight shortcomings in the fulfillment of assumptions. Normal distribution assumption was usually fulfilled; however, plots for homogeneity of variance showed some changes in the variance over the range of predicted values. Fixed effects were studied by partial Wald-type F-tests. The most appropriate method of Kenward and Roger (1997) to approximate the denominator degrees of freedom was relinquished because of disproportionately high computing time and the 'between-within-method' was used instead (Schluchter and Elashoff, 1990).

The influence of covariates such as household size and size of cultivated land was further studied in univariate models, where the time-use for the respective agricultural activities were regressed on different covariates. Hence, multiple linear regressions were performed where all regressors entered the model linearly without interaction. Fixed main effects for community, mechanization, gender and interaction of gender and mechanization, as well as random intercepts for households were constituent components of the model. The terms in the model were successively removed from by backwards-elimination. The criterion for keeping or removing a covariate was the p-value in a partial Wald-F-test at α=10%. All three response variables of the three multiple linear regression models, time spent on land preparation, weeding and harvesting/processing were square-root transformed to fulfill homogeneity of variance. The multiple linear regressions were fitted using the MIXED procedure.

4.4. Results

In this part, the research hypotheses will be tested. Section 4.4.1 addresses hypothesis 1 and 2, section 4.4.2 addresses hypothesis 3 and section 4.4.3 addresses hypothesis 4.

4.4.1. Are land preparation activities gendered? To which extent benefit different gender from mechanization?

In section 4.2, two hypotheses were developed: 1) land preparation is predominantly a male activity and 2) mechanized land preparation benefits males relatively more than females. The F-tests show a significant effect of gender and mechanization based on model (1). This means that the composition of overall time-use differs depending on gender and mechanization (Table 12). There is no significant interaction of mechanization and gender on the overall daily composition of time-use.

Effect	Description	Numerator DF	Denominator DF‡	F-value	p-value
μ_k	Effect of time-use category (tuc)	8	20	756.53	<0.0001
τ_{kj}	tuc-specific Effect mechanization (M)	16	48	1.90	0.0450
γ_{kl}	tuc-specific Effect of gender (G)	24	72	17.29	<0.0001
$(\tau\gamma)_{kjl}$	tuc-specific Interaction of M and G	48	136	1.15	0.2597

Table 12. Partial Wald-F-tests for fixed effects for land preparation
Tests based on model (1); k=1 to 8 are 8 additive log-ratios of tuc with 'personal care' as common denominator
‡Denominator Degrees of freedom are adjusted according to the 'Between-Within-Method'

However, there is a significant interaction of gender and mechanization ($p < 0.0001$) for the single time-use activity 'land preparation on the own farm' based on model (1). This interaction was further studied in pairwise t-tests (Figure 13).

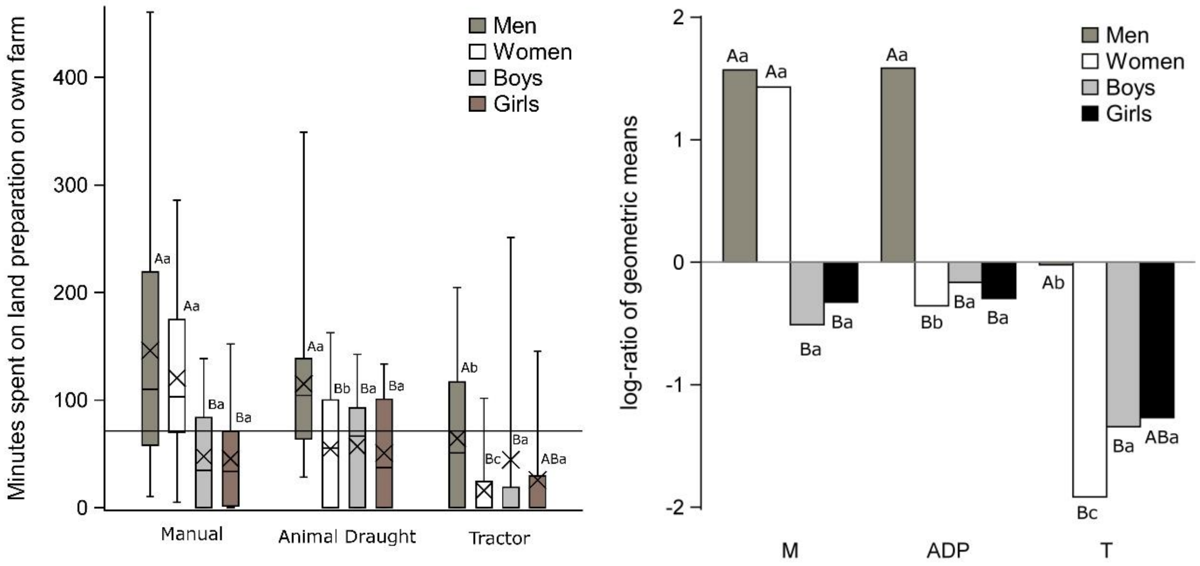


Figure 13. Boxplots (left) and descriptive log-ratios of geometric (right) of time-use on land preparation on own farm
By mechanization and gender. In the right figure, each bar represents the log-transformed ratio of the mean of each group compared to the overall mean of all 12 groups. Log-ratios larger or lower than zero represent above or below average time-use. Pairwise comparisons are based on estimates from model (1). Lower case letter refer to differences by mechanization within the same gender at α=10%. Capital letters refer to differences of different gender within the same mechanization type at α=10%.

In tractor-households, men spent significantly less time on land preparation (arithmetic mean of 64 minutes) compared animal-households (115 minutes, p=0.0072) and manual-households (146 minutes, p=0.0081), the latter two did not differ significantly (p=0.9671). Women in manual-households spent 120 minutes on land preparation while their counterparts in animal-households (54 minutes, p=0.0040) and tractor-households spent significantly less time (16 minutes, p<0.0001). Time spent is significantly lower for women in tractor-households compared to animal-households (p=0.0063). The reduction of time-use can be observed despite tractor- and animal-households cultivating more land (Table 10Table 10. Sample characteristics). Time spent did not differ significantly between mechanization types for boys and girls. Within tractor and animal-households, men spent the significantly highest amount of time (64 and 115 minutes) compared to women (16 minutes, p=0.0005 for tractor-households and 54 minutes, p=0.0011 for animal-households) and children. However, in manual-households the contribution of men (146 minutes) and women (120 minutes) did not differ (p=0.8211) and both spent significantly more time than their children.

The numbers presented so far cannot prove causality (mechanization leading to less time spent on land preparation), although economic theory would suggest this. The difference between time-uses may also occur because households differ with regard to other variables (and differed already before some became mechanized). In Table 13, some factors that might also be correlated with time spent on land preparation are controlled for using multiple linear regression. Controlling these factors, the interaction factor (gender*mechanization) remains highly significant. This suggests that mechanization has more influence on the time spent than factors such as cultivated land size, household size, or hired labor. However, many of these variables differ between the mechanization groups (Table 10) and, consequently, a regression on these variables without mechanization shows significant slopes (for example, a negative slope for cultivated land size, data not shown).

Effect	Estimate	DF	F-value	p-value
Community	- ‡	55.3	3.80	**0.0151**
Gender (M, F, B, G)	-	133	12.53	**<.0001**
Mechanization (M, A, T)	-	64.3	7.53	**0.0012**
Gender*mechanization	-	133	2.48	**0.0266**
Off-farm income	-0.00005 (0.000027)	54.5	3.45	**0.0687**
Costs per ha	0.000418 (0.000266)	53.4	2.47	0.1217
Pregnancy	1.2098 (1.0592)	52.5	1.30	0.2586
Household size	-0.1348 (0.1858)	53.7	0.53	0.4714
Tropical livestock unit	-0.03753 (0.06147)	51.3	0.37	0.5442
Distance market	-0.01339 (0.02223)	49.2	0.36	0.5498
Hired labour	-0.01016 (0.02537)	48.5	0.16	0.6906
Months with food shortage	0.07680 (0.2438)	50.4	0.10	0.7541
Education	0.02409 (0.1075)	46.2	0.05	0.8326
Crop diversity	0.06811 (0.3980)	45.4	0.03	0.8649
Land cultivated	-0.00992 (0.1181)	47.4	0.01	0.9334

Table 13. Multiple linear regression of covariates on time-use for land preparation
With parameter estimates for slopes and standard error in parentheses and F-tests. Multiple linear regression on square-root transformed time spent in land preparation. Covariates were removed in back-wards elimination. Threshold of deletion were p-values below 10%. The model contains a random intercept for each household.
‡ Parameter estimates for qualitative factors are not shown for brevity.

The hypothesis that land preparation is a male activity can only partially be confirmed. In manual-households, men and women equally contribute to land preparation. A gender differentiation emerges only with the use of different forms of mechanization (by animal draught and tractors). In animal-households, women spend less time on land preparation activities compared to manual-households, while men spent a comparable amount of time. When using tractors, both men and women work less but men work more than women. In general, the time spent on land preparation is the lowest for all household members when tractors are used and children contribute less time irrespective of mechanization. The hypothesis that men benefit most from mechanization cannot be confirmed. Men do benefit from mechanization but women benefit relatively more. Children who spend less time on land preparation than their adults are little affected by mechanization.

4.4.2. Is time "saved" used differently by gender?

The previous section has shown that agricultural mechanization is associated with less time spent on land preparation. Clearly, this time must be spent on some other activities. In this section, the hypothesis is tested that males and females use this time differently. In the previous section, no time saving effects for children were found, who are thus omitted in this section.

Estimates of time spent on different activities from model (1) were compared in pairwise t-tests between men and women and between the three mechanization categories. Table 5 presents the difference compared to manual-households. Table 14 suggests that women in animal-households spent significantly less time (65 minutes) on land preparation compared to manual-households. This was not the case for men who thus have no extra time that could be spent on other activities. It is not clear for which activities the additional time that women in animal-households have is used. Potentially, time is spent on domestic work, which is 114 minutes higher but the difference is slightly above significance. In tractor-households, both men and women spent less time on land preparation activities compared to manual-households. The extra time seems to be used for off-farm work by women. Men in tractor-households spent more time on domestic work compared to animal-households but this is compared to low base, and compared to manual-households, no significant difference was found. This suggests that additional time saved might be distributed across all other time-use categories such as leisure and transport, and therefore, stays below detection level. Still, the hypothesis that males and females in mechanized households use their extra time differently can be confirmed.

	Animal		Tractor	
	women	men	women	men
Crop farming (land preparation)	-65^{a} (0.004)	-31^{a} (0.967)	-105^{b} (< 0.0001)	-81^{b} (0.008)
Crop farming (others)	0^{a} (0.453)	8 (0.129)	-8^{b} (0.094)	-1 (0.617)
Rural livelihood activities	-10 (0.757)	-33 (0.631)	10 (0.779)	-30 (0.899)
Off-farm and seasonal labour	8^{a} (0.399)	1 (0.747)	39^{b} (0.087)	27 (0.664)
Transportation	-24 (0.677)	43 (0.398)	-19 (0.254)	43 (0.562)
Education	0 (0.983)	0 (0.893)	0 (0.923)	4 (0.534)
Domestic	114 (0.123)	-15^{a} (0.553)	28 (0.409)	3^{b} (0.131)
Leisure	-18 (0.613)	31 (0.216)	22 (0.828)	35 (0.473)

Table 14. Difference of time-use relative to manual-households
By mechanization and gender. How to read the table, example row 'Crop farming (others)': Women in animal-households spent the same and women in tractor-households spent 8 minutes less time compared to manual-households. While the first difference is statistically not significant (p=0.453), the second is at α=10%. A pairwise comparison between time-use of women in animal- and tractor-households was significant at α=10%, therefore the two values carry different lower case letters a and b.

Figure 14 presents a framework of how different activities are perceived by respondents. Following this framework, animal- and tractor-households spent less time on hard and non-enjoyable activities but more on enjoyable activities (such as child care, cleaning and cooking). Thus, despite not finding a significant difference with regard to time spent on leisure, respondents seem to have a higher life quality with regard to these criteria.

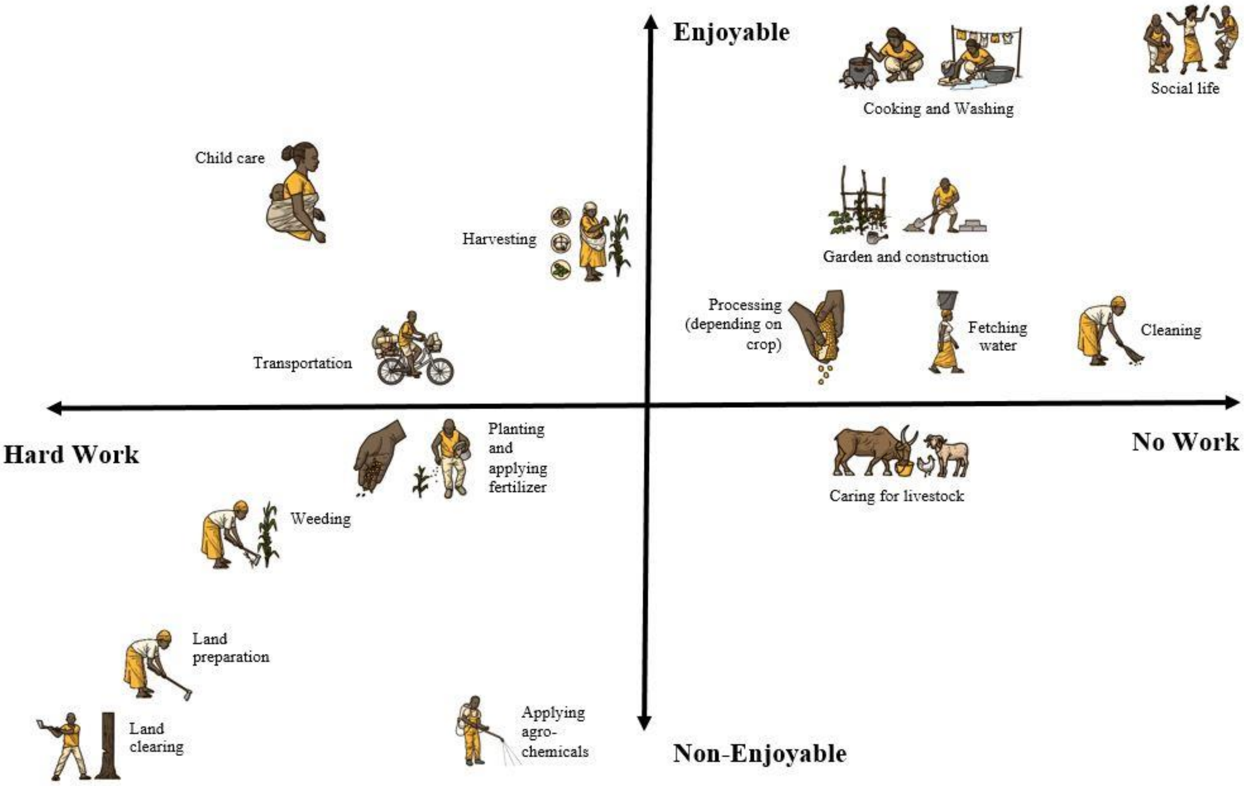


Figure 14. Matrix of activities by enjoyableness and drudgery

4.4.3. What happens during the next farming steps?

In this section, the hypothesis is tested whether females spent more time on weeding and harvesting/processing on mechanized farms compared to non-mechanized farms. This could be the case when mechanized households cultivate more land, which increases the need for weeding and harvesting/processing, which might be primarily female tasks. The argument that mechanization leads to land expansion cannot be thoroughly analyzed in this study as it is based on cross-sectional data but seems plausible based on economic theory and previous studies. Adu-Baffour et al. (2018), for example, have shown that Zambian farm households mechanizing land preparation can double the amount of land cultivated. In this paper, the sampled tractor-households cultivated significantly (p=0.0053) more land (6.7 ha) than animal-households (3.9 ha); and animal-households cultivated significantly more land than manual-household (2.1 ha). The larger amount of land cultivated may be correlated with more time spent on weeding (and harvesting/processing). Indeed, Table 1Table 15 suggests a significant effect of the interaction of mechanization and gender on the daily time-use composition during weeding based on model (1).

Effect	Description	Numerator DF	Denominator DF‡	F-value	p-value
μ_k	Effect of time-use category (tuc)	8	20	1313.09	<0.0001
τ_{kj}	tuc-specific Effect mechanization (M)	16	48	2.09	0.0249
γ_{kl}	tuc-specific Effect of gender (G)	24	72	83.82	<0.0001
$(\tau\gamma)_{kjl}$	tuc-specific Interaction of M and G	48	128	2.53	<0.0001

Table 15. Partial Wald-F-tests for fixed effects for weeding
Tests based on model (1)

However, there is no significant gender differences in pairwise t-tests at α=10% between manual-households and animal-households for the single time-use category of 'weeding on the own farm' (Figure 15). In tractor-households, men work significantly less than women and men and boys work significantly less compared to their counterparts in manual- and animal-households. This suggests that a gender differentiation for weeding activities only appears with the use of tractors and men especially benefit from this. However, neither girls nor women are negatively affected in terms of time-use.

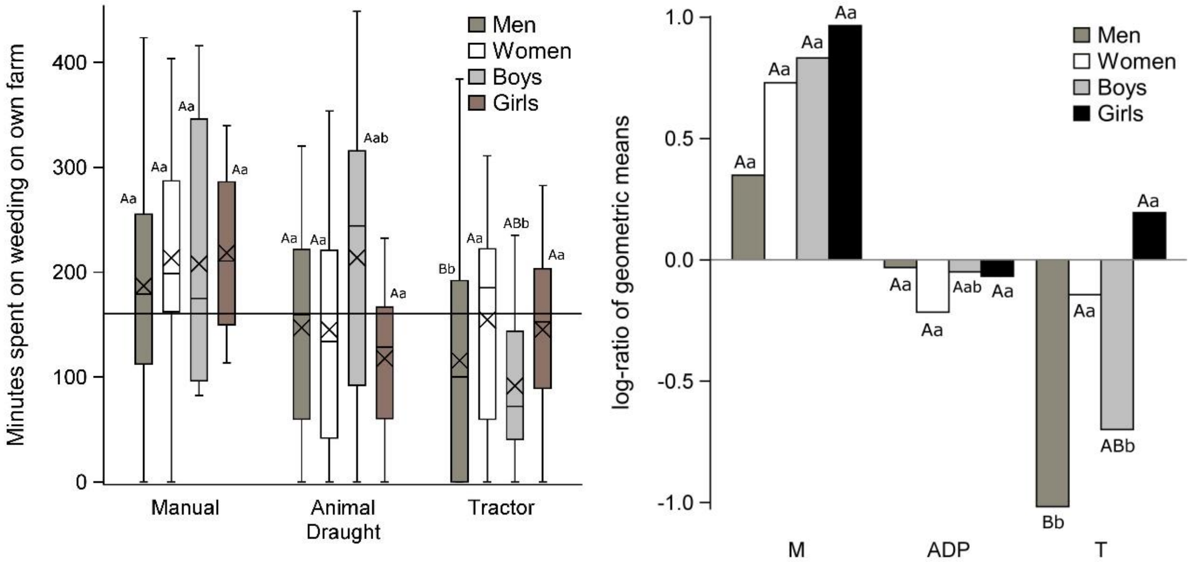


Figure 15. Boxplots (left) and descriptive log-ratios of geometric means (right) of time-use for weeding on own farm
By mechanization and gender

Despite cultivating more land, animal- and tractor-households do not spend more time on weeding. The time spent on weeding was highest for manual-households (204 minutes), significantly higher than animal-households (152 minutes, p=0.075). Time spent on weeding for animal-households did not differ significantly from tractor-households (130 minutes, p=0.308). However, time spent on weeding may also be influenced by other factors such as the use of herbicides and laborers.

Effect	Estimate	DF	F-value	p-value
Community	-	58.1	3.47	**0.0216**
Gender (M, F, B, G)	-	132	0.97	0.4103
Mechanization (M, A, T)	-	48.9	0.68	0.5133
Land cultivated	-0.4946 (0.09852)	57.5	25.20	**<0.0001**
Tropical livestock unit	-0.1130 (0.07136)	55.7	2.51	0.1191
Months with food shortage	-0.3685 (0.2661)	55.8	1.92	0.1717
Crop diversity	-0.4423 (0.4426)	54.0	1.00	0.3221
Off-farm income	-0.00004 (0.000038)	54.1	0.87	0.3550
Fertilizer per ha	-0.00323 (0.003439)	51.2	0.88	0.3516
Education	0.1128 (0.1302)	50.8	0.75	0.3905
Hired labour	-0.01037 (0.01338)	50.5	0.60	0.442
Pregnancy	-0.5667 (1.4045)	46.6	0.16	0.688
Household size	-0.06683 (0.2605)	46.8	0.07	0.7987
Distance market	-0.00840 (0.03059)	45.0	0.08	0.7848
Pesticides per ha	-0.01235 (0.05238)	43.7	0.06	0.8147
Gender * mechanization	-	122	0.44	0.8474
ADP weeding	-0.1296 (1.3958)	43.5	0.01	0.9265
Cost per land	0.000035 (0.000632)	40.8	0.00	0.9560

Table 16. Multiple linear regression of covariates on time-use for weeding
With parameter estimates for slopes, standard errors in parentheses and F-tests. Multiple linear regression on square-root transformed time spent on weeding.

Table 16 shows that when controlling for covariates, the effect of mechanization on time spent on weeding becomes insignificant as the size of cultivated land has a larger influence on time spent on weeding. The relationship between cultivated land size and time-use for weeding is negative. For subsistence farming households with little land, weed control may be more essential than for households with large landholdings. Thus, the hypothesis that mechanization of land preparation is associated with increased female labor for weeding must be rejected.

However, this may still be the case for harvesting/ processing. Table 17 shows that there were no tuc-specific effects of mechanization and the interaction of mechanization and gender on the overall daily time-use composition during harvesting/processing.

Effect	Description	Numerator DF	Denominator DF‡	F-value	p-value
μ_k	Effect of time-use category (tuc)	8	20	1659.88	<0.0001
τ_{kj}	tuc-specific Effect mechanization (M)	16	48	0.46	0.9560
γ_{kl}	tuc-specific Effect of gender (G)	24	72	14.88	<0.0001
$(\tau\gamma)_{kjl}$	tuc-specific Interaction of M and G	48	136	1.18	0.2251

Table 17. Partial Wald-F-tests for fixed effects for harvesting/processing
Tests based on model (1)

In pairwise t-tests on the single time-use category of harvesting/processing no gender differences were found in manual-households based on model (1). In animal-households, girls work significantly less than all other household members, while boys work less in tractor-households (Figure 2Figure 16).

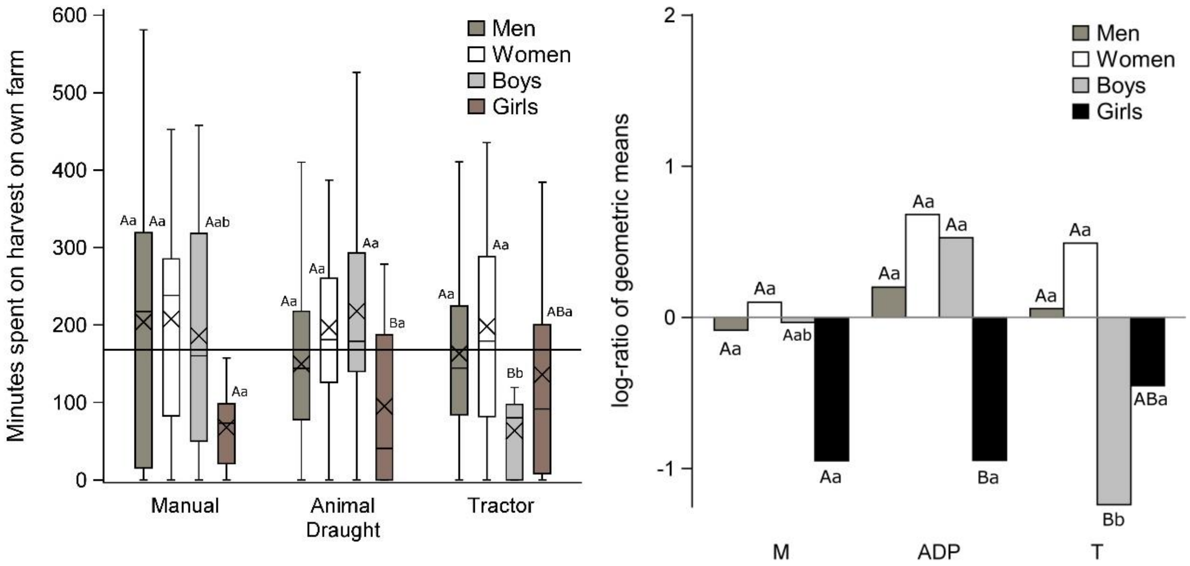


Figure 16. Boxplots (left) and descriptive log-ratios of geometric means (right) of time-use for harvesting/processing on own farm
By mechanization and gender

Table 18 shows that factors other than mechanization have a bigger influence on time spent on harvesting/processing. This includes livestock owned based on tropical livestock units – potentially, households with more livestock spend more time caring for animals and have less time for harvesting/processing. Another factor is the use of hired labor: households hiring more labor spent less time on harvesting/processing. Finally, households with more months of food shortage spent less time on harvesting/processing (even after yields were dropped from the regression), a phenomenon that may show that households who suffered food shortages consume most of the harvest directly rather than processing for sale. It may also be that such households have less energy to work.

Effect	Estimate	DF	F-value	p-value
Gender (M, F, B, G)	-	132	4.70	**0.0037**
Tropical livestock unit	-0.1366 (0.07523)	55.6	3.30	**0.0748**
Hired labour	-0.03548 (0.01575)	56	5.07	**0.0282**
Months with food shortage	-0.7116 (0.3024)	56.2	5.54	**0.0221**
Community	-	53.1	1.61	0.1981
Pregnancy	-2.0299 (1.4822)	51.6	1.88	0.1768
Off-farm income	-0.00004 (0.000039)	53	0.99	0.3232
Mechanization (M, A, T)	-	48.9	0.05	0.9481
Gender*mechanization	-	125	0.99	0.4362
Cultivated land	0.1362 (0.1679)	49.5	0.66	0.4213
Crop diversity	-0.4907 (0.5282)	45.4	0.86	0.3578
Distance market	-0.02771 (0.03342)	45.3	0.69	0.4113
Education	0.07261 (0.1518)	45	0.23	0.6374
Fertilizer per ha	-0.00186 (0.004419)	42.9	0.18	0.6764
Cost per ha	0.000296 (0.000666)	42.3	0.20	0.6584
Household size	0.07607 (0.3025)	43.1	0.06	0.8026
Yield	0.000069 (0.000459)	41.5	0.02	0.8819

Table 18. Multiple linear regression of covariates on time-use for harvesting/processing

With parameter estimates for slopes, standard errors in parentheses and F-tests. Multiple linear regression on square-root transformed time spent on harvesting/processing.

The hypothesis that agricultural mechanization during land preparation increases female labor needed for harvesting/processing has to be rejected.

4.5. Discussion and conclusion

New technologies, policies and practices can affect the intra-household allocation of time in smallholder farming households, which may put more vulnerable household members at a disadvantage. Understanding time-use effects is important to target policy interventions. However, exploring such effects has been difficult because 1) a lack of suitable data collection methods and 2) the structure of time-use data, which cannot be addressed with conventional statistical methods. This study showed that using a pictorial smartphone application called Timetracker provides sufficiently good and comprehensive data to study such concerns. Furthermore, the study has shown that compositional data analysis can be used to address the specific challenges of time-use data. Solving these two challenges, the study then explored the time-use effects of agricultural mechanization.

The results confirm existing literature that some farming activities such as land preparation are gendered (Alesina, 2011; Baanante et al., 1999). However, in this study the gender differentiation for land preparation activities (and weeding) only emerges with mechanization. No evidence could be found that harvesting/processing are gendered activities. This echoes findings from Doss et al. (2001) and Palacios-Lopez et al. (2017) who question stylized facts on the gender division of agriculture. The study finds that men and women benefit from agricultural mechanization with regard to time-use and that women benefit relatively more than men. It remains unclear whether this is a sign of empowerment or dis-empowerment as once they are not working on the fields, women may have less influence on farming, including farm income. This resonates with Alesina et al. (2011), who found that historically plough-based societies were less dependent on female labor compared to hoe-based societies and still have lower rates of female participation in work and society today. No significant evidence of time benefits for children during land preparation was found. Adu-Baffour et al., (2018), having a larger sample and focusing on the whole of Zambia, found that children benefit from agricultural mechanization.

Time "saved" due to mechanization seems to be distributed across various activities, with some evidence that women in animal-households use the "saved" time on domestic chores, which may be a sign of dis-empowerment; women in tractor-households use the extra time for off-farm work, which may be a sign of empowerment. In tractor-households, men spent significantly more time on domestic work but this is compared to a very low base. No negative second-round effects of increased time-use for weeding by women was found, despite mechanized household cultivating more land. A reason might be that mechanized land preparation reduces weed pressure

(Nyamangara et al., 2014). Also, households with more land spent less time on weeding as the intensity of labor use may decrease with farm size (Sen, 1952; Wineman and Jayne, 2017). During harvesting/processing, no effects of mechanization during land preparation on time-use were found.

As mentioned above, the study faces some limitations. Given that the focus has been on finding a reliable way to collect time-use data and how to analyze such data, the sample remained small. In subsequent studies, larger sample sizes should be envisaged. Ideally, future studies can find ways to use a randomized control trial (RCT) approach to establish causality. However, using RCTs is challenging given the ado-hoc adoption of tractor service and the diversity of agronomic conditions of farmers. Future study should at least envision having a larger sample and using propensity score matching. Another limitation is that the extrapolation of the daily data to the entire farm season remains difficult. Given these limitations, the study remains cautious with regard to policy implications. While no evidence of agricultural mechanization negatively affecting woman and children was found, this may be different in other situations depending on the tasks and crops are mechanized, the use of accompanying inputs as well as the existing gender roles and how they can be re-negotiated (Alderman, 1995; Doss, 2013; Fisher et al., 2000).

The study provides proof-of-concept that using picture-based smartphone apps can help to collect data on research areas that are difficult to measure and analyze but that are potentially highly relevant. Thus, the study opens the field to more studies focusing on agricultural development and time-use in rural areas. For example, this study found a high share of time spent on mobility and transportation (see appendix 2), which is often neglected by studies focusing on time-use in agriculture, although reducing such time-use may allow farmers to spend more time on their fields. Similarly, the time-use effects of technologies for home economics such as improved cook stoves, electronic household items and processed food, which may all help to reduce time poverty among women and loosen constraints to participate in paid work, may be interesting to study.

4.6. Acknowledgments

We are grateful to all the farm families participating in the study and the financial support from the "Program of Accompanying Research for Agricultural Innovation" (PARI), which is funded by the German Federal Ministry of Economic Cooperation and Development (BMZ).

4.7. References

Adu-Baffour, F., Daum, T., Birner, R. (2018). Can Big Companies' Initiatives to Promote Mechanization Benefit Small Farms in Africa? A Case Study from Zambia. ZEF-Discussion Papers No. 262. Center for Development Research (ZEF).

Agarwal, B. (1981). Agricultural mechanisation and labour use: a disaggregated approach. Int'l Lab. Rev., 120, 115.

Alderman, H., Chiappori, P. A., Haddad, L., Hoddinott, J., and Kanbur, R. (1995). Unitary versus collective models of the household: is it time to shift the burden of proof? The World Bank Research Observer, 10(1), 1-19.

Alesina, A. F., Giuliano, P., and Nunn, N. (2011). On the origins of gender roles: Women and the plough (No. 17098). National Bureau of Economic Research.
Arora, D. (2015). Gender Differences in Time-Poverty in Rural Mozambique. Review of Social Economy, 73(2), 196-221.
Atchison, J. (1986). The Statistical Analysis of Compositional Data. Springer.
Baanante, C., Thompson, T. P., and Acheampong, K. (1999). Labour contributions of women to crop production activities in three regions of West Africa: an analysis of farm-survey data. Institute of African Studies: Research Review, 15(1), 80-100.
Bacon-Shone, J. (2011). A short history of compositional data analysis. In Pawlowsky-Glahn, V., and Buccianti, A. (Eds.). (2011). Compositional data analysis: Theory and applications (p. 3–11). John Wiley and Sons.
Benin, S. (2015). Impact of Ghana's agricultural mechanization services center program. Agricultural economics, 46(S1), 103-117.
Bergman-Lodin, J., Paulson, S., and Mugenyi, M. S. (2012). New seeds, gender norms and labor dynamics in Hoima District, Uganda. Journal of Eastern African Studies, 6(3), 405-422.
Bell, A., Ward, P., Tamal, E. H., and Killilea, M. (2019). Assessing recall bias and measurement error in high-frequency social data collection for human-environment research. Population and Environment, 1-21.
Bryceson, D. F. (2019). Gender and generational patterns of African deagrarianization: Evolving labour and land allocation in smallholder peasant household farming, 1980-2015. World Development, 113, 60-72.
Blackden, M. and Wodon, Q. (2006). Gender, Time-use, and Poverty in Sub-Saharan Africa. World Bank Working Paper 73.
Carletto, C., Gourlay, S., and Winters, P. (2015a). From guesstimates to GPStimates: Land area measurement and implications for agricultural analysis. Journal of African Economies, 24(5), 593-628.
Carletto, C., Jolliffe, D., and Banerjee, R. (2015b). From tragedy to renaissance: improving agricultural data for better policies. The Journal of Development Studies, 51(2), 133-148.
Chastin, S. F., Palarea-Albaladejo, J., Dontje, M. L., and Skelton, D. A. (2015). Combined effects of time spent in physical activity, sedentary behaviors and sleep on obesity and cardio-metabolic health markers: a novel compositional data analysis approach. PloS one, 10(10), e0139984.
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) https://doi.org/10.1177/1525822X18797303.
Daum, T., Buchwald, H., Gerlicher, A., and 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.
Daum, T., and Birner, R. (2017). The neglected governance challenges of agricultural mechanisation in Africa-insights from Ghana. Food Security, 9(5), 959-979.
Diao, X., Cossar, F., Houssou, N., and Kolavalli, S. (2014). Mechanisation in Ghana: Emerging demand and the search for alternative supply models. Food Policy.
Doss, C. (2013). Intrahousehold bargaining and resource allocation in developing countries. The World Bank Research Observer, 28(1), 52-78.
Doss, C. R. (2001). Designing agricultural technology for African women farmers: Lessons from 25 years of experience. World Development, 29(12), 2075-2092.
Evers, B., and Walters, B. (2001). The Model of a Gender-Segregated Low-Income Economy Reconsidered: Evidence from Uganda. Review of Development Economics, 5(1), 76-88.
Farnworth, C. R., Baudron, F., Andersson, J. A., Misiko, M., Badstue, L., and Stirling, C. M. (2016). Gender and conservation agriculture in East and Southern Africa: towards a research agenda. International Journal of Agricultural Sustainability, 14(2), 142-165.
Fisher, M., R. Warner, and W. Masters. 2000. Gender and agricultural change: Crop-livestock integration in Senegal. Society and Natural Resources 13 (3): 203-222.
Fraval, S., Hammond, J., Wichern, J., Oosting, S. J., De Boer, I. J., Teufel, N., ... and Giller, K. E. (2018). Making the most of imperfect data: A critical evaluation of standard information collected in farm household surveys. Experimental Agriculture, 1-21.
Gupta, N., Mathiassen, S. E., Mateu-Figueras, G., Heiden, M., Hallman, D. M., Jørgensen, M. B., and Holtermann, A. (2018). A comparison of standard and compositional data analysis in studies addressing group differences in sedentary behavior and physical activity. International Journal of Behavioral Nutrition and Physical Activity, 15(1), 53.
Indaba Agricultural Policy Research Institute. 2016. Rural Agricultural Livelihoods Survey 2015 Survey Report. https://images.agri-profocus.nl/upload/post/RALS_2015_Survey_Report__Finalv-_edited1456827249.pdf (accessed January 8, 2018).

ILO (International Labor Organization). (2019). Child labor in agriculture. https://www.ilo.org/ipec/areas/Agriculture/lang--en/index.htm (retrieved on February 16 2019).
Jayne, T. S., Chamberlin, J., Traub, L., Sitko, N., Muyanga, M., Yeboah, F. K., ... and Kachule, R. (2016). Africa's changing farm size distribution patterns: the rise of medium-scale farms. Agricultural Economics, 47(S1), 197-214.
Johnston, D., Stevano, S., Malapit, H. J., Hull, E., and Kadiyala, S. (2018). Time-use as an explanation for the agri-nutrition disconnect? Evidence from rural areas in low and middle-income countries. Food policy.
Kenward, M. G., and Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. Biometrics, 983-997.
Kumar, S. K. (1994). Adoption of hybrid maize in Zambia: Effects on gender roles, food consumption, and nutrition (Vol. 100). IFPRI.
Lambrecht, I., Schuster, M., Samwini, S. A., and Pelleriaux, L. (2017). Changing gender roles in agriculture? Evidence from 20 years of data in Ghana. Agricultural Economics.
Martín-Fernández, J. A., Barceló-Vidal, C., and Pawlowsky-Glahn, V. (2003). Dealing with zeros and missing values in compositional data sets using nonparametric imputation. Mathematical Geology, 35(3), 253-278.
Mukhopadhyay, M. (1984). Silver shackles: Women and development in India. Oxfam.
Nyamangara, J., Mashingaidze, N., Masvaya, E. N., Nyengerai, K., Kunzekweguta, M., Tirivavi, R., and Mazvimavi, K. (2014). Weed growth and labor demand under hand-hoe based reduced tillage in smallholder farmers' fields in Zimbabwe. Agriculture, ecosystems and environment, 187, 146-154.
Palacios-Lopez, A., Christiaensen, L., and Kilic, T. (2017). How much of the labor in African agriculture is provided by women?. Food policy, 67, 52-63.
Pearson, K. (1897). Mathematical contributions to the theory of evolution - on a form of spurious correlation which may arise when indices are used in the measurement of organs. Proceedings of the royal society of london, 60(359-367), 489-498.
Quisumbing, A. R., Brown, L. R., Feldstein, H. S., Haddad, L., and Peña, C. (1995). Women: The key to food security. Food policy statement, 21.
Sen, A. K. (1962). An aspect of Indian agriculture. Economic Weekly, 14(4-6), 243-246.
Schluchter, M. D., and Elashoff, J. T. (1990). Small-sample adjustments to tests with unbalanced repeated measures assuming several covariance structures. Journal of Statistical Computation and Simulation, 37(1-2), 69-87.
Takeshima, H. (2017). Overview of the evolution of agricultural mechanization in Nepal: A focus on tractors and combine harvesters (Vol. 1662). Intl Food Policy Res Inst.
Theis, S., Lefore, N., Meinzen-Dick, R., and Bryan, E. (2018). What happens after technology adoption? Gendered aspects of small-scale irrigation technologies in Ethiopia, Ghana, and Tanzania. Agriculture and Human Values, 1-14.
Wang, X., Yamauchi, F., and Huang, J. (2016). Rising wages, mechanization, and the substitution between capital and labor: evidence from small scale farm system in China. Agricultural economics, 47(3), 309-317.
Wineman, A., and Jayne, T. (2018). Factor market activity and the inverse farm size-productivity relationship in Tanzania. Feed the Future Innovation Lab for Food Security Policy. Research Paper 79.
Zanello, G., Srinivasan, C. S., and Nkegbe, P. (2017). Piloting the use of accelerometry devices to capture energy expenditure in agricultural and rural livelihoods: Protocols and findings from northern Ghana. Development Engineering, 2, 114-131.

5. Discussion

This thesis aimed to improve the collection of data underlying applied socioeconomic research in developing countries. This aim was formulated against the background of poor quality survey data relying on recall questions and the paucity of some socioeconomic data that is important but difficult to measure. It has been argued that both poor and lacking data can lead to misguided policy recommendations and actions - with adverse effects on farmers and more vulnerable population groups such as women and children. This can lead to agricultural development trajectories that are socially unequal and unsustainable. Against this background, this thesis set out to explore the potential of smartphone applications to collect data from agricultural households in developing countries. To assess the potential, a pictorial smartphone application called Timetracker was developed that allowed respondents to record data themselves in real time to reduce recall bias and improve data accuracy. Following the three papers presented as part of this endeavor, this final chapter distills some of the major methodological and empirical contributions of this thesis (see section 5.1.). The chapter then reflects on the limitations of the methodological and empirical approach chosen and discusses ways to address these limitations (see section 5.2.). The chapter then summarizes how the three research questions of this thesis have to be answered (see section 5.3) before taking a broader perspective and discussing potential future research pathways for using smartphone applications to collect data from agricultural households (see section 5.4.). Section 5.5 provides some concluding remarks.

5.1. Contributions to the literature

The purpose of this chapter is to discuss the contribution of smartphone application in its current form to the methodological portfolio of applied socioeconomic researchers in developing countries (see 5.1.1.) and how the findings of the three papers presented contribute to the existing empirical debates and knowledge gaps (5.1.2.)

5.1.1. Methodological contributions

This thesis was largely formulated against the background that the existing methods for collecting socioeconomic data, such as household surveys, are prone to recall bias, especially in developing countries. This thesis presented a method that allows time-use and nutrition data to be recorded in real time. This reduces recall bias, which is inherent to household surveys relying on seasonal, weekly or 24-hour recall. Studies have shown that such recall bias can be large, thereby affecting data quality for both time-use and nutrition data.

For example, with regard to time-use data, Arthi et al. (2018) showed that rural respondents in Tanzania who are asked about their time spent farming reported a value four times higher when asked using postharvest questionnaires compared to being asked on a weekly basis. Other studies have shown that the recall bias is high, even with much shorter recall periods, such as 24-hour recall periods (Chatzitheochari et al. 2017; Juster et al. 2003). Time-use researchers thus consider time-use diaries to be the gold standard (Juster et al. 2003). Time-use diaries can be paper-based, but an increasing number of researchers also use electronic time-use diaries (Fernee and Sonck, 2014; Minnen et al., 2014). Diaries are text-based, however, and thus leave-behind or self-reporting time-use diaries that are filled out by respondents cannot be used when the respondents lack literacy. To address this problem, some researchers have developed picture-based time-use diaries. Masuda et al. (2014) provided proof-of-concept that pictorial diary sets can be used to collect time-use data in developing countries. However, their diary sets remained cumbersome (involving a beeper, a booklet and activity stickers) and were based on coarse 30-minute time slots.

Building on the efforts of other researchers, this thesis developed a smartphone-based time-use diary, which works only with pictures and is thus applicable in developing countries where literacy levels can be low. The Timetracker minimizes selection bias by ensuring that no and low literate respondents and respondents without clock-based concepts of time can use the application, including elders and children. Collecting time-use data from adolescents is of interest for various research questions, for example, questions related to the high prevalence of child labor in many developing countries (ILO, 2019). For this thesis, all of the respondents sampled where willing and able to carefully record data for 15 days distributed across one farming season.

Given the real-time recording of time-use, the smartphone application reduces recall bias compared to recall-based data collection methods, as by design, time is not actually recalled in the application. Additionally, the real-time recording allows us to reach higher data granularity compared to time-use diaries, which are typically based on 15- or 30-minute intervals, where short activities are often ignored or grossly overstated (Buvinic and King, 2018; Chatzitheochari et al., 2017; Kelly et al., 2015). In general, the findings of this thesis suggest that well-designed tools, such as the Timetracker, can expand the data collection method toolbox of time-use researchers in developing countries, notwithstanding some limitations that are discussed in section 5.2.

With regard to the collection of nutrition data, recall bias poses a similar to that of time-use data (Beegle et al., 2016; Shim et al., 2014). Several researchers have therefore worked with nutrition diaries. Such diaries reduce recall bias but require literacy and numeracy and well-trained and motivated respondents (Shim et al., 2014). As these preconditions cannot always be fulfilled in developing countries, Beegle et al. (2012) found high levels of underreporting for illiterate households and for urban respondents (with higher opportunity costs) completing such diaries. They thus question the superiority of diaries over recall questions. This thesis offers an alternative view and suggests that picture-based and well-designed smartphone diaries may help to overcome some of the drawbacks of food and nutrition diaries. However, the food and nutrition data collected with the help of the smartphone application as part of this thesis were not validated. Thus, this thesis carefully regards any statements on the quality of the food and nutrition data collected and leaves the validation open for future studies. In general, however, higher reporting frequencies are associated with higher data quality, which should promote optimism (Bell et al., 2016; Bell et al., 2019).

5.1.2. Empirical contributions

In addition to contributing to the methods toolbox of researchers working in developing countries (and potentially also developed countries), this thesis also makes some empirical contributions, particularly with regard to two aspects: a) recall bias in time-use and b) gender roles in agricultural development. Both contributions will be discussed in this section.

Recall bias

This study presented a tool for data collection that minimizes recall bias. However, researchers may, for various reasons, continue to use other research methods to collect data, including household surveys relying on recall questions. Using recall-based data collection methods may be justified depending on the context, but when relying on recall questions, it is important to be aware and understand the nature of the associated recall bias. For example, it is important to understand which time-use categories or which food groups are likely to be overestimated and which ones are likely to be underestimated. This study contributes to a better understanding of the direction and magnitude of recall bias with regard to time use. This was done by comparing the data recorded using the app with data collected through 24-hour recall questions. When asking the 24-hour recall questions, only time-use data (no nutrition data) were collected. Thus, this section focuses on recall bias with regard to time use.

This thesis shows that regular and externally structured activities, such as going to school, are recalled very well, a fact that helps to validate the Timetracker application and confirms previous studies (Juster et al. 2003). This thesis finds that socially desirable and drudgery activities, such as farming, are reported as taking significantly longer. For example, when using 24-hour recall questions, harvesting was reported as taking between 29% and 47% longer compared to the data captured with the real-time recording of the Timetracker. The likely overestimation of time spent farming echoes work by Arthi et al. (2018), who found that time spent farming decreases by a factor of 4 when reducing the recall period from seasonal to weekly. In contrast to farming, other activities are reported to have lasted for shorter periods based on 24-hour recall. These include social activities and sleeping. This confirms much of the literature on time-use data. According to Hofferth (1999) and Juster et al. (2003), social activities, for example, are often underestimated because respondents perceive these activities to be less socially desirable.

In this thesis, child care activities were recalled as lasting longer (or being more frequent) when recalled compared to what was recorded with the Timetracker. This confirms studies from Europe and the United States showing that child care activities are often overestimated (Bianchi et al., 2012; Juster, 2003). Importantly, with regard to child care activities, time-use data collection faces various limitations; however, this will be discussed in more detail in section 5.2. Child care activities can be of a primary or secondary nature. In the latter case, they are performed simultaneously with other activities, and such activities are often overestimated (Juster et al. 2003), which would explain the findings of this thesis. However, secondary child care activities may also be underestimated when they are not perceived as work by the respondents and society. Then, they may simply not be recorded – neither when using time-use diaries nor tools such as the Timetracker. In Malawi, Lentz et al. (2018) found that child care work is “dramatically underreported” (p. 1).

The example of child care activities points towards an important aspect of this study. Recall bias depends on the values and norms of societies. Thus, the external validity of the findings on recall bias may be limited, and generalizations that are too brave with regard to the magnitude and direction of recall bias should be avoided. The direction and order of magnitude of recall bias can be very different in other sociocultural contexts. Given that the direction and magnitude of recall bias may differ by gender, education, age and sociocultural factors, studies comparing different population groups (such as male or females, rural compared to urban populations, and pastoralists compared to farmers) need to be especially cautious of recall bias.

In general, this thesis points to the existence of a strong recall bias when collecting time-use data in developing countries using recall questions, even when using relatively short recall periods, such as 24-hour recall questions, which is a much shorter recall period than that of postharvest questions typically used by agricultural economics. This suggests that much of the agriculture and development economics work using time-use data should be revisited, as the underlying data may be biased. This includes work that uses time-use data to calculate labor productivity, to explore power relations and asymmetries between gender, to model the decision making of farmers, or to obtain a more nuanced understanding of poverty and wellbeing.

Gender roles in agricultural development

This study makes an contribution to the literature on time-use and gender roles in smallholder farming households in developing countries. While most studies as part of this literature focus on adult males and females, this thesis provides additional insights into gender differences among children. This study confirms previous literature highlighting strong gender patterns with regard to domestic work and child care activities and helps make female work more visible (Arora, 2015). Interestingly, this thesis finds, for example, that while women spend much more time on domestic work than men, boys and girls spend similar amounts of time on such activities. This study shows that some farming time-use activities are gendered, for example, land preparation. However, when stratifying households by mechanization type, this gender differentiation seems to emerge only with mechanization and cannot be detected for households relying on manual labor. No clear gender patterns were found for farming activities, such as weeding and harvesting. This echoes the findings from Doss et al. (2001) and Palacios-Lopez et al. (2017), who questioned stylized facts on the gender divisions in agriculture.

This thesis also assessed the effects of agricultural mechanization on gender roles and the intrahousehold allocation of time, which has been a concern in the literature (e.g., Doss, 2001; Fischer et al., 2018). However, with agriculture being less gendered then assumed, the results suggest that both men and women benefit from agricultural mechanization with regard to time-use. In fact, women benefit relatively more than men. Land preparation becomes a male-dominated activity. Thus, agricultural transformation seems to change gender patterns. It remains unclear from the data collected whether this is a sign of empowerment or dis-empowerment, as once they are not working in the fields anymore, women may have less influence on farming practices, including decisions about how farm income is used. This resonates the findings from Alesina et al. (2011), who argue that traditionally, plow-based societies were less dependent on female work power than hand hoe-based societies and thus have lower rates of female participation in work and society. Looking at which activities the time "saved" from mechanized land preparation is used for does not provide a clear answer to the questions whether the shorter amount of time spent on land preparation by females in mechanized households is a sign either of empowerment or dis-empowerment. There is some (weak) evidence that women in animal-drawn-power-using households spent more time on domestic chores than women in manual labor-using households, which may be a sign of disempowerment. In tractor-using households, the extra time available seems to be used for off-farm work by women and domestic work (such as household chores and care for children) by men, out of which the latter is compared to a very small information base; however, this may be a sign of empowerment in women. More qualitative data are needed to better interpret the time-use changes (see also the discussion on limitation of this thesis in section 5.2.).

Gender studies in agriculture often present narrow temporal snapshots, for example, on male and female activities. The findings of this thesis suggest that gender roles are changing with agrarian change. This thesis confirms some of the studies that look at gender patterns as evolving over time and not as static, some of which find dramatic changes in gender roles with agricultural development. For example, Tavenner et al. (2019) found that increasing commercialization resulted in a decline in female control over productive assets, while farm diversification had positive effects on female control. However, the study neglected time-use and substitution effects, for example, with regard to off-farm labor and income. Lambrecht et al. (2018), who looked at gender roles over time in Ghana, equally neglected time-use aspects, mainly given the paucity of data on such aspects. Neither of the two studies focused on children. This thesis did but find weak gender patterns for farm work by children. This study shows no significant time benefits for children related to agricultural mechanization. This contradicts a previous study with data from across Zambia (and not only the Eastern Province), which showed that children do benefit largely from agricultural mechanization (Adu-Baffour et al., 2018).

5.2. Limitations and remedies

This study was the first to use a smartphone application to collect data from smallholder farming families within the field of agricultural economics. In addition to making a methodological contribution, any such new endeavors are likely to witness limitations, and this study is no exception. The limitations of this study can be grouped around four main areas: lack of validation and extrapolation; missing information on intensity of efforts and meaning; challenges in recording secondary activities, short activities and small food groups; and limitations due to a small sample size and cross-sectional data.

Lack of validation

Chapter 3 points to a high data quality due to the careful entering of data by respondents. This may not always be the case. This thesis does assess the accuracy of the data by comparing the data collected with the smartphone application with those of 24-hour recall questions, an approach that is then used to compare the differences in the two methods with previous studies on recall bias from the time-use literature. In addition, random cross checks were used throughout the data collection phase to validate the data collected. However, more attention could have been paid to data validation. Future studies may test different forms of data validation. Using fitness trackers would allow to validate whether recorded time-use activities fit with the recorded physical activity levels. Using smartphone position sensors would equally allow data validation. For example, researchers could elicit whether respondents are truly cycling (which would be shown by a moving GPS signal). Validation could be used ex post during data cleaning but also during data collection. If a mismatch between the recorded and actual activity seems likely, popup windows could open that ask the respondent whether the recorded activity is correct. Cameras could also be used to validate data; for example, respondents may be allowed to take pictures of their food to validate the nutrition data. It is important to keep in mind the tradeoffs with regard to battery life when using the GPS functions or cameras of smartphones. In addition to using technical means to validate data, findings from the behavioral economics literature may be used to “nudge” respondents toward truthful reporting. For example, it has been shown that people filling out their tax return forms to sign at the beginning rather than at the end decreases dishonest self-reporting (Shu et al., 2012).

These forms of data validation can be detected when respondents perform some activities but record others. However, respondents may change their behaviors by knowing that they are recording data. For example, they may work more or drink less alcohol. This may be conscious or subconscious. In both cases, it leads to a behavior that does not reflect the normal behavior of the respondents. This effect has been termed the Hawthorne effect. The Hawthorne effect was first observed in the Western Electric Company Hawthorne Works during the early 1930s as part of a research project on how lighting affects labor productivity. Surprisingly, the researchers at that time found that both more and less light improve labor productivity. Eventually, they concluded that changes in productivity are likely not because of the changing light intensities but because the workers knew that they were under observation. The existence and magnitude of the Hawthorne effect is much debated (Friedman & Gokul, 2014) and has not been studied with regard to time-use data, as having a control group that is observed without its knowledge is challenging. Nevertheless, future studies may try to create such a control group. Importantly, the Hawthorne effect may diminish over time as respondents become used to data recording (or become tired or forget to alter their behaviors). One could compare data recorded at the beginning of data collection with data reported at the end of data collection to explore whether there are significant changes, for example, with regard to the reported work time or time spent on social activities.

Extrapolation

In the introduction chapter of this thesis, it was criticized that most studies calculating labor productivity do this per worker (e.g., yield or value per worker) or per days worked (e.g., yield or value per day worked) without taking into account how much farmers actually work on specific days. The data collected here provide a more accurate picture of time-use on specific days. Additionally, data were recorded for three-day periods at five different times during the farming season to capture seasonality. This allows us to see the exact amount of work on these days, which, for example, enables comparisons of time use by different household members and different types of households (for example, differently mechanized households). However, recording data only at specific points in time makes the extrapolation of data over the entire farming season challenging. This can be problematic when the timing of activities differs for different household members and different types of households. For example, a study participant may do land preparation for 100 minutes per day for a period of 20 days (a total of 2000 minutes), whereas another participant may spend 50 minutes per day for 50 days (a total of 2,500 minutes). Choosing only three days for data recording may lead to the false impressions that the second person spends less time on land preparation (which is true per day but not per season).

Future researchers interested in the labor contribution per season thus need to find ways to extrapolate their data. This could be done by using a combination of Timetracker data and survey data or by participatory extrapolation with the respondents themselves. Alternatively, data may be recorded throughout the entire season. Studies that aim to collect data over extended periods must consider potential respondent fatigue. While Glorieux and Minnen (2009) find no proofs that

7-day time-use diaries are leading to less reliable data than 2-day diaries but time periods much longer than seven days may well face such challenges. One solution to avoid respondent fatigue may be to make the app usage more attractive by working with game-design elements, allowing participants to collect airtime (phone credit) or provide extension information through the app that is valuable to the respondents. However, in this regard, there is a danger that some of such elements might alter the behaviors of respondents.

Missing information on intensity of efforts and meaning

One limitation of this study and the developed Timetracker application is that they do not capture the intensity of efforts: a farmer who is hoeing his or her land with much rigor and one who is hoeing with little effort are treated the same. This shortcoming could be addressed by combining the use of the Timetracker with fitness trackers, which measure physical activities (see also Zanello et al., 2017). Other types of burdens may be more difficult to assess. For example, supervision of children is often not so much burdensome because it requires attention "every minute" but because it requires attention "every other minute", a fragmentation of time-use that can be exhaustive.

In addition, the data do not allow us to make conclusions with regard to meaning or well-being, as the data do not provide any indication of whether activities are perceived as enjoyable or not, for example, whether taking care of children is perceived as work, as leisure or as a social activity, although child care may be perceived as each of them depending on the context. Given the lack of such contextualization, it is difficult to conclusively determine whether a reduction in time spent on domestic work and child care is good or bad. Another example, using the time spent on unpaid domestic work as a proxy for female empowerment, may be misleading: a woman who spends less time on domestic work but is forced to work in a textile factory on a low salary that she has to give to her husband may actually be worse off. Similarly, fetching water from a public well may be, depending on the society, much more than work but a rare opportunity for social exchange. In some societies, fetching water may be one of the few windows for women to leave the house.

Some additional information on wellbeing and agency could be collected with popup windows asking participants whether they decided to perform the respective activities themselves. This would give a better understanding of the agency over time-use that the respondents have. However, the questions may be difficult to understand by respondents, and the respondents may not be fully aware of the deeper societal power dynamics leading to different gender patterns. It would be easier to ask whether respondents enjoyed the activity, which could be done using simple icons (Fernee and Sonck, 2014). This would allow us to calculate the portion of daily time that respondents spend "happy". Again, such an approach is not without limitations, as the perception of happiness may be socially constructed: are we truly happy when we do certain activities, or have we only learned to be happy when doing such activities? For example, in patriarchal structures, women may have learned to be content doing domestic chores.

The lack of contextualization is problematic in the third paper. It was shown that women in mechanized households spent less time on agricultural activities. It remained unclear, however, whether this is a sign of empowerment or the opposite – whether it is signaling that women are losing control over the production process. In this case, the above mentioned pop-up windows on happiness during activities or agency over time-use may not provide enough background information. More qualitative data collection methods may be needed to adequately understand the power dynamics and implications of such trends.

Challenges recording secondary activities, short activities and small food groups

Adequately capturing secondary time-use activities has long occupied time-use researchers. For the developed world, there has been a perception that secondary activities, such as child care, tend to be overestimated (Bianchi et al., 2012; Juster, 2003). Lentz et al. (2018) showed that this cannot be generalized to developing countries. In Malawi, they found that, albeit using promoting techniques, child care work is "dramatically underreported" (p. 1). They argue that child care work is invisible to those carrying it out as a secondary activity. This study may face similar limitations, and thus secondary activities should equally be studied carefully. Importantly, only primary actives are considered in chapters 4 and 5.

One way to address the potential underreporting of secondary activities, such as some types of care, would be to use pop-up windows asking respondents with whom they did an activity (e.g., elders, children and infants). Mullan (2010), using data from the United Kingdom, argued that secondary care can be detected by such an approach. It may not be trivial, however, to clearly define the boundaries of persons involved in an activity in multigenerational households located in a rural village with several houses (and thus other adults and children) in close proximity. For example, one could argue that all adults (and even older children) in such a village constantly jointly supervise children, but with so many grandparents, relatives, friends and older children, who should this care activity be attributed to?

Relatedly, respondents may not record very short and urgent activities. Recording such activities may not be feasible or not convenient in some situations (for example, when a baby needs immediate help). Thus, activities such as child care may also be underreported in the Timetracker because it can involve multiple short and potentially urgent activities, which may not always have been recorded. With regard to the nutrition data, smaller and nonvisible food groups (such as oil and sugar) may have not always been recorded, especially by respondents who did not cook the food themselves. It is debatable whether striving to record even the shortest activities is a realistic goal or whether one should rely on qualitative data to assess such activities. For food and nutrition data, such smaller and nonvisible food groups may be highly relevant, especially when assessing micronutrient gaps. However, collecting such data may be feasible given additional training and raising awareness among the respondents.

Limitations due to small sample size and cross-sectional data

This study can be considered a proof-of-concept study. Therefore, the sample size is relatively small (2790 data days from 186 respondents from 62 households). This is a limitation, especially for the empirical section of this thesis. Another limitation is the cross-sectional nature of the collected data. This is a limitation in chapter 2, where two data collection methods are compared: Timetracker and 24-hour recall questions. Respondents are asked to recall questions for a time period during which they used the Timetracker and were already sensitized about time-use. In chapter 4, the time-use between differently mechanized households is compared. The dissimilarity between differently mechanized types of households may be due to differences that existed before some become mechanized. While the thesis controls for some other factors affecting time-use, this remains a form of selection bias (selection into the use of tractors) that cannot be adequately addressed with cross-sectional data and a small sample size. Future studies aiming to establish a causal relationship between different interventions and time-use should envision using longitudinal data, potentially using panel data or randomized control trials, as well as larger sample sizes, which would allow the use of propensity score matching methods (to address observable selection bias).

5.3. Research questions and answers

Before elaboration on the future potential of using smartphone applications for data collection (see section 5.4.), this subchapter revisits the three research questions formulated in the outline of this thesis and summarizes how they need to be answered.

First, this thesis aimed to explore and test whether smartphone applications can be used to collect data from rural households in developing countries, focusing on time-use and nutrition data. The answer to this question is yes. The results presented, mainly as part of the first paper, a proof-of-concept that smartphone applications, such as the Timetracker application developed here, can be used to collect data in rural areas of developing countries when certain preconditions, such as addressing the challenges of low and non-literate users, are fulfilled. Research participants were both willing and able to record data using the smartphone application given that the Timetracker application was picture-based, intuitively designed and carefully introduced to them.

Second, this thesis aimed to assess the accuracy of the data collected with the help of smartphone applications such as the Timetracker vis-à-vis data collection using 24-hour recall questionnaires. Again, the answer is positive. The results show, mainly drawing on the second paper, that smartphone application allowed us to obtain better quality data on time-use, notwithstanding some limitations. The collected food and nutrition data were not validated.

Third, this thesis aimed to explore whether having such accurate and detailed data allows the exploration of socioeconomic aspects of agricultural development and rural transformation that are otherwise difficult to study. In particular, this thesis aimed to assess whether the data collected can be used to understand the effects of agricultural mechanization on the intrahousehold allocation of time-use within smallholder farming households in Zambia. Again, the answer is a yes: having such data at the individual level and for a whole farming season does allow for a better understanding of such research questions.

Having answered the three research questions, the next chapter will reflect on the future potential of smartphone applications to collect data from agricultural households in developing countries.

5.4. Future potential of smartphone application for data collection

The previous sections of the discussion have highlighted the methodological and empirical contributions of this thesis, discussed some of the limitations faced (and the remedies to address them) and summarized the answers to the three research questions. This section goes beyond the status quo and deliberates on the future potential to use smartphone applications, such as the one developed as part of this study, for data collection in developing countries. The section highlights opportunities for real-time data collection, the flexible and open-ended nature of the Timetracker application, and the advantages of combining self-recorded data with sensor-recorded data, which may open interesting transdisciplinary research pathways. After highlighting these potentials, the author of this thesis takes a step back and reflects on the ethical and data safety implications of such future endeavors.

Thus far, the emphasized advantage of real-time data recording was the reduction in recall bias. In addition to this advantage, having detailed real-time data from different households and from different members of such households is a value in itself. Bell et al. (2016), who developed a text-based application to record socioeconomic data based on a microtask for micropayment model, argued that real-time recording helps to "transform data collection in rural areas, providing near-real-time windows into the development of markets, the spread of illnesses, or the diffusion of ideas and innovations." (p. 1). If data are not only recorded in real-time but also transferred in real-time by transmitting the data online as opposed to offline, as done in this thesis, this would offer additional opportunities. In particular, the continuous real-time availability of data may allow policymakers and development practitioners to better monitor new policies and development programs (Daum et al. 2018) and environmental hazards and risks (Paul et al., 2017). For example, real-time food and nutrition data would allow for timelier reactions to changing levels of malnutrition.

Future researchers should see the Timetracker application developed as a building kit, which can be adopted and modified depending on the context and research questions. For example, depending on the country, the illustrations representing the different time-use and nutrition categories can be readjusted. Additionally, the activities can be aggregated and disaggregated depending on the study focus. For example, while for this thesis, one livestock activity was used, other studies may want to disaggregate this activity into its sub-activities, such as herding, feeding, rearing, and milking. The Timetracker building kit can also be extended to collect different types of data. To date, the Timetracker has been used to collect primarily time-use and nutrition data. For nutrition, a plug-in was used that opened whenever the activity "eating and drinking" was terminated. Similarly, other plug-in windows could be designed. For example, respondents could provide details on agricultural input use while doing respective agricultural activities (for example, on the type of feed when feeding animals). Likewise, they could record the money spent when the activity "buying groceries" is terminated. Such consumption data are usually difficult to recall (Bell et al., 2019; Deaton 2013). In addition, the Timetracker app could allow participants to take photos. Taking photos can be used to collect additional data, such as nutrition data. Quinn et al. (2011) showed that camera phones could be used to study the spread of pests. With real-time recording and transmission of data, such information would allow policymakers and development practitioners to act more quickly and in a more targeted manner (as discussed in the previous section). When expanding the Timetracker application to allow the recording of additional types of data, researchers should carefully balance the value of the additional data with the risk of "overloading" the application, which can burden respondents and may alter their willingness and ability to record data. An "overloading" of the application may also alter their daily routines and behaviors and therefore affect the quality of the time-use data collected.

The Timetracker may also be linked with data from the motion, environmental and position sensors of smartphones or other external sensors. Such data could be used for validation (see section 5.2.), but sensors could also be used to collect additional data. For example, when studying migrations patterns, routes of pastoralists, land use dynamics or social networks, data from the positions sensors accelerometers of smartphones may be used (see also Minnen et al., 2014). Position sensor data can also be used to measure agricultural land sizes. Knowing plot sizes is key for calculating agricultural productivity, but the farmers themselves often do not exactly know the size of their farms (Carletto et al., 2013).

Combining time-use data with accelerometers (such as external fitness trackers) can help to validate data (see 5.2.) and provides additional data on physical activity levels, which is an increasingly important topic given high obesity rates in developed countries and the emerging double burden of nutrition in developing countries (Ng and Popkin, 2012; Popkin, 2001; Steyn and Mchiza, 2014). In this regard, detailed time-use data may be used to train artificial intelligence software (deep learning) to understand patterns in the data recorded with device-based methods such accelerometers. In the future, this may allow to obtain more contextual information on time-use activities only by looking at the data recorded with accelerometers. Combining time-use diaries and device-based methods such as accelerometers may therefore help to benefit from the potentials of both methods, namely that device-based such as fitness trackers methods are more accurate, while time-use diaries provide more information about the context of time-use (Deyaert et al., 2017).

Using data from the sensors of smartphones has the advantage that data are recorded en passant without burdening the respondents. Because this type of data recording requires little attention from respondents, it may also help to reduce the Hawthorne effect (as discussed in section 5.2.). Combining self-recorded data using smartphone applications with different types of sensor-recorded data, which can be smartphone-integrated or external, may open interesting future research pathways. Such pathways could be of a transdisciplinary nature, for example, by combining socioeconomic and agronomic disciplines. Eventually, such an approach may lead to a revival of the village study approach, which was popular in the 1960s and 1970s (Dasgupta, 1978; Lipton and Moore, 1972). In these studies, researchers from different domains, such as agricultural economics, sociology, geography and agronomy, studied the same village from different perspectives over an extended period of time. In the digital era, such a transdisciplinary approach may comprise building digital twins of farmers and villages, which can then be used to model different future scenarios (see also section 1.2). Tools such as the developed smartphone application may then be a key element. This would resonate with the growingly popular concepts of “citizen science”. As argued by Buytaert et al. (2016), citizen science in general involves the “participation of the general public (i.e., nonscientists) in the generation of new scientific knowledge” (p. 1). Citizen science can range from using “citizens as sensors” (crowd-sourcing) to collaborating with citizens not only for data collection but also for problem formulation and data analysis (collaborative science) (Paul et al., 2017). This thesis has remained at the first level of citizen science, using “citizens as sensors”, but future research, such as research on the revival of the village study approach discussed above, may be more collaborative depending on the buy-in of respondents. While this thesis had to lend smartphones to respondents, future researchers may then benefit from rapidly rising smartphone ownership rates even in rural areas of developing countries.

While the previous paragraphs stressed the possibilities of using smartphone applications for real-time data collection and the combination of self-recorded and sensor-recorded data, not everything that is feasible is also ethical. As the data collected that may be collected are personal and in some cases highly sensitive, ethical and data safety implications must be critically reflected as well. In particular, future studies envisioning work with sensors must be scrutinized for ethical downfalls. Besides standard good research practices, such as ensuring anonymity, special attention is required to ensure upfront and continuous consent to the data collection process. While this thesis collected and transferred data offline, future studies collecting and transferring data online must ensure data safety, possibly by encrypting the data.

In summary, this thesis showed that smartphone applications such as the Timetracker can be used to collect data in rural areas of developing countries, thereby not only improving existing data collection methods but also creating new and transdisciplinary future research opportunities. This holds promise for the field of applied agricultural economics and other disciplines studying rural areas in developing countries, such as researchers focusing on health, governance and education.

5.5. Concluding remarks

This study has highlighted the importance of accurate data for researchers, private actors and policymakers. It has been argued that socioeconomic data often suffer from poor quality due to recall bias and that both poor and lacking data can lead to misguided policy recommendations and actions - with adverse effects on more vulnerable population groups, such as women and children. This study set out to assess whether digital tools, such as smartphone applications, can be used to enhance the accuracy of socioeconomic data collection. For this, a smartphone application called Timetracker was presented, which allowed the collection of data on time-use and nutrition. Such data are key for several research strands and needed, for example, to explore the drivers and effects of technology adoption, to reveal gender-based power relations and asymmetries and for targeting policies and programs. The Timetracker is picture-based and allows real-time data recording, thereby minimizing selection and recall bias. The results suggest that using well-tailored smartphone applications for data collection can greatly enhance the accuracy of socioeconomic data. This study has used such data, some of which are difficult to collect with conventional data collection methods, to explore the effects of agricultural mechanization on the intrahousehold allocation of time-use. The results empathize the relevance of time-use data for designing agricultural development trajectories that are socially equal and sustainable. Overall, this thesis suggests that there is a large and still untapped potential to use smartphone apps to collect data self-recorded data in agricultural households in rural areas of low-income countries and to study complex agricultural systems.

5.6. References

Adu-Baffour, F., Daum, T., and Birner, R. 2019. Can small farms benefit from big companies' initiatives to promote mechanization in Africa? A case study from Zambia. Food Policy, 84, 133-145.

Alesina, A., Giuliano, P., and Nunn, N. 2013. On the origins of gender roles: Women and the plough. The Quarterly Journal of Economics, 128(2), 469-530.

Arora, D. 2015. Gender Differences in Time-Poverty in Rural Mozambique. Review of Social Economy, 73(2), 196-221.

Arthi, V., Beegle, K., de Weerdt, J., Palacios-López, A. 2018. Not your average job: Measuring farm labor in Tanzania. Journal of Development Economics, 130, 160–72.

Beegle, K., De Weerdt, J., Friedman, J., and Gibson, J. 2012. Methods of household consumption measurement through surveys: Experimental results from Tanzania. Journal of Development Economics, 98(1), 3-18.

Bell, A., Ward, P., Tamal, E. H., and Killilea, M. 2019. Assessing recall bias and measurement error in high-frequency social data collection for human-environment research. Population and Environment, 1-21.

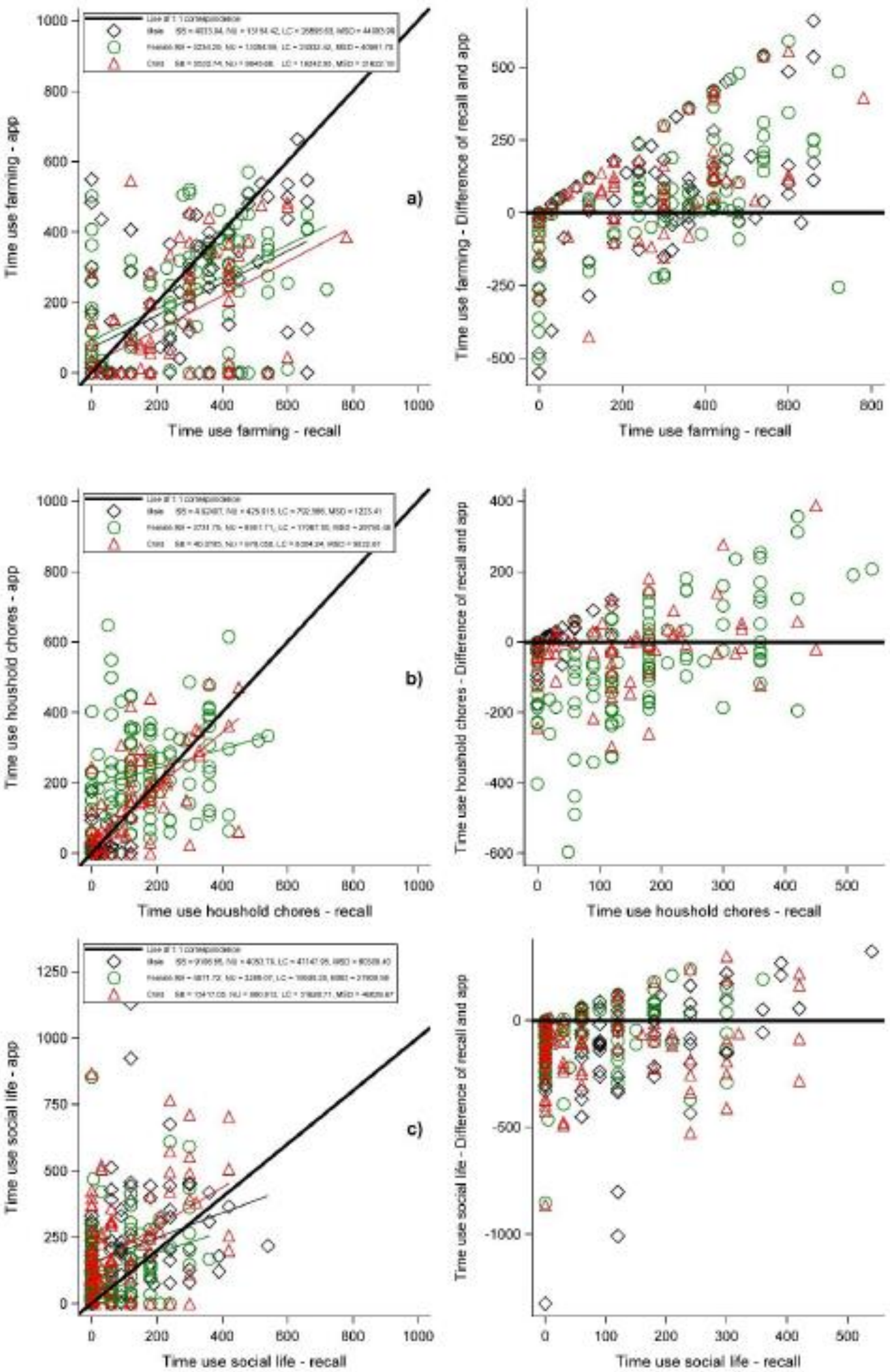
Bell, A., Ward, P., Killilea, M., Tamal, M., 2016. Real-time social data collection in rural Bangladesh via a 'Microtasks for Micropayments' platform on android smartphones. PLoS ONE 11 (11), e0165924.

Bianchi, S. M., Sayer, L. C., Milkie, M. A., Robinson, J. P. 2012. Housework: Who did, does or will do it, and how much does it matter? Social Forces, 91, 55–63.
Buvinic, B., King, E. 2018. Invisible No More? A Methodology and Policy Review of How Time-use Surveys Measure Unpaid Work. Data2X.
Buytaert, W., Dewulf, A., De Bièvre, B., Clark, J., and Hannah, D. M. 2016. Citizen science for water resources management: toward polycentric monitoring and governance? Journal of Water Resources Planning and Management, 142 (4).
Carletto, C., Savastano, S., Zezza, A., 2013. Fact or artifact: The impact of measurement errors on the farm size–productivity relationship. Journal of Development Economics, 103, 254–261.
Chatzitheochari, S., Fisher, K., Gilbert, E., Calderwood, L., Huskinson, T., Cleary, A., Gershuny, J. 2017. Using new technologies for time diary data collection: Instrument design and data quality findings from a mixed-mode pilot survey. Social Indicators Research, 137, 379–390.
Dasgupta, B. 1978. (Eds.). Village Studies in the Third World. Hindustan Publishing Corporation Printing Press, Delhi.
Daum, T., Birner, R., Buchwald H., Gerlicher, A. 2017. The Potential of Apps to Study Smallholder Farming Systems and More. Rural 21, 51 (4)
Deyaert, J., Harms, T., Weenas, D., Gershuny, J., and Glorieux, I. (2017). Attaching metabolic expenditures to standard occupational classification systems: perspectives from time-use research. BMC public health, 17(1), 620.
Deaton, A. 2013. The great escape: health, wealth, and the origins of inequality. Princeton University Press. Princeton.
Doss, C. R. 2001. Designing agricultural technology for African women farmers: Lessons from 25 years of experience. World Development, 29(12), 2075-2092.
Fernee, H., Sonck, N., 2014. Measuring smarter: time-use data collected by smartphones. Electronic International Journal of Time-use Research, 11 (1), 94–96.
Fischer, G., Wittich, S., Malima, G., Sikumba, G., Lukuyu, B., Ngunga, D., and Rugalabam, J. 2018. Gender and mechanization: Exploring the sustainability of mechanized forage chopping in Tanzania. Journal of Rural Studies, 64, 112-122.
Friedman, J., Gokul, B. 2014. Quantifying the Hawthorne Effect. http://blogs.worldbank.org/impactevaluations/quantifying-hawthorne-effect (retrieved on February 16 2019).
Glorieux, I., and Minnen, J. 2009. How many days? A comparison of the quality of time-use data from 2-day and 7-day diaries. Electronic International Journal of Time Use Research, 6(2), 314-327.
Hofferth, S. 1999. Family reading to young children: Social desirability and cultural biases in reporting. National Research Council Workshop on measurement and research on time-use. National Research Council. Washington, DC.
ILO (International Labor Organization). 2019. Child labor in agriculture. www.ilo.org/ipec/areas/Agriculture/lang--en/index.htm (retrieved on February 16th 2019).
Juster, F. T., Ono, H., Stafford, F. P. 2003. An assessment of alternative measures of time-use. Sociological Methodology, 33, 19–54.
Kelly, P., Thomas, E., Doherty, A., Harms, T., Burke, Ó., Gershuny, J., Foster, C., 2015. Developing a method to test the validity of 24-hour time-use diaries using wearable cameras: a feasibility pilot. PloS One 10 (12), e0142198.
Lambrecht, I., Schuster, M., Asare Samwini, S., and Pelleriaux, L. 2018. Changing gender roles in agriculture? Evidence from 20 years of data in Ghana. Agricultural Economics, 49(6), 691-710.
Lentz, E., Bezner Kerr, R., Patel, R., Dakishoni, L., and Lupafya, E. 2018. The Invisible Hand that Rocks the Cradle: On the Limits of Time-use Surveys. Development and Change, 50(2), 301-328.
Lipton, M. and Moore, M. 1972. The methodology of village studies in less developed countries. Institute of Development Studies. University of Sussex. Brighton.
Masuda, Y. J., Fortmann, L., Gugerty, M. K., Smith-Nilson, M., Cook, J. 2014. Pictorial approaches for measuring time-use in rural Ethiopia. Social Indicators Research, 115, 467–482.
Minnen, J., Glorieux, I., van Tienoven, T., Daniels, S., Weenas, D., Deyaert, J., Van den Bogaert, S., Rymenants, S.. 2014. Modular online time-use survey (MOTUS)—Translating an existing method in the 21st century. Electronic International Journal of Time-Use Research, 11, 73–93.
Mullan, Killian. 2010. "Valuing parental childcare in the United Kingdom." Feminist Economics, 16 (3), 113–139.
Ng, S. W., and Popkin, B. M. 2012. Time-use and physical activity: a shift away from movement across the globe. Obesity Reviews, 13(8), 659-680.
Palacios-Lopez, A., Christiaensen, L., and Kilic, T. 2017. How much of the labor in African agriculture is provided by women? Food Policy, 67, 52-63.
Paul, J. D., Buytaert, W., Allen, S., Ballesteros-Cánovas, J. A., Bhusal, J., Cieslik, K., Clark, J., Dugar, S., Hannah, D. M., Stoffel, M., Dewulf, A., Dhital, M. R., Liu, W., Lal Nayaval J., Neupane, B., Schiller,

A., Smith, P., Supper, R. 2018. Citizen science for hydrological risk reduction and resilience building. Water, 5(1), e1262.
Popkin, B. M. 2001. The nutrition transition and obesity in the developing world. The Journal of Nutrition, 131(3), 871-873.
Shim, J. S., Oh, K., and Kim, H. C. 2014. Dietary assessment methods in epidemiologic studies. Epidemiology and Health, 36.
Shu, L. L., Mazar, N., Gino, F., Ariely, D., and Bazerman, M. H. 2012. Signing at the beginning makes ethics salient and decreases dishonest self-reports in comparison to signing at the end. Proceedings of the National Academy of Sciences, 109(38), 15197-15200.
Steyn, N. P., and Mchiza, Z. J. 2014. Obesity and the nutrition transition in Sub-Saharan Africa. Annals of the New York Academy of Sciences, 1311(1), 88-101.
Tavenner, K., van Wijk, M., Fraval, S., Hammond, J., Baltenweck, I., Teufel, N., Kihoro, E., de Haan, N., van Etten, J., Steinke, J., Baines, D. Carpena, P., Skirrow, T., Rosenstock, T., Lamanna, C., Ng'endo, M., Chesterman, S., Namoi, N. Manda, L. 2019. Intensifying Inequality? Gendered Trends in Commercializing and Diversifying Smallholder Farming Systems in East Africa. Frontiers in Sustainable Food Systems, 3 (10).
Quinn, J., Leyton-Brown, K, Mwebaze, E., 2011. Modelling and monitoring crop disease in developing countries. In: Presented at the Twenty-Fifth AAAI Conference on Artificial Intelligence, San Francisco, 7–11 Aug 2011.
Zanello, G., Srinivasan, C. S., and Nkegbe, P. 2017. Piloting the use of accelerometry devices to capture energy expenditure in agricultural and rural livelihoods: Protocols and findings from northern Ghana. Development Engineering, 2, 114-131.

APPENDIX

Appendix 1. Mean squared standard deviation around the 1:1 line (left) and recall error by size of recall estimate (right).



Appendix 2. Time-use by gender and mechanization across seasons.

Activities and Season	Males			Females			Boys			Girls		
	M	ADP	T	M	ADP	T	M	ADP	T	M	ADP	T
Land Preparation												
Crop farming (land preparation)	146	115	64	120	54	16	48	57	44	45	51	26
Crop farming (others)	38	46	38	56	56	48	30	9	36	41	23	45
Rural livelihood	166	133	136	44	34	55	21	81	25	26	37	55
Off-farm work and seasonal labour	5	6	32	0	8	39	4	1	7	9	5	0
Transportation	142	186	186	108	83	89	113	137	140	152	154	124
Education	0	0	4	0	0	0	46	138	60	105	74	71
Domestic	26	11	29	234	349	262	182	114	170	151	144	162
Personal care	597	589	596	590	585	622	657	607	638	588	623	621
Leisure	312	343	347	276	258	298	330	285	308	309	319	326
Weeding												
Crop farming (weeding)	187	147	116	213	145	154	208	214	92	218	118	145
Crop farming (others)	46	82	86	50	75	43	38	57	61	47	81	52
Rural livelihood	51	77	91	18	6	30	37	9	29	0	16	22
Off-farm work and seasonal labour	62	15	2	14	13	33	4	3	1	20	35	15
Transportation	117	220	223	107	100	108	126	118	169	114	144	147
Education	0	0	0	0	0	0	1	10	0	0	4	3
Domestic	19	12	19	200	290	185	121	138	103	169	110	136
Personal care	605	574	609	613	590	637	634	642	675	622	625	651
Leisure	345	305	288	216	212	241	264	238	303	242	297	264
Harvesting/processing												
Crop farming (harvesting/processing)	204	150	163	208	197	198	186	218	64	68	95	136
Crop farming (others)	18	9	1	6	1	1	0	0	3	1	4	1
Rural livelihood activities	69	58	69	17	15	14	6	11	34	10	16	7
Off-farm work and seasonal labour	0	21	36	0	0	10	0	0	0	5	68	0
Transportation	164	272	159	86	134	67	111	157	146	124	87	89
Education	0	0	0	8	3	1	78	41	85	103	52	55
Domestic	19	17	29	246	278	235	169	78	116	201	157	172
Personal care	656	618	651	630	654	655	659	600	701	641	627	661
Leisure	305	288	326	232	198	252	225	328	287	281	280	313