THE CHOICE OF PAYMENT VEHICLE IN DISCRETE CHOICE EXPERIMENTS: LABOUR VERSUS MONEY

A thesis submitted to the University of Manchester
for the degree of Doctor of Philosophy
in the Faculty of Humanities

2021

CAMILLA KNUDSEN

School of Social Sciences

Department of Economics

LIST OF CONTENTS

WORD COUNT: 34,641

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

There is a growing literature advocating for the use of non-monetary payment vehicles (PVs) in stated preference (SP) studies. It is typically argued that the use of a monetary PV underestimates willingness to contribute in subsistence-oriented communities where households often lack access to waged labour opportunities. Many studies use labour as a non-monetary PV which leads to estimates of willingness to work (WTW) in place of, or alongside, estimates of willingness to pay (WTP).

This thesis comprises a series of chapters investigating issues around the use of labour as the PV. The empirical analysis is based on two discrete choice experiments – one concerns fortified flour in Kenya and one concerns water scarcity in India. Respondents in both studies were randomly assigned to treatments where the PV was either money or labour. In the Kenyan study, respondents were further randomly assigned to a treatment where choices were either hypothetical or consequential.

Chapter 1 uses data from the Kenyan field study to investigate issues regarding monetisation of WTW. To obtain a monetary measure of welfare, which is a requirement in cost benefit analysis, estimates of WTW are often monetised *ex post*. A fundamental challenge, however, is the need to apply an opportunity cost of time (i.e. a rate for converting labour to money). Most studies use some proportion of the wage rate as a proxy for the opportunity cost of time but there is no consensus about the most appropriate conversion rate to use. Estimates of WTP (based on a monetary PV) are often used as the benchmark against which the performance of one or more conversion rates is evaluated.

Previous (comparative) studies of PV effects are based on hypothetical SP studies. As a result, the benchmark against which monetised WTW is assessed is hypothetical WTP. The Kenyan study design allows a comparison of the performance of six commonly employed conversion rates using hypothetical WTP (which is standard in the literature) and consequential WTP (which is unique to this study) as the benchmark. The results in Chapter 1 indicate that the best performing rate when hypothetical WTP is used as the benchmark performs poorly when consequential WTP is used as the benchmark thus demonstrating the importance of the choice of benchmark.

In Chapter 2, data from the Kenyan field study is used to test for differences in hypothetical bias (HB) between a monetary and a labour PV. A recurrent finding in the WTW literature is that monetised WTW exceeds WTP. One possible explanation is higher levels of HB for labour PVs. Chapter 2 investigates this claim and finds that hypothetical bias is 26 to 31 percentage points higher when a labour PV is used instead of money.

Chapter 3 uses data from the Indian study to test for gender-based differences in the opportunity cost of time in the context of a patriarchal family structure. The results show that women, ceteris paribus, are WTW more than men while men are WTP more than women. The estimated shadow value of time is thus lower for women than for men which suggests that the choice of PV impacts not only welfare values in general but also the social importance attached to men and women in welfare valuation.

DECLARATION

I declare that no portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

COPYRIGHT STATEMENT

ACKNOWLEDGEMENTS

I would like to thank my supervisors Dan Rigby and Prasenjit Banerjee for their feedback and guidance throughout my PhD studies. A special thanks to Dan Rigby for his kindness and support during the writing-up period where I had to balance PhD and parenting responsibilities.

The support of Akansha Yadav and Nilamadhab (Nil) Digal was indispensable for the success of my field trips to India. Akansha and Nil ensured that my travels were both safe and enjoyable. I am also grateful for the input from Sudhansu Behera and the rest of the ICRG team at IPE Global. Finally, I extend a big thank you to the enumerators at SG Foundation for their hard work in the data collection process.

During my field trips to Kenya, I was fortunate to be waited on hand and foot by Efrancia Mobegi and Erick Nyairo. I owe them a massive thank you for making my trips safe and comfortable. I am also grateful for the high-quality work delivered by Benjamin Omwaga, Daniel Nyambane and Linnet Janganya during the data collection process.

Being away from home occasionally comes with challenges and homesickness. I want to thank Ditte, Isabelle, Maria, Mia and Nadia for the visits, the video calls and the virtual hugs. I am truly lucky to have friends like you.

My parents and my sister have been an incredible help and source of moral support. I want to thank my dad for showing me, by example, that acquiring new skills and knowledge is an enjoyable task and that hard work pays off. I want to thank my mum for worrying whether I eat well (and enough!) and for all the mailed gifts (I truly am spoiled). I want to thank my sister for being the kindest and smartest cookie that I know. She is my favourite sparring partner and I hope to be as supportive during her PhD studies as she has been through mine.

One very special person was key for the progress and completion of this thesis. I want to thank Tomkeen Onyambu Mobegi for having listened to my ranting and for having dealt with my (not so) occasional outbursts. He believed in me all through and never stopped encouraging me and cheering me on despite my capriciousness. I could have never wished for a better co-parent and partner in life.

Last but not least, I feel like I need to thank Kendrick Johannes Mobegi Onyambu. Although he had me sleep deprived for months, his smiles were always a reminder as to why I should never give up.

I would like to gratefully acknowledge the financial support received from the School of Social Sciences and from the (then) Department for International Development (now replaced by the Foreign, Commonwealth & Development Office).

DEDICATION

To my one and only, Tomkeen, for his love, support, smiles and pancakes

To my son, Kendrick, for being my favourite distraction

INTRODUCTION

SETTING THE SCENE

The lack of markets for nonmarket goods, including many environmental goods and services, means that their values are not revealed in market prices. To elicit information about the preferences that individuals hold for nonmarket goods, researchers have developed a series of nonmarket valuation techniques. The two main types of nonmarket valuation techniques are *revealed* and *stated* preferences methods. Researchers using revealed preference methods observe actual choices in real-life scenarios while researchers using stated preference methods rely on choices that individuals make in hypothetical scenarios. Revealed preferences methods avoid a range of potential biases, including hypothetical bias, but their use is limited because they can be applied only to individuals who use or purchase a particular good and to combinations of characteristics that are already available in the market.

To obtain estimates of the economic value of a nonmarket good or service, stated preference studies (e.g. contingent valuation and discrete choice experiments) commonly include a monetary payment vehicle. A monetary payment vehicle is required to calculate 'willingness to pay' which is the marginal rate of substitution between a non-monetary attribute and money. There is a growing literature, however, advocating for the use of non-monetary payment vehicles, of which the most popular is labour contributions, leading to estimates of 'willingness to work' in place of, or alongside, estimates of 'willingness to pay'. This has been primarily in relation to low-income communities in developing countries (e.g. Hagedoorn et al., 2020) but also includes high-income countries (e.g. Ando et al., 2020).

In a developing country context, the argument typically put forward is that monetary prices underestimate willingness to contribute when low-income groups are asked to state their willingness (and ability) to pay using money as the medium of exchange. If labour markets are such that households lack the opportunity to engage in waged labour, stated preference choices will be made in relation to these budget constraints. Low-income households may, therefore, be willing but not be able to contribute money while they are both willing and able to contribute in-kind or with labour. This argument is supported by a finding in several contingent valuation studies that labour contributions prompt fewer zero bids than monetary contributions (e.g. Asrat et al., 2004; Hung et al., 2007; Kamuanga et al., 2001). Most discrete choice experiments similarly find that participants are willing to contribute more labour than money when labour contributions are monetised *ex post* using the local wage rate (or a fraction thereof) as conversion rate (e.g. Abramson et al., 2011; Rai et al., 2015).

In addition to a lack of ability to provide cash payments due to labour market imperfections, low-income households may also lack experience in exchanging money for goods and services which is likely to cause increases in hypothetical bias. Gibson et al. (2016) argue that hypothetical markets involving monetary payments, such as those often utilised in stated preference surveys, are likely to be more unrealistic to poor households who are less integrated in labour, goods and credit markets.

To address the issues of using a monetary payment vehicle, researchers have adopted a range of non-monetary payment vehicles including crops such as rice and maize (Shyamsundar and Kramer, 1996; Sutton et al., 2008), everyday household items (Hossack and An, 2015) and meals to labourers (Diafas et al., 2017). The most commonly used non-monetary payment vehicle, however, is time contributions. In many studies, time contributions involve volunteering time towards a project or programme that provides the good or service being valued but it can also be increases in housework time (Eom and Larson, 2006) or an unrelated unskilled work task (Gibson et al., 2016; Hoffmann, 2018). Table 1 presents an overview of stated preference studies using time contributions in addition to, or instead of, money as the payment vehicle.

Table 1: Papers using money and/or time as the payment vehicle

	Country	Good/Service	Money	Time
Abramson et al. (2011)	Zambia	Water services	X	X
Ahlheim et al. (2017)	Vietnam	Landslide protection	X	X
Ando et al. (2020)	USA	Stormwater management	X	X
Arbiol et al. (2013)	Philippines	Disease prevention		X
Asrat et al. (2004)	Ethiopia	Soil conservation	X	X
Casiwan-Launio et al. (2011)	Philippines	Fishery reserve protection	X	X
Davies et al. (2014)	United Kingdom	Long-term ecological data		X
Durán-Medraño et al. (2017)	Spain	Forest management		X
Echessah et al. (1997)	Kenya	Tsetse control	X	X
Eom and Larson (2006)	South Korea	Water quality	X	X
Gibson et al. (2016)	Cambodia	Drinking water	X	X

continued on next page

continued from previous page

	Country	Good/Service	Money	Time
Hagedoorn et al. (2020)	Vietnam	Ecosystem-based adaptation	X	X
Hardner (1996)	Ecuador	Water purification		X
Hung et al. (2007)	Vietnam	Forest fire prevention	X	X
Kamuanga et al. (2001)	Burkina Faso	Tsetse control	X	X
Kassahun et al. (2020)	Ethiopia	Irrigation	X	X
Lankia et al. (2014)	Finland	Recreational quality	X	X
Lee and Wang (2017)	Taiwan	Land use programme	X	X
Meginnis et al. (2020)	Uganda	Disease prevention	X	X
Navrud et al. (2012)	Vietnam	Flood prevention		X
Navrud and Vondolia (2020)	Ghana	Flood insurance	X	X
O'Garra (2009)	Fiji	Traditional fishing ground	X	X
Pokou et al. (2010)	Côte d'Ivoire	Tsetse control	X	X
Pondorfer and Rehdanz (2018)	Papua Ny Guinea	Evacuation route	X	X
Rai and Scarborough (2013)	Nepal	Invasive species mitigation	X	X
Rai and Scarborough (2015)	Nepal	Invasive species mitigation	X	X
Rai et al. (2015)	Nepal	Watershed services	X	X
Schiappacasse et al. (2013)	Chile	Forest restoration	X	X
Susilo et al. (2017)	Indonesia	Mangrove restoration		X
Swallow and Woudyalew (1994)	Ethiopia	Tsetse control	X	X
Tilahun et al. (2015)	Ethiopia	Forest conservation	X	X
Tilahun et al. (2017)	Ethiopia	Invasive species mitigation	X	X
Vasquez (2014)	Guatemala	Water services	X	X
Vondolia et al. (2014)	Ghana	Irrigation	X	X
Vondolia and Navrud (2018)	Ghana	Flood insurance	X	X

This thesis explores a number of hypotheses related to using labour as an alternative to money as the payment vehicle in stated preference studies. To this end, two discrete choice experiments were designed and administered – one in India and one in Kenya. Respondents in both studies were randomly assigned to treatments with either money or labour as the payment vehicle. In Kenya, respondents were further randomly assigned to a treatment where choices were either hypothetical or consequential.

While previous studies using labour contributions as the payment vehicle (see Table 1) are hypothetical in nature, the work in this thesis provides an original and significant contribution to the literature by examining the difference between hypothetical and consequential labour (and monetary) payments in a split-sample discrete choice experiment. Here, 'consequential' refers to the fact that respondents were asked to make *real* purchasing choices i.e. paying for the selected product, if any, using *actual* out-of-pocket money, in the case of a monetary payment vehicle, or by *actually* working (i.e. sorting seeds), in the case of a labour payment vehicle. The discrete choice experiment can be considered a lab-in-the-field experiment since behaviour is investigated in a more naturalistic environment (i.e. in the neighbourhood where respondents either live or work) rather than in a laboratory setting, whilst maintaining a high level of experimenter control by using a standardised survey design across treatments. While many of the studies in Table 1 include similar fieldwork components, the discrete choice experiment developed as part of this thesis is innovative in that it includes a treatment in which choices are consequential.

A consequential discrete choice experiment design poses some practical problems since the selected products will have to be made available to respondents after the experiment. To ensure that all product combinations can be provided, the type of good used for this study is a private good (a bag of flour) which could be either purchased or produced by the researchers. Previous studies using labour payment vehicles often concern public goods, club goods or common goods (see Table 1) and an interesting avenue for future research is, therefore, to examine the extent to which the differences in hypothetical and consequential willingness to pay and willingness to work estimates can be generalised to other types of goods.

THESIS OVERVIEW

Chapter 1 investigates issues around monetisation of estimates of willingness to work. To enable assessment of welfare effects (e.g. by comparing benefits and costs), many studies monetise willingness to work estimates *ex post*. Most studies use a proportion of area average or sample average wage rates (typically one third) to convert labour into money but there is no consensus about the most appropriate conversion rate to use. To evaluate the performance of one or more conversion rates, a common approach in the literature is to compare estimates of monetised willingness to work to estimates of willingness to pay. Since previous labour payment vehicle studies are hypothetical stated preference studies, the benchmark against which monetised willingness to work is assessed, thus far in the literature, is hypothetical willingness to pay.

Chapter 1 uses data from the Kenyan field study to generate estimates of not only hypothetical willingness to pay and hypothetical willingness to work but also consequential willingness to pay and consequential willingness to work. Six conversion rates (based on wage rate data) that are commonly employed in the literature are used to generate monetised estimates of (hypothetical and consequential) willingness to work. The performance of the six conversion rates is then evaluated using either hypothetical willingness to pay (as is standard in the literature) or consequential willingness to pay (which is unique to this study) as the benchmark.

The results show that there is no statistically significant difference between hypothetical willingness to work when it is monetised, using one third of the national average wage rate, and hypothetical willingness to pay. This result provides support for the use of one third of the wage rate (which can be traced back to Cesario, 1976) as the opportunity cost of time. Chapter 1 argues, however, that the relevant benchmark is consequential willingness to pay, rather than hypothetical willingness to pay, since consequential willingness to pay, in theory, is unaffected by hypothetical bias. When consequential willingness to pay is used as the benchmark, the value of monetised willingness to work, when it is monetised using one third of the national average wage rate, exceeds willingness to pay by a factor of almost 2 and one third of the national average wage rate ranks only fourth of the six conversion rates employed. The results thus cast doubt on the use of hypothetical willingness to pay as a suitable benchmark.

Chapter 2 examines the effect of the payment vehicle (money versus labour) on hypothetical bias. The literature provides competing claims regarding the impact of a non-monetary payment vehicle on hypothetical bias. Gibson et al. (2016) hypothesise that contingent monetary markets may seem less realistic if households have limited experience with money as the medium of exchange which can amplify hypothetical bias. The use of labour as the payment vehicle in subsistence economies, where transactions occur primarily through barter or work exchange, might thus reduce hypothetical bias. In contrast, Ando et al. (2020) speculate that the recurrent finding that monetised willingness to work exceeds willingness to pay is caused by a labour payment vehicle exhibiting higher levels of hypothetical bias than a monetary payment vehicle.

To evaluate the competing claims regarding the impact of the payment vehicle on hypothetical bias, Chapter 2 uses data from the Kenyan field study to generate and compare four sets of estimates: hypothetical willingness to pay, consequential willingness to pay, hypothetical willingness to work and consequential willingness to work. Hypothetical bias is then estimated for each of the two payment vehicles as the ratio of hypothetical to consequential willingness to pay/work. The results show that hypothetical bias is 26 to 31 percentage points higher when respondents are asked to pay with labour instead of money thus providing evidence supporting the claim that the use of a labour payment vehicle increases hypothetical bias.

Chapter 3 uses data from the Indian field study to evaluate willingness to pay and willingness to work as measures of welfare from a gendered perspective. In cost-benefit analysis, monetary estimates of benefits are typically estimated as the sum of individuals' stated willingness to pay. It can be argued, however, that the sum of willingness to pay only measures the sum of social benefits in cases where money is equally important to all individuals on the margin (see e.g. Fleurbaey and Abi-Rafeh, 2016; Nyborg, 2014). In all other cases, more weight will be given to preferences of individuals with a lower marginal utility of money (typically more wealthy individuals).

Due to the patriarchal family structure of many households in rural India, women have limited financial independence. In this context, it is hypothesised that women have a higher marginal utility of money, which likely translates into a lower opportunity cost of time, compared to men. To investigate this issue, estimates of willingness to pay and willingness to work are calculated separately for male and female respondents. The results show that women are willing to work more than men, ceteris paribus, while men are willing to pay more than women. This finding suggests that money is more important to women on the margin while time is more important to men on the margin. Neither of the two payment vehicles, therefore, provide an accurate measure of the sum of social benefits. More weight will be given to preferences of men if willingness to pay is used as the measure of welfare and more weight will be given to preferences of women if willingness to work is used as the measure of welfare. These findings suggest that the choice of payment vehicle is non-trivial and necessarily based on value judgments with respect to the social importance of men and women.

The objective of this thesis is thus twofold: (1) to provide support to decision makers in developing country settings who use cost-benefit analysis to evaluate development projects, programmes or policies; and (2) to expand methodological knowledge on the choice of payment vehicle in stated preference studies. The first objective is aimed at policymakers and decision takers while the second objective is aimed, primarily, at environmental economists. In addition to the two main objectives, Chapters 2 and 3 also concern the substantive issues of fortified flour as a means to reduce malnutrition in Kenya (Chapter 2) and improved water infrastructure in rural India (Chapter 3).

Chapter 2 investigates the demand for fortified flour of low-income communities in rural and urban Kenya. Malnutrition is a pressing issue in Kenya where 24% of the population suffer from undernourishment. Industrial fortification of foodstuffs has been suggested as a means of tackling malnutrition and many donor schemes are focused on increasing the capacity to fortify. The results in Chapter 2 demonstrate high demand for fortified flour thus encouraging further investments.

Chapter 3 investigates the demand for improved water supply via a discrete choice experiment where respondents are asked to trade off different uses of water. The study is part of 'Infrastructure for Climate Resilient Growth' which is a 43-month programme funded by the UK's Government Department for International Development (now replaced by Foreign, Commonwealth & Development Office). The programme aims to improve the design and implementation of natural resource management works under MGNREGA (India's flagship national rural employment scheme).

It is estimated that more than 132 million people in rural India lack access to basic water services yet there are more works related to e.g. rural connectivity completed under MGNREGA than works related to water conservation. The results in Chapter 3 demonstrate high demand for improved water services thus encouraging decision makers to reconfigure MGNREGA to implement more water-related projects.

To summarise, the objective of this thesis is twofold: (1) to provide support to decision makers in developing country settings who use cost-benefit analysis to evaluate development projects, programmes or policies; and (2) to expand methodological knowledge on the choice of payment vehicle in stated preference studies. The first objective is aimed at policymakers and decision takers while the second objective is aimed, primarily, at environmental economists. Both objectives concern the relative merits of using labour as an alternative to money as the payment vehicle in stated preference surveys but with a focus on either the impact of the choice of payment vehicle on decision strategies or on methodological issues.

ETHICAL APPROVAL

The projects were approved under University of Manchester ethical application 5243 (Kenya) and 2637 (India).

THESIS STRUCTURE

This thesis is presented in journal format. Chapters 1 to 3 are written in the format that they are either submitted (Chapter 2) or intended to be submitted (Chapter 1 and Chapter 3) for publishing in peer-reviewed journals. For ease of readability, the pagination sequence has been adapted to flow throughout the thesis and tables and figures are numbered sequentially.

AUTHORSHIP

Chapters 1 and 2 are co-authored with Dan Rigby. Chapter 3 is co-authored with Akansha Yadav, Dan Rigby and Prasenjit Banerjee (co-authors are presented in alphabetical order). Camilla Knudsen is the lead author on all three chapters.

All chapters and sections in this thesis are written by Camilla Knudsen with comments and suggestions by co-authors. The surveys are designed by Camilla Knudsen with input from co-authors. Data collection in Kenya was organised and managed by Camilla Knudsen with help from Benjamin Omwaga and Linnet Janganya who acted as enumerators and Daniel Nyambane, Erick Nyairo and Tomkeen Mobegi who provided practical and logistical support. Data collection in India was organised and managed by Camilla Knudsen and Akansha Yadav with help from enumerators from Saunta Gaunta Foundation and the ICRG team at IPE Global Limited, including Nilamadhab Digal and Sudhansu Behera, who provided practical and logistical support. Data analysis and econometric modelling was conducted by Camilla Knudsen with comments and suggestions by co-authors.

CHAPTER 1

THE MONETARY VALUE OF WILLINGNESS TO WORK

1.1 ABSTRACT

Labour payment vehicles are increasingly used in stated preference studies to estimate welfare values in developing countries. These studies yield willingness to work (WTW) measures of welfare which are often monetised *ex post*. There is no consensus about the appropriate conversion rate (i.e. value of time) to use in the monetisation. Many studies employ split-sample designs using both money and labour payment vehicles to allow comparison of willingness to pay (WTP) and monetised WTW. Estimates of WTP then act as the benchmark against which the performance of (one or more) conversion rates can be evaluated.

This paper uses a discrete choice experiment to investigate the performance of a set of conversion rates typically used to monetise WTW, where performance is assessed in terms of the closeness of the value of monetised WTW to WTP. Unlike previous studies, we perform this exercise using both hypothetical and consequential choice data. Respondents are randomly assigned to one of four treatments where the payment vehicle is either money or labour and choices are either hypothetical or consequential. Six conversion rates that are commonly employed in the literature are then applied *ex post* to generate monetised estimates of WTW. This allows us to compare the performance of the conversion rates using first hypothetical WTP (which is standard in the literature) and second consequential WTP (which is unique to our study) as the benchmark. We argue that consequential WTP is the relevant benchmark since this value is free of hypothetical bias. Our results indicate that the best performing rate when hypothetical WTP is used as the benchmark performs rather poorly (by ranking 4th) when consequential WTP is used as the benchmark thus casting doubt on the use of hypothetical WTP as the benchmark.

1.2 INTRODUCTION

There is a growing literature advocating for the use of labour as the payment vehicle (PV) in stated preference (SP) studies. The argument commonly put forward is that monetary prices underestimate willingness to contribute when low-income groups are asked about their willingness (and ability) to pay. To limit hypothetical bias, SP studies often ask respondents to consider their budget constraints before responding to discrete choice experiment (DCE) or contingent valuation (CV) type questions. In circumstances where households lack opportunity to engage in waged labour, responses will, therefore, be made in relation to these budget constraints (Gibson et al., 2016). Consequently, money-constrained households may be willing but unable to contribute money while they are both willing and able to contribute labour. This argument is supported by a finding in several CV studies that labour contributions lower the number of zero bids compared to monetary contributions (e.g. Asrat et al., 2004; Brouwer et al., 2009; Echessah et al., 1997; Hung et al., 2007; Kamuanga et al., 2001; Swallow and Woudyalew, 1994).

To enable assessment of welfare effects, estimates of willingness to work (WTW) are often converted into their monetary equivalent *ex post.* Such task, however, is complicated by the need to apply an opportunity cost of time. Many studies use a conversion rate that is based on some constant fraction (most commonly one third) of the wage rate (see Section 1.3). Czajkowski et al. (2019) and Lloyd-Smith et al. (2019) find, however, only weak correlation between individuals' value of time (estimated using SP methods) and self-reported wage rates. Their results cast doubt on the use of wage rate data as a proxy for the value of time.

Despite the concerns with respect to identifying an opportunity cost of time using wage rate data, it remains the most common practice for monetising WTW. With a few exceptions (e.g. Navrud et al., 2012), most studies produce both willingness to pay (WTP) and WTW estimates which enables them to compare WTP and monetised WTW. WTP is then the benchmark against which the performance of one or more conversion rates can be evaluated. Meginnis et al. (2020) find, for example, that "*using the market wage rate, we find higher levels of willingness to contribute time than money*" and conclude that "*this suggests that using the market wage rate is not an appropriate means to translate labour hours into monetary values*" (Meginnis et al., 2020, p. 9).

In this paper, we investigate the issues around monetisation of WTW in a split-sample discrete choice experiment (DCE) concerning fortified flour in Kenya. Respondents were randomly assigned to one of four treatments where the PV was either money or labour and purchase choices were either hypothetical or consequential. This study design allows us to use a set of conversion rates that are commonly employed in the WTW literature to generate monetised values of not only hypothetical WTW (as is standard in the literature) but also consequential WTW (which, we believe, is unique to this study). We can furthermore compute both hypothetical and consequential estimates of WTP.

The objective of this paper is to examine the reliability of inferences drawn from hypothetical data (i.e. where the benchmark for assessing the performance of competing conversion rates is hypothetical WTP) by making comparisons to inferences drawn from consequential data (i.e. where the benchmark for assessing the performance of competing conversion rates is consequential WTP). To this end, we use six conversion rates (based on wage rate data) that are commonly employed in the literature to monetise hypothetical WTW. We find that the best performing conversion rate, when the benchmark is hypothetical WTP, is one third of the national average wage rate. When hypothetical WTW is monetised using one third of the national average wage, we find that there is no significant difference between monetised WTW and hypothetical WTP.

In contrast to Czajkowski et al. (2019) and Lloyd-Smith et al. (2019), whose findings call into question the income-based approach to valuing time, the above result provides support for the use of one third of the wage rate as a measure of the opportunity cost of time. We find, however, that model fit is reduced when individual- or site- specific wage rates, instead of generic wage rates, are employed to monetise WTW. This is consistent with Czajkowski et al. (2019) and Lloyd-Smith et al. (2019) who find that individuals value their time (a lot) differently than implied by their wage rates.

When the benchmark is consequential WTP, instead of hypothetical WTP, we find that hypothetical WTW, when monetised using one third of the national average wage, exceeds WTP by 92% and that it ranks only fourth of the six conversion rates employed.

Since inferences drawn from hypothetical choice data do not translate to inferences drawn from consequential choice data, our findings cast doubt on the use of hypothetical WTP as the benchmark for evaluating the performance (and viability) of conversion rates. The best performing conversion rate, when the benchmark is consequential WTP, is the sample average wage rate. Despite being top-ranked, however, the sample average wage rate generates monetised WTW values that are 31% higher than WTP. This is, however, of a much lower magnitude than other WTW papers who find that monetised WTW is up to 20 times larger than WTP (Hagedoorn et al., 2020).

An alternative approach to monetising WTW assumes that welfare can be estimated using either of the two PVs (Eom and Larson, 2006). The implicit assumption is that the underlying preference structure is the same for both PVs (i.e. that both labour and money are suitable PVs) and that willingness to contribute more (less) labour relative to money is due to a low (high) value of time. Following such approach, joint estimation of the marginal utilities of labour and money can be utilised to elicit the shadow value of time. Using our unique dataset, we are able to estimate and compare four values of the shadow value of time which vary according to the nature of the choice tasks: (1) both labour and monetary payments are consequential, (2) both labour and monetary payments are hypothetical, (3) labour "payments" are consequential but monetary payments are hypothetical, and (4) labour "payments" are hypothetical but monetary payments are consequential. The latter is particularly interesting because it is essentially the rate at which WTW can be converted to (HB-deflated) WTP.

Both of the abovementioned approaches (*ex post* conversion and joint estimation) are complicated by potential differences in hypothetical bias (HB) between the two PVs. If HB is more (less) of a problem when labour is used as an alternative to money, the monetised value of WTW will be biased upwards (downwards) and the associated shadow value of time will be biased downwards (upwards). A final observation, therefore, relates to differences in the scale of hypothetical bias (HB) between the two PVs. We find that the ratio of monetised WTW and WTP is higher in the hypothetical treatments (i.e. when both WTW and WTP are hypothetical) compared to the consequential treatments (i.e. when both WTW and WTP are consequential). This can be attributed to differences in HB between the two PVs. While we find that HB affects both PVs, HB is significantly higher when labour is used as the PV which suggests that the recurrent finding that monetised WTW exceeds WTP (see e.g. Abramson et al., 2011) is (at least partially) due to this HB-PV asymmetry.

In summary, this paper investigates the performance of six commonly employed conversion rates with respect to their ability to generate monetised WTW values that match WTP. First we use hypothetical WTP as the benchmark (this is the standard approach in the literature). We then use consequential WTP as the benchmark to assess the reliability of inferences drawn from the use of hypothetical WTP. We are further testing for differences in the scale of HB between money and labour PVs.

To the best of our knowledge, this is the first study to compare monetised WTW and WTP using data from a consequential DCE. Our results are an important addition to the WTW literature as we provide the first bit of evidence about the rate at which WTW derived from hypothetical choice tasks can be converted to HB-deflated WTP.

The remainder of this paper is structured as follows. We present an overview of commonly employed conversion rates in Section 1.3. In Section 1.4, we provide an introduction to the dataset and a description of the conversion rates that we employ. The conceptual framework is outlined in Section 1.5 and the results of the empirical analysis are presented in Section 1.6. In Section 1.7, we discuss the implications of the results and conclude.

1.3 COMMONLY EMPLOYED CONVERSION RATES

To enable assessment of welfare effects, many SP studies monetise labour "payments" *ex post*. Such conversion, however, is complicated by the need to apply an opportunity cost of time. Most studies use either individual specific, area average or sample average wage rates to convert labour into money (thus assuming that wage rates are a proxy for the value of time) but there is no consensus in the literature about the most appropriate conversion rate to use (see Palmquist et al., 2010 for a review of theoretical and empirical time valuation studies). Table 1.1 provides an overview of the most commonly applied conversion rates along with a list of SP studies in which they have been employed. The list of SP studies includes both CV and DCE type questions and between- and within- subjects designs. Of the papers using within-subjects designs, money and labour contributions are either treated as substitutes (e.g. Meginnis et al., 2020) or as complements (e.g. Kassahun et al., 2020).

Table 1.1: Overview of conversion rates employed in the literature

Conversion rate	
Area average wage rate (based on secondary data)	Abramson et al. (2011)
	Gibson et al. (2016)
	Hagedoorn et al. (2020)
	Hoffmann (2018)
	Meginnis et al. (2020)
	Navrud et al. (2012)
	Navrud and Vondolia (2020)
	O'Garra (2009)
	Rai and Scarborough (2015)
	Rai et al. (2015)
	Schiappacasse et al. (2013)
	Tilahun et al. (2015)
	Tilahun et al. (2017)
	Vasquez (2014)
	Vondolia et al. (2014)
1/3 of area average wage rate (based on secondary data)	O'Garra (2009)
	Schiappacasse et al. (2013)
Minimum wage rate	Vondolia et al. (2014)
	Susilo et al. (2017)
	Tilahun et al. (2015)
	Tilahun et al. (2017)
Sample average wage rate	Ando et al. (2020)
	Lankia et al. (2014)
1/3 of sample average wage rate	Ando et al. (2020)
Average expected wage rate	Arbiol et al. (2013)
	Susilo et al. (2017)
1/3 of average expected wage rate	Arbiol et al. (2013)
Average income (from livelihood activities)	Tilahun et al. (2015)
	Tilahun et al. (2017)
1/3 of average income (from livelihood activities)	Casiwan-Launio et al. (2011)
Individual-specific wage rate	Hagedoorn et al. (2020)
1/3 of individual-specific wage rate	Hagedoorn et al. (2020)

As shown in Table 1.1, the most popular type of conversion rate is area average wage rates from secondary data sources. Within this category, studies have used both average wage rates at the local (village or city) level (e.g. Hagedoorn et al., 2020), at province level (e.g. Gibson et al., 2016), and at national level (e.g. O'Garra, 2009). Other types of conversion rates are minimum wage rates, based on legislative acts, and sample average wage rates. A couple of studies have also used the average expected wage rate by asking respondents to state the wage rate that they would expect to get paid if they were to get paid for the contributed labour time. Instead of using wage rate data, a few studies have used information about income from livelihood activities. This is relevant in subsistence-oriented communities where households do not engage in waged labour activities.

It is often hypothesised that individuals substitute leisure time, rather than time already spent in work, when responding to SP type questions with a labour PV. To this end, many studies use a fraction (most commonly one third) of the wage rate as the opportunity cost of leisure time in addition to or instead of the full wage rate (see Table 1.1). Using one third of the wage rate to value leisure time is a result which can be traced back to Cesario (1976). In a recent study, Whittington and Cook (2019) recommend instead the use of 50% of after-tax wages as the opportunity cost of time in low- and middle-income countries but suggest that a sensitivity analysis is undertaken to ensure that the conclusion of a cost-benefit analysis does not change for value of time estimates between 25% and 75% of the after-tax wage rate.

Individuals are likely to value their time heterogeneously depending e.g. on employment status, wage rates and alternative uses of time (Ahlheim et al., 2017; Feather and Shaw, 1999 Lloyd-Smith et al., 2019). To incorporate some of this heterogeneity, Hagedoorn et al. (2020) use individual-specific wage rate data, instead of generic wage rates, to convert labour payments into money. Lloyd-Smith et al. (2019) find that respondents value their free time heterogeneously which is consistent with the use of an individual-specific approach to monetising labour contributions. They find, however, that respondents' opportunity cost of time is only weakly correlated with self-reported wage rates which casts doubt on the use of wage rates (whether generic or individual-specific) as a suitable proxy for the value of time. Bockstael et al. (1987) further find that wage rates are an inappropriate proxy when individuals work fixed hours since fixed work schedules hinder free substitution between work and leisure and thus between time and money.

1.4 THE DATASET

1.4.1 SURVEY DESCRIPTION

The empirical analysis is based on a unique dataset from a DCE field study where purchase choices are either hypothetical or consequential and the payment vehicle is either money or labour. Respondents were randomly assigned to one of the four treatments (hypothetical money, hypothetical labour, consequential money or consequential labour).

The DCE concerns (2 kg) bags of flour which are described by a price attribute (money or labour) and three non-price attributes: *plant*, *sifted* and *fortified*. The attributes *sifted* and *fortified* take one of two levels: sifted[1] or unsifted and fortified[2] or unfortified. The attribute *plant* indicates the type of plant from which the root or grains are used to produce the flour. The attribute takes one of four levels: 100% maize, 50% maize and 50% sorghum, 50% maize and 50% cassava or 50% maize and 50% millet.

We used Ngene software (ChoiceMetrics, 2018) to generate a D-efficient fractional factorial design based on priors derived from a pilot study conducted in January 2019. The Modified Federov algorithm was applied to generate 80 choice tasks which were split into 10 blocks. The choice tasks are identical across treatments except the payment vehicle which is either monetary (in Kenyan Shillings) or labour (specified as an amount of time that the respondents would commit to sorting seeds[3]). In each choice task, respondents are asked to select their preferred alternative amongst two different types of flour and an opt-out option.

Respondents were recruited from two low-income communities in Kenya – Kibera (an urban slum on the edge of Nairobi) and Nyamira/Kisii (neighboring counties in Western Kenya in which the majority of the population live in rural areas growing food crops for subsistence). The survey was conducted one-on-one in public places in March and April 2019 by trained enumerators who spoke the local languages (Swahili/Sheng and Kisii). In the consequential treatments, enumerators informed respondents (prior

[1] Unsifted maize flour is made from whole kernel maize while sifted maize flour has had the husk and the germ removed.
[2] Flour that is (industrially) fortified has been enriched with micronutrients (e.g. iron, zinc, folic acid, vitamin A and B vitamin) during the processing of crop.
[3] Seed sorting was successfully used as a non-monetary PV in Hoffmann, 2018.

to the choice tasks) that one of the choice tasks would be randomly selected (after the choice tasks) by rolling of an eight sided dice. The randomly selected choice task would be consequential i.e. respondents who selected one of the purchase options in the randomly selected consequential choice task would be expected to purchase the selected flour. Respondents who selected the no purchase option in the randomly selected choice task would not be purchasing any flour.

Respondents in the consequential treatment were further informed that payment (with money or labour) and collection of the flour would take place in a well-known venue in the local area in the days following the experiment. To keep the transaction costs consistent across respondents, participants in all treatments were entitled to a small gift which was collectable only at the same days/times/venues as the work/payment for the selected flour.

1.4.2 DATA SETUP

We generate four price interaction terms: money (in KSh) and labour (in hours) are each interacted with indicator variables for the hypothetical and the consequential treatments. Prices are described in terms of either money or labour and choices are either hypothetical or consequential. For each observation (i.e. choice task) in the dataset, three out of four price metrics are, therefore, coded as zero i.e. monetary (labour) prices when the payment vehicle is labour (money) and hypothetical (consequential) prices when the treatment is consequential (hypothetical). To keep the analysis simple, but without loss of generality[4], we merge the four levels for the attribute *plant* and define a new variable *flour* that acts as an alternative-specific constant (ASC) for the two purchase options. The non-price attributes are dummy-coded while the price attributes are treated as continuous variables. The coding scheme is demonstrated in Table A1.2 in Appendix A1.

To enable comparison of labour and monetary contributions, we generate two additional variables (not shown in Table A1.2) where hypothetical and consequential labour contributions are converted to money. We perform this exercise using six different conversion rates (see Section 1.4.3). Three of these are based on secondary

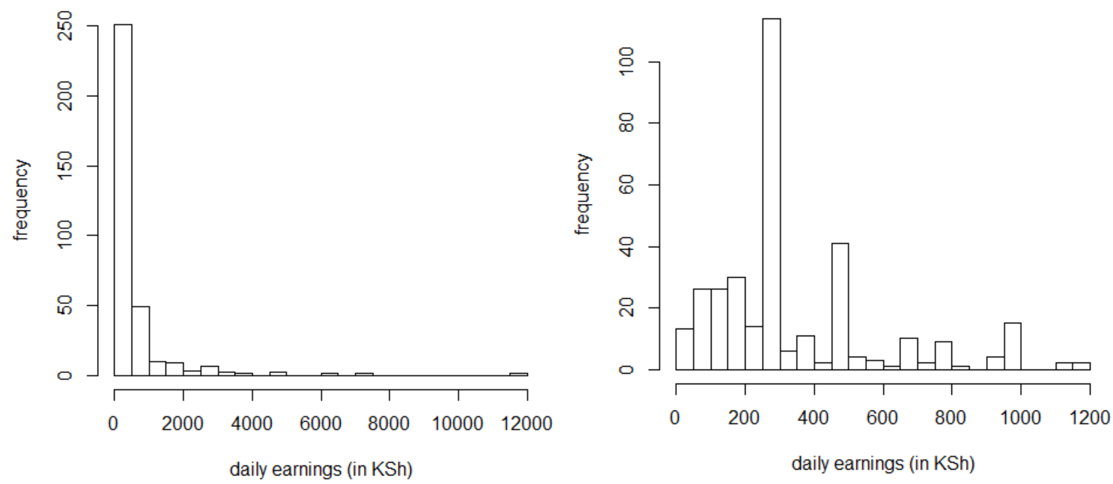
[4] We cannot reject a null hypothesis of equality of the coefficients on *maize*, *sorghum*, *cassava* and *millet* at the 5% level (see estimation output in Table A1.1 in Appendix A1)

data and three are derived from primary data. The procedure for obtaining an estimate of the sample hourly wage rate is described below.

Respondents were able to report their level of income either per day, per week or per month. Answers provided per week or per month, were converted to daily values assuming a five day working week and 22 working days per month. As shown in Figure 1.1 (a), the distribution of the reported level of daily income is skewed with 54 of 336 of the respondents reporting zero earnings and a few respondents reporting daily levels of income up to above 10,000 KSh.

Figure 1.1: Frequency plots of daily earnings

(a) before correcting for outliers (b) after correcting for outliers



We use the common 1.5×IQR (interquartile range) rule of thumb for identifying observations that are potential outliers (Hoaglin et al., 1986) and replace extreme values with the median value (300 KSh). Following this approach, 32 respondents with daily earnings above 1214 KSh are labelled as outliers. We further replace zero earnings with the median value to avoid zero prices when converting labour to money. Frequency plots of earnings before and after correcting for outliers are shown in Figure 1.1.

Respondents were also asked to indicate the number of hours spent on a typical day (i) sleeping, (ii) working (paid), and (iii) doing productive non-work activities such as cooking and cleaning. We use this individual-level data to compute an hourly wage rate by dividing daily earnings by the number of hours spent working.

1.4.3 CONVERSION RATES

To obtain monetised estimates of labour "payments", we use six different conversion rates, all of which have been previously applied in the literature (see Table 1.1). The conversion rates are shown and described in Table 1.2. Labour "payments" were monetised (after the survey but before model estimation) by multiplying the labour time presented in the choice tasks by one of the six types of conversion rates. Rates 1 to 4 are generic conversion rates while rate 5 and rate 6 are site- and individual-specific, respectively.

Table 1.2: Conversion rates

	Value	Description
rate 1	185 KSh	185 KSh is the average hourly wage rate in Kenya for the lowest paid industry (agriculture, forestry and fishing) in the private sector (Kenya National Bureau of Statistics, 2020) assuming 250 working days per year and 8 hours work per day.
rate 2	62 KSh	One third of rate 1.
rate 3	42 KSh	Average sample hourly wage rate.
rate 4	14 KSh	One third of rate 3.
rate 5	121 KSh or 68 KSh	121 KSh is the minimum hourly wage in Nairobi, Mombasa and Kisumu for general labour (Government of Kenya, 2018). We use this conversion rate in Kibera. 68 KSh is the minimum hourly wage in areas outside the cities of Nairobi, Mombasa and Kisumu and the town councils of Mavoko, Ruiru and Limuru for general labour (Government of Kenya, 2018). We use this conversion rate in Kisii/Nyamira.
rate 6	N/A	Sample individual hourly wage rates.

1.5 CONCEPTUAL FRAMEWORK

The conceptual framework for analysis of the choice data is based on random utility theory. In any given choice task, respondent *n* is assumed to select flour *j* if the utility gained from flour *j* is at least as high as the utility gained from each of the other (purchase or no-purchase) options in the choice task. The utility of flour *j* for respondent

n is modelled as a linear function of flour j's attributes X_j and a vector of preference parameters β_n indicating the desirability of the attributes (Hensher et al., 2015):

$$U_{nj} = \beta_n X_j + \varepsilon_{nj} = \sum_{k=1}^{K} \beta_{nk} A_{jk} + \beta_{money} M_j + \beta_{labour} L_j + \varepsilon_{nj} \quad \text{(Eq. 1)}$$

The utility specification in Eq. 1 includes one price attribute (either money M_j or labour L_j) and K=3 non-price attributes (represented by A_{jk}) of which two describe the characteristics of the flour and one is an alternative-specific constant (see Section 1.4.2). The marginal utilities of money and labour are denoted by β_{money} and β_{labour}, respectively, and ε_{nj} is a random error term. Taste heterogeneity is accommodated by allowing the preference parameter for the alternative-specific constant to vary across respondents (hence the subscript on β_n).

Assuming that ε_{nj} is independently and identically distributed type 1 extreme value, the probability of respondent n choosing flour i over flour j, conditional on knowing β_n, is given by the conditional logit formula in Eq. 2 (Holmes et al., 2017; McFadden, 1974).

$$P_{ni|\beta_n} = \frac{\exp(\beta_n X_i)}{\sum_{j=1}^{J} \exp(\beta_n X_j)} \; \forall \; j \neq i \in J \quad \text{(Eq. 2)}$$

To obtain respondent-specific preference parameters for the alternative-specific constant, we estimate a mixed logit model (MIXL) using 1000 Halton draws from the normal distribution. Following Carson and Czajkowski (2019), the negative of the price attributes (including any interaction terms) are included in the model as random variables with a log–normal distribution and standard deviations restricted to be zero (using the *mixlogit* command with the *ln(#)* option in Stata 14.0) (Hole, 2007b). This ensures that all respondents have non-positive price coefficients. The mean coefficients reported for the price variables M_j and L_j in the results section are, therefore, the natural logarithm of the negative of β_{money} and β_{labour}, respectively. Estimates of willingness to pay (work) for attribute k are then calculated as the marginal utility of the attribute relative to the exponential of the estimated mean coefficient for money (labour).

1.6 RESULTS

Summary statistics (mean value, minimum value and maximum value) are presented in Table 1.3. There is an equal number of men and women in the sample. The average respondent lives in a household with 2.6 individuals including 2 children and is 34 years old. Mean daily earnings is 362 KSh (≈ 3.6 USD at the time of the survey). Respondents reported working between 0 and 16 hours with the average respondent working 7.3 hours on a typical work day. Some respondents reported spending up to 12 hours per day on productive non-work activities (e.g. housework and subsistence farming) while the average respondent spends 3.1 hours per day on such activities. A bit more than half of the sample completed high school as their highest level of education while 28% had only completed primary school. The statistics are presented separately for the treatments using a payment vehicle that is either money or labour as well as for the treatment where choices are either hypothetical or consequential in Table B1.1 in Appendix B1. As expected (due to random allocation of respondents to treatments), we find no statistically significant difference in means across treatments for any of the variables.

Table 1.3: Summary statistics

	mean	minimum	maximum
Female (1 if female, 0 otherwise)	0.50 (0.50)	0	1
Age (years)	34 (13)	18	80
Daily earnings (KSh)*	362 (247)	11	1200
Household size	5.6 (3.7)	1	25
Children (<16 years) in household	2.0 (1.9)	0	12
Hours spent on a typical work day			
Sleeping	7.6 (1.4)	2	13
Working (paid)	7.3 (3.9)	0	16
Productive non-work activities	3.1 (2.0)	0	12
Highest level of education completed			
No formal qualifications	0.01 (0.08)	0	1
Primary school	0.28 (0.45)	0	1
High school	0.52 (0.50)	0	1
Vocational training	0.14 (0.35)	0	1
Higher education	0.05(0.22)	0	1
Number of respondents	336		

* after correcting for outliers (see Section 3.2)
Standard deviations in parentheses

To enable estimation of the shadow value of time, we pool the data from the four treatments and estimate treatment-specific[5] utility weights for each of the price variables. Parameter estimates are presented in the first column of Table 1.4 (MIXL–0). MIXL-0 is the model in which labour "prices" are retained in the units of hours in which they were presented in the choice tasks.

Table 1.4: MIXL estimation (*ex post* conversion of labour in MIXL1 to MIXL-6)

	MIXL–0	MIXL–1	MIXL–2	MIXL–3	MIXL–4	MIXL–5	MIXL–6
		rate 1	rate 2	rate 3	rate 4	rate 5	rate 6
flour (ASC)	2.37***	2.37***	2.37***	2.37***	2.37***	2.30***	2.09***
	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)	(0.15)
sifted	-0.14*	-0.14*	-0.14*	-0.14*	-0.14*	-0.13*	-0.11
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
fortified	0.66***	0.66***	0.66***	0.66***	0.66***	0.64***	0.56***
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.07)
money*conseq	-3.35***	-3.35***	-3.35***	-3.35***	-3.35***	-3.35***	-3.35***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)
money*hyp	-3.92***	-3.92***	-3.92***	-3.92***	-3.92***	-3.93***	-3.96***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.07)
labour*conseq	0.92***						
	(0.06)						
labour*hyp	0.12*						
	(0.06)						
[a]labour*conseq		-4.30***	-3.20***	-2.82***	-1.72***	-3.66***	-3.17***
		(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
[a]labour*hyp		-5.10***	-4.00***	-3.62***	-2.52***	-4.37***	-3.89***
		(0.06)	(0.06)	(0.06)	(0.06)	(0.07)	(0.08)
SD							
flour (ASC)	1.42***	1.42***	1.42***	1.42***	1.42***	1.55***	1.84***
	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.11)	(0.12)
observations	2688	2688	2688	2688	2688	2688	2688
LL	-2017	-2017	-2017	-2017	-2017	-2063	-2182
AIC	4051	4051	4051	4051	4051	4141	4381
BIC	4106	4106	4106	4106	4106	4197	4437

[a] monetised labour
Standard errors in parentheses
* p<0.10; ** p<0.05; *** p<0.01
hyp = hypothetical
conseq = consequential

[5] This is equivalent to estimating additive interaction terms.

The mean parameter for *flour* is positive and significant which means that the respondents typically prefer the purchase option over the no purchase option. Respondents further prefer flour that is fortified and not sifted (although the attribute *sifted* is significant at the 10% level only). Due to the reparameterisation of the price variable coefficients (see Section 1.5), (the exponential of) the price parameters are constrained to strictly positive values.

The mean parameters for the price variables can be used to calculate the shadow value of time which is the marginal rate of substitution between labour time and money. As shown in Table 1.4, (minus the exponential of) the marginal utilities of labour time and money are higher (i.e. less negative) when choices are hypothetical. This means that respondents are more sensitive to payment (using either payment vehicle) when choices are consequential. We can, therefore, calculate four different values of the shadow value of time which vary according to the nature of the choice tasks (hypothetical or consequential) (see Eq. 3 to Eq. 6).

$$
\text{mean shadow value of time} = \begin{cases}
\exp\left(\beta_{labour*conseq}\right) \Big/ \exp\left(\beta_{money*conseq}\right) = 72 \text{ KSh} & \text{(Eq. 3)} \\[2ex]
\exp\left(\beta_{labour*hyp}\right) \Big/ \exp\left(\beta_{money*hyp}\right) = 57 \text{ KSh} & \text{(Eq. 4)} \\[2ex]
\exp\left(\beta_{labour*hyp}\right) \Big/ \exp\left(\beta_{money*conseq}\right) = 32 \text{ KSh} & \text{(Eq. 5)} \\[2ex]
\exp\left(\beta_{labour*conseq}\right) \Big/ \exp\left(\beta_{money*hyp}\right) = 127 \text{ KSh} & \text{(Eq. 6)}
\end{cases}
$$

If we accept the premise that consequential choices provide incentives for truthful preference revelation, the ‘true’ average shadow value of time is 72 KSh per hour. Using values from the hypothetical treatments, the shadow value of time is only 57 KSh per hour. This is due to a difference in the scale of hypothetical bias (HB) between the two payment vehicles. The level of HB when the payment vehicle is labour is

$\exp(\beta_{labour*conseq}) / \exp(\beta_{labour*hyp}) = 2.23$ which exceeds the level of HB when money is used as the payment vehicle: $\exp(\beta_{money*conseq}) / \exp(\beta_{money*hyp}) = 1.77$ (a null hypothesis of equality is rejected at the 5% significance level). The shadow value of time is, therefore, lower when purchase choices are hypothetical compared to when purchase choices are consequential because, even if respondents overstate both hypothetical WTW and hypothetical WTP, the bias is higher, in relative terms, when the payment vehicle is labour.

When labour "prices" are hypothetical and monetary prices are consequential, the shadow value of time is 32 KSh. This value has been HB-deflated with respect to the monetary payment vehicle (PV) but not with respect to the labour PV and is thus biased downwards compared to the 'true' shadow value of time. It is an interesting value (and unique to this study) because it is essentially the rate that would convert hypothetical labour to consequential money by correcting not only for HB in general but also for differences in HB between the two PVs.

The shadow value of time when labour "prices" are consequential and monetary prices are hypothetical is 127 KSh. Since we do not expect that SP practitioners are interested in converting consequential labour to hypothetical money, this value is included in the interest of completeness.

In the last six columns of Table 1.4 (MIXL–1 to MIXL–6), labour "payments" are converted to monetary values (before model estimation) using one of six different conversion rates (see Section 1.4.3). This approach is the one commonly employed in the WTW literature thus far (see Table 1.1). When a flat rate is used (rates 1-4), (the exponential of) the labour "price" interaction terms are scaled (down) while the mean parameters for the remaining variables, estimated standard deviations and log likelihoods are unchanged compared to MIXL–0. When site- or individual- specific rates are used (rates 5 and 6, MIXL–5 and MIXL–6), mean parameters and standard deviations change (marginally) compared to MIXL–0. There is no difference in the sign (and significance) of the mean parameters and standard deviations across the models (except the attribute *sifted* which is insignificant at the 10% level in MIXL–6). The statistics reported in the lower part of Table 1.4 (log likelihood, Akaike information criterion and Bayesian information criterion) indicate that MIXL–0 to MIXL-4 fit the data better than

MIXL–5 and MIXL–6. Model fit is thus reduced when heterogeneity (at the level of the site or individual) is introduced as opposed to the use of generic conversion rates.

We can use the mean parameter estimates from MIXL–0 in Table 1.4 to calculate WTP and WTW separately for the hypothetical and the consequential treatments. The estimates are presented in the first two columns of Table 1.5. WTP and WTW are calculated as the ratio of the non-price variable's coefficient to the exponential of the relevant price variable's coefficient and the delta method is utilised to obtain 95% confidence intervals (Hole, 2007a). We omit the attribute *sifted* due to the statistical insignificance of the estimated coefficient for this variable.

As shown in the first column of Table 1.5, the average respondent is willing to work 126 minutes for a bag of unsifted, unfortified flour in the hypothetical treatment but only 56 minutes in the consequential treatment thus demonstrating the scale of the HB (as discussed above). Similarly, when the PV is money (see the second column in Table 1.5), the average respondent in the hypothetical treatment is willing to pay 119 KSh for a bag of unsifted, unfortified flour which is significantly higher than mean WTP in the consequential treatment (67 KSh).

In the last six columns of Table 1.5, we use the mean parameters from MIXL–1 to MIXL–6 in Table 1.4 to calculate estimates of WTP based on monetised labour "payments" (WTP_{labour}) for each of the six conversion rates. The subscripts 'money' and 'labour' are used to distinguish between willingness to pay based on monetary prices (WTP_{money}) and willingness to pay based on monetised labour "prices" (WTP_{labour}), i.e. monetised WTW.

The last six columns in the upper panel of Table 1.5 show estimates of monetised WTW derived from hypothetical choice tasks (denoted as WTP^{hyp}_{labour}). These values are equivalent to the values generated by the WTW literature thus far i.e. the papers listed in Table 1.1 each generate one or more of these values depending on which of the conversion rates they apply. WTP^{hyp}_{labour} for a bag of unsifted, unfortified flour range between 29 KSh using the sample average hourly wage rate and 387 KSh using the national average hourly wage rate.

Table 1.5: WTW, WTP_{money} and WTP_{labour}

			rate 1	rate 2	rate 3	rate 4	rate 5	rate 6
	WTW	WTP_{money}	WTP_{labour}					
HYP								
flour (ASC)	126	119	387	129	88	29	182	102
	[108;144]	[104;135]	[331;443]	[110;148]	[75;101]	[25;34]	[154;210]	[83;121]
fortified	35	33	108	36	25	8	51	27
	[27;44]	[25;42]	[82;135]	[27;45]	[19;31]	[6;10]	[38;63]	[19;35]
consequential								
flour (ASC)	56	67	174	58	40	13	89	50
	[50;63]	[60;75]	[153;194]	[51;65]	[35;44]	[12;15]	[77;100]	[43;58]
fortified	16	19	49	16	11	4	25	13
	[12;20]	[14;23]	[37;60]	[12;20]	[8;14]	[3;5]	[19;31]	[10;17]

95% confidence interval in square brackets
HYP = hypothetical treatments
CONSEQ = consequential treatments
WTP is in KSh and WTW is in minutes

The common approach in the WTW literature for evaluating the performance of one or more conversion rates is to assess the closeness of WTP_{labour}^{hyp} to WTP_{money}^{hyp} (i.e. WTP derived from hypothetical choice tasks with monetary prices). Here, we follow this approach before assessing its viability using choice data from the consequential treatment. The rate which produces an estimate of WTP_{labour}^{hyp} closest to WTP_{money}^{hyp} is one third of the national average hourly wage rate but both one third of the national average hourly wage rate and individual-specific wage rates generate WTP_{labour}^{hyp} estimates that are not significantly different from WTP_{money}^{hyp}.

The lower panel in Table 1.5 shows estimates of WTP_{labour} derived from consequential choice tasks (WTP_{labour}^{conseq}). These values are, to the best of our knowledge, unique to this paper. WTP_{labour}^{conseq} for a bag of unsifted, unfortified flour range between 13 KSh using one third of the sample average hourly wage rate and 174 KSh using the national average wage rate. As in the upper panel, the rate which produces an estimate of WTP_{labour}^{conseq} closest to WTP_{money}^{conseq} (i.e. WTP derived from consequential choice tasks with monetary prices) is one third of the national average wage rate. Individual-specific conversion rates continue to perform second-best with respect to generating WTP_{labour}^{conseq} estimates close to WTP_{money}^{conseq} but, in contrast to the upper panel, the difference between

WTP^{conseq}_{labour} and WTP^{conseq}_{money} is now significant at the 1% level for the attribute *flour* and at the 10% level for the attribute *fortified*.

As demonstrated in Table 1.5, WTP_{labour} is not, in either consequential or hypothetical treatments, consistently higher or lower than WTP_{money}. The relationship between WTP_{labour} and WTP_{money} is further explored in Table 1.6 which presents ratios of WTP_{labour} to WTP_{money} (see also Table B1.2 in Appendix B1 which presents the difference between WTP_{labour} and WTP_{money} and the associated 95% confidence intervals). In Table 1.6, the ratio is equal to 1 if WTP_{labour} equals WTP_{money}. If the ratio is higher (lower) than 1, then WTP_{labour} is higher (lower) than WTP_{money}. By definition, if both labour and monetary prices are hypothetical (consequential), the ratios in the upper (middle) panel will be equal to 1 when the conversion rate equals the shadow value of time estimated in Eq. 4 (Eq.3). Similarly, if labour "prices" are hypothetical and monetary prices are consequential, the ratios in the lower panel will be equal to 1 when the conversion rate equals the shadow value of time estimated in Eq. 5.

The upper panel in Table 1.6 shows the ratio of hypothetical WTP_{labour} (WTP^{hyp}_{labour}) to hypothetical WTP_{money} (WTP^{hyp}_{money}). We find that WTP^{hyp}_{labour} exceeds WTP^{hyp}_{money} using the national average wage rate and site-specific minimum wage rates (although the difference is significant at the 10% level only for the attribute *fortified* using site-specific wage rates) and that WTP^{hyp}_{labour} is lower than WTP^{hyp}_{money} using the sample average wage rate or one third thereof. WTP for a bag of unsifted, unfortified flour based on monetised hypothetical labour "prices" using e.g. site-specific minimum hourly wage rates is 52% higher than mean WTP for a bag of unsifted, unfortified flour based on consequential monetary prices. When one third of the national average wage rate or individual-specific wage rates are used to monetise hypothetical labour "prices", WTP^{hyp}_{labour} is not statistically significantly different from WTP^{hyp}_{money}.

The ratio of consequential WTP_{labour} (WTP^{conseq}_{labour}) to consequential WTP_{money} (WTP^{conseq}_{money}) is shown in the middle panel of Table 1.6. As in the upper panel, we find that WTP^{conseq}_{labour} exceeds WTP^{conseq}_{money} using the national average wage rate and site-specific minimum wage rates (although the difference is insignificant at the 10% level for the attribute *fortified* using site-specific wage rates) and that WTP^{conseq}_{labour} is lower than WTP^{conseq}_{money} using the sample average wage rate or one third thereof. In contrast to the upper panel, we find that WTP^{conseq}_{labour} is significantly lower than consequential WTP^{conseq}_{money} using one third of the national average wage rate and individual-specific wage rates.

The results in the lower panel of Table 1.6 are unique (thus far) to this study. In the lower panel, we present the ratio of hypothetical WTP_{labour} (WTP^{hyp}_{labour}) to consequential WTP_{money} (WTP^{conseq}_{money}). WTP for a bag of unsifted, unfortified flour based on monetised hypothetical labour "prices" using e.g. one third of the sample average wage rate is 1 – 0.44 = 56% lower than mean WTP for a bag of unsifted, unfortified flour based on consequential monetary prices. We find that WTP^{hyp}_{labour} for a bag of unsifted, unfortified flour exceeds WTP^{conseq}_{money} using all six conversion rates except one third of the sample average wage rate. As shown, the rate which produces an estimate of hypothetical WTP_{labour} closest to consequential WTP_{money} (i.e. the rate that generates a ratio closest to 1) is the sample average wage rate. When the sample average wage rate is used to monetise labour "prices", WTP^{hyp}_{labour} is 31% higher than WTP^{conseq}_{money} (for both *flour* and *fortified*).

Table 1.6: WTP_{labour} / WTP_{money}

	rate 1	rate 2	rate 3	rate 4	rate 5	rate 6
$WTP^{hyp}_{labour}/WTP^{hyp}_{money}$						
flour (ASC)	3.24***	1.08	0.74***	0.25***	1.52***	0.85
	(0.26)	(0.09)	(0.06)	(0.02)	(0.15)	(0.10)
fortified	3.24***	1.08	0.74***	0.25***	1.51*	0.81
	(0.26)	(0.09)	(0.06)	(0.02)	(0.27)	(0.15)
$WTP^{conseq}_{labour}/WTP^{conseq}_{money}$						
flour (ASC)	2.58***	0.86**	0.59***	0.20***	1.32***	0.75***
	(0.17)	(0.06)	(0.04)	(0.01)	(0.11)	(0.07)
fortified	2.58***	0.86**	0.59***	0.20***	1.31	0.71**
	(0.17)	(0.06)	(0.04)	(0.01)	(0.23)	(0.13)
$WTP^{hyp}_{labour}/WTP^{conseq}_{money}$						
flour (ASC)	5.75***	1.92***	1.31***	0.44***	2.70***	1.51***
	(0.43)	(0.14)	(0.10)	(0.03)	(0.26)	(0.17)
fortified	5.75***	1.92***	1.31***	0.44***	2.68***	1.44
	(0.43)	(0.14)	(0.10)	(0.03)	(0.47)	(0.27)

Standard errors in parentheses
We test H_0: WTP_{labour} / WTP_{money} = 1
* p<0.10; ** p<0.05; *** p<0.01

A final interesting observation from Table 1.6 relates to the comparison of ratios between the hypothetical and the consequential treatments (i.e. a comparison between the values in the upper panel and the middle panel). The WTP_{labour} / WTP_{money} ratios are larger in the upper panel (which displays results from hypothetical treatments) compared to the middle panel (which displays results from consequential treatments) for all attributes and for all conversion rates. When a generic wage rate is used (rates 1-4), the ratio of WTP^{hyp}_{labour} to WTP^{hyp}_{money} is 26%[6] higher than the ratio of WTP^{conseq}_{labour} to WTP^{conseq}_{money} and when site- or individual- specific rates are used (rates 5-6) the ratio of WTP^{hyp}_{labour} to WTP^{hyp}_{money} is 14-16% higher than the ratio of WTP^{conseq}_{labour} to WTP^{conseq}_{money}. The difference using rates 1-4 is significant at the 5% level while the difference using rates 5 and 6 is not statistically significant ($p > 0.20$).

The differences in WTP_{labour} / WTP_{money} ratios between hypothetical and consequential treatments can be attributed to differences in HB between the two PVs. As articulated above, HB is present both when the PV is money and when the PV is labour but HB is higher when labour is used as the PV. This means that both WTP_{money} and WTP_{labour} are biased upwards in the hypothetical treatments but the bias does not affect WTP_{money} and WTP_{labour} by the same proportions (the bias associated with WTP_{labour} is higher). Consequently, the WTP ratios in the upper panel of Table 6 are higher than the HB-deflated WTP ratios in the middle panel.

1.7 DISCUSSION

This paper is based on a (at the time of writing) unique dataset from a field study with four treatments in which the payment vehicle (PV) is either money or labour and purchase choices are either hypothetical or consequential. The study design allows an investigation of the impact of the choice of PV on estimates of WTP based on monetised labour "prices" (WTP_{labour}), i.e. monetised WTW, relative to WTP based on monetary prices (WTP_{money}). Specifically, we consider the impact of (i) the use of different rates to convert WTW values into WTP values, and (ii) differences in the level of HB between money and labour PVs.

---

[6] This is equivalent to the ratio of HB when labour is used as the PV to HB when money is used as the PV i.e. $\left(\exp\left(\beta_{labour*conseq}\right) / \exp\left(\beta_{labour*hyp}\right)\right) / \left(\exp\left(\beta_{money*conseq}\right) / \exp\left(\beta_{money*hyp}\right)\right) = 1.26$

The argument typically put forward in the literature advocating for the use of non-monetary PVs is that a monetary PV underestimates welfare values of low-income households in subsistence economies due to liquidity constraints and/or lack of experience with money as the medium of exchange. Labour payments have been suggested as a suitable alternative to monetary payments. A fundamental challenge, however, is the need to apply an opportunity cost of time to obtain monetary measures of welfare which is required in most cost benefit analyses. In this paper, six commonly used rates for converting labour "prices" into their monetary equivalents were applied and compared across subgroups of the sample where choices are either hypothetical or consequential. The use of consequential treatments in addition to the standard hypothetical money- and/or labour- PV treatments allows a deeper investigation regarding the performance of the competing rates to monetise WTW.

1.7.1 HYPOTHETICAL WTP_{LABOUR} AND HYPOTHETICAL WTP_{MONEY}

In this section, we compare and discuss the difference between hypothetical WTP based on monetised labour "prices" (WTP^{hyp}_{labour}), i.e. monetised WTW, and hypothetical WTP based on monetary prices (WTP^{hyp}_{money}) for six conversion rates that are commonly employed in the WTW literature. As this section concerns hypothetical treatments only, it reflects the analysis typically conducted in the non-monetary PV literature.

Using the area average wage rate (at the national level) from secondary data (the most popular conversion rate in the literature – see Table 1.1), we find that WTP^{hyp}_{labour} (=387 KSh) is 224% higher than WTP^{hyp}_{money} (=119 KSh) which suggests that respondents, on average, value their time at a rate which is (about 70%) lower than the average wage rate. We similarly find that WTP^{hyp}_{labour} exceeds WTP^{hyp}_{money} when labour is converted to money using site-specific minimum wage rates (although the difference for the attribute *fortified* is significant at the 10% level only). As discussed by Gibson et al. (2016), if respondents lack unlimited access to labour markets, they are less likely to consider market wages as a foregone opportunity and their opportunity cost of time is predictably lower. Given that respondents in our study were sampled from low-income areas where unemployment rates are high, it is not surprising that WTP_{labour}, evaluated at the minimum wage rate or the national average wage rate (albeit for the lowest paid industry), exceeds WTP_{money}.

If instead we apply one third of the national average wage rate as the conversion rate (a rate which can be traced back to Cesario, 1976), we find no statistically significant difference between WTP^{hyp}_{labour} and WTP^{hyp}_{money}. This result provides support for the use of one third of the wage rate as a measure of the value of time and is consistent with the literature suggesting that individuals substitute leisure time, rather than time already spent in work, when considering their willingness to contribute labour. To this end, many studies (e.g. Ando et al., 2020; Arbiol et al., 2013; O'Garra, 2009; Schiappacasse et al., 2013) use one third of the (sample/local/national) average wage rate as the opportunity cost of leisure time. We find, however that WTP^{hyp}_{labour} is lower than WTP^{hyp}_{money} using one third of the sample average wage rate i.e. it is only one third of the national average wage rate which performs well by generating values of WTP^{hyp}_{labour} that match WTP^{hyp}_{money}.

Using the sample average wage rate (or one third of the sample average wage rate as discussed above), we find that WTP^{hyp}_{labour} is lower than WTP^{hyp}_{money} which suggests that respondents value their time at a rate that is higher than the sample average wage rate. We speculate, that this may be due to time constraints related to non-work activities and non-wage income (e.g. subsistence farming or hustling) which increase respondents' opportunity cost of time. It is also possible that respondents have underreported their daily earnings.

If we apply individual-specific wage rates to convert labour "prices" into money, we find no statistically significant difference between WTP^{hyp}_{labour} and WTP^{hyp}_{money}. This suggests that individual-level wage rate data performs well when performance is measured in terms of the closeness of WTP^{hyp}_{labour} to WTP^{hyp}_{money}. We find, however, that the use of individual- (or site-) specific wage rates, instead of generic wage rates, reduces model fit. This is consistent with Czajkowski et al. (2019) and Lloyd-Smith et al. (2019) who find that individuals' value of time (estimated using SP methods) are only weakly correlated with self-reported wage rates.

In summary, we have compared the mean values of hypothetical WTP based on monetised labour "prices" and hypothetical WTP based on monetary prices using six different conversion rates of which four are generic and two are site or individual-specific. We find that the relationship between WTP^{hyp}_{labour} and WTP^{hyp}_{money} is highly sensitive to the conversion rates used and that WTP^{hyp}_{labour} is not consistently higher (or lower) than WTP^{hyp}_{money}. This is in contrast to Hagedoorn et al. (2020) who, in a DCE study concerning ecosystem-based adaptation measures in Vietnam, find that WTP^{hyp}_{labour} exceeds WTP^{hyp}_{money} for all the conversion rates used[7].

Thus far the comparisons we have made are akin to those made in the extant literature on monetised WTW and WTP. That is, *ex post* conversion of hypothetical estimates of WTW to WTP at differing rates to compare the resulting WTP^{hyp}_{labour} values to WTP^{hyp}_{money}. In the following two subsections, we analyse the issue of converting labour to money more deeply through the use of consequential money and labour PV treatments.

1.7.2 CONSEQUENTIAL WTP_{LABOUR} AND CONSEQUENTIAL WTP_{MONEY}

When both labour and monetary payments are consequential, we find that the relative difference between WTP_{labour} and WTP_{money} is larger than when both labour and monetary payments are hypothetical. This finding can be attributed to differences in hypothetical bias (HB) between the two PVs. We find that the level of HB is 26% higher when the PV is labour compared to when the PV is money. Since the effect of HB on WTP_{labour} is higher than the effect on WTP_{money}, the relative difference between WTP_{labour} and WTP_{money} is larger in the hypothetical treatment compared to the consequential treatment. If HB affected WTP_{labour} and WTP_{money} by the same proportions, the bias would cancel out and the ratio of WTP^{hyp}_{labour} to WTP^{hyp}_{money} would equal the HB-deflated ratio of WTP^{conseq}_{labour} to WTP^{conseq}_{money}.

This finding has important implications for the evaluation of welfare effects. If, for example, the national average hourly wage rate is applied to monetise hypothetical labour payments, we find that WTP_{labour} is 224% higher than WTP_{money}. However, when HB is removed from the equation (by taking differences in HB between the two PVs into

[7] Hagedoorn et al. (2020) use six different conversion rates: (1) the average market wage rate, (2) individual-specific wage rates, (3) one third of the individual-specific wage rates, (4) half of the individual-specific wage rates, (5) an individual-specific weighted value of time which is a function of paid and unpaid work hours and leisure time, and (6) like (5) but with different weights.

account or by using consequential choice data), WTP_{labour} is “only” 158% higher than WTP_{money}. Decision-makers relying on monetised hypothetical labour payments, will therefore not only overestimate benefits due to the presence of HB in general (as discussed in the next section) but also because of the PV-HB asymmetry.

Our results confirm the untested claim (made e.g. by Ando et al., 2020 and Eom and Larson, 2006) that the recurrent finding that WTP^{hyp}_{labour} exceeds WTP^{hyp}_{money} (see e.g. Abramson et al., 2011; Casiwan-Launio et al., 2011; Hagedoorn et al., 2020; Meginnis et al., 2020) is (at least partly) due to differences in HB between the two PVs. The level of HB is estimated to be 26% higher when the PV is labour which suggests that monetised values of hypothetical labour payments should be knocked down by a factor 0.80 to account for this HB-PV asymmetry[8]. Such correction, however, accounts only for differences in HB between the two PVs and not for HB in general. We address the latter in the next subsection by comparing WTP^{hyp}_{labour} and WTP^{conseq}_{money}.

1.7.3 HYPOTHETICAL WTP_{LABOUR} AND CONSEQUENTIAL WTP_{MONEY}

A common approach in the WTW literature for assessing the performance of one or more conversion rates is to compare WTP^{hyp}_{labour} to WTP^{hyp}_{money}. We argue, however, that the relevant benchmark is WTP derived from consequential choice tasks (WTP^{conseq}_{money}) rather than WTP derived from hypothetical choice tasks (WTP^{hyp}_{money}). If we accept the premise that consequentiality of the choice tasks induces truth telling, then WTP^{conseq}_{money} can be thought of as the ‘true’ monetary measure of welfare. In this section, we compare and discuss the difference between WTP^{hyp}_{labour} and WTP^{conseq}_{money} for the same six conversion rates as in Section 1.7.1. In the next section, we will then compare the performance of the different conversion rates across the two benchmarks (WTP^{hyp}_{money} and WTP^{conseq}_{money}). To the best of our knowledge, this is the first labour PV study to examine the ability of a set of (commonly used) conversion rates to generate HB-deflated monetised WTW values.

We find that WTP^{hyp}_{labour} for a bag of unsifted, unfortified flour exceeds WTP^{conseq}_{money} using all six conversion rates except when using one third of the sample average wage rate. WTP^{hyp}_{labour} for the attribute *fortified* exceeds WTP^{conseq}_{money} using the national average

[8] $\left(\exp\left(\beta_{money*conseq}\right) / \exp\left(\beta_{money*hyp}\right)\right) / \left(\exp\left(\beta_{labour*conseq}\right) / \exp\left(\beta_{labour*hyp}\right)\right) = 0.80$ (based on coefficients from MIXL–0 in Table 1.4)

wage rate, one third of the national average wage rate, the sample average wage rate and site-specific minimum wage rates while WTP^{hyp}_{labour} is lower than WTP^{conseq}_{money} using one third of the sample average wage rate and WTP^{hyp}_{labour} is not significantly different from WTP^{conseq}_{money} using individual-specific wage rates.

The conversion rate which produces an estimate of hypothetical WTP^{hyp}_{labour} closest to the estimate of WTP^{conseq}_{money} is the sample average wage rate. Of the six conversion rates that we apply, it seems, therefore, that the sample average wage rate is the best performing rate with respect to converting hypothetical WTW to the 'true' estimate of (consequential) WTP. Despite being the best performing conversion rate, however, the sample average wage rate still generates WTP^{hyp}_{labour} estimates (for both attributes) that are 31% higher than WTP^{conseq}_{money} thus overestimating the 'true' value of WTP.

1.7.4 CONVERSION RATE PERFORMANCE

Thus far we have discussed the performance of the different conversion rates with respect to converting WTP_{labour} to WTP_{money} when both labour and monetary prices are hypothetical (Section 1.7.1), when both labour and monetary prices are consequential (Section 1.7.2) and when labour "prices" are hypothetical but monetary prices are consequential (Section 1.7.3). In this section we compare the performance of the different conversion rates with respect to converting (1) WTP^{hyp}_{labour} to WTP^{hyp}_{money} (this approach is standard in the WTW literature) and (2) WTP^{hyp}_{labour} to WTP^{conseq}_{money} (this approach, we believe, is unique to our study). In particular, we are comparing the performance of six commonly employed conversion rates across two different benchmarks: WTP^{hyp}_{money} (the commonly applied benchmark) and WTP^{conseq}_{money} (our proposed benchmark). We argue that WTP^{conseq}_{money} is the relevant benchmark because, unlike WTP^{hyp}_{money}, it eliminates HB (both differences in HB between labour and money PVs, as demonstrated in Section 1.7.2, and HB in general).

When both labour and monetary prices are hypothetical, one third of the national average wage rate and individual-specific wage rates generate estimates of WTP^{hyp}_{labour} that are not significantly different from WTP^{hyp}_{money}. Following the common approach in the literature, one third of the national average wage rate and individual-specific wage rates are thus the best performing rates with respect to producing WTP^{hyp}_{labour} which match WTP^{hyp}_{money}. If instead we use WTP^{conseq}_{money} as the benchmark, then one third of the national average wage rate overestimates WTP_{money} by 92% and performs rather poorly by ranking 4th out of the 6 rates employed. Individual-specific wage rates continue to perform best or second-best by producing estimates of WTP^{hyp}_{labour} that are (i) 51% higher than WTP^{conseq}_{money} for *flour* (second-best) and, (ii) not significantly different from WTP^{conseq}_{money} for *fortified* (best). The best performing conversion rate when WTP^{conseq}_{money} is used as the benchmark is the sample average wage rate. The sample average wage rate generates estimates of WTP^{hyp}_{labour} that are 31% higher than the 'true' value of WTP (WTP^{conseq}_{money}) for both attributes. Following the common approach in the literature, however, WTP^{hyp}_{labour} is (incorrectly) assumed to be underestimated by a factor 0.74.

In summary, when using WTP^{conseq}_{money} as the benchmark instead of WTP^{hyp}_{money}, the sample average wage rate moves from being ranked 4th to 1st while one third of the national average wage rate drops from 1st to 4th and individual-specific wage rates continue to be second-best. Our results demonstrate why the choice of benchmark is important: if the common benchmark (WTP^{hyp}_{money}) is used to evaluate the performance of e.g. one third of the national average wage rate, we (incorrectly) conclude that labour and money PVs generate WTP values of the same magnitude while WTP derived from monetised labour "payments", i.e. monetised WTW, in fact overestimates the 'true' value of WTP (i.e. WTP^{conseq}_{money}) by 92%.

1.7.5 THE SHADOW VALUE OF TIME

For the purpose of assessing welfare effects e.g. by comparing benefits and costs, the conversion of labour to money appears to be a relevant and important exercise. It can be argued, however, that if labour is used as the PV due to low-income groups having little or no access to waged labour, and money therefore is deemed an inappropriate metric, then the conversion of labour to money is somewhat ill-founded. It is furthermore common practice in many WTW papers (e.g. Ando et al., 2020; Gibson et al., 2016; Navrud and Vondolia, 2020) to collect WTP data in addition to WTW data in order to compare estimates of WTP_{labour} and WTP_{money} (this is also the approach which we used to assess the performance of a set of conversion rates in Sections 1.7.1 to 1.7.3 albeit using consequential data in place of hypothetical data in Sections 1.7.2 and 1.7.3).

Following the same argument, however, WTP_{money} seems to be an unfitting benchmark if the motivation for using labour as the PV in the first instance was the unsuitability of a monetary PV.

As such, the comparison between WTP_{labour} and WTP_{money} seems only justified in circumstances where both PVs are deemed appropriate. In such situations, if both PVs are used, the shadow value of time can be estimated which, by definition, is the value at which WTP_{labour} equals WTP_{money}. Unique to our paper is the ability to calculate the shadow value of time when both labour and monetary payments are consequential (72 KSh per hour), when both labour and monetary payments are hypothetical (57 KSh per hour) and when labour "payments" are hypothetical but monetary payments are consequential (32 KSh per hour). The latter is of particular interest because it is the rate at which hypothetical labour payments can be converted to consequential monetary payments. While the 'true' shadow value of time is 72 KSh, hypothetical labour payments will have to be monetised using a (much) lower conversion rate to account not only for HB in general but also for differences in HB between the two PVs.

32 KSh is approximately three quarters of the sample average wage rate or one sixth of the national average wage rate which suggests that the opportunity cost of time is not one third of the wage rate as it is commonly presumed. As discussed in Section 1.7.3, the conversion rate which produces an estimate of WTP_{labour}^{hyp} that most closely matches WTP_{money}^{conseq} is the sample average wage rate (42 KSh) but despite being the best performing conversion rate, the sample average wage rate still generates WTP_{labour}^{hyp} estimates that are 31% higher than WTP_{money}^{conseq} thus overestimating the 'true' value of WTP.

1.8 CONCLUSION

This paper has investigated the performance of six commonly employed conversion rates with respect to their ability to generate monetised WTW values that match WTP. First we used hypothetical WTP as the benchmark (this is the standard approach in the literature) and found that one third of the national average wage rate performed well by generating monetised WTW values that are not significantly different from WTP. We then used consequential WTP as the benchmark to assess the reliability of inferences drawn from the use of hypothetical WTP. When consequential WTP was used as the benchmark, we found that the value of WTW, when monetised using one third of the national average wage rate, exceeded WTP by a factor of almost 2 and that one third of the national average wage rate ranked only fourth of the six conversion rates employed. The results thus cast doubt on the use of hypothetical WTP as a suitable benchmark. We further found that monetised values of hypothetical WTW overestimate the 'true' value of WTP due to not only hypothetical bias (HB) in general but also differences in HB between the two types of payment vehicles. Our results are an important addition to the WTW literature as they provide the first bit of evidence about the rate at which hypothetical WTW can be converted to HB-deflated WTP.

A1 APPENDIX: DATA SETUP

Table A1.1: MIXL estimation (unrestricted model)

maize	2.14***
	(0.15)
sorghum	2.22***
	(0.14)
cassava	1.92***
	(0.14)
millet	2.26***
	(0.16)
sifted	-0.18**
	(0.08)
fortified	0.66***
	(0.08)
money*conseq	-0.03***
	(0.00)
money*hyp	-0.02***
	(0.00)
labour*conseq	-2.40***
	(0.12)
labour*hyp	-1.28***
	(0.07)
SD	
maize	1.12***
	(0.16)
sorghum	1.14***
	(0.14)
cassava	0.83***
	(0.18)
millet	1.18***
	(0.16)
observations	2688
LL	2090

Standard errors in parentheses
* p<0.10; ** p<0.05; *** p<0.01
hyp = hypothetical
conseq = consequential

Table A1.2: Coding scheme

							hypothetical		consequential	
id	task	alternative	choice	flour (ASC)	sifted	fortified	money	labour	money	labour
1	1	1	0	1	0	1	120	0	0	0
1	1	2	1	1	1	0	40	0	0	0
1	1	3	0	0	0	0	0	0	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
4	1	1	1	1	1	0	0	0	40	0
4	1	2	0	1	1	1	0	0	160	0
4	1	3	0	0	0	0	0	0	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
7	1	1	0	1	1	1	0	4	0	0
7	1	2	1	1	1	0	0	0.25	0	0
7	1	3	0	0	0	0	0	0	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
11	1	1	0	1	1	1	0	0	0	2.5
11	1	2	0	1	0	0	0	0	0	3
11	1	3	1	0	0	0	0	0	0	0

B1 APPENDIX: RESULTS

Table B1.1: Summary statistics by treatment type

	MPV	LPV	difference	HYP	CONSEQ	difference
Female (1 if female, 0 otherwise)	0.49	0.51	-0.02	0.51	0.49	0.01
	(0.50)	(0.50)	(0.05)	(0.50)	(0.50)	(0.05)
Age (years)	35	33	1	34	35	-1
	(14)	(13)	(1)	(13)	(14)	(1)
Daily earnings (KSh)*	366	358	8	368	356	12
	(239)	(255)	(27)	(243)	(250)	(27)
Household size	5.5	5.8	-0.2	5.8	5.5	0.2
	(3.4)	(3.9)	(0.4)	(3.6)	(3.7)	(0.4)
Children (<16 years) in household	2.0	2.0	0.0	2.1	1.8	0.2
	(1.8)	(2.1)	(0.2)	(1.9)	(1.9)	(0.2)
Hours spent on a typical day						
Sleeping	7.6	7.7	-0.1	7.5	7.7	-0.3
	(1.4)	(1.4)	(0.2)	(1.4)	(1.4)	(0.2)
Working (paid)	7.5	7.0	0.4	7.0	7.5	-0.5
	(3.9)	(3.8)	(0.4)	(4.0)	(3.7)	(0.4)
Productive non-work activities	3.0	3.2	-0.2	3.2	2.9	0.3
	(2.0)	(2.0)	(0.2)	(2.0)	(2.0)	(0.2)
Highest level of education completed						
No formal qualifications	0.01	0.01	0.00	0.01	0.01	0.00
	(0.08)	(0.08)	(0.01)	(0.08)	(0.08)	(0.01)
Primary school	0.29	0.28	0.01	0.26	0.30	-0.04
	(0.45)	(0.45)	(0.05)	(0.44)	(0.46)	(0.05)
High school	0.49	0.55	-0.07	0.50	0.53	-0.03
	(0.50)	(0.50)	(0.05)	(0.50)	(0.50)	(0.05)
Vocational training	0.16	0.13	0.03	0.16	0.12	0.04
	(0.36)	(0.33)	(0.04)	(0.37)	(0.33)	(0.04)
Higher education	0.06	0.04	0.03	0.07	0.03	0.03
	(0.25)	(0.19)	(0.02)	(0.25)	(0.18)	(0.02)
Number of respondents	171	165		164	172	

* after correcting for outliers (see Section 3.2)
Means of each variable with standard deviations in parentheses in MPV, LPV, HYP and CONSEQ columns
Standard errors in parentheses in the 'difference' columns
MPV = money as payment vehicle
LPV = labour as payment vehicle
HYP = hypothetical treatment
CONSEQ = consequential treatment

Table B1.2: WTP_{labour} – WTP_{money} (in KSh)

	rate 1	rate 2	rate 3	rate 4	rate 5	rate 6
$WTP^{hyp}_{labour} - WTP^{hyp}_{money}$						
flour (ASC)	268	10	-31	-90	63	-17
	[215;321]	[-10;30]	[-48;-15]	[-104;-75]	[31;94]	[-42;7]
fortified	75	3	-9	-25	17	-6
	[54;95]	[-3;8]	[-14;-4]	[-32;-19]	[2;32]	[-18;5]
$WTP^{conseq}_{labour} - WTP^{conseq}_{money}$						
flour (ASC)	106	-9	-28	-54	21	-17
	[87;125]	[-17;-2]	[-35;-21]	[-61;-47]	[8;35]	[-28;-6]
fortified	30	-3	-8	-15	6	-5
	[22;38]	[-5;0]	[-10;-5]	[-19;-11]	[-2;13]	[-11;0]
$WTP^{hyp}_{labour} - WTP^{conseq}_{money}$						
flour (ASC)	320	62	21	-38	115	35
	[266;373]	[44;79]	[9;33]	[-45;-31]	[86;143]	[14;55]
fortified	90	17	6	-11	32	8
	[67;112]	[12;23]	[2;9]	[-14;-8]	[18;45]	[-1;17]

95% confidence interval in square brackets

CHAPTER 2

DO NON-MONETARY PRICES REDUCE HYPOTHETICAL BIAS?

2.1 ABSTRACT

It is well established that stated preference studies are vulnerable to hypothetical bias. This paper uses a discrete choice experiment to investigate two competing claims in the literature: (i) using a non-monetary payment vehicle reduces hypothetical bias; (ii) using a non-monetary payment vehicle amplifies hypothetical bias. Respondents are randomly assigned to a treatment where choices are either hypothetical or consequential and the payment vehicle is either monetary or non-monetary (labour). This 2 x 2 design allows us to detect differences in hypothetical bias between the two payment vehicles. We find evidence supporting the second claim: hypothetical bias is 26 to 31 percentage points higher when respondents are asked to pay with labour instead of money.

The discrete choice experiment concerns industrial fortification of flour, with the attributes including whether the flour is fortified (a pressing issue in Kenya where 24% of the population suffer from malnutrition). Our results demonstrate high demand for increased access to nutrient rich food in the form of industrially fortified maize flour thus encouraging investments in food fortification programmes.

2.2 INTRODUCTION

This paper concerns hypothetical bias (HB) in stated preference studies (Penn and Hu, 2018; Penn and Hu, 2019). The primary research question it addresses is whether HB is reduced by the use of non-monetary prices. We report the results of a discrete choice experiment (DCE) in Kenya incorporating both hypothetical and consequential treatments, using both monetary and non-monetary (labour) prices. The design allows the scale of HB, using both labour and monetary payment vehicles (PVs), to be estimated and the competing claims regarding the impact of non-monetary PVs on HB to be evaluated.

The results contribute to the literatures on HB and on the impact of the choice of PV on welfare estimates. There is a growing literature investigating and/or advocating the use of non-monetary PVs. This has been primarily in relation to low-income communities in developing countries (see for example Gibson et al., 2016; Hagedoorn et al., 2020; Meginnis et al., 2020; Navrud and Vondolia, 2020; Pondorfer and Rehdanz, 2018; Rai and Scarborough, 2015) but also includes high-income countries where volunteering time has been used as a PV (e.g. Ando et al., 2020; Davies et al., 2014; Durán-Medraño et al., 2017; Lankia et al., 2014). Within this literature, competing claims have been made that (i) "*the problem* [i.e. using monetary prices to measure welfare in developing countries] *may be further amplified though increased hypothetical biases*" (Gibson et al., 2016, p. 698), and (ii) "*hypothetical bias is more of a problem when cost is expressed in terms of time rather than money*" (Ando et al., 2020, p. 14).

Ando et al. (2020) conclude their paper, which features monetary and volunteer time contributions, by recommending future investigation of the PV-HB issue via a DCE study with a field study component in which time commitments are not hypothetical. This paper uses a design of the kind proposed: respondents were randomly assigned to one of four treatments where decisions were either hypothetical or consequential and the PV was either money or labour. We do so regarding an enhanced-health foodstuff in a developing country. The paper is (to our knowledge) the first investigation of the impact of the choice of PV (money or labour payment) on HB.

Non-monetary PVs, primarily labour, have been advocated for use in developing countries, in place of monetary payments. The argument typically put forward is that a significant part of consumption of low-income groups in developing countries comes from subsistence production, or barter, and that the use of monetary prices will underestimate willingness to contribute:

> *An important, yet overlooked aspect of these WTP studies is that cash is the only form of payment examined, even though these cash payments are unlikely to be meaningful in partially monetized economies typical of rural areas of the developing world*
>
> (Abramson et al., 2011, p. 2)

This argument is supported by a finding in several contingent valuation studies that labour contributions prompt fewer zero bids than monetary contributions contributions (e.g. Asrat et al., 2004; Brouwer et al., 2009; Echessah et al., 1997; Hung et al., 2007; Kamuanga et al., 2001; Swallow and Woudyalew, 1994). Many studies in which willingness to work (WTW) is monetised *ex post* generate willingness to pay (WTP) values (often far) greater than those generated using a standard, monetary PV (e.g. Abramson et al., 2011; Casiwan-Launio et al., 2011; Hagedoorn et al., 2020; Rai et al., 2015; Vasquez, 2014). HB is a possible cause of these higher valuations when WTW rather than WTP is used (see discussions in Ando et al., 2020; Diafas et al., 2017; Eom and Larson, 2006; Tilahun et al., 2017; Vondolia et al., 2014). In contrast, Gibson et al. (2016) speculate that if contingent monetary markets are less realistic in subsistence economies then the use of labour payments might reduce HB. This paper seeks to provide evidence to inform evaluation of these competing claims.

The case study is a DCE concerning fortified flour in Kenya. A challenge in undertaking a consequential DCE is that the multi-attribute multi-level framework leads to a high number of potential composite goods which must be available to respondents *ex post*. This serves to both keep the number of DCE-HB studies in developing countries low and for those conducted to have few attributes and levels. For example Chowdhury et al., 2011 and Meenakshi et al., 2012 both feature a single non-cost attribute (variety of maize or potato). This simplicity may be problematic since there is likely to be greater confidence in inferences drawn from such studies the more closely they resemble typical DCEs and, ultimately, real market transactions. This study includes a relatively high number of non-cost attributes (three) and levels yielding a total of 16 product types. While this poses practical problems (the 16 flour types had to be produced by the researchers) we contend that it means the DCE more closely mimics both real markets and typical DCEs to which our results speak.

Two of the three non-cost attributes were familiar to the respondents while the third attribute (flour fortification) was unfamiliar to 85% of our sample prior to the study. Empirical evidence in developed countries suggests that familiarity with the good (or attribute) reduces HB (e.g. Schläpfer and Fischhoff, 2012). We test for the impact of attribute-familiarity on HB in a developing country context.

A final objective of the study is to investigate the substantive issue of the value of maize flour fortification; a significant issue since almost 25% of the Kenyan population suffer from malnutrition. The EU has contributed €3.2 million to a 7 year project (European Commission, 2017) to strengthen the capacity of mills to fortify flour and thereby improve access to nutrient rich food in Kenya. Such food fortification programmes are now common with USAID having funded the SPRING industrial fortification project (2012-2017) in Uganda and the UK Department for International Development (DFID) funding a £66 million project to reduce undernourishment in Pakistan involving the installation of micro-feeders for fortification (DFID, 2019). These investments concern the supply side; the aim of this study is to investigate the demand side. Investments in fortification should be demand-driven because such investments will only be effective if consumers are willing to pay a price sufficient to maintain and run the donated equipment as well as source the micronutrients required.

In summary, first we test whether labour contributions, as opposed to monetary payments, reduce HB. Second, we test for the impact of attribute-familiarity on HB. Third, we test for differences in error variance across treatments. These hypotheses are investigated via a DCE field study, with consequential treatments, conducted in Kenya concerning the value of industrially fortified flour.

The rest of the paper is structured as follows. Section 2.3 provides a description of the case study and field sites. Section 2.4 presents the experimental design, recruitment process and survey procedures. Section 2.5 describes the modelling framework. Section 2.6 defines the testable hypotheses. Section 2.7 presents the empirical results. Section 2.8 concludes with a discussion.

2.3 CASE STUDY

Malnutrition is a significant problem in Kenya where 11.7 million people (24% of the population) suffer from undernourishment (FAO, 2018). The effects of malnutrition include vitamin A deficiency (the leading cause of blindness in children worldwide (UNICEF, 2019) and iron-deficiency anaemia (the most common type of anaemia) (WHO, 2014).

The Government of Kenya has introduced a number of initiatives to reduce undernutrition (and realise Sustainable Development Goal 2)[9] including industrial fortification of packaged wheat and maize flour (Government of Kenya, 2012). Industrial food fortification[10] refers to the process of enriching processed foods with micronutrients such as iron, zinc, folic acid, vitamin A and B vitamins (WHO and FAO, 2006). It is widely regarded as a cost-effective means to improve nutritional intake without requiring people to change their food habits (Horton et al., 2008).

A number of studies have examined preferences for bio-fortified foods both in Kenya and other developing countries (see e.g. Birol et al., 2011; Chowdhury et al., 2011; De Groote and Kimenju, 2008; De Groote et al., 2011; González et al., 2009; Meenakshi et al., 2012) but there is no study, to the best of our knowledge, estimating the premium, that Kenyan consumers are willing to pay for industrially fortified maize flour. Maize accounts for about 36% of total calories consumed (Mohajan, 2014) meaning that the addition of nutrients to maize flour has large potential public health implications.

Although the Kenyan Government has introduced mandatory industrial fortification of packaged wheat and maize flour it is common practice amongst small-scale farmers (who produce more than 60% of the country's agricultural output) to bring crops to small-scale local ('posho') mills for processing, thereby bypassing fortification. Posho mills typically process whole kernel maize (to produce 'unsifted' flour) while packaged flour is produced from maize after removal of the husk and the germ ('sifted' flour). Posho mills lack the capacity to fortify.

Respondents were recruited from two low-income communities in Kenya – Kibera and Nyamira/Kisii. Kibera is an urban slum on the edge of Nairobi where a majority of the population (who have emigrated from all over Kenya) live in small rented shacks (see Figure A2.1 in Appendix A2). Most residents are self-employed or casual labourers without regular incomes. Nyamira and Kisii are neighbouring counties in Western Kenya in which the majority of the population live in rural areas with small land holdings growing food crops (e.g. maize, bananas, beans and potatoes) for subsistence (see

[9] Target 2.2: "*By 2030, end all forms of malnutrition, including achieving, by 2025, the internationally agreed targets on stunting and wasting in children under 5 years of age, and address the nutritional needs of adolescent girls, pregnant and lactating women and older persons*." (United Nations, 2015)

[10] Industrial fortification is the addition of nutrients during the processing of crops as opposed to bio-fortification which refers to the process of increasing nutrient levels in crops during plant growth.

Figure A2.2 in Appendix A2). Unemployment rates are high (above 50%) in both Nyamira, Kisii and Kibera.

2.4 EXPERIMENTAL DESIGN AND PROCEDURE

2.4.1 SURVEY DESIGN

We designed a DCE with three non-price attributes and a price attribute (either money or labour) as specified in Table 2.1. The last column in Table 2.1 specifies the variable names used in the results section. The design yielded 4x2x2 = 16 different product types. As two of the treatments were consequential, the researchers produced and bagged the flour types using ingredients from various outlets (including one outlet in Tanzania as unsifted, fortified maize flour is unavailable in Kenya) (see Figures A2.3 and A2.4 in Appendix A2).

Table 2.1: Attributes

Attributes	Levels	variable names
crop	100% maize	maize
	50% maize, 50% sorghum	sorghum
	50% maize, 50% cassava	cassava
	50% maize, 50% millet	millet
sifted	yes	sifted
	no	
fortified	yes	fortified
	no	
price	40; 60; 80; 100; 120; 140; 160; 180; 200; 220 (KSh)	money
	0.25; 0.5; 0.75; 1; 1.5; 2; 2.5; 3; 3.5; 4 (hours)	labour

Respondents were randomly assigned to one of four treatments (see Table 2.2) where choices were either hypothetical or consequential and the PV was either money or labour.

Table 2.2: Treatments

	Hypothetical	Consequential
Money	WTP^{HYP}	WTP^{CONSEQ}
Labour	WTW^{HYP}	WTW^{CONSEQ}

WTP = willingness to pay
WTW = willingness to work
HYP = hypothetical
CONSEQ = consequential

Each respondent completed 8 choice tasks. Examples of choice tasks with monetary and labour prices, respectively, can be seen in Appendix A2, Figures A2.5 and A2.6. Ngene software (ChoiceMetrics, 2018) was used to generate a D-efficient fractional factorial design using priors derived from a pilot study. Constraints were included to avoid dominated alternatives and the Modified Federov algorithm was applied to generate 80 choice situations which were blocked into 10 blocks.

The survey comprised the following sections: (1) a general introduction; (2) a conversation about food fortification to elicit the respondent's understanding prior to the survey and to subsequently ensure that the respondent knows what food fortification is and is aware of common arguments for and against it; (3) introduction to the DCE including two practice rounds; (4) eight choice tasks and follow-up questions if the respondent selected either 'flour' or 'no flour' in all the tasks; (5) rolling of an eight-sided dice in the consequential treatments to determine the binding choice task; (6) sociodemographic questions and questions about food preferences; (7) thank you and completion of voucher with details about flour and/or gift collection details.

2.4.2 RECRUITMENT

Respondents were approached in public places (on the street, in cafes, at market places) in March and April 2019 by trained, local, enumerators who spoke the local languages (see the participation information sheet in Appendix A2.8, pre-screening survey in Appendix A2.9, interview procedure in Appendix A2.10 and debriefing note in Appendix A2.11). The survey interview was conducted in person on tablets using an offline survey app (Lighthouse Studio, Sawtooth Software). To avoid surveying respondents who had received information by people who had already participated, the enumerators would not accept requests to be surveyed and frequently changed their location each day.

2.4.3 SURVEY PROCEDURES

In the consequential treatments, respondents were informed (prior to the choice tasks) that one of the eight choices would be randomly selected (after the choice tasks) to be binding by rolling an eight sided dice. It was explained at the outset that:

- Respondents who selected the no purchase option in the binding choice task would not be purchasing any flour.
- Respondents who selected one of the purchase options in the binding choice task would be expected to purchase the selected flour at the specified price (money[11] or labour).

The labour price was specified as an amount of time that the respondent would commit to perform a task. The task was described as the sorting of seeds (see Figure A2.7 in Appendix A2) – a task regarded as not requiring excessive strength, skill or education (seed sorting was successfully used as a non-monetary PV in Hoffmann, 2018).

It was explained that payment (with money or labour) and collection of the flour would take place in the days following the experiment. The location would be a well-known venue in the local area (e.g. a school or community centre) within 15 minutes walking distance of the experiment location. More than one time/day was offered including one weekday and either a Saturday or Sunday. All participants were entitled to a small (non-monetary) gift for taking part. To keep transaction costs (i.e. travel and time costs associated) consistent across all treatments the gift was only collectable at the same days/times/locations as the work/payment for selected flour options. Steps were taken to ensure only the respondent could work/pay for their chosen flour and/or collect the gift.

2.5 MODELLING FRAMEWORK

2.5.1 UTILITY SPECIFICATION

The analysis of choice data from the DCE is based on random utility theory. We assume respondent n chooses the flour j that yields the highest utility in any given choice task. Each flour is described by a set of observed attributes X_j accompanied by a vector of preference parameters β indicating the marginal utility of the attributes (Hensher et al., 2015). As shown in Table 2.1, we have K = 3 attributes describing the characteristics of the flour (represented by A_{jk}) and a price attribute (either money M_j or labour L_j). The

[11] Unlike most consequential DCEs, e.g. Aoki et al. (2010) and Johansson-Stenman and Svedsäter (2012), we asked respondents to make out-of-pocket payments. This avoids potential issues relating to the endowment effect (Loureiro et al., 2003).

utility for respondent *n* from flour *j* is then specified as a linear and additively separable function of the attributes:

$$U_{nj} = \beta X_j + \varepsilon_{nj} = \sum_{k=1}^{K} \beta_k A_{jk} + \beta_{money} M_j + \beta_{labour} L_j + \varepsilon_{nj} \qquad \text{Eq. 1}$$

where ε_{nj} is a random error term capturing factors that are outside the model such as measurement error, β_{money} is the marginal utility of money and β_{labour} is the marginal utility of labour. Willingness to pay (WTP) and willingness to work (WTW) for attribute *k* is then defined as the marginal utility of the attribute relative to (the negative of) the marginal utility of money and labour, respectively:

$$WTP = -\frac{\beta_k}{\beta_{money}} \text{ ; } WTW = -\frac{\beta_k}{\beta_{labour}} \qquad \text{Eq. 2}$$

2.5.2 (HETEROSCEDASTIC) CONDITIONAL LOGIT MODEL

The distribution of the random term ε_{nj} is assumed to be independently and identically distributed following a type 1 extreme value distribution. This assumption leads to the conditional logit (CL) model (McFadden, 1974) in which the probability of respondent *n* choosing flour *i* over flour *j* is given by:

$$Pr_{ni} = Pr\left(\beta X_i + \varepsilon_{ni} > \beta X_j + \varepsilon_{nj}\right) \ \forall\ j \neq i \in J = \frac{\exp(\lambda_n \beta X_i)}{\sum_{j=1}^{J} \exp(\lambda_n \beta X_j)} \qquad \text{Eq. 3}$$

In the homoscedastic model the scale parameter λ, which is inversely related to the variance of the random term ($\lambda = \pi/\sqrt{6\text{var}[\varepsilon]}$), is assumed to be constant with λ normalised to 1. The heteroscedastic conditional logit (HCL) model allows for unequal error variance across individuals, treatments etc. (DeShazo and Fermo, 2002; Hensher et al., 1999; Hole, 2006).

In the HCL model, the scale term is a function of individual or treatment-specific characteristics Z_n and parametrised as $\exp(\theta Z_n)$ which ensures that λ_n always is positive and that the model nests the homoscedastic model when $\theta = 0$.

2.5.3 GENERALISED MULTINOMIAL LOGIT MODEL

Heterogeneity in both preferences and error variance can be accommodated in infinite mixture (mixed logit) models, for example Fiebig et al. (2010)'s generalised multinomial logit (G-MNL) model. In the G-MNL model, the utility obtained by respondent *n* from flour *j* is given by:

$$U_{nj} = \lambda_n \beta X_j + \gamma \eta_n X_j + (1-\gamma)\lambda_n \eta_n X_j + \varepsilon_{nj} \qquad \text{Eq. 4}$$

where β is the vector of mean utility weights in the population and η_n is respondent n's deviation from the mean thus capturing residual taste heterogeneity. λ_n captures scale heterogeneity and is parametrised as $\exp(\bar{\lambda} + \theta Z_n + \tau\varepsilon_0)$ where $\bar{\lambda}$ is a normalising constant (Gu et al., 2013), θZ_n has the same interpretation as in the HCL model, ε_0 is a standard normally distributed scalar and τ indicates the degree to which the vector of preference parameters is scaled proportionately across individuals.

The parameter γ in Eq. 4 specifies how the variance of residual taste heterogeneity varies with scale. Fiebig et al. (2010) impose γ to be between 0 and 1 where $\gamma = 0$ implies that η_n is proportional to λ_n whereas $\gamma = 1$ indicates that η_n is independent of scale. Following Keane and Wasi (2013) we will allow γ to take any value (Gu et al., 2013).

2.6 HYPOTHESES

We test a number of hypotheses which relate to the impact of the PV upon HB. First, we test if the magnitude of HB differs when the PV is money (HB^{money}) versus when the PV is labour (HB^{labour}).

Hypothesis 1 H_0: $HB^{money} = HB^{labour}$

H_A: $HB^{money} \neq HB^{labour}$

This is motivated by the increasing use of non-monetary PVs and the recurrent finding that the monetised value of the WTW values generated are greater than the equivalent WTP measures generated with a monetary PV (e.g. Hagedoorn et al., 2020). Most studies use either individual-specific or area-average wage rates, or a fraction thereof, to monetise WTW and compare to WTP values (e.g. Abramson et al., 2011;

Arbiol et al., 2013; Casiwan-Launio et al., 2011; Gibson et al., 2016; Tilahun et al., 2015; Vasquez, 2014). Lloyd-Smith et al. (2019), however, find that the opportunity cost of leisure time (which they estimate on survey data using a stochastic payment card elicitation technique) is only weakly correlated with the wage rate. Lew and Larson (2005) further argue that the value of time is stochastic and that the wage-based approach, which assumes a constant value of time, therefore is biased. Following Larson et al. (2004), we adopt an alternative approach where, instead of imposing a (somewhat arbitrary) conversion rate, the marginal utilities of time and money are estimated jointly (see Eq. 1 in Section 2.5.1).

Second, we test for the impact of attribute-familiarity on HB.

Hypothesis 2A H_0: $HB^{money}_{familiar} = HB^{money}_{unfamiliar}$

H_A: $HB^{money}_{familiar} \neq HB^{money}_{unfamiliar}$

Hypothesis 2B H_0: $HB^{labour}_{familiar} = HB^{labour}_{unfamiliar}$

H_A: $HB^{labour}_{familiar} \neq HB^{labour}_{unfamiliar}$

Previous studies have investigated the effect of experience with the good being valued on HB. In a meta-analysis, Schläpfer and Fischhoff (2012) find that experience with the good reduces HB. In a study about bio-fortified potatoes in Uganda, Chowdhury et al. (2011) similarly finds that the magnitude of HB is minimal for the familiar potato varieties while WTP for the unfamiliar varieties are overstated by a factor of more than two. List (2001), in contrast, finds no significant difference in the degree of HB for dealers and non-dealers of sports cards in an experimental auction (but he finds that a cheap talk script successfully eliminates HB for inexperienced participants). To inform the debate about the impact of familiarity on HB, we conduct pairwise tests of equality of HB between the unfamiliar attribute *fortified* and the familiar attributes *maize*, *sorghum*, *cassava*, *millet* and *sifted*. We perform these tests using both money (Hypothesis 2A) and labour (Hypothesis 2B) to test whether any HB-familiarity effect is modified by the choice of PV.

Third, we test for differences in error variance across treatments.

Hypothesis 3 H_0: error variance is constant across treatments

H_A: error variance is *not* constant across treatments

In their hypothetical DCE studies, Vondolia and Navrud (2018) find that non-monetary PVs (labour and in-kind) exhibit higher levels of error variance whilst Gibson et al. (2016) find no difference in error variance between money and labour treatments. To inform the debate about the impact of the PV on error variance, we test for equality of the scale term across the two subgroups using labour and money as the PV, respectively. The scale term is parametrised as $\exp(\theta Z_n)$ (see Section 2.5.2) so a test for $\theta=0$ is a test for error variance being constant across treatments. We further test for equality of the scale term across hypothetical and consequential treatments which allows us to evaluate if it is the hypothetical/consequential nature of the DCE or the type of PV that contributes more to increased error variance.

2.7 RESULTS

Descriptive statistics are shown in Table 2.3. The average respondent is 34 years old and living in a household with 5.6 household members including 2 children. An equal number of men and women were interviewed. The mean daily income is 569 KSh but as indicated by the large standard deviation, the level of income varies greatly across respondents. The median daily income (not reported in Table 2.3) is 300 KSh. More than half of the respondents thus live on less than 3 USD a day (at the time of the experiment, 1 USD = 101 KSh). High school is the highest level of education completed for over half of the sample while about a quarter of the sample have completed primary school only as their highest level of education. Food fortification, which is the attribute of interest, is unfamiliar to a majority of the respondents as only 3% of the sample has a good understanding of what it means while the remaining sample has either a vague understanding (12%) or no understanding (84%).

Table 2.3: Descriptive statistics

Female (1 if female, 0 otherwise)	0.50
	(0.50)
Age (years)	34
	(13)
Daily income (KSh)	569
	(1048)
Household size	5.6
	(3.7)
Children (<16 years) in household	2.0
	(1.9)
Highest level of education completed (%)	
No formal qualifications	0.01
Primary school	0.28
High school	0.52
Vocational training	0.14
Higher education	0.05
How familiar is the respondent with food fortification (%)	
Never heard about it	0.59
Heard about it but does not know what it is	0.25
Has a vague understanding of what it is	0.12
Has a good understanding of what it is	0.03
Number of respondents	
Rural	202
Urban	134

Means of each variable with standard deviations in parentheses unless stated otherwise
* p<0.10; ** p<0.05; *** p<0.01

The study design included work or payment for flour, and this work/payment (and/or the collection of the participation gift) occurred on subsequent days to the experiment. The degree of attrition, and its variation over treatments, is therefore of interest and shown in Table 2.4.

Table 2.4: Attrition

	sample size	buy	opt-out	attrition (no show)	flour collected	flour not collected
labour, hypothetical	82			3		
money, hypothetical	89			4		
labour, consequential	91	56	35	5	53	3
money, consequential	92	44	48	6	39	5
total	354			18	92	8

Of the 91 respondents in the consequential treatment with labour as the PV, 56 respondents selected a flour purchase option in the binding choice task and were thus expected to collect the selected flour in exchange for the designated hours of work. Five of the 91 did not attend of which 3 were due to work (the other two were expected to collect only the gift).

The values in Table 2.4 indicate that the logistical arrangements (see Section 2.4.3) led to low attrition rates in all treatments. In total, 18 of 354 respondents did not collect their flour and/or gift. To maintain the consequential nature of the responses, the 18 no-shows are excluded from the analysis thus reducing the sample size to 336 (336 respondents x 8 choice tasks = 2688 observations). We conclude that attrition rates are low and of a similar scale across treatments, occurring also in the hypothetical treatments.

In the empirical analysis, we pool the data from the four treatments and estimate a number of choice models to investigate treatment effects and test the hypotheses articulated above. The three non-price attributes are dummy-coded while the price attributes (money and labour) are treated as continuous variables. Missing values (monetary prices when labour is the PV and labour "prices" when money is the PV) are coded with a zero value to enable joint estimation of the marginal utilities of labour and money. The parameter estimates are presented in Tables 2.5 and 2.6.

CL-1 in Table 2.5 is a conditional logit model including only the flour attributes (see Table 2.1). No alternative-specific constant is included as estimation of all the four attribute levels describing the type of crop preclude estimation of an alternative-specific constant. As expected, the utility weights for maize, sorghum, cassava and millet are positive while the utility weights for the two price attributes are negative. Respondents thus prefer a bag of flour (irrespective of the type of crop) compared to the no purchase option and dislike higher prices. The sifted and fortified parameters indicate that the average effect is indifference with respect to the processing of the flour (sifted or unsifted) but the average effect is positive regarding fortification of the flour.

Table 2.5: Choice models (without treatment-specific interaction terms)

	CL-1	HCL-1
maize	1.641***	2.856***
	(0.109)	(0.233)
sorghum	1.689***	2.866***
	(0.100)	(0.219)
cassava	1.443***	2.464***
	(0.108)	(0.220)
millet	1.740***	3.017***
	(0.114)	(0.243)
sifted	-0.122*	-0.267**
	(0.070)	(0.113)
fortified	0.569***	0.749***
	(0.069)	(0.114)
money (KSh)	-0.020***	-0.033***
	(0.001)	(0.002)
labour (hours)	-1.367***	-2.710***
	(0.055)	(0.212)
Scale θ		
MPV x HYP		-0.760***
		(0.108)
LPV x HYP		-1.225***
		(0.127)
LPV x CONSEQ		-0.307***
		(0.095)
LL	-2257	-2186
Observations	2688	2688

Standard errors in parentheses
* p<0.10; ** p<0.05; *** p<0.01
MPV = money as payment vehicle
LPV = labour as payment vehicle
HYP = hypothetical
CONSEQ = consequential

HCL-1 is a heteroscedastic conditional logit model which allows investigation of scale differences across treatments (Carlsson and Johansson-Stenman, 2010). Respondents in the consequential treatment (CONSEQ) with monetary PV (MPV) are used as the reference group (i.e. $\lambda_{\text{MPV-CONSEQ}}$ is normalised to 1). As λ_n is parametrised as $\exp(\theta Z_n)$ (see Section 2.5.2), a test for θ=0 is a test for error variance being constant across treatments. If θ is significantly smaller (greater) than 0, then λ_n is significantly smaller (greater) than 1. As shown in Table 2.5, the estimated scale parameter θ is negative and highly significant for all the treatment groups indicating that respondents in the reference group (MPV-CONSEQ) have lower error variance (since λ is inversely related to the variance of the error term). Respondents thus appear to be making more consistent choices when choices are consequential and when the PV is money. When comparing the magnitude of the scale terms, it appears that the hypothetical nature of the DCE contributes more to increased error variance than a labour PV since $\theta_{\text{MPV-HYP}}$ is greater (in absolute value) than $\theta_{\text{LPV-CONSEQ}}$ (a test of equality gives $\chi^2 = 13.35$).

Table 2.6: Choice models (with treatment-specific interaction terms)

	HCL-2	G-MNL-1
HYP x maize	1.950***	12.178***
	(0.192)	(2.581)
HYP x sorghum	2.012***	11.689***
	(0.187)	(2.475)
HYP x cassava	1.908***	11.789***
	(0.198)	(2.506)
HYP x millet	2.030***	12.016***
	(0.204)	(2.578)
HYP x sifted	-0.172*	-0.521
	(0.105)	(0.415)
HYP x fortified	0.753***	1.073**
	(0.106)	(0.527)
HYP x money	-0.018***	-0.122***
	(0.001)	(0.026)
HYP x labour	-1.294***	-9.519***
	(0.106)	(1.978)
CONSEQ x maize	2.706***	46.189***
	(0.241)	(11.618)
CONSEQ x sorghum	2.670***	43.047***
	(0.216)	(10.647)
CONSEQ x cassava	2.193***	40.168***
	(0.225)	(10.253)
CONSEQ x millet	2.878***	46.644***
	(0.249)	(11.923)
CONSEQ x sifted	-0.241*	-3.535
	(0.130)	(2.456)
CONSEQ x fortified	0.479***	3.965**
	(0.127)	(1.852)
CONSEQ x money	-0.033***	-0.527***
	(0.002)	(0.128)
CONSEQ x labour	-2.760***	-50.486***
	(0.210)	(11.692)

continued on next page

Scale θ		
LPV	-0.316***	-0.440***
	(0.083)	(0.109)
τ		1.526***
		(0.108)
γ		0.647***
		(0.097)
standard deviations		
HYP x maize		3.946***
		(0.764)
HYP x sorghum		4.158***
		(0.804)
HYP x cassava		3.746***
		(0.688)
HYP x millet		4.561***
		(0.858)
HYP x sifted		2.833***
		(0.526)
HYP x fortified		5.077***
		(0.951)
CONSEQ x maize		11.361***
		(3.010)
CONSEQ x sorghum		11.687***
		(3.036)
CONSEQ x cassava		10.230***
		(3.366)
CONSEQ x millet		12.027***
		(3.050)
CONSEQ x sifted		6.917***
		(1.753)
CONSEQ x fortified		9.592***
		(2.519)
LL	-2123	-1710
Observations	2688	2688

Standard errors in parentheses
* p<0.10; ** p<0.05; *** p<0.01
LPV = labour as payment vehicle
HYP = hypothetical
CONSEQ = consequential

In HCL-2 and G-MNL-1, we estimate treatment-specific utility weights (rather than additive interaction terms) for the hypothetical (HYP) and the consequential (CONSEQ) subsamples, respectively. In G-MNL-1, we further estimate correlated random parameters for all non-price attributes to test for taste heterogeneity. We used 500 draws and multiple starting points to ensure convergence at the global maximum (Gu et al., 2013). The final model is estimated using starting values based on estimation of an uncorrelated G-MNL model.

In G-MNL-1 all mean parameters, with the exception of *sifted*, are significant and with the expected signs. In both the hypothetical and the consequential treatments, respondents prefer a bag of flour (maize, sorghum, cassava or millet) over the no purchase option, they obtain positive utility from flour that is fortified (although *HYP x fortified* and *CONSEQ x fortified* are significant at the 5% level only) and they dislike higher prices both in terms of money and labour. As in Table 2.5, the average respondent appears to be indifferent with regards to whether the flour is sifted or unsifted.

The scale parameter τ is positive and significant in G-MNL-1 confirming scale heterogeneity in the data. This is in addition to differences in scale between the two PVs. The estimated scale parameter θ is negative and highly significant ($p < 0.01$) in both HCL-2 and G-MNL-1 for the LPV group (i.e. the subgroup of respondents who were presented with choice tasks where labour was used as the PV). As in HCL-1 respondents thus appear to be making more consistent choices (the error variance is lower because $\lambda_{LPV} = \exp(-0.440) = 0.644 < 1$) when the PV is money.

The parameter γ is estimated at 0.647 which implies that the variance of residual taste heterogeneity is varying less than proportionally with scale. Finally, standard deviations for the random parameters are all significant which means that respondents have heterogeneous preferences with respect to the non-price attributes.

2.7.1 WILLINGNESS TO PAY AND WILLINGNESS TO WORK

In Table 2.7, we calculate WTP and WTW using the utility weights from G-MNL-1 in Table 2.6 and the delta method is utilized to estimate standard errors (Hole, 2007a). Estimates of WTP and WTW are calculated separately for respondents in the hypothetical and consequential treatments and HB is defined as the ratio of the two (Murphy et al., 2005). We calculate WTP, WTW and HB for each of the 16 composite products generated by our study design. The monetary values in the WTP columns are essentially estimated market prices for the different types of flour (2kg bags). The difference in HB when the PV is labour versus when the PV is money and the associated standard errors (in parentheses) are presented in the last column of Table 2.7. As shown, HB is 26 to 31 percentage points higher when the PV is labour and this difference is significant for all flour combinations at least at the 5% level.

Table 2.7: WTP, WTW and HB by product type

	WTP			WTW			
	HYP	CONSEQ	HB^{money}	HYP	CONSEQ	HB^{labour}	HB^{labour} - HB^{money}
maize							
unsifted, unfortified	100	88	1.14**	77	55	1.40***	0.26**
	(5)	(3)	(0.07)	(5)	(4)	(0.13)	(0.10)
unsifted, fortified	108	95	1.14**	84	60	1.40***	0.26**
	(5)	(4)	(0.07)	(5)	(5)	(0.14)	(0.11)
sifted, unfortified	95	81	1.18**	73	51	1.45***	0.27***
	(4)	(4)	(0.07)	(4)	(2)	(0.11)	(0.10)
sifted, fortified	104	89	1.18***	80	55	1.45***	0.27**
	(4)	(3)	(0.05)	(5)	(3)	(0.11)	(0.11)
sorghum							
unsifted, unfortified	96	82	1.17**	74	51	1.44***	0.27***
	(5)	(3)	(0.07)	(4)	(3)	(0.12)	(0.10)
unsifted, fortified	104	89	1.17**	80	56	1.44***	0.27**
	(5)	(3)	(0.07)	(5)	(4)	(0.13)	(0.11)
sifted, unfortified	91	75	1.22**	70	47	1.50***	0.28***
	(5)	(4)	(0.09)	(5)	(2)	(0.11)	(0.10)
sifted, fortified	100	83	1.21***	77	52	1.49***	0.28***
	(5)	(3)	(0.07)	(5)	(2)	(0.11)	(0.11)
cassava							
unsifted, unfortified	96	76	1.27***	74	48	1.56***	0.29**
	(4)	(3)	(0.07)	(4)	(4)	(0.14)	(0.12)
unsifted, fortified	105	84	1.26***	81	52	1.55***	0.29**
	(6)	(4)	(0.09)	(5)	(5)	(0.16)	(0.12)
sifted, unfortified	92	70	1.33***	71	44	1.63***	0.31***
	(4)	(3)	(0.08)	(4)	(2)	(0.12)	(0.11)
sifted, fortified	101	77	1.31***	78	48	1.61***	0.30***
	(5)	(3)	(0.09)	(5)	(3)	(0.14)	(0.12)
millet							
unsifted, unfortified	98	89	1.11	76	55	1.37***	0.26**
	(5)	(3)	(0.07)	(5)	(4)	(0.12)	(0.10)
unsifted, fortified	107	96	1.11	83	60	1.37***	0.26**
	(6)	(4)	(0.07)	(6)	(5)	(0.14)	(0.10)
sifted, unfortified	94	82	1.15**	72	51	1.41***	0.26***
	(4)	(3)	(0.07)	(4)	(2)	(0.10)	(0.10)
sifted, fortified	103	89	1.15**	79	56	1.42***	0.27***
	(5)	(2)	(0.06)	(5)	(3)	(0.11)	(0.10)

Standard errors in parentheses
* p<0.10; ** p<0.05; *** p<0.01
WTP is in KSh and WTW is in minutes
In the HB^{PV} columns, we test H_0: $HB^{PV} = 1$

Figure 2.1 shows calculations of WTP and WTW along with their 95% confidence interval for each of the attributes. WTP and WTW for the attributes maize, sorghum, cassava and millet are for 2kg bags of unsifted, unfortified flour and these estimates are thus equivalent to the estimates in Table 2.7 rows 1, 5, 9 and 13, respectively. Respondents in the hypothetical treatment with money as the PV are, on average, WTP more for all four crop types although the difference is insignificant for millet (at the 10% level) and significant at the 5% level only for maize and sorghum as indicated by the overlapping confidence intervals (see Figure 2.1a). Similarly, as shown in Figure 2.1b, the average respondent in the hypothetical treatment with labour as the PV is WTW more for all four crop types and this difference is significant at the 1% level for both maize, sorghum, cassava and millet. With the exception of millet when money is used as the PV, we thus find substantial evidence of HB. There is, however, no significant difference in WTP and WTW between hypothetical and consequential treatments for the attributes *sifted* and *fortified* and thus no evidence of HB, for the average respondent, when considering these two attributes in isolation.

Figure 2.1: WTP and WTW by attribute

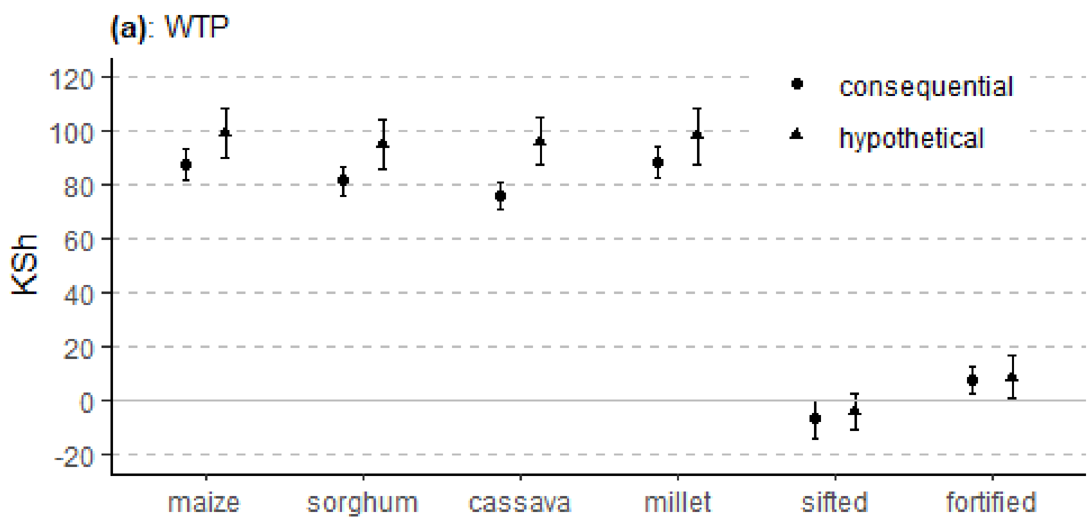
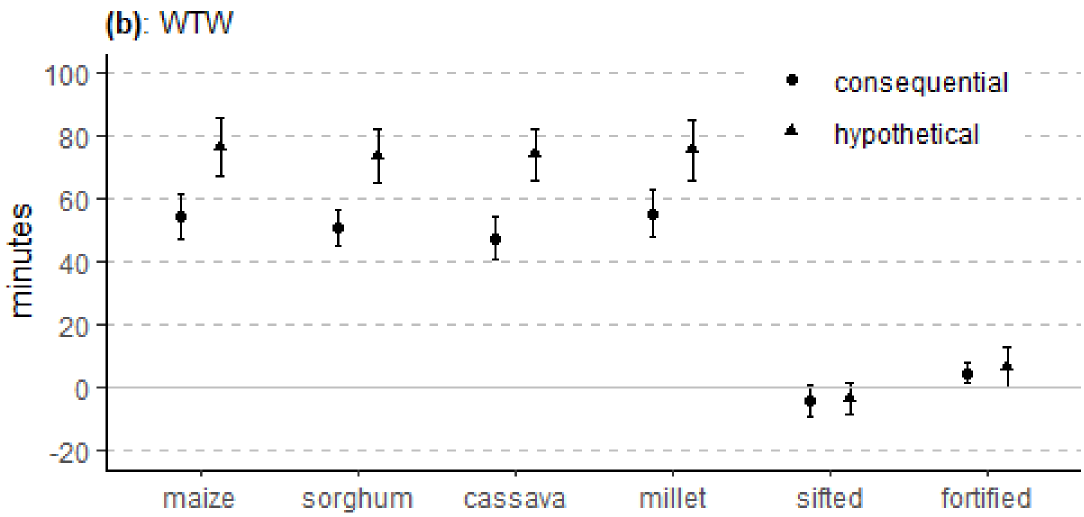


Figure 2.1: WTP and WTW by attribute (*continued from previous page*)



We further test for equality of HB across attributes (Hypotheses 2A and 2B) but reject all pairwise tests, for both PVs, between the unfamiliar attribute *fortified* and each of the familiar non-cost attributes (*maize*, *sorghum*, *cassava*, *millet*, *sifted*). Table 2.8 presents HB calculated for each non-cost attribute (familiar and unfamiliar) along with the associated 95% confidence intervals. As shown, the confidence intervals for the unfamiliar attribute *fortified* and each of the familiar non-cost attributes are overlapping for both PVs. Attribute familiarity thus seems to have no effect on the magnitude of HB.

Table 2.8: HB by attribute

	MPV	LPV
maize	1.14	1.40
	[1.00;1.27]	[1.15;1.64]
sorghum	1.17	1.44
	[1.03;1.31]	[1.21;1.67]
cassava	1.27	1.56
	[1.13;1.40]	[1.29;1.82]
millet	1.11	1.37
	[0.97;1.25]	[1.13;1.60]
sifted	0.64	0.78
	[-0.62;1.89]	[-0.79;2.35]
fortified	1.17	1.44
	[-0.11;2.44]	[-0.19;3.06]

95% confidence interval in square brackets
LPV = labour as payment vehicle
MPV = money as payment vehicle

2.7.2 WILLINGNESS TO PAY FOR FORTIFIED FLOUR

The average respondent in the consequential treatment with money as the PV is WTP a premium of 8 KSh for 2kg fortified flour. At the time of the study, the market price for maize flour was 85–110 KSh per 2 kg flour depending on the season, weather events, brand (for packaged flour) and perceived quality of the crops (for unpackaged flour). The average respondent is thus WTP a premium of about 8% of the market price for fortified flour.

Perhaps of greater value than an average effect is information on the potential market share at different price premiums. To this end, we use respondent-specific coefficients derived from the G-MNL model in Table 2.6 (G-MNL-1) to produce individual-level estimates of WTP for the attribute *fortified*. In Table 2.9, we report the market penetration at different price points. The results are reported for the consequential treatment with money as the PV in aggregate and disaggregated by location (rural and urban).

Table 2.9: Share of respondents WTP for the attribute *fortified*

	% WTP		
premium	all	rural	urban
0 KSh	0.63	0.66	0.59
5 KSh	0.57	0.60	0.51
10 KSh	0.43	0.46	0.37
15 KSh	0.33	0.36	0.28
20 KSh	0.27	0.27	0.27
25 KSh	0.23	0.24	0.22
30 KSh	0.19	0.20	0.18

The results in Table 2.9 demonstrate high levels of preference heterogeneity for fortified flour (as also evident from the significant standard deviations in G-MNL-1, Table 2.6). 63% of the respondents in the consequential treatment with money as the PV are WTP a premium for fortified flour. Over half of the sample would buy fortified flour with a price premium of 5 KSh or less and one third of the respondents are WTP at least 15 KSh premium per 2 kg fortified flour. It is further evident from Table 2.9 that our results suggest greater market penetration for rural respondents compared to urban respondents at all price points.

2.8 DISCUSSION AND CONCLUSIONS

This paper has reported the design and implementation of a consequential multi-attribute DCE field study concerning fortified flour in Kenya. Unlike previous consequential DCEs in developing countries, we included a high number of attributes and levels yielding a total of 16 composite products. While our design came with some practical challenges around having to make the composite products available to respondents *ex post*, we argue that there is greater confidence in inferences drawn from a multi-attribute consequential DCE which more closely resembles typical (hypothetical) DCEs and market exchange.

To investigate the impact of the choice of the PV on HB, respondents were randomly assigned to one of four treatments where purchase decisions were either hypothetical or consequential and the PV was either money or labour. This is (to our knowledge) the first study examining the magnitude of HB in a DCE with labour contributions and the first study investigating the impact of the choice of PV (money or labour prices) on HB. Our study yields several findings that are relevant for researchers using DCEs as well as for aid agencies and policymakers in developing countries.

We find that respondents, on average, are WTP approximately 8% of the market price for fortified flour. More than half of the rural respondents (66%) are WTP a premium for having their flour fortified. This is a relevant finding for aid agencies who fund food fortification programmes aimed at improving access to nutrient rich food amongst the rural poor. Our results demonstrate high demand for fortified flour in rural (and urban) communities thus encouraging efforts to strengthening the capacity of posho mills to fortify flour.

When comparing hypothetical and consequential purchase choices, we find that consumers in Kenya overstate their WTP for 14 out of 16 flour types by 14 to 33 percent. This is lower than the median level of HB of 39 percent reported in a recent meta-analysis by Penn and Hu (2018) and much lower than other HB studies in developing countries. Alemu and Olsen (2018) find that participants in Kenya are WTP three to six times more when purchase decisions are hypothetical compared to when purchase decisions are consequential. Similarly, Chowdhury et al. (2011) find that respondents in Uganda overstate their WTP by a factor of up to nearly 5 for bio-fortified potatoes. They find that the scale of HB is lower for more familiar potato varieties – a finding consistent with a meta-analysis by Schläpfer and Fischhoff (2012) concluding that experience with the product reduces HB. We find, however, no statistically significant difference in the scale of HB between the familiar attributes (*maize*, *sorghum*, *cassava*, *millet*, *sifted*) and the unfamiliar attribute (*fortified*) for either payment vehicle (PV). This suggests that attribute (un)familiarity has no effect on the magnitude of HB and we thus fail to reject Hypotheses 2A and 2B. Of course a categorisation of attributes as (un)familiar is crude and the degree of familiarity will differ across respondents within a study and across attributes between studies.

We compute HB as the ratio of consequential to hypothetical contributions (money or labour) for each of the 16 flour types generated by our study design and find that HB is 26 to 31 percentage points higher when the PV is labour compared to when the PV is money. A common finding in the stated preference literature in developing countries (see Hagedoorn et al., 2020 for a recent review) is that the monetised value of hypothetical labour contributions exceeds estimated WTP. A possible cause put forward e.g. by Ando et al. (2020) is that labour payments cause higher levels of HB. We find evidence supporting this claim thus rejecting our Hypothesis 1 – HB is consistently higher when consumers are asked to contribute labour instead of money and the difference is significant at least at the 5% level. This finding has important implications for researchers using stated preference methods in developing countries because monetary payments seem to outperform labour payments with respect to HB and thus serve as a cautionary note for those advocating for the use of labour as PV.

When looking at choice consistency across treatments, we find further support for the use of monetary payments. Respondents make more consistent choices when the PV is money which means we reject Hypothesis 3 – error variance varies across treatments. This is consistent with a DCE study in Ghana where Vondolia and Navrud (2018) find that non-monetary PVs (labour and in-kind) exhibit higher degrees of uncertainty. Gibson et al. (2016), in contrast, find no difference in choice consistency between labour and monetary treatments in a DCE study about improved drinking water quality in Cambodia.

The PV-HB tests in a developing country setting are in part motivated by the argument that labour as the PV is particularly well suited to more subsistence-oriented economies. Given this literature, we would have liked to test for variations in HB, using different PVs, as a function of the extent of respondents' integration with monetised economic activities. In particular, whether the magnitude of HB, using money and labour as PVs, differs between an urban area (Kibera, Nairobi) and a mixed subsistence-cash local economy in a rural area (Nyamira/Kisii). While we find a numerical difference in HB between the two PVs for the rural subsample of similar magnitude to the aggregate sample, estimates are too imprecise (likely due to sample size issues and associated standard errors) and the difference is only (very) weakly significant ($0.10 < p < 0.11$) for all 16 flour types (the results are presented in Appendix B2). Investigation of such differences, via studies with the necessary sample sizes and statistical power, seems worthy of consideration.

The primary research question addressed in this paper is whether HB is reduced by the use of non-monetary prices. This is a pertinent question given the increasing use of labour as PV in stated preference studies. We find that HB is higher when the PV is labour compared to when a monetary PV is used. This suggests that the recurrent finding that the monetised value of hypothetical labour contributions exceeds estimated WTP is (at least partly) due to differences in HB. How much of the gap, identified in many papers, between monetised WTW and WTP is due to differences in HB cannot be resolved by one study. Given that papers (see e.g. Hagedoorn et al., 2020) have found monetised labour values up to 20 times larger than WTP, a difference of a different order of magnitude to the HB difference identified in this paper, there are likely to be additional factors contributing to this difference between WTP values derived from the two PVs.

A2 APPENDIX: SURVEY INFORMATION

Figure A2.1: Kibera

Figure A2.2: Kisii

Figure A2.3: Flour production

Figure A2.4: Research assistants and produced flour bags

Figure A2.5: Sample choice task (labour as the payment vehicle)

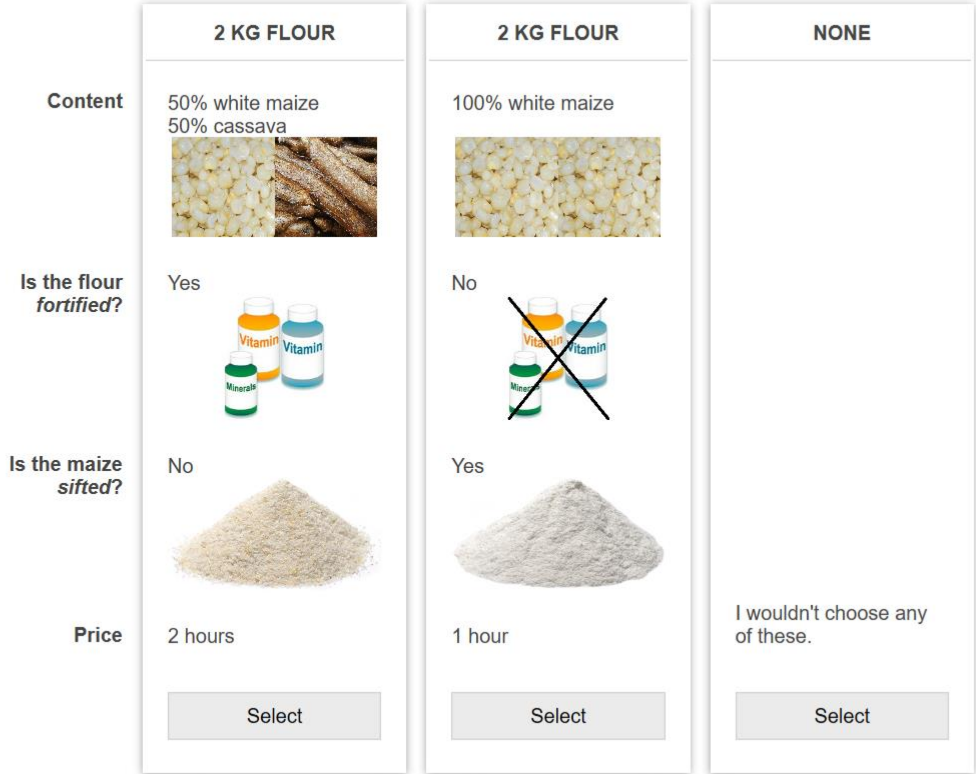


Figure A2.6: Sample choice task (money as the payment vehicle)

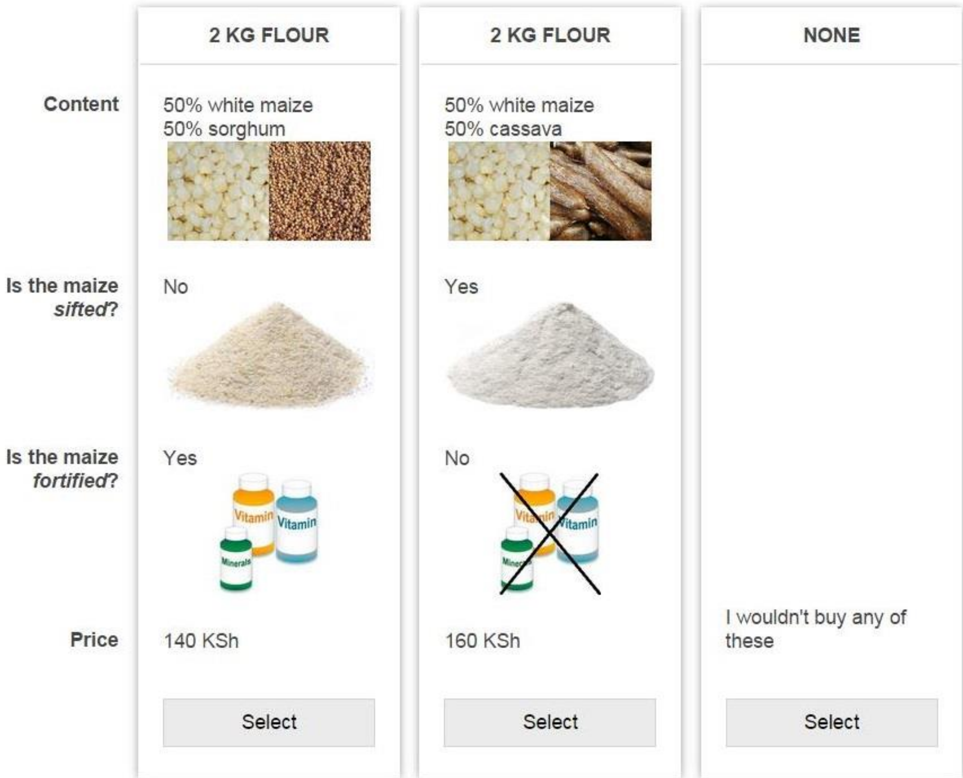


Figure A2.7: Seed sorting

Appendix A2.8: Participation information sheet

This PIS should be read in conjunction with the University of Manchester Privacy Notice.

You are being invited to take part in a research study that is part of a student project. Before you decide whether to take part, it is important for you to understand why the research is being conducted and what it will involve. Please ask if there is anything that is not clear or if you would like more information.

Who will conduct the research?
This research is conducted by Camilla Knudsen (PhD student in the School of Social Sciences at the University of Manchester) in collaboration with her supervisors Prof Dan Rigby and Dr Prasenjit Banerjee

What is the purpose of the research?
To investigate what kind of ugali people prefer.

Who takes part in the survey?
About 200 adults from Kibera Sub-County and about 200 adults from Nyamira County

Why have I been invited to participate?
You have been invited to participate because your answers to the pre-screen questionnaire indicate that you meet the inclusion criteria for the study. You have been approached following a systematic random sampling strategy.

What would I be asked to do if I took part?
In the survey you will be asked

- questions about you and your household
- questions about how much and what kind of ugali you eat
- to make some choices about different types of flour

You may be invited to take part in a follow-up event but this is completely voluntary and you do not have to decide if you want to participate in the follow-up event at this stage.

What will happen to my personal information?
Data will be securely stored in electronic format at the University of Manchester and only the research team will have access to this information. Your data will be identified by a unique survey number, and not by your name. However, for practical reasons, you will be asked to show your personal ID and we will note down your name on a piece of paper along with your unique survey number. This piece of paper will be destructed as soon as the follow-up event has taken place. You will also be asked if would be happy to provide your mobile number which we will use to send you at most two reminders about the follow-up event. However, if you wish not to receive any reminders then you will not have to provide your mobile number. If you do submit your mobile number then it will be kept safe along with your name and the details will be destroyed as soon as the follow-up event has taken place.

What happens if I do not want to take part or if I change my mind?
It is up to you to decide whether or not to take part. If you do decide to take part you will be given this information sheet to keep and be asked to give verbal consent. If you decide to take part you are still free to withdraw at any time without giving a reason and without detriment to yourself. However, it will not be possible to remove your data from the project once it has been anonymised and forms part of the dataset as we will not be able to identify your specific data. This does not affect your data protection rights.

Will my data be used for future research?
When you agree to take part in a research study, the information about you may be provided to researchers running other research studies in this organisation. This information will not identify you and will not be combined with other information in a way that could identify you.

Will I be paid to participate in the research?
No but you will be able to select a small gift from a list of available gifts. The selected gift will have to be collected at [location] between [time] and [time] on [date], [date] or [date].

What is the duration of the research?
If you agree to participate, it will take approximately 20 minutes to complete the survey. The follow-up event, if you are invited and if you agree to attend, will take up to 4 hours.

Will the outcomes of the research be published?
The results may be presented at scientific conferences or in a journal article.

What if I want to make a complaint?
If you have a minor complaint then you need to talk to the interviewer or contact the researchers in the first instance (see contact details below).

Camilla Knudsen
Telephone: +254 727 365 308
Email: camilla.knudsen@manchester.ac.uk

Dan Rigby
Telephone: +44 (0) 161 275 4808
Email: dan.rigby@manchester.ac.uk

If you wish to make a formal complaint or if you are not satisfied with the response you have gained from the researchers in the first instance then please contact the Research Governance and Integrity Manager (see contact details below).

The Research Governance and Integrity Manager
Research Office
Christie Building
University of Manchester

Oxford Road
Manchester
M13 9PL
Email: research.complaints@manchester.ac.uk
Telephone: +44 (0) 161 275 2674

Appendix A2.9: Pre-screening survey

- How old are you?

The participant should be at least 18 years old

- How often do you eat ugali?

The participant must be consuming ugali at least a couple of times per month.

- Do you visit this area frequently?

The participant must either live nearby (i.e. within a walking distance of 15 minutes) or visit the area frequently (i.e. at least once a week) because of for example work or to visit relatives.

Appendix A2.10: Interview procedure

You should always wear the name tag provided and carry the following with you:

- Tablet (always keep the tablet in your bag until you begin the survey)
- Permission letter from local authorities
- Copies of the Participant Information Sheet
- Duplicate book for creating vouchers + writing tools
- Designated folder for storing vouchers until collected by or delivered to the researcher

Recruiting participants

You should approach participants (according to the strategy agreed with researchers) and do the following:

- Greet the person
- State your name
- Tell the person that you are looking to interview him/her about ugali
- Tell the person that you have a short pre-screen questionnaire that will determine if the he/she is eligible to take part
- Tell the person that if he/she is eligible then you will provide them with further information about the survey and he/she can decide if they want to take part

If a person agrees to answer the pre-screen questionnaire, take him/her through the questions. If a person does not satisfy the criteria of one of the screening questions, briefly explain to the person why he/she is ineligible and thank the person for his/her time.

If a person fulfils the criteria of the pre-screen questionnaire, give the person a copy of the PIS (to keep) and ask him/her if he/she prefers to read it on his/her own or if he/she prefers to have it read aloud by the interviewer. Ask the person if he/she has any questions. Allow the person to take up to 30 minutes, if needed, to consider if he/she wants to take part while you recruit/interview someone else.

<u>Interviewing participants</u>

You should guide the participant through the survey (following the procedures discussed during the training sessions) and record answers on the tablet. You should always allow the participant to see the screen of the tablet (also if the person is illiterate).

When you write the voucher (either for a gift only or for a gift and collection/payment of flour), ask the person for his/her ID. Add the person's full name on the voucher. If the person asks, repeat the message from the PIS about what will happen to the voucher. Tell the person that he/she has to collect the flour (and/or gift) himself/herself. Give back the ID and keep our copy of the voucher in the designated folder.

<u>Collection/payment of flour and collection of participation gift</u>

When a person arrives to collect and pay/work for flour and/or collect the gift, confirm the details on the voucher with our copy of the voucher and double-check the name with the person's ID. Keep collected vouchers in the designated folder (participants should not keep them).

If a participant is collecting a gift only, give him/her the selected gift and a copy of the debriefing note. Ask the participant if he/she wishes to have the note read aloud. Read the note aloud, if required. Thank the participant for his/her participation.

If a participant is also collecting and paying/working for flour, give him/her the selected gift and ask the following:

"You agreed to work/pay X hours/KSh for a bag of 2 kg flour (add specifications of the flour). Are you still willing to work/pay this amount of hours/KSh to receive that bag of flour?"

If the participant has changed his/her mind about the purchase, give him/her a copy of the debriefing note. Ask the participant if he/she wishes to have the note read aloud. Read the note aloud, if required. Thank the participant for his/her participation and add the following note on the voucher: "opt out".

If the participant continues to be willing to pay/work X KSh/hours, do the following:

- If the participant committed to pay a certain amount, collect the agreed amount and give the flour. Give him/her a copy of the debriefing note. Ask the participant if he/she wishes to have the note read aloud. Read the note aloud, if required. Thank the participant for his/her participation and add the following note on the voucher: "collected"
- If the participant committed to work a certain amount of hours, initiate the work task and keep time (practical details to be determined on the ground). Ensure that the participant is aware that he/she will have to work the total amount of agreed time in order to get the flour. Give respondents the opportunity to leave and come back (the time should be paused while the respondent is away). Once the agreed

amount of time is completed, give the flour. Give him/her a copy of the debriefing note. Ask the participant if he/she wishes to have the note read aloud. Read the note aloud, if required. Thank the participant for his/her participation and add the following note on the voucher: “collected”

Appendix A2.11: Debriefing note

Thank you for having taken the time to participate in this research. You can read more about the aims of the study below, if you are interested.
You will have been making choices about different types of flour that had either a monetary price (i.e. a certain amount of KSh that you would be required to pay) or a non-monetary price (i.e. a certain amount of time that you would be required to work). In addition, you will have been asked to make either hypothetical choices (i.e. where no actual purchase takes place) or non-hypothetical choices (i.e. where one of the choices is selected at random to be enforced and an actual purchase may take place).
Participants were randomly assigned to one of these different versions of the survey.

What is the aim of the study?
The aim of the survey is twofold:
1. To see if people are willing to pay/work a premium for flour that has been fortified.
2. To see if there is a difference between the average willingness to pay/work between people who answered the hypothetical version and people who answered the non-hypothetical version (i.e. to see if people in the hypothetical version overestimate their willingness to pay/work when their answers are purely hypothetical).

What if I have questions or want to make a complaint?
If you have any questions about the research or if you have a minor complaint then you need to contact the student researcher in the first instance (see contact details below).

Camilla Knudsen
Telephone: +254 727 365 308
Email: camilla.knudsen@manchester.ac.uk

If you wish to make a formal complaint or if you are not satisfied with the response you have gained from the researchers in the first instance then please contact the Research Governance and Integrity Manager (see contact details below).

The Research Governance and Integrity Manager
Research Office
Christie Building
University of Manchester
Oxford Road
Manchester
M13 9PL
Email: research.complaints@manchester.ac.uk
Telephone: +44 (0) 161 275 2674

B2 APPENDIX: ESTIMATION WITH SITE-SPECIFIC INTERACTION TERMS

We investigate whether the magnitude of HB, using money and labour as PVs, differs between an urban area (Kibera, Nairobi) and a mixed subsistence-cash local economy in a rural area (Nyamira/Kisii). These PV-HB tests between rural and urban populations are in part motivated by the argument that labour as the PV is particularly well suited to more subsistence-oriented economies. The urban-rural split is, however, not just important in relation to the investigation of the effects of different PVs but in the consideration of HB in general. For example, an important consideration when applying benefit-transfer techniques is whether the presence and scale of HB is stable across locations. Ehmke et al. (2008) argue that most studies make 'panhuman' assumptions when focusing on economic and sociodemographic drivers of HB, paying little attention to cultural differences. They tested this in a multinational HB contingent valuation study and found significant differences in the scale of HB between China, France, Niger and the United States, thus rejecting (untested) assumptions about the panhuman nature of HB. The analysis presented below extends this logic further by testing pan-national assumptions about the scale of HB.

In Table B2.1, we treat *sifted* and *fortified* as random variables and interact all attributes with site-specific indicator variables. This enables us to estimate WTP, WTW and HB for the rural and the urban subgroup separately (see Table B2.2 and Figure B2.1).

Table B2.1: Site-specific G-MNL model

	G-MNL-2
HYP x maize x rural	6.514***
	(1.696)
HYP x sorghum x rural	6.358***
	(1.625)
HYP x cassava x rural	6.178***
	(1.596)
HYP x millet x rural	7.313***
	(1.905)
HYP x sifted x rural	-1.558***
	(0.525)
HYP x fortified x rural	2.303***
	(0.723)
HYP x money x rural	-0.068***
	(0.017)
HYP x labour x rural	-4.811***
	(1.236)
HYP x maize x urban	7.405***
	(2.058)
HYP x sorghum x urban	7.184***
	(1.952)
HYP x cassava x urban	7.598***
	(2.050)
HYP x millet x urban	7.096***
	(1.951)
HYP x sifted x urban	0.057
	(0.499)
HYP x fortified x urban	0.025
	(0.643)
HYP x money x urban	-0.066***
	(0.017)
HYP x labour x urban	-7.367***
	(2.030)
CONSEQ x maize x rural	14.568***
	(3.907)
CONSEQ x sorghum x rural	16.154***
	(4.227)
CONSEQ x cassava x rural	13.746***
	(3.577)
CONSEQ x millet x rural	17.272***
	(4.512)
CONSEQ x sifted x rural	-2.342**
	(0.916)
CONSEQ x fortified x rural	1.508*
	(0.816)

continued on next page

CONSEQ x money x rural	-0.210***
	(0.053)
CONSEQ x labour x rural	-17.744***
	(4.647)
CONSEQ x maize x urban	7.237***
	(1.854)
CONSEQ x sorghum x urban	5.589***
	(1.431)
CONSEQ x cassava x urban	5.687***
	(1.436)
CONSEQ x millet x urban	6.493***
	(1.706)
CONSEQ x sifted x urban	0.188
	(0.419)
CONSEQ x fortified x urban	0.416
	(0.468)
CONSEQ x money x urban	-0.071***
	(0.018)
CONSEQ x labour x urban	-7.367***
	(1.852)
Scale θ	
LPV	-0.436**
	(0.189)
τ	1.423***
	(0.139)
γ	0.414***
	(0.119)
Standard deviations	
HYP x sifted x rural	1.668***
	(0.416)
HYP x fortified x rural	3.422***
	(0.741)
HYP x sifted x urban	2.114***
	(0.654)
HYP x fortified x urban	3.600***
	(1.001)
CONSEQ x sifted x rural	2.481***
	(0.837)
CONSEQ x fortified x rural	3.908***
	(1.068)
CONSEQ x sifted x urban	1.422***
	(0.530)
CONSEQ x fortified x urban	2.055***
	(0.554)
LL	-1842
Observations	2688

Standard errors in parentheses
* p<0.10; ** p<0.05; *** p<0.01

Table B2.2: WTP, WTW and HB by product type and by site

	WTP			WTW			
	HYP	CONSEQ	HB^{money}	HYP	CONSEQ	HB^{labour}	HB^{labour} - HB^{money}
RURAL							
maize							
unsifted, unfortified	96	69	1.39***	81	49	1.65***	0.26
	(6)	(5)	(0.14)	(7)	(3)	(0.18)	(0.16)
unsifted, fortified	130	76	1.70***	110	54	2.02***	0.32
	(8)	(5)	(0.16)	(9)	(4)	(0.22)	(0.20)
sifted, unfortified	73	58	1.26	62	41	1.50**	0.24
	(6)	(6)	(0.16)	(7)	(4)	(0.21)	(0.14)
sifted, fortified	107	65	1.64***	91	46	1.95***	0.31
	(8)	(5)	(0.19)	(8)	(4)	(0.25)	(0.19)
sorghum							
unsifted, unfortified	94	77	1.22**	79	55	1.45***	0.23*
	(5)	(5)	(0.10)	(7)	(3)	(0.14)	(0.14)
unsifted, fortified	128	84	1.52***	108	60	1.81***	0.28
	(8)	(5)	(0.13)	(9)	(4)	(0.19)	(0.17)
sifted, unfortified	71	66	1.08	60	47	1.28	0.20
	(7)	(5)	(0.14)	(7)	(3)	(0.17)	(0.12)
sifted, fortified	105	73	1.44***	89	52	1.71***	0.27
	(8)	(6)	(0.16)	(9)	(4)	(0.22)	(0.17)
cassava							
unsifted, unfortified	91	65	1.40***	77	46	1.66***	0.26*
	(5)	(4)	(0.12)	(6)	(2)	(0.16)	(0.16)
unsifted, fortified	125	72	1.73***	106	52	2.05***	0.32
	(9)	(5)	(0.17)	(9)	(3)	(0.23)	(0.20)
sifted, unfortified	68	54	1.26	58	39	1.49**	0.24*
	(7)	(5)	(0.17)	(7)	(3)	(0.21)	(0.14)
sifted, fortified	102	61	1.67***	86	44	1.98***	0.31
	(9)	(5)	(0.21)	(9)	(4)	(0.27)	(0.19)
millet							
unsifted, unfortified	108	82	1.32***	91	58	1.56***	0.25*
	(6)	(5)	(0.11)	(8)	(2)	(0.14)	(0.15)
unsifted, fortified	142	89	1.59***	120	64	1.89***	0.30*
	(9)	(5)	(0.13)	(10)	(3)	(0.19)	(0.18)
sifted, unfortified	85	71	1.20*	72	50	1.42***	0.22*
	(6)	(5)	(0.12)	(7)	(3)	(0.15)	(0.14)
sifted, fortified	119	78	1.52***	101	56	1.81***	0.28
	(8)	(5)	(0.15)	(9)	(4)	(0.20)	(0.17)

continued on next page

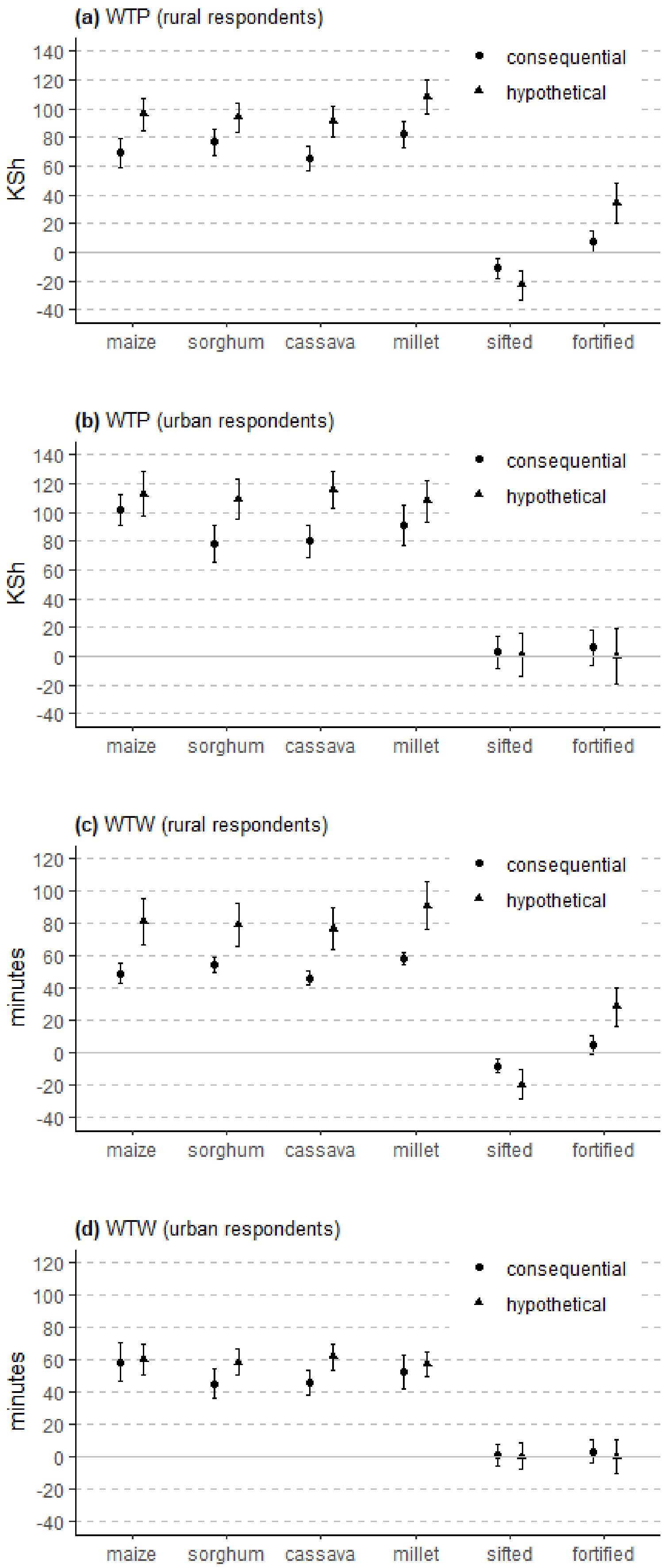
URBAN							
maize							
unsifted, unfortified	113	102	1.11	60	59	1.02	–0.08
	(8)	(5)	(0.10)	(5)	(6)	(0.13)	(0.13)
unsifted, fortified	113	108	1.05	61	62	0.97	–0.08
	(10)	(7)	(0.12)	(7)	(7)	(0.15)	(0.12)
sifted, unfortified	114	105	1.09	61	60	1.00	–0.08
	(9)	(7)	(0.11)	(6)	(7)	(0.15)	(0.12)
sifted, fortified	114	110	1.03	61	64	0.95	–0.08
	(12)	(9)	(0.14)	(8)	(7)	(0.17)	(0.12)
sorghum							
unsifted, unfortified	109	79	1.39***	59	46	1.29*	–0.11
	(7)	(7)	(0.15)	(4)	(5)	(0.17)	(0.16)
unsifted, fortified	110	85	1.30*	59	49	1.20	–0.10
	(11)	(8)	(0.18)	(7)	(6)	(0.19)	(0.15)
sifted, unfortified	110	81	1.36**	59	47	1.25	–0.10
	(9)	(8)	(0.17)	(6)	(5)	(0.19)	(0.16)
sifted, fortified	111	87	1.27	59	50	1.17	–0.10
	(13)	(10)	(0.21)	(8)	(6)	(0.22)	(0.14)
cassava							
unsifted, unfortified	116	80	1.45***	62	46	1.34**	–0.11
	(7)	(6)	(0.13)	(4)	(4)	(0.14)	(0.17)
unsifted, fortified	116	86	1.35**	62	50	1.25	–0.10
	(10)	(9)	(0.18)	(6)	(5)	(0.19)	(0.16)
sifted, unfortified	117	83	1.41**	62	48	1.30*	–0.11
	(8)	(8)	(0.17)	(5)	(5)	(0.18)	(0.16)
sifted, fortified	117	89	1.32	63	51	1.22	–0.10
	(12)	(10)	(0.21)	(8)	(6)	(0.22)	(0.15)
millet							
unsifted, unfortified	108	91	1.18	58	53	1.09	–0.09
	(7)	(7)	(0.12)	(4)	(5)	(0.13)	(0.14)
unsifted, fortified	109	97	1.11	58	56	1.03	–0.08
	(11)	(9)	(0.15)	(6)	(6)	(0.16)	(0.13)
sifted, unfortified	109	94	1.16	58	54	1.07	–0.09
	(8)	(8)	(0.14)	(5)	(6)	(0.15)	(0.13)
sifted, fortified	109	100	1.09	58	58	1.01	–0.08
	(13)	(10)	(0.17)	(8)	(7)	(0.18)	(0.12)

Standard errors in parentheses
* p<0.10; ** p<0.05; *** p<0.01
WTP is in KSh and WTW is in minutes
In the HB^{PV} columns, we test H_0: $HB^{PV} = 1$

From Table B2.2 and Figure B2.1, it appears that HB is driven primarily by the rural subsample. HB for the urban subsample is insignificant ($p > 0.05$) for 11 of the 16 flour types when money is used as the PV and for 15 out of 16 flour types when labour is used as the PV. The rural subsample, on the other hand, exhibits levels of HB that are statistically significant ($p < 0.05$) for almost all flour types using both PVs (12 and 15 out of 16 flour types when money and labour is used as the PV, respectively). Rural areas (with low degrees of marketisation) thus seem to exhibit higher levels of HB compared to urban areas (with high degrees of marketisation) when the PV is money, as hypothesised by Gibson et al. (2016), but this marked difference in HB persists when labour is used as the PV. Our results suggest that differences in HB between urban and rural populations are not explained by varying degrees of marketisation i.e. while rural respondents may be less familiar with money as a PV compared to urban respondents this is not the main driver of HB. Our results further suggest that the scale of HB is far from stable across locations. HB thus has implications not only for stated preference practitioners but also for policy makers worldwide who rely on benefit-cost analysis to support their decision making.[12]

[12] See for example the requirement for environmental impact assessments in the European Union as outlined in Directive 2011/92/EU and Directive 2001/42/EC; in Australia as outlined in the *Environment Protection and Biodiversity Conservation Act 1999; and in the United States as outlined in Chapter 2 of the Environmental Protection Agency's Guidelines for Preparing Economic Analyses.*

Figure B2.1: WTP and WTW by attribute and by site



CHAPTER 3

DO NON-MONETARY PRICES FAVOUR WOMEN?

3.1 ABSTRACT

This paper investigates the effects of using a non-monetary price attribute (labour) versus a monetary price attribute in a split-sample discrete choice experiment concerning improved water infrastructure in rural India. As well as sample average effects we test for gender effects in the context of gendered control of financial assets in rural India. Respondents were randomly assigned to one of two treatments using either money or work as the payment vehicle (PV). At the sample level, we find no significant effect of the PV on marginal utilities suggesting that aggregate welfare measures and configuration of optimal water projects is unaffected by the choice of the PV. At a more disaggregate level, however, we find that women are willing to *work more* but *pay less* for improvements in water provision compared to men. In societies where it is predominantly men who earn and control financial resources, this result can be attributed to gender differences in the opportunity cost of time. We argue, however, that 'willingness to pay' and 'willingness to work' as measures of welfare are likely to assign unequal social importance to men and women which suggests that the choice of PV is non-trivial and necessarily based on value judgments.

3.2 INTRODUCTION

Stated preference (SP) studies are increasingly used to estimate economic values of non-market goods and services in developing countries. Bennett and Birol (2010) argue that such advancement in public policymaking is particularly important in developing countries given that the effect of improved decision making is likely to be greater compared to developed countries. This study adds to the literature that estimates non-market values in the developing world.

The focus of the study is the value of improvements in water supply in rural India and specifically the relative value of different benefits which could flow from a national programme of water infrastructure investments. The results thus speak to the configuration of the projects within that programme.

The results presented and discussed are from a discrete choice experiment (DCE) designed and administered across nine rural Indian villages. To increase the realism of the DCE, we use a pivoted design in which the choice tasks are customised based on individual responses to a set of pre-DCE questions. As well as a focus on the substantive issue of the value of improvements in water supply, the study speaks to the methodological issue of the choice of payment vehicle (PV) in SP studies conducted with the rural poor in developing countries.

3.2.1 CONTEXT AND AIMS

In 2015, more than 132 million people in rural India did not have access to basic water services (World Bank, 2018c). The problem is particularly acute during the summer season when many natural water sources dry out. Water scarcity is a life threatening issue to millions of people in India and other developing countries yet there are more works related to rural connectivity completed under MGNREGA (Mahatma Gandhi National Rural Employment Guarantee Act – India's flagship national rural employment scheme) than works related to water conservation (Government of India, 2018). The aim of this study is to inform decision makers about the economic values of the benefits that water projects could deliver and also how that varies with different project configurations.

To estimate the economic value of different water attributes, a DCE is conducted in nine rural villages in Chhattisgarh and Odisha. To enable estimation of willingness to pay (WTP), DCEs often asks respondents to make trade-offs between nonmarket goods and money such that marginal rates of substitution can be calculated. It has been argued, however, that monetary PVs are inappropriate in rural developing settings because of the subsistence nature of many households in such communities (see for example Abramson et al., 2011; Gibson et al., 2016; Meginnis et al., 2020). Due to liquidity constraints and lack of experience in monetary markets, economic values are likely to be underestimated when rural households are asked about their willingness and ability to pay using money as the medium of exchange. To address this problem, researchers have adopted a range of non-monetary PVs of which the most popular is labour contributions. In line with this relatively new strand of literature, we randomly assigned respondents to one of two treatments using either money or work as the PV.

We jointly estimate the marginal utilities of time and money and find that there is no significant effect of the PV on the coefficients for the non-cost attributes which suggests that welfare measures can be estimated using either of the two PVs. However, further analysis shows that women, compared to men, are more payment sensitive when the PV is money and less payment sensitive when the PV is labour. From this result it can be inferred that the relative importance of time and money differs between men and women. The choice of PV is, therefore, controversial because willingness to pay (WTP) and willingness to work (WTW), as measures of welfare, are likely to assign different weight to the preferences of men and women. Choosing money or labour as the PV requires value judgment because it is essentially a choice between two different distributional weights. Given MGNREGA's commitment to empower women by ensuring high female participation, and the fact that MGNREGA projects are delivered using the labour of beneficiaries, the consideration of gender and labour as the means of payment for improvements in water provision is particularly pertinent.

The aim of this study is to: (1) investigate demand for improved water infrastructure in India, (2) compare estimates of WTP and WTW for those improvements and (3) examine gender differences in the marginal value of the improvements and how the choice of PV affects those differences.

3.3 LITERATURE REVIEW

3.3.1 NON-MONETARY PAYMENT VEHICLES

An increasing number of studies in the environmental valuation literature argue that monetary payment vehicles are inappropriate when a significant part of the population is engaged primarily in subsistence activities (see e.g. Hagedoorn et al., 2020; Kassahun et al., 2020; Meginnis et al., 2020). It is estimated that 70 percent of the nearly 900 million rural residents in India depend primarily on agriculture for their livelihood (UNFAO, 2018). Many of these households have limited access to off-farm labour and since 82 percent of the rural farmers are small and marginal, they have little or no opportunity for generating income. According to data by the World Bank, more than 25 percent of the rural population in India lived on less than 1.90 US dollars a day in 2011 (World Bank, 2018a). Trade involving money will be limited amongst such low-income households and the exchange of many goods and services is therefore likely to take place using alternative forms of transaction such as barter or work exchange. In this context, it is often argued that monetary estimates of WTP provide a poor measure of welfare. Surveys that ask liquidity-constrained respondents about their willingness (and ability) to pay using money as the medium of exchange risk underestimating the demand of the good or service under consideration. When considering their budget constraints, low-income households will be able to contribute very little or nothing for a good with high benefits. In a contingent valuation study, Brouwer et al. (2009) find, for example, that rural households in Bangladesh are willing to pay substantially less for flood risk protection compared to the costs of the damage. 60 percent of the respondents refused to contribute money to an embankment project but 40 percent of these respondents indicated a willingness to contribute either labour or in-kind instead. The use of a non-monetary payment vehicle in addition to (or instead of) a monetary one is therefore likely to reduce zero bids and increase willingness to contribute. Other contingent valuation studies finding higher levels of acceptability for labour contributions compared to cash contributions are Echessah et al. (1997), Kamuanga et al. (2001), Asrat et al. (2004) and Hung et al. (2007).

In a meta-analysis of 21 contingent valuation studies about water service improvements, Abramson et al. (2011) find that WTP is significantly lower in rural areas compared to urban areas as well as in smaller rural communities compared to larger rural communities. The authors argue that monetary exchange markets are limited in areas where population densities are low and differences in WTP estimates are therefore likely to reflect inability and potentially inexperience in providing cash payments rather than an absence of demand for improved water services. Given that the provision of water services in rural areas of developing countries is considerably behind that of urban areas (WHO/UNICEF, 2017), it does indeed seem unlikely that the demand for water service improvements should be lower in rural areas.

In an attempt to address the problems of using monetary PVs in low-income economies, researchers have adopted a range of non-monetary payment vehicles including crops such as rice and maize (Shyamsundar and Kramer, 1996; Sutton et al., 2008), everyday household items (Hossack and An, 2015) and meals to labourers (Diafas et al., 2017). However, the most common non-monetary PV in stated preference studies is labour contributions. All households are faced with decisions about how to allocate time between productive (paid or non-paid) and leisure activities which makes time contributions a popular alternative to money contributions. A respondent's willingness to allocate time towards receiving a certain good is, ceteris paribus, expected to increase with the perceived benefits of that good.

In order to assess welfare effects by comparing benefits and costs, many studies convert labour contributions into monetary contributions. Such task requires an estimate of the opportunity cost of contributed time. Most studies use either area or sample average wage rates or a fraction thereof (see e.g. Abramson et al., 2011; Casiwan-Launio et al., 2011; Arbiol et al., 2013; Vasquez, 2014). A common finding when using this approach is that the average respondent is willing to contribute more labour than money.

While the literature comparing monetary and non-monetary PVs is growing, only a few papers have addressed the impact of the PV on male and female respondents separately. Gender discrimination exists in many contexts in rural India for example in relation to property rights (Bhalotra et al., 2019), education (Kaul, 2018) and nutrition (Aurino, 2017). In recognition of this vulnerability, many projects and programmes include provisions aiming to empower women. MGNREGA stipulates, for example, that at least one-third of the beneficiaries should be women. Information about how PVs can affect women compared to men is relevant to decision makers who are looking to implement gender-equitable or female-favoured policies.

Existing studies that examine the impact of the PV across gender focus mostly on household heads. In an open-ended contingent valuation study amongst household heads in rural Kenya, Echessah et al. (1997) found that male-headed households, ceteris paribus, are willing to contribute more labour but less money for tsetse control compared to female-headed households. The authors do not explore this finding any further except in an analysis of the average male- and female- headed household, they note that the number of available adult labourers is lower in female-headed households. This observation is then used to calculate monetary contributions as a percentage of income and labour contributions as a percentage of available labour leading to the conclusion that female-headed households, mutatis mutandis, are more committed to tsetse control. Similarly, Tilahun et al. (2015) find that male-headed households, ceteris paribus, are more likely to accept higher bids in a contingent valuation study about forest conservation, especially in terms of labour, compared to female-headed households.

Larson et al. (2015) conduct a CV study in Botswana asking respondents to accept or decline wildlife conservation jobs with varying wages and varying time commitments. They estimate the (dis)utility associated with the work as well as the shadow value of time. The reservation wage (minimum wage required to accept the work) is the shadow value of time adjusted for the (dis)utility from the work activity. They find that women have higher disutility for almost all the different categories of work and a higher shadow value of time compared to men.

This paper provides the first study that compares monetary and non-monetary contributions across gender when a significant proportion (half) of the sampled population is not the head of their household. We demonstrate that this comparison is relevant and argue that the choice of PV is related to the long-standing discussion about equity concerns in cost-benefit analysis. According to the World Bank (2018b), the labour force participation of the female population in India was 28.7 percent in 2017 while that of the male population was 81.7 percent. Women are therefore less likely to have access to finances for personal consumption compared to men. As a result, women are likely to have a higher marginal utility of money than men because, at the margin, income is assumed to be more important to a poor individual than to a rich individual (Layard et al., 2008). If we accept this premise, then the discussion about equity concerns in cost-benefit analysis, and economic valuation in general, applies to gender as well.

3.3.2 EQUITY CONCERNS IN ECONOMIC VALUATION

Welfare effects are often assessed using cost-benefit analysis (CBA) which requires monetary estimates of both the costs and the benefits of a particular project or policy. Benefits are typically calculated as the sum of individuals' monetary-equivalents (i.e. the amount of money that will produce the same level of welfare as the project or policy in question). Elicitation of an individual's stated willingness to pay for a good (or service) is a popular approach amongst economists to estimating the monetary-equivalent for that good.

If the benefits of a project outweigh the costs, then the project is thought to be welfare improving from a utilitarian perspective. However, CBA has been extensively criticised for being insensitive to distributional considerations (see e.g. Blackorby and Donaldson, 1990 and Fleurbaey and Abi-Rafeh, 2016). The rationale for this criticism is typically based on differences in the marginal social value of income. The argument commonly put forward is that a dollar is worth more to a poor individual than to a rich individual. In a study using more than 200,000 observations of self-reported happiness from over 50 countries, Layard et al. (2008) confirm that the marginal utility of income indeed decreases as income increases. Thus, only in cases where money is equally important to all individuals on the margin, does the sum of WTP actually measure the sum of social benefits (Nyborg, 2014). In all other cases, more weight will be given to preferences of individuals with lower marginal utility of money.

To mitigate the issue of differences in the social marginal utility of income, a long-established scholarly literature argue that distributional weights should be incorporated into CBA (Adler, 2016). The concept of distributional weights is straightforward – the weights should reflect the marginal utility of income such that an individual's monetary equivalent can be adjusted to measure the impact on social welfare. Distributional weights play "*the same role for utility comparisons across individuals as the exchange rate plays for comparison of goods prices expressed in different currencies*" (Nyborg, 2014, p. 126). The implementation of distributional weights, however, is much less straightforward. It is difficult, if not impossible, to compare the importance of money, at the margin, across individuals. Such interpersonal comparisons involve value judgments which includes dealing with conflicting interests in terms of who is more or less deserving (Fleurbaey and Abi-Rafeh, 2016). However, as pointed out by many proponents of distributional weights, unweighted CBA is itself a value judgment. Individuals with lower marginal utility of income (most likely richer individuals) are considered more socially important in unweighted CBAs.

Unweighted CBAs rely on the criterion of Kaldor-Hicks efficiency. A project or policy is considered to be efficient if the 'winners' (i.e. those with positive monetary equivalents) can compensate the 'losers' (i.e. those with negative monetary equivalents) such that the project or policy is a potential Pareto improvement. Some scholars argue that such redistribution is better dealt with through taxation while other scholars argue that it is practically impossible (for an overview see Johansson-Stenman, 2005 and Nurmi and Ahtiainen, 2018). In this study, where the comparison of interest is not rich and poor per se but rather intra-household comparisons between men and women, it can be argued that redistribution is implausible because gender roles are socially prescribed rather than politically determined.

3.4 EXPERIMENTAL DESIGN AND PROCEDURE

A discrete choice experiment (DCE) was designed to elicit the willingness of rural households in India to contribute money or labour towards improved water availability. The study was conducted during the months of January and February 2018 in nine villages (across four districts) in two states of India (see Figures A3.1 and A3.2 in Appendix A3). The study sites were selected within blocks identified as being some of the most backward in the country and with extreme drought frequencies (Samaj Pragati Sahayog, 2016).

3.4.1 RECRUITMENT AND INTERVIEWING

Respondents were recruited by community facilitators who work with local NGOs. The community facilitators provided a familiar face to the respondents which is important because of the limited interaction of rural households with individuals from outside communities. The community facilitators also served as the main point of contact for obtaining authorisation from village leaders and they identified and organised suitable survey locations.

Respondents were recruited one day prior to the survey and if they agreed to participate they were invited to the survey location during a morning or an afternoon session. Only individuals who at least occasionally undertake manual labour and who experience frequent water shortages were recruited to take part. The survey location was either an office building or a local school and was always within walking distance from the village. The interviews were conducted by research assistants who are fluent in the local languages (Hindi or Oriya). To avoid potential gender biases, the research assistants were both male and female and they interviewed both men and women.

Respondents were interviewed individually and it was ensured that the survey was conducted in an isolated environment without anyone listening or interfering. The interviews were conducted in conversational form because many of the questions required probing. Conversation also allows for better explanation of the hypothetical scenarios in the choice experiment and may therefore have reduced hypothetical bias (Hardner, 1996). Pictures from the interviews are displayed in Figure A3.3 in Appendix A3.

To compensate respondents for their time, they were gifted a box of sweets after completing the interview. Monetary compensation was avoided due to the sensitive nature of cash in rural poor areas (Hossack and An, 2015). It is believed that cash payments would have attracted a lot of unwanted attention from villagers who were not recruited and potentially offended village leaders and other elite members of the community.

3.4.2 ATTRIBUTES

A reconnaissance visit to Chhattisgarh and Odisha in October 2017 revealed that the most important attributes relating to water use are irrigation for agricultural production, drinking water, water for domestic use, water for livestock, and water for vegetable cultivation. Designing a discrete choice experiment with these attributes, however, is not straightforward because (i) *different* water uses may come from the *same* water sources, and (ii) the *same* water uses may come from *different* water sources. Both (i) and (ii) differ across households and also across seasons which complicates the design even further. There is thus no straightforward way of defining attribute levels that are generic across respondents. One household may, for example, fetch drinking water from a well which means attribute levels could be in terms of trips taken, quantity collected or time spent but another household may rely on government supplied water which means attribute levels rather should be in terms of e.g. frequency of delivery.

After multiple dialogues with community members, community facilitators and local engineers, four attributes related to water use and availability were selected (see Table 3.1). To make the choice experiment more realistic to the individual respondent, we used a pivot design in which the choice tasks were customised based on individual responses to a set of pre-DCE questions. This means that respondents were presented with choice tasks which featured relevant attributes only (i.e. between one and four water-related attributes) as described in detail below.

The first attribute, *Months*, indicates the number of months where water is unavailable from a self-selected water source (i.e. the type of water source (e.g. pond, lake, well) was allowed to vary across individuals). Respondents were asked to think about the water sources currently used by the household. Of the water sources that occasionally dry out, respondents were asked to select the source that they would most want to hold water longer than usual. Subsequently, they were asked to indicate what they use the source for and the months in which water is typically unavailable from that source. To avoid multi-collinearity, respondents were not allowed to select sources that they used for irrigation of agricultural fields or kitchen gardens, if any.

Table 3.1: Attributes

attributes	description	respondents presented with choice tasks featuring this attribute
Months	number of months where water is unavailable from a self-selected water source	all respondents using a water source which occasionally dries out (except sources used for irrigation)
DoubleCrop	dummy variable (0 = single cropping, 1 = double cropping)	all farmers currently practising single cropping
yield	kharif crop yield (tonnes)	all farmers
VegYes	dummy variable (0 = no vegetable production, 1 = vegetable production)	all respondents living in households with private kitchen gardens (but not currently growing any vegetables in it)
VegDouble	dummy variable (0 = current vegetable production, 1 = double vegetable production)	all respondents living in households which grow vegetables in private kitchen gardens
Labour	person-days of unpaid manual labour (one-off contribution)	
Money	rupees (one-off contribution)	

The second and third attribute, *DoubleCrop* and *Yield*, relate to water used for irrigation of agricultural fields. More than 80 percent of the sample is engaged in subsistence agriculture of which 89 percent practise single cropping. Single cropping refers to farmers who harvest only once per year. This crop is called kharif crop and it is cultivated and harvested during monsoon season (June to October). The attribute *DoubleCrop* is a dummy variable indicating whether the household will be able to practise double cropping. Double cropping refers to farmers who can harvest twice per year. The second crop is called rabi crop and it is cultivated and harvested during winter or spring (October to March). Due to weather conditions, farmers whose agricultural fields are solely rainfed (89 percent of the sample) are generally unable to practise double cropping.

The third attribute *Yield* specifies the kharif crop yield. Almost all the farmers in the sample (98 percent) grow rice as their kharif crop. Both *DoubleCrop* and *Yield* were included because increases in kharif crop yield and the ability to grow rabi crop are likely to depend on two different types of irrigation systems. Supplemental or protective irrigation systems (e.g. canals) are designed such that water is spread over a large area to protect farmers against crop failure (Jurriëns and Mollinga, 1996). The aim of such irrigation systems is not to supply farmers with water so as to maximise their yields but rather to provide enough to prevent crop failure. Supplemental or protective irrigation systems are likely to reduce the risk of crop failure during kharif season but because rainfall remains the principal source of moisture under these systems, they are less likely to enable double cropping. Conventional irrigation systems (e.g. farm ponds) are more effective in enabling double cropping because, unlike protective irrigation systems, irrigation water is the principal source of moisture while rainfall is only supplementary.

All farmers in the sample were presented with choice tasks which featured *Yield* as an attribute but only farmers who are not already practising double cropping were presented with choice tasks featuring both *Yield* and *DoubleCrop*.

The fourth attribute, *VegYes* or *VegDouble*, is a dummy-coded variable indicating whether the household can cultivate vegetables (if they currently do *not* grow vegetables) or whether they can double their production (if they currently *do* grow vegetables). Only households who currently grow or have the ability to grow vegetables privately in kitchen gardens will be presented with choice tasks where *VegYes* or

VegDouble is included as an attribute. Respondents who cultivate vegetables as part of a community (self-help) group will not be given choice tasks with *VegDouble* included as an attribute because different community group members may enjoy different shares of the increased production. Similarly, to avoid multi-collinearity, respondents who cultivate vegetables on agricultural land were not presented with choice tasks which featured *VegDouble* as an attribute.

In addition to the attributes related to water use, the choice tasks featured either a monetary (in rupees) or a labour (in days) payment vehicle. Both types of payments were described as one-off contributions.

3.4.3 SURVEY PROCEDURES

Two versions of the survey were designed and respondents were randomly assigned to one the two. The only difference between the two versions is the payment vehicle (labour or money). Respondents were asked to imagine a hypothetical scenario in which an international NGO is planning to design a project that can improve the availability of water in the village. In the money treatment, they were told that the project would not generate employment but that households would have to pay a one-off contribution in order to get access to the benefits of the project. In the labour treatment, respondents were told that the NGO would cover all material costs but that households would have to contribute manual labour for the project to be implemented. Respondents could only volunteer labour on behalf of themselves (i.e. it was explained that the work, if any, would have to be carried out by the respondent and not by other members of the household).

Each respondent was presented with eight choice tasks in which they were asked to choose between an improved hypothetical scenario (WITH project) at some cost (labour or money) and the current scenario (WITHOUT project) at zero costs. An example of a choice task with labour (money) as the payment vehicle is shown in Figure A3.4 (Figure A3.5) in Appendix A3.

The experimental design of the attribute levels in the choice tasks was generated using the software Ngene (version 1.1.2). A pivot design was used where attribute levels in the *WITH project* alternative are pivoted around the respondent's current attributes levels that are displayed in the *WITHOUT project* alternative. The *WITHOUT project* alternative is thus respondent-specific and the attributes for this alternative are invariant across choice tasks. The *WITH project* alternative represents an upgraded scenario where at least one of the non-payment attributes have improved but this will come at some cost in the form of either money or labour contributions. The survey was designed using Sawtooth Software Lighthouse Studio (version 9.5.3) and administered on tablets using an offline survey app.

3.5 THEORETICAL FRAMEWORK AND MODEL SPECIFICATION

3.5.1 UTILITY SPECIFICATION

The analysis of choice data from the choice experiment is based on random utility theory. The conceptual foundation for random utility theory was developed by Louis Thurstone who noted that "*an observer is not consistent in his comparative judgments from one occasion to the next*" (Thurstone, 1927, p. 274). Choice behaviour is thus assumed to be stochastic and the utility U obtained by individual *n* from alternative *j* can be partitioned into a deterministic component $X_{njk}\beta$ and a random component ε_{nj} (McFadden, 1974).

$$U_{nj} = X_{njk}\beta + \varepsilon_{nj} \qquad \text{Eq. 1}$$

The deterministic component of the utility function in Eq. 1 is defined as a linear function of observed attributes (see Table 3.1) accompanied by a set of preference parameters β indicating the marginal utility of the attributes (Hensher et al., 2015; Hole, 2006). X_{njk} denotes the attribute level of the k^{th} attribute relating to the j^{th} alternative plus any relevant interactions between x_{jk} and choice-specific or individual-specific characteristics.

The utility function in Eq. 1 can be estimated either separately for the two treatments (i.e. each model is estimated with one cost variable which is either money or labour) or together in a pooled model. The pooled model approach infers an empirical value of contributed time by estimating a joint model with two cost variables (money and labour). The marginal disutility of labour relative to the marginal disutility of money then provides an estimate of the implicit value of contributed time.

A basic utility specification where the preference parameters for money and labour are estimated jointly is shown in Eq. 2. An alternative-specific constant for the status quo (ASC_{SQ}) is included to reflect any (dis)utility associated with this alternative. A statistically significant positive (negative) alternative specific constant for the status quo alternative suggests that respondents like (dislike) the current state of affairs independent of the attributes describing that alternative.

$$U_{nj} = \beta_1 Months_j + \beta_2 DoubleCrop_j + \beta_3 Yield_j + \beta_4 VegYes_j + \beta_5 VegDouble_j + \beta_6 Money_j + \beta_7 Labour_j + \beta_8 ASC_{SQ} + \varepsilon_{nj} \qquad \text{Eq. 2}$$

The joint estimation approach assumes that the underlying preference structure is the same for both labour and monetary PVs. To test for potential utility differences between the two PVs, we include a set of interaction terms between a dummy variable indicating the type of PV (LPV=1 if labour is used as the PV, LPV=0 if money is used as the PV) and each of the non-cost attributes (see Eq. 3). If these interaction terms are statistically significant, either individually or jointly, it means that utility differences between the two PVs exist and that the utility function in Eq. 2 is misspecified (Gibson et al., 2016).

$$\begin{aligned} U_{nj} = {} & \beta_1 Months_j + \beta_2 DoubleCrop_j + \beta_3 Yield_j + \beta_4 VegYes_j \qquad \text{Eq. 3} \\ & + \beta_5 VegDouble_j + \beta_6 Money_j + \beta_7 Labour_j + \beta_8 ASC_{SQ} \\ & + \beta_9 LPV \times Months_j + \beta_{10} LPV \times DoubleCrop_j + \beta_{11} LPV \times Yield_j \\ & + \beta_{12} LPV \times VegYes_j + \beta_{13} LPV \times VegDouble_j + \varepsilon_{nj} \end{aligned}$$

In the utility specifications in Eq. 2 and Eq. 3 we estimate generic coefficients (i.e. coefficients that are shared across alternatives) for all attributes. Hess and Rose (2009) argue, however, that in the case of pivot style data, marginal utilities might differ between the status quo which is a ‘real world’ alternative and the improved scenario which is a hypothetical alternative. To investigate this, we estimate a model with separate coefficients for the status quo and the improved scenario. Since attribute levels for *DoubleCrop*, *VegYes*, *VegDouble*, *Money* and *Labour* are not shared across alternatives[13], separate coefficients are only estimated for the two (continuous) variables attributes *Months* and *Yield*.

3.5.2 CHOICE MODELS

The distribution of the random term is unknown but a standard assumption in the literature is that ε_{nj} is independently and identically distributed following a type 1 extreme value distribution. This assumption leads to the conditional logit (CL) model (McFadden, 1974) in which the probability of respondent *n* choosing alternative *i* over alternative *j* is given by Eq. 4

$$P_{ni} = \Pr\left(\beta X_{nik} + \varepsilon_{ni} > \beta X_{njk} + \varepsilon_{nj}\right) \forall\, j \neq i \in J = \frac{\exp(\lambda\beta X_{nik})}{\sum_{j=1}^{J} \exp(\lambda\beta X_{njk})} \qquad \text{Eq. 4}$$

The scale term λ is both inversely related to the error variance ($\lambda=\pi/\sqrt{6\mathrm{var}[\varepsilon]}$) and confounded with the deterministic component of the utility function which means that differences in the vector of estimated preference parameters can be a result of differences in marginal utilities and/or it can be a result of differences in error variance. Respondents with higher error variance (i.e. respondents whose choices are more affected by factors not captured by the model) will, ceteris paribus, have estimated utility coefficients that are smaller in magnitude than respondents with lower error variance (Hess and Train, 2017). This kind of heterogeneity is referred to as scale heterogeneity because it impacts estimated parameters for the observed variables in the same way (Louviere and Eagle, 2006). In most applications, however, including the CL model, the error variance is assumed to be constant and λ is therefore normalised to 1. An alternative model is the heteroscedastic conditional logit (HCL) model which allows for unequal error variance across individuals (DeShazo and Fermo, 2002; Hensher et al., 1999; Hole, 2006). This is illustrated in Eq. 5 where the subscript *n* has been added to the scale parameter.

---

[13] For *DoubleCrop*, *VegYes*, *VegDouble*, it is only the improved scenario that can take non-zero values. For *Money* and *Labour*, the improved scenario always takes non-zero values.

$$P_{ni} = \frac{\exp(\lambda_n \beta X_{nik})}{\sum_{j=1}^{J} \exp(\lambda_n \beta X_{njk})} \qquad \text{Eq. 5}$$

In the HCL model, the scale term is parametrised as $\exp(\theta Z_n)$ which ensures that λ_n always is positive and that Eq. 5 equals Eq. 4 when the parameter θ is zero. Z_n is a vector of e.g. individual and treatment-specific characteristics.

In Section 3.6, we estimate HCL models to test for differences in error variance between (i) money and labour PVs, and (ii) respondents who were presented with choice tasks featuring only one water-related attribute versus respondents who were presented with choice tasks featuring at least two water-related attributes (as suggested by Hess and Rose, 2009).

3.6 RESULTS

Summary statistics are presented in Table 3.2 for the aggregate sample (first column) and disaggregated by treatment (second and third column). As shown (in the fourth column), there is no statistically significant difference in means between the two treatments for any of the variables. Nearly half of the surveyed respondents are women. The average respondent lives in a household with 9.4 individuals, including 4.1 children under the age of 18, and is 36.5 years old. 38% of the sample hold no formal qualifications and a majority of the sample belong to a caste classified as backward (45%, scheduled tribe, 14% scheduled caste and 36% other backward caste). The remaining 5% of the sample belong to either Brahmin or other forward castes. The average respondent lives in a household that earns 319 rupees per day (≈ 5 USD at the time of the interview) and 16% of the surveyed respondents live in a household where at least one household member migrates seasonally (e.g. to big cities) in search of work. Nearly half of the surveyed respondents (188 of 387) are head of their household of which 41 are female-headed households (not shown in Table 3.2). A majority of the sample (86%) are farmers with 2 acres of land, on average, which is most commonly rainfed (only 11% of the farmers have access to irrigation systems). 89% of the farmers practise single cropping (i.e. kharif crop) and the average yield is 1.28 tonnes per season.

Table 3.2: Summary statistics (by treatment)

	aggregate	LPV	MPV	LPV-MPV
all respondents				
gender (1 = female, 0 = male)	0.48	0.49	0.47	0.02
	(0.50)	(0.50)	(0.50)	(0.05)
age	36.5	37.1	36.0	1.2
	(9.0)	(9.4)	(8.6)	(0.9)
education (1 = yes, 0 = no)	0.62	0.58	0.65	-0.07
	(0.49)	(0.49)	(0.48)	(0.05)
household size	9.4	9.5	9.3	0.1
	(2.0)	(2.1)	(2.0)	(0.2)
children in household (17 years or below)	4.1	4.1	4.1	0.0
	(1.4)	(1.4)	(1.3)	(0.14)
scheduled tribe (1 = yes, 0 = no)	0.45	0.45	0.45	0.00
	(0.50)	(0.50)	(0.50)	(0.05)
scheduled caste (1 = yes, 0 = no)	0.14	0.16	0.13	0.03
	(0.35)	(0.37)	(0.34)	(0.04)
other backward caste (1 = yes, 0 = no)	0.36	0.35	0.37	-0.02
	(0.48)	(0.50)	(0.48)	(0.05)
migration (1 = yes, 0 = no)	0.16	0.16	0.16	0.00
	(0.37)	(0.37)	(0.37)	(0.04)
household daily earnings (rupees)	319	320	317	3
	(107)	(108)	(107)	(11)
household head (1 = yes, 0 = no)	0.49	0.52	0.46	0.06
	(0.50)	(0.50)	(0.50)	(0.05)
farmer (1 = yes, 0 = no)	0.86	0.87	0.84	0.03
	(0.35)	(0.34)	(0.37)	(0.04)
respondents	387	186	201	
non-farmers excluded				
land (acres owned)	2.0	1.9	2.1	-0.2
	(2.0)	(1.9)	(2.1)	(0.2)
irrigation (1 = yes, 0 = no)	0.11	0.12	0.11	0.02
	(0.32)	(0.33)	(0.31)	(0.04)
yield (tonnes)	1.28	1.15	1.40	-0.25
	(1.40)	(1.29)	(1.49)	(0.15)
cropping (1 = single, 0 = double)	0.89	0.91	0.86	0.05
	(0.32)	(0.28)	(0.34)	(0.03)
respondents	331	162	169	

Means of each variable with standard deviations in parentheses in the first 3 columns
Standard errors in parentheses in the last column (LPV-MPV)
* p<0.10; ** p<0.05; *** p<0.01
LPV = labour as payment vehicle
MPV = money as payment vehicle

A total of 481 respondents were interviewed of which 387 are included in the analysis. 50 respondents were practice interviews and thus dropped from the dataset. 12 respondents provided conflicting responses (e.g. they reported to use a source from which water occasionally dries out but when asked to state the months in which water is temporarily unavailable, they stated none). Following the common IQR (interquartile range) method for identifying outliers, 20 respondents are dropped due to having reported extreme wage rates and 6 large-scale farmers (with high yields) are excluded from the analysis. 2 respondents who always selected the improved scenario are dropped because they misunderstood the conditions of the choice experiment (one selected the improved scenario at any cost because the payment was not real and one agreed to any number of labour days because he/she would have other household members do the work). 4 (female) respondents are dropped because they always selected the status quo alternative due to being uncomfortable with making decisions.

3.6.1 ESTIMATION OF CHOICE MODELS

Results from conditional logit estimation of the pooled model in Eq. 2 are presented in the first column of Table 3.3 (CL-1). All water-related variables as well as the payment variables are significant at least at the 5% level and with the expected signs. Respondents obtain positive marginal utility from an increase in kharif crop yield (*Yield*) and from an increase in the number of months that water is available from a self-selected water source (*Months*)[14]. Respondents who are currently practising single cropping prefer double cropping (*DoubleCrop*) and respondents who are currently cultivating vegetables in a private garden would like to double their production (*VegDouble*). Respondents who own a private garden but are not currently cultivating vegetables would like to do so if water was available (*VegYes*). The estimated coefficients for the price variables (*Money* and *Labour*) are negative thus indicating that respondents dislike higher prices (rupees or labour days). The alternative specific constant for the status quo (ASC_{SQ}) is negative (but significant at the 10% level only) which indicates that respondents, on average, dislike their current situation.

[14] We include the negative of the attribute *Months* for ease of interpretation when calculating WTP estimates in Section 3.6.2

Table 3.3: Conditional logit estimation

	CL-1	CL-2
Months	0.196***	0.179***
	(0.039)	(0.049)
Yield (tonnes)	0.461***	0.497***
	(0.068)	(0.089)
DoubleCrop	0.410***	0.391***
	(0.092)	(0.117)
VegYes	0.618***	0.577***
	(0.098)	(0.124)
VegDouble	0.336**	0.220
	(0.136)	(0.195)
Money (1/1000 rupees)	-0.446***	-0.438***
	(0.023)	(0.025)
Labour (days)	-0.025***	-0.026***
	(0.001)	(0.002)
ASC_{SQ}	-0.247*	-0.251*
	(0.129)	(0.129)
LPV × Months		0.035
		(0.059)
LPV × Yield		-0.086
		(0.134)
LPV × DoubleCrop		0.042
		(0.149)
LPV × VegYes		0.089
		(0.160)
LPV × VegDouble		0.210
		(0.251)
Log Likelihood	-1687	-1686
Observations	3096	3096

Standard errors in parentheses
* p<0.10; ** p<0.05; *** p<0.01
LPV = labour as payment vehicle

The second model in Table 3.3 (CL-2) includes a set of interaction terms: each of the five water-related variables are interacted with a dummy variable indicating the type of payment vehicle (LPV=1 if labour is used as the payment vehicle, LPV=0 if money is used as the payment vehicle). The five interaction terms are insignificant (both jointly and individually) which indicates that there is no effect of the payment vehicle on marginal utilities. Since there is no difference in choice behaviour when a labour PV is used compared to when a money PV is used, a scale factor exists which scales the marginal utility of money to the marginal utility of labour. This scale factor is essentially an estimated value of the opportunity cost of contributed time i.e. the marginal rate of substitution between labour time and money. We find that the estimated value of time is 60 rupees per day[15].

To test for differences in error variance between money and labour PVs, we estimate a heteroscedastic conditional logit model (HCL-1) in Table 3.4. HCL-1 is also testing for scale differences between respondents who were presented with choice tasks featuring only one attribute[16] and respondents who were presented with choice tasks featuring at least two attributes. The scale parameter for respondents who answered choice tasks with at least two attributes and a monetary PV is normalised to 1. Since the scale term is parametrised as $\exp(\theta Z_n)$, we can test for differences in scale by testing if $\theta=0$.

As shown in Table 3.4, the estimated scale parameter for the subgroup which was asked to "pay" with labour (θ_{LPV}) is statistically insignificant which means that there is no difference in error variance between the two PVs. The scale parameter for the subgroup which was presented with only one water-related attribute in the choice tasks (*OneAttributeOnly*), is positive and significant at the 1% level. This means that we reject a null hypothesis of $\lambda_{OneAttributeOnly}$ being equal to 1. Since $\theta_{OneAttributeOnly}$ is positive, $\lambda_{OneAttributeOnly}$ is larger than 1 thus indicating that individuals who respondent to single attribute choice tasks make more deterministic choices (i.e. their error variance is lower).

To test if marginal utilities differ across alternatives, HCL-1 includes alternative-specific coefficients for the attributes *Months* (Months × ASC_{SQ} and Months × $ASC_{project}$) and *Yield* (Yield × ASC_{SQ} and Yield × $ASC_{project}$). Hess and Rose (2009) suggest that in the case of pivot style data, marginal utilities might differ between the status quo which is a 'real world' alternative and the improved scenario which is a hypothetical alternative (see Section 3.5.1). To investigate this, we estimate a model with separate coefficients for the status quo and the improved scenario. However, since attribute levels for *DoubleCrop*, *VegYes*, *VegDouble*, *Money* and *Labour* are not shared across alternatives (i.e., it is only the improved scenario that can take non-zero values), separate

[15] $\beta_{labour}/(\beta_{money} \times \frac{1}{1000})$

[16] 26 respondents were presented with choice tasks featuring only *Months*, 10 respondents were presented with choice tasks featuring only *Yield*, and 3 respondents were presented with choice tasks featuring only *VegYes* or *VegDouble*.

coefficients are only estimated for the two (continuous) variables *Months* and *Yield*. A null hypothesis of equality of Months × ASC_{SQ} and Months × $ASC_{project}$ is rejected ($p < 0.01$) while we fail to reject a null hypothesis of equality of Yield × ASC_{SQ} and Yield × $ASC_{project}$ ($p > 0.10$). Since Months × $ASC_{project}$ > Months × ASC_{SQ}, it appears that respondents are more sensitive to the number of months in which water is (un)available for the alternative representing the status quo than for the alternative representing an improved scenario.

In model HCL-2, we include interaction terms between each of the payment variables (*Labour* and *Money*) and a dummy variable indicating if the respondent is female (Female=1 if female, Female=0 if male). The estimated coefficient for the interaction term between *Money* and *Female* is negative and significant which suggests that women are more payment sensitive compared to men when the PV is money. The estimated coefficient for the interaction term between *Labour* and *Female* is positive and significant which suggests that women are less payment sensitive compared to men when the PV is labour. Female respondents are thus WTW more, on average, than male respondents while male respondents are WTP more than female respondents for marginal improvements in the water attributes. Subsequently, the value of time estimated for women (40 rupees per day) is lower than that for men (68 rupees per day). A null hypothesis of equality of the value of time across gender is rejected at the 1% level. We also include interaction terms between each of the water-related variables and a dummy variable indicating if the respondent is female. As shown in Table B3.1 in Appendix B3, these interaction terms are insignificant which suggests that there are no gender differences in the marginal value of improvements in water supply.

HCL-2 further includes interaction terms between each of the cost variables and a dummy variable indicating if the respondent is schooled (Educ=1 if the respondent is schooled, Educ=0 if the respondent is unschooled). The interaction term between *Educ* and *Money* is insignificant which means there is no difference in the marginal utility of money between schooled and unschooled individuals. The interaction term between *Educ* and *Labour*, on the other hand, is positive and significant indicating that schooled individuals are less payment sensitive than unschooled individuals.

As in models CL-1 and CL-2, the estimated coefficients for the water-related attributes are significant and with the expected signs in both HCL-1 and HCL-2. The alternative specific constant for the status quo (ASC_{SQ}) and the price variable coefficients are negative and highly significant (including the ASC_{SQ} coefficient which is now significant at the 1% level). This is not surprising in areas of rural India where water is a scarce resource and households depend on water for their livelihoods.

Table 3.4: Heteroscedastic conditional logit estimation

	HCL-1	HCL-2
Months × $ASC_{project}$	0.271***	0.287***
	(0.045)	(0.047)
Months × ASC_{SQ}	0.100***	0.116***
	(0.038)	(0.040)
Yield × $ASC_{project}$ (tonnes)	0.369***	0.381***
	(0.088)	(0.089)
Yield × ASC_{SQ} (tonnes)	0.332***	0.364***
	(0.118)	(0.121)
DoubleCrop	0.422***	0.436***
	(0.087)	(0.090)
VegYes	0.627***	0.640***
	(0.097)	(0.100)
VegDouble	0.370***	0.350***
	(0.126)	(0.129)
Money (1/1000 rupees)	-0.415***	-0.406***
	(0.025)	(0.047)
Labour (days)	-0.022***	-0.028***
	(0.002)	(0.003)
Money × Female (1/1000 rupees)		-0.184***
		(0.042)
Labour × Female (days)		0.004**
		(0.002)
Money × Educ (1/1000 rupees)		0.061
		(0.043)
Labour × Educ (days)		0.004**
		(0.002)
ASC_{SQ}	-0.381***	-0.423***
	(0.112)	(0.119)
Scale θ		
LPV	0.146	0.122
	(0.108)	(0.108)
One Attribute Only	0.432***	0.332**
	(0.136)	(0.140)
Log Likelihood	-1671	-1649
Observations	3096	3096

Standard errors in parentheses
* p<0.10; ** p<0.05; *** p<0.01
LPV = labour as payment vehicle

3.6.2 WILLINGNESS TO PAY AND WILLINGNESS TO WORK

The parameter estimates from HCL-2 can be used to calculate WTP and WTW for marginal improvements in each of the water-related attributes. Marginal WTP (WTW) is calculated as the water-related attribute coefficient relative to the (negative of) the marginal utility of money (labour). Estimates of marginal WTP and marginal WTW are presented in Table 3.5 for the aggregate sample. The Krinsky-Robb method is used to construct 95% confidence intervals (Hole, 2007a).

Table 3.5: WTP and WTW (aggregate sample)

	WTP	WTW
months × $ASC_{project}$	655	12.3
	[444;875]	[8.4;16.6]
months × ASC_{SQ}	242	4.6
	[59;422]	[1.1;7.9]
yield × $ASC_{project}$	890	16.8
	[495;1272]	[9.5;24.0]
yield × ASC_{SQ}	801	15.1
	[257;1349]	[4.9;25.6]
double crop	1018	19.2
	[661;1463]	[12.7;27.3]
veg yes	1512	28.5
	[1070;1976]	[20.6;37.5]
veg double	893	16.8
	[306;1535]	[5.9;28.7]

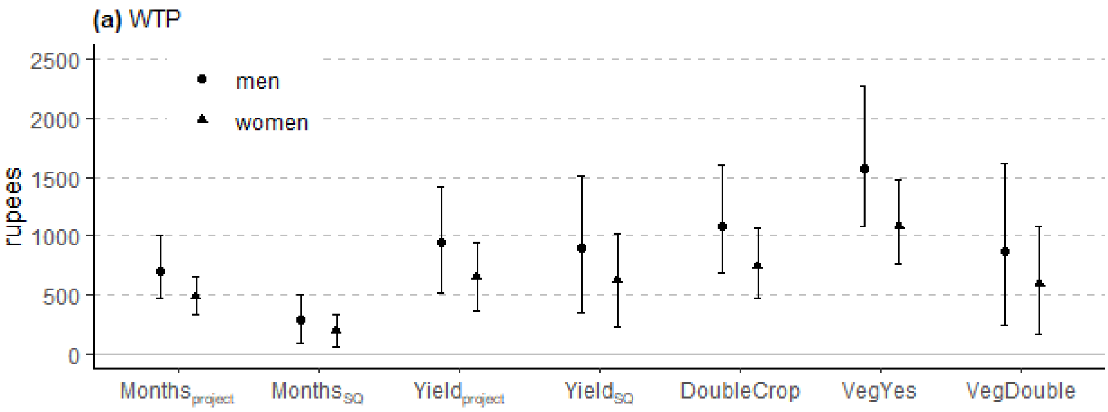
95% confidence interval in brackets
WTP = willingness to pay (rupees)
WTW = willingness to work (days)

Table 3.5 demonstrates high demand for improved water availability. The average respondent is, for example, willing to pay 1512 rupees (≈ 24 USD at the time of the interview) to gain access to water which can enable vegetable cultivation and 890 rupees (≈ 14 USD) to increase kharif crop yield by one tonne. Similarly, if the PV is labour days, the average respondent is willing to work 28.5 days to be able to grow vegetables and 16.8 days per tonne increase in kharif crop yield.

WTP and WTW for the attribute *Months*[17] is significantly higher for the hypothetical alternative representing an improved scenario than for the status quo alternative. Respondents are thus treating the improved scenario alternative systematically different than the status quo alternative. Hess and Rose (2009) suggest that this may be because respondents act differently to the attributes of an alternative that they are experiencing in real life compared to the attributes of an alternative that is hypothetical only. It may also be because variation in the attributes in the improved scenario cause individuals to react differently than the invariant attributes nature of the status quo alternative.

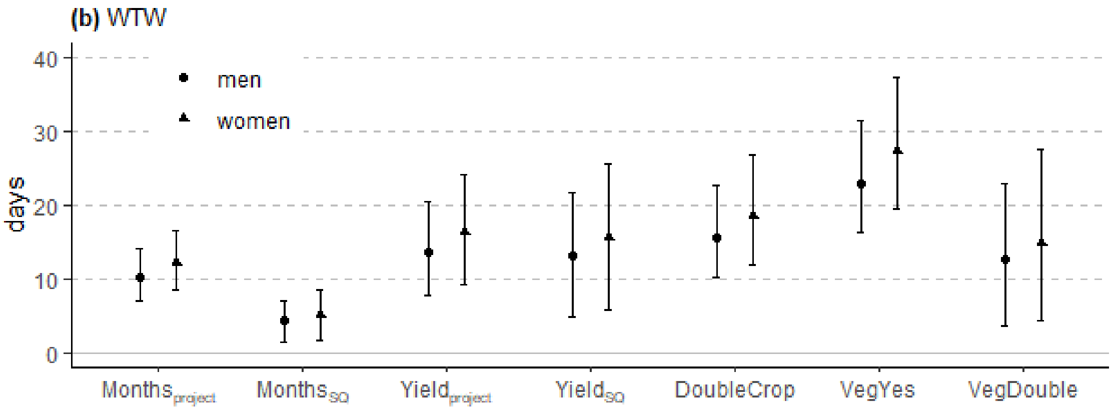
WTP and WTW (along with the associated 95% confidence intervals) for marginal improvements in the attributes are illustrated separately for men and women in Figure 3.1. As shown, men are WTP more than women for improved water availability while women are WTW more than men (as also discussed in Section 3.6.1). Despite the overlapping confidence intervals in Figure 3.1 (due to the size of the standard errors of the water-related attribute coefficients which are generic across gender), the difference in WTP and WTW between men and women is significant at least at the 5% for all the attributes (except the difference in WTW for *VegDouble* which is significant at the 10% level only).

Figure 3.1: WTP and WTW (disaggregated by gender)



[17] The attribute *Months* is presented to respondents as the number of months where water is unavailable from a self-selected water source. To obtain WTP and WTW estimates, we include the negative of this variable in the estimation thus assuming that there is no disparity between estimates of WTP/WTW and willingness to accept (money or labour).

Figure 3.1: WTP and WTW (disaggregated by gender) (*continued from previous page*)



3.6 DISCUSSION AND CONCLUSIONS

In this paper, we reported the design and administration of a discrete choice experiment (DCE) concerning water scarcity in rural India. To investigate the demand for improved water availability, respondents were randomly assigned to one of two treatments where the payment vehicle is either money or labour. Both money and labour were employed as payment vehicles (using a between subjects design) to enable estimation of the shadow value of time. By estimating gender-specific coefficients for each of the two types of payments, we were able to estimate two values of the shadow value of time – one for men (68 rupees/day) and one for women (40 rupees/day). This is (to our knowledge) the first study examining how the choice of PV affects gender-based differences in WTP and WTW and, ultimately, welfare valuation.

We used a pivoted design to construct the choice tasks in the DCE. The survey was administered on tablets which allowed for a dynamic questionnaire where questions and attributes in the choice tasks adapted to the individual throughout the survey. This relatively complex design, we argue, increased the realism of the DCE but it comes with a number of challenges with respect to analysing the data. We find that when the model is specified to have estimated coefficients that are different for the status quo and the improved scenario (as suggested by Hess and Rose, 2009), the estimated value of both WTP and WTW for improvements in the attribute *Months* is higher for the improved scenario alternative than for the status quo alternative. We further find that respondents presented with only one water-related attribute in the choice tasks have an estimated scale term larger than 1 which indicates a lower error variance compared to respondents presented with at least two water-related attributes in the choice tasks. An easy explanation arises in that the cognitive load of choice tasks featuring only one (water-related) attribute is lower than that of multi-attribute choice tasks.

To test for potential utility differences between money and labour payments, we included a set of interaction terms between a dummy variable indicating the type of PV and each of the water-related attributes. The insignificance of these interaction terms indicates that there are no utility differences between the two PVs which suggest that welfare valuation (for the aggregate sample) is unaffected by the choice of PV. The fact that both PVs (money and labour) yield similar results (i.e. the preference structure is the same for both PVs) enables estimation of the shadow value of time (i.e. the marginal rate of substitution between money and labour).

We estimate the shadow value of time for the aggregate sample and disaggregated by gender. The shadow value of time for the average respondent in the sample is 60 rupees/day based on CL-2 or 53 rupees/day based on HCL-1. This is roughly one third of the wage rate offered under MGNREGA and the result is therefore in concurrence with previous papers that have used the economic value of leisure time, commonly assumed to be one-third of the value of work time (Cesario, 1976), when generating monetised estimates of WTW (e.g. O'Garra, 2009).

In the context of gender inequality and socio-cultural norms in rural India, we hypothesise that the estimated shadow value of time is lower for women than for men. The patriarchal family structure of many households in rural India means that women are engaged primarily in unpaid care work activities (e.g. housework and child care) and thus have limited financial independence. Subsequently, the marginal utility of money is likely to be higher for women than for men which is likely to translate to a lower opportunity cost of time.

We find that the estimated shadow value of time for women is 40 rupees/day which is significantly lower than 68 rupees/day which is the estimated shadow value of time for men. This gender gap in the value of time translates to differences in welfare values (men are WTP more while women are WTW more for marginal improvements in water supply) and raises some concerns about the use of both money and labour as the PV for valuing goods and services in developing countries. Since the marginal utilities of money and time are not constant across gender, more weight will be given to preferences of men (women) when the PV is money (labour). Our finding echoes discussions in the longstanding literature about equity concerns in cost-benefit analysis (e.g. Fleurbaey and Abi-Rafeh, 2016; Nyborg, 2014) but while previous studies focus on income inequality, we focus on gender inequality.

Estimates of WTP and WTW are typically used to make inferences about the social benefits of alternative projects. Our results demonstrate, however, that such inference is complicated by gender-based differences in WTP and WTW (men are WTP more while women are WTW more) for marginal improvements in water supply. While we find no gender-based differences in the marginal utilities of the water-related attributes, many labour PV studies monetise WTW estimates *ex post* using a generic conversion rate which is typically some proportion of the wage rate (e.g. Navrud and Vondolia, 2020). If the same conversion rate is used for men and women, then men will appear to be the biggest benefactors, ceteris paribus, if WTP is employed as the measure of social welfare while women will appear to be the biggest benefactors if the measure of social welfare is based on WTW.

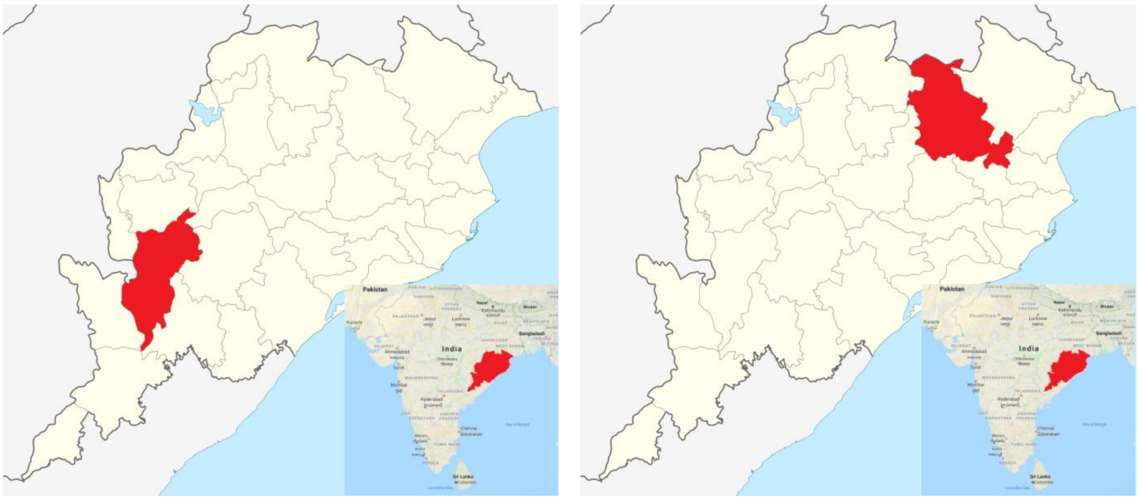
Our results suggest that the conclusion of a cost-benefit test can be affected by the choice of payment vehicle and the gender of the project beneficiaries. Choosing money or labour as the payment vehicle therefore requires value judgment because it is essentially a choice between two different distributional weights. In our sample, decision makers who rely on WTP are implicitly attaching more weight to men while decision makers using WTW attach more weight to women. Our study is relevant to decision-makers since there is a growing interest in the potential role of social protection programmes, such as MGNREGA, to empower women. To that end, the choice of payment vehicle is an avenue through which the decision making process can favour one gender over the other.

Our findings are further relevant to decision makers as we have estimated the economic value of the benefits from four different uses of water which provide relevant information about the benefits that water projects could deliver. The average respondent is willing to pay 890 rupees for increased access to irrigation systems which will increase kharif crop yield by one tonne and 1018 rupees for increased access to irrigation systems which enable double cropping. Respondents are further willing to pay 1512 rupees for improved water supply which enable vegetable cultivation or, if the respondent is already cultivating vegetables, 893 rupees to double the production. Finally, the average respondent is willing to pay 655 rupees per extra month that water is available from a self-selected water source. The uses of the self-selected water-source vary across respondents but 86% report using it for household purposes, 79% report using it for bathing, 70% report using it for drinking water and 65% report using it for livestock. The above estimates of willingness to pay demonstrate high demand for improvements in water supply thus encouraging decision makers to consider a reconfiguration of social protection programmes, including MGNREGA, to implement more water-related projects.

A3 APPENDIX: SURVEY INFORMATION

Figure A3.1: Sampled districts in Odisha

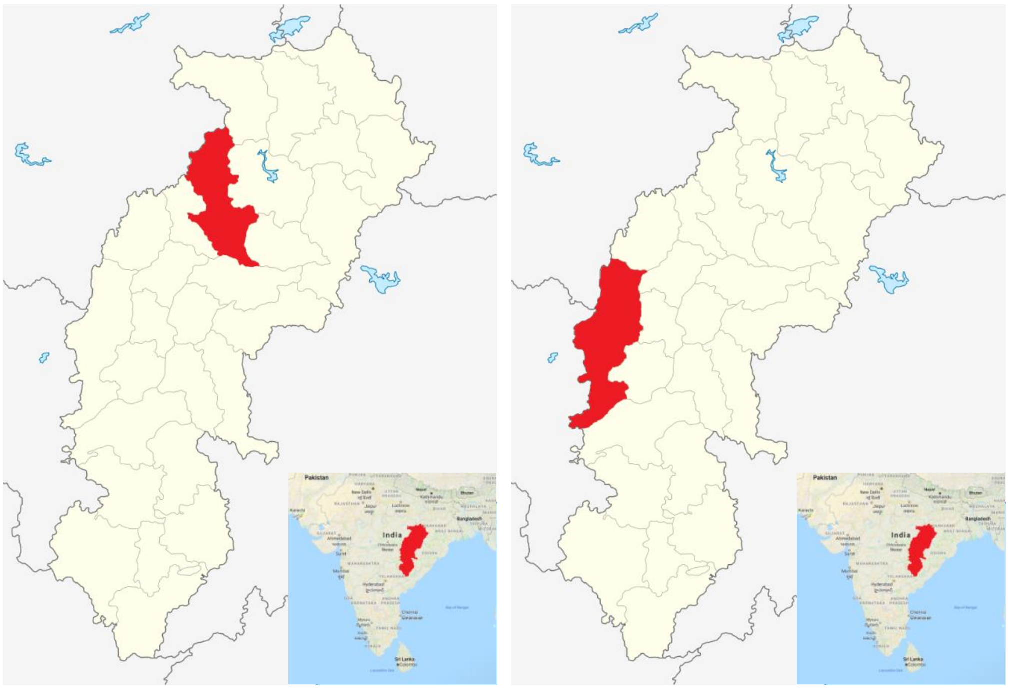
(a) Kalahandi (b) Kendujhar



Source: Adapted from Wikimedia Commons (2013); Wikimedia Commons (2016b)

Figure A3.2: Sampled districts in Chhattisgarh

(a) Bilaspur (b) Rajnangaon



Source: Adapted from Wikimedia Commons (2016a); Wikimedia Commons (2016b)

Figure A3.3: Interviews

Figure A3.4: Sample choice task (labour as payment vehicle)

	WITHOUT project	WITH project
Water Unavailability (Pond)	3 months	0 months
Cropping Pattern	Single Cropping	Double Cropping
Kharif Crop Yield	50 kilos	80 kilos
Vegetable Cultivation	No	No
Unpaid Work	None	60 person-days
	Select	Select

Figure A3.5: Sample choice task (money as payment vehicle)

	WITHOUT project	WITH project
Water Unavailability (Pond)	3 months	0 months
Cropping Pattern	Single Cropping	Double Cropping
Kharif Crop Yield	50 kilos	80 kilos
Vegetable Cultivation	No	No
One-time contribution	None	2000 rupees
	Select	Select

B3 APPENDIX: RESULTS

Table B3.1: Conditional logit model (with gender interaction terms)

Months	0.255***
	(0.048)
Yield (tonnes)	0.563***
	(0.088)
DoubleCrop	0.404***
	(0.117)
VegYes	0.682***
	(0.127)
VegDouble	0.414**
	(0.166)
Money (1/1000 rupees)	-0.396***
	(0.027)
Labour (days)	-0.029***
	(0.002)
ASC_{SQ}	-0.273**
	(0.131)
Female × Months	-0.090
	(0.061)
Female × Yield	-0.333
	(0.238)
Female × DoubleCrop	0.011
	(0.151)
Female × VegYes	-0.108
	(0.162)
Female × VegDouble	-0.338
	(0.270)
Female × Money (1/1000 rupees)	-0.168***
	(0.048)
Female × Labour (days)	0.007***
	(0.003)
Log Likelihood	-1658
Observations	3096

CONCLUSION

This thesis is a collection of three papers which investigate issues around the use of labour as the payment vehicle (PV), as an alternative to money, in stated preference (SP) studies. Labour is increasingly used as the PV in studies which aim to investigate preferences of low-income groups in developing countries. The argument commonly put forward is that a monetary PV, leading to estimates of willingness to pay (WTP), will underestimate willingness to contribute of households engaged primarily in subsistence-based activities. Since these households are limited by extremely tight budget constraints but arguably less constrained in terms of time, a labour PV, leading to estimates of willingness to work (WTW), is often considered a more appropriate utility measure than WTP.

The results in this thesis contribute to at least two strands of the literature on SP methods. The first is hypothetical bias (HB) which arises when respondents misrepresent their willingness to pay (WTP) for a good because of the hypothetical nature of both the payment and the provision of the good in question. The second is the impact of the payment vehicle (money or labour) on welfare estimates. Welfare estimates are key inputs in cost-benefit analysis and thus essential for evaluating the performance of development projects and programs.

The empirical analysis is based on two discrete choice experiments (DCEs). Chapters 1 and 2 use data from a DCE concerning fortified flour in Kenya and Chapter 3 builds on data from a DCE concerning water scarcity in India. Respondents in both studies were randomly assigned to treatments where the PV was either money or labour. In the DCE administered in Kenya, respondents were further randomly assigned to a treatment where purchase choices were either hypothetical or consequential. This 2x2 design enabled an investigation of differences in hypothetical bias between the two PVs. While there is a large body of literature attempting to determine the direction as well as the magnitude of HB and also a growing literature examining the potential of ex ante and ex post techniques to mitigate or eliminate this bias (Loomis, 2014), this thesis provides a significant contribution to the literature by testing the impact of differing means of payment upon HB.

Chapter 1 has investigated issues around the choice of conversion rate for monetising WTW. To enable assessment of welfare effects (e.g. by comparing benefits and costs), many studies monetise WTW estimates *ex post*. The common approach in the WTW literature for assessing the performance of one or several competing conversion rates is to employ a split-sample design using both labour and money PVs. Performance is then evaluated based on the closeness of monetised WTW to WTP. While previous labour PV studies are hypothetical SP studies, which means the benchmark for assessing the performance of conversion rates is hypothetical WTP, Chapter 1 reported the results from a comparative study using both hypothetical and consequential WTP as the benchmark. The results showed that the best performing conversion rate (one third of the national average wage rate) when hypothetical WTP is used as the benchmark performs poorly (by ranking fourth of the six conversion rates employed) when consequential WTP is used as the benchmark. The findings thus cast doubt on the use of hypothetical WTP as the benchmark and call into question the use of one third of the wage rate (a commonly employed conversion rate in the labour PV literature for monetising WTW) as an appropriate proxy for the value of contributed time.

Monetisation of WTW is complicated by at least two factors. First, if hypothetical bias (HB) is more (less) of a problem when labour is used as an alternative to money, the monetised value of WTW will be biased upwards (downwards). Chapter 2 investigated this issue by testing for differences in the scale of HB between the two PVs. The results showed that HB is 26-31 percentage points higher when respondents are asked to pay with labour instead of money. Second, to convert labour "payments" to monetary values, it is necessary to apply an opportunity cost of time. The opportunity cost of time is affected by a wide range of factors including wage rate, family status and alternative uses of time. Chapter 3 focused on gender heterogeneity in the value of time in the context of a patriarchal family structure. The results showed that the value of time is lower for women than for men which suggests that women have a higher marginal utility of money and/or a lower marginal utility of time. Subsequently, more weight will be given to preferences of women in an SP study if the PV is labour while more weight will be given to men if the PV is money. This findings suggest that the choice of PV is based on value judgments with respect to the social importance of men and women.

In summary, this thesis set out to investigate some of the issues around the use of labour versus money as the payment vehicle in stated preference studies in low-income communities. It concluded that money outperforms labour with respect to hypothetical bias but that a monetary payment vehicle implicitly attaches more social importance to male respondents. The research in this thesis provides the first bit of evidence about (1) the reliability of inferences drawn from the use of hypothetical WTP as the benchmark for evaluating the performance of one or several competing rates for monetising WTW, (2) differences in the scale of hypothetical bias between a labour and a monetary payment vehicle, and (3) the impact on welfare evaluation of gender-based differences in the value of time. Since the results are based on one study only i.e. one country (Kenya or India) and one good/service (fortified flour or improved water supply), this research opens up an interesting avenue for future research. To examine the extent to which the results in this thesis can be generalised, future studies may, for example, want to explore how the findings are affected by the type of good (private versus public) and the degree of marketisation in the surveyed population.

BIBLIOGRAPHY

Abramson, A., Becker, N., Garb, Y. & Lazarovitch, N. (2011). Willingness to pay, borrow, and work for rural water service improvements in developing countries. *Water Resources Research,* 47**,** 1-12.

Adler, M. D. (2016). Benefit–Cost Analysis and Distributional Weights: An Overview. *Review of Environmental Economics and Policy,* 10(2)**,** 264-285.

Ahlheim, M., Frör, O., Nguyen, D., Rehl, A., Siepmann, U. & Van Dinh, P. (2017). Working Paper: Labour as a utility measure reconsidered. Hohenheim Discussion Papers in Business, Economics and Social Sciences, No. 03-2017.

Alemu, M. H. & Olsen, S. B. (2018). Can a Repeated Opt-Out Reminder mitigate hypothetical bias in discrete choice experiments? An application to consumer valuation of novel food products. *European Review of Agricultural Economics,* 45(5)**,** 749-782.

Ando, A. W., Cadavid, C. L., Netusil, N. R. & Parthum, B. (2020). Willingness-to-volunteer and stability of preferences between cities: Estimating the benefits of stormwater management. *Journal of Environmental Economics and Management,* 99**,** 102274.

Aoki, K., Shen, J. & Saijo, T. (2010). Consumer reaction to information on food additives: Evidence from an eating experiment and a field survey. *Journal of Economic Behavior & Organization,* 73(3)**,** 433-438.

Arbiol, J., Borja, M., Yabe, M., Nomura, H., Gloriani, N. & Yoshida, S. (2013). Valuing Human Leptospirosis Prevention Using the Opportunity Cost of Labor. *International Journal of Environmental Research and Public Health,* 10(5)**,** 1845-1860.

Asrat, P., Belay, K. & Hamito, D. (2004). Determinants of farmers' willingness to pay for soil conservation practices in the southeastern highlands of Ethiopia. *Land Degradation & Development,* 15(4)**,** 423-438.

Aurino, E. (2017). Do boys eat better than girls in India? Longitudinal evidence on dietary diversity and food consumption disparities among children and adolescents. *Economics & Human Biology,* 25**,** 99-111.

Bennett, J. & Birol, E. (2010). Choice Experiments in Developing Countries: Implementation, Challenges and Policy Implications. Cheltenham: Edward Elgar.

Bhalotra, S., Chakravarty, A., Mookherjee, D. & Pino, F. (2019). Property Rights and Gender Bias: Evidence from Land Reform in West Bengal. *American Economic Journal: Applied Economics,* 11(2)**,** 205-237.

Birol, E., Asare-Marfo, D., Karandikar, B. & Roy, D. (2011). A latent class approach to investigating farmer demand for biofortified staple food crops in developing countries: The case of high-iron pearl millet in Maharashtra, India. *IDEAS Working Paper Series from RePEc*.

Blackorby, C. & Donaldson, D. (1990). A review article: The case against the use of the sum of compensating variations in cost-benefit analysis. *Canadian Journal of Economics,* 23(3)**,** 471.

Bockstael, N. E., Strand, I. E. & Hanemann, W. M. (1987). Time and the Recreational Demand Model. *American Journal of Agricultural Economics,* 69(2)**,** 293-302.

Brouwer, R., Akter, S. & Brander, L. (2009). Economic valuation of flood risk exposure and reduction in a severely flood prone developing country. *Environment and Development Economics,* 14**,** 397-417.

Carlsson, F. & Johansson-Stenman, O. (2010). Scale factors and hypothetical referenda: A clarifying note. *Journal of Environmental Economics and Management,* 59(3)**,** 286-292.

Carson, R. T. & Czajkowski, M. (2019). A new baseline model for estimating willingness to pay from discrete choice models. *Journal of Environmental Economics and Management,* 95**,** 57-61.

Casiwan-Launio, C., Shinbo, T. & Morooka, Y. (2011). Island Villagers' Willingness to Work or Pay for Sustainability of a Marine Fishery Reserve: Case of San Miguel Island, Philippines. *Coastal Management,* 39(5)**,** 459-477.

Cesario, F. J. (1976). Value of Time in Recreation Benefit Studies. *Land Economics,* 52(1)**,** 32-41.

ChoiceMetrics (2018). Ngene 1.2 User Manual & Reference Guide.

Chowdhury, S., Meenakshi, J. V., Tomlins, K. I. & Owori, C. (2011). Are Consumers in Developing Countries Willing to Pay More for Micronutrient-Dense Biofortified Foods? Evidence from a Field Experiment in Uganda. *American Journal of Agricultural Economics,* 93(1)**,** 83-97.

Czajkowski, M., Giergiczny, M., Kronenberg, J. & Englin, J. (2019). The Individual Travel Cost Method with Consumer-Specific Values of Travel Time Savings. *Environmental and Resource Economics,* 74(3)**,** 961-984.

Davies, A. L., Colombo, S. & Hanley, N. (2014). Improving the application of long-term ecology in conservation and land management. *Journal of Applied Ecology,* 51(1)**,** 63-70.

De Groote, H. & Kimenju, S. C. (2008). Comparing consumer preferences for color and nutritional quality in maize: Application of a semi-double-bound logistic model on urban consumers in Kenya. *Food Policy,* 33(4)**,** 362-370.

De Groote, H., Kimenju, S. C. & Morawetz, U. B. (2011). Estimating consumer willingness to pay for food quality with experimental auctions: the case of yellow versus fortified maize meal in Kenya. *Agricultural Economics,* 42(1)**,** 1-16.

DeShazo, J. R. & Fermo, G. (2002). Designing Choice Sets for Stated Preference Methods: The Effects of Complexity on Choice Consistency. *Journal of Environmental Economics and Management,* 44(1)**,** 123-143.

DFID. (2019). *Supporting Nutrition in Pakistan (SNIP)* [Online]. Available: https://devtracker.dfid.gov.uk/projects/GB-1-204023 [Accessed 26 May 2019].

Diafas, I., Barkmann, J. & Mburu, J. (2017). Measurement of Bequest Value Using a Non-monetary Payment in a Choice Experiment-The Case of Improving Forest Ecosystem Services for the Benefit of Local Communities in Rural Kenya. *Ecological Economics,* 140**,** 157-165.

Durán-Medraño, R., Varela, E., Garza-Gil, D., Prada, A., Vázquez, M. X. & Soliño, M. (2017). Valuation of terrestrial and marine biodiversity losses caused by forest wildfires. *Journal of Behavioral and Experimental Economics,* 71**,** 88-95.

Echessah, P. N., Swallow, B. M., Kamara, D. W. & Curry, J. J. (1997). Willingness to contribute labor and money to tsetse control: Application of contingent valuation in Busia District, Kenya. *World Development,* 25(2)**,** 239-253.

Ehmke, M. D., Lusk, J. L. & List, J. A. (2008). Is Hypothetical Bias a Universal Phenomenon? A Multinational Investigation. *Land Economics,* 84(3), 489-500.

Eom, Y.-S. & Larson, D. M. (2006). Valuing housework time from willingness to spend time and money for environmental quality improvements. *Review of Economics of the Household,* 4(3), 205-227.

European Commission. (2017). *Strengthening the Kenya National Food Fortification Programme to improve the health and nutritional status of poor and vulnerable groups* [Online]. Available: https://bit.ly/37QYeLM [Accessed 26 May 2019].

FAO. (2018). *The State of Food Security and Nutrition in the World: Building Climate Resilience for Food Security and Nutrition* [Online]. Available: http://www.fao.org/3/I9553EN/i9553en.pdf [Accessed 17 November 2020].

Feather, P. & Shaw, W. D. (1999). Estimating the Cost of Leisure Time for Recreation Demand Models. *Journal of Environmental Economics and Management,* 38(1), 49-65.

Fiebig, D. G., Keane, M. P., Louviere, J. & Wasi, N. (2010). The Generalized Multinomial Logit Model: Accounting for Scale and Coefficient Heterogeneity. *Marketing Science,* 29(3), 393-421.

Fleurbaey, M. & Abi-Rafeh, R. (2016). The Use of Distributional Weights in Benefit–Cost Analysis: Insights from Welfare Economics. *Review of Environmental Economics and Policy,* 10(2), 286-307.

Gibson, J. M., Rigby, D., Polya, D. A. & Russell, N. (2016). Discrete Choice Experiments in Developing Countries: Willingness to Pay Versus Willingness to Work. *Environmental & Resource Economics,* 65(4), 697-721.

González, C., Johnson, N. & Qaim, M. (2009). Consumer Acceptance of Second-Generation GM Foods: The Case of Biofortified Cassava in the North-east of Brazil. *Journal of Agricultural Economics,* 60(3), 604-624.

Government of India. (2018). *R6.3 Work Category Wise Analysis for FY: 2016-2017* [Online]. Available: https://bit.ly/3rwzald [Accessed 6 June 2018].

Government of Kenya. (2012). *Legal Notice No. 62 - Food, Drugs and Chemical Substances Act* [Online]. Available: https://bit.ly/2WKinN4 [Accessed 17 November 2020].

Government of Kenya. (2018). *Legal Notice No. 2 - The Labour Institutions Act* [Online]. Available: https://bit.ly/37M5mZU [Accessed 1 December 2020].

Gu, Y., Hole, A. R. & Knox, S. (2013). Fitting the generalized multinomial logit model in Stata. *Stata Journal,* 13(2), 382-397.

Hagedoorn, L. C., Koetse, M. J., van Beukering, P. J. H. & Brander, L. M. (2020). Time equals money? Valuing ecosystem-based adaptation in a developing country context. *Environment and Development Economics,* 25(5), 482-508.

Hardner, J. J. (1996). Measuring the value of potable water in partially monetized rural economies. *Water Resources Bulletin,* 32(6), 1361-1366.

Hensher, D., Louviere, J. & Swait, J. (1999). Combining sources of preference data. *Journal of Econometrics,* 89(1), 197-221.

Hensher, D. A., Rose, J. M. & Greene, W. H. (2015). Chapter 4: Families of discrete choice models. In: Applied Choice Analysis (2nd ed.). Cambridge: Cambridge University Press.

Hess, S. & Rose, J. M. (2009). Should Reference Alternatives in Pivot Design SC Surveys be Treated Differently? *Environmental and Resource Economics,* 42(3)**,** 297-317.

Hess, S. & Train, K. (2017). Correlation and scale in mixed logit models. *Journal of Choice Modelling,* 23**,** 1-8.

Hoaglin, D. C., Iglewicz, B. & Tukey, J. W. (1986). Performance of Some Resistant Rules for Outlier Labeling. *Journal of the American Statistical Association,* 81(396)**,** 991-999.

Hoffmann, B. (2018). Do non-monetary prices target the poor? Evidence from a field experiment in India. *Journal of Development Economics,* 133**,** 15-32.

Hole, A. (2006). Small-sample properties of tests for heteroscedasticity in the conditional logit model. *Economics Bulletin,* 3(18).

Hole, A. R. (2007a). A comparison of approaches to estimating confidence intervals for willingness to pay measures. *Health Economics,* 16(8)**,** 827-840.

Hole, A. R. (2007b). Fitting Mixed Logit Models by Using Maximum Simulated Likelihood. *The Stata Journal,* 7(3)**,** 388-401.

Holmes, T. P., Adamowicz, W. L. & Carlsson, F. (2017). Chapter 5: Choice Experiments. In: A Primer on Nonmarket Valuation (2nd ed. Vol. 13). Dordrecht: Springer Netherlands.

Horton, S., Mannar, V. & Wesley, A. (2008). Best Practice Paper - Food Fortification with Iron and Iodine. Copenhagen Consensus Centre - Working Paper.

Hossack, F. & An, H. (2015). Does payment type affect willingness-to-pay? Valuing new seed varieties in India. *Environment and Development Economics,* 20(3)**,** 407-423.

Hung, L. T., Loomis, J. B. & Thinh, V. T. (2007). Comparing money and labour payment in contingent valuation: the case of forest fire prevention in Vietnamese context. *Journal of International Development,* 19(2)**,** 173-185.

Johansson-Stenman, O. (2005). Distributional Weights in Cost-Benefit Analysis: Should We Forget about Them? *Land Economics,* 81(3)**,** 337-352.

Johansson-Stenman, O. & Svedsäter, H. (2012). Self-image and valuation of moral goods: Stated versus actual willingness to pay. *Journal of Economic Behavior & Organization,* 84(3)**,** 879-891.

Jurriëns, M. & Mollinga, P. (1996). Scarcity by Design: Protective Irrigation in India and Pakistan. *Icid jounal,* 45(2)**,** 31-53.

Kamuanga, M., Swallow, B. M., Sigue, H. & Bauer, B. (2001). Evaluating contingent and actual contributions to a local public good: Tsetse control in the Yale agro-pastoral zone, Burkina Faso. *Ecological Economics,* 39(1)**,** 115-130.

Kassahun, H. T., Jacobsen, J. B. & Nicholson, C. F. (2020). Revisiting money and labor for valuing environmental goods and services in developing countries. *Ecological Economics,* 177**,** 106771.

Kaul, T. (2018). Intra-household allocation of educational expenses: Gender discrimination and investing in the future. *World Development,* 104**,** 336-343.

Keane, M. & Wasi, N. (2013). Comparing Alternative Models of Heterogeneity in Consumer Choice Behavior. *Journal of Applied Econometrics,* 28(6)**,** 1018-1045.

Kenya National Bureau of Statistics. (2020). *Economic Survey 2020, Table 3.7: Average Wage Earnings per Employee, 2015- 2019* [Online]. Available: https://bit.ly/38APl8b [Accessed 21 November 2020].

Lankia, T., Neuvonen, M., Pouta, E. & Sievanen, T. (2014). Willingness to contribute to the management of recreational quality on private lands in Finland. *Journal of Forest Economics,* 20(2)**,** 141-160.
Larson, D. M., Pienaar, E. F. & Jarvis, L. S. (2015). Wildlife conservation, labor supply and time values in rural Botswana. *Environment and Development Economics,* 21(2)**,** 135-157.
Larson, D. M., Shaikh, S. L. & Layton, D. F. (2004). Revealing Preferences for Leisure Time from Stated Preference Data. *American Journal of Agricultural Economics,* 86(2)**,** 307-320.
Layard, R., Mayraz, G. & Nickell, S. (2008). The marginal utility of income. *Journal of Public Economics,* 92(8)**,** 1846-1857.
Lee, C. H. & Wang, C. H. (2017). Estimating Residents' Preferences of the Land Use Program Surrounding Forest Park, Taiwan. *Sustainability,* 9(4)**,** 1-19.
Lew, D. K. & Larson, D. M. (2005). Accounting for stochastic shadow values of time in discrete-choice recreation demand models. *Journal of Environmental Economics and Management,* 50(2)**,** 341-361.
List, J. A. (2001). Do Explicit Warnings Eliminate the Hypothetical Bias in Elicitation Procedures? Evidence from Field Auctions for Sportscards. *American Economic Review,* 91(5)**,** 1498-1507.
Lloyd-Smith, P., Abbott, J. K., Adamowicz, W. & Willard, D. (2019). Decoupling the Value of Leisure Time from Labor Market Returns in Travel Cost Models. *Journal of the Association of Environmental and Resource Economists,* 6(2)**,** 215-242.
Loomis, J. B. (2014). 2013 WAEA Keynote Address: Strategies for Overcoming Hypothetical Bias in Stated Preference Surveys. *Journal of Agricultural and Resource Economics,* 39(1)**,** 34-46.
Loureiro, M. L., Umberger, W. J. & Hine, S. (2003). Testing the initial endowment effect in experimental auctions. *Applied Economics Letters,* 10(5)**,** 271-275.
Louviere, J. & Eagle, T. (2006). Confound it! That Pesky Little Scale Constant Messes Up Our Convenient Assumptions. *Proceedings of the Sawtooth Software Conference***,** 211-228.
McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. *In:* Zarembka, P. (ed.) *Frontiers in Econometrics.* New York: Academic Press.
Meenakshi, J. V., Banerji, A., Manyong, V., Tomlins, K., Mittal, N. & Hamukwala, P. (2012). Using a discrete choice experiment to elicit the demand for a nutritious food: Willingness-to-pay for orange maize in rural Zambia. *Journal of Health Economics,* 31(1)**,** 62-71.
Meginnis, K., Hanley, N., Mujumbusi, L. & Lamberton, P. H. L. (2020). Non-monetary numeraires: Varying the payment vehicle in a choice experiment for health interventions in Uganda. *Ecological Economics,* 170**,** 106569.
Mohajan, H. (2014). Food and Nutrition Scenario of Kenya (Vol. 2).
Murphy, J. J., Allen, P. G., Stevens, T. H. & Weatherhead, D. (2005). A Meta-analysis of Hypothetical Bias in Stated Preference Valuation. *Environmental and Resource Economics,* 30(3)**,** 313-325.
Navrud, S., Tuan, T. H. & Tinh, B. D. (2012). Estimating the welfare loss to households from natural disasters in developing countries: a contingent valuation study of flooding in Vietnam. *Global Health Action,* 5**,** 1-11.

Navrud, S. & Vondolia, G. K. (2020). Farmers' preferences for reductions in flood risk under monetary and non-monetary payment modes. *Water Resources and Economics,* 30**,** 100151.

Nurmi, V. & Ahtiainen, H. (2018). Distributional Weights in Environmental Valuation and Cost-benefit Analysis: Theory and Practice. *Ecological Economics,* 150**,** 217-228.

Nyborg, K. (2014). Project evaluation with democratic decision-making: What does cost–benefit analysis really measure? *Ecological Economics,* 106**,** 124-131.

O'Garra, T. (2009). Bequest Values for Marine Resources: How Important for Indigenous Communities in Less-Developed Economies? *Environmental & Resource Economics,* 44(2)**,** 179-202.

Palmquist, R. B., Phaneuf, D. J. & Smith, V. K. (2010). Short Run Constraints and the Increasing Marginal Value of Time in Recreation. *Environmental & Resource Economics,* 46(1)**,** 19-41.

Penn, J. & Hu, W. (2019). Cheap talk efficacy under potential and actual Hypothetical Bias: A meta-analysis. *Journal of Environmental Economics and Management,* 96**,** 22-35.

Penn, J. M. & Hu, W. (2018). Understanding Hypothetical Bias: An Enhanced Meta-Analysis. *American Journal of Agricultural Economics,* 100(4)**,** 1186-1206.

Pokou, K., Kamuanga, M. J. B. & N'Gbo, A. G. M. (2010). Farmers' willingness to contribute to tsetse and trypanosomosis control in West Africa: The case of northern Côte d'Ivoire. *Biotechnology, Agronomy, Society and Environment,* 14(3)**,** 441-450.

Pondorfer, A. & Rehdanz, K. (2018). Eliciting Preferences for Public Goods in Nonmonetized Communities: Accounting for Preference Uncertainty. *Land Economics,* 94(1)**,** 73-86.

Rai, R. K. & Scarborough, H. (2013). Economic value of mitigation of plant invaders in a subsistence economy: incorporating labour as a mode of payment. *Environment and Development Economics,* 18(2)**,** 225-244.

Rai, R. K. & Scarborough, H. (2015). Nonmarket valuation in developing countries: incorporating labour contributions in environmental benefits estimates. *Australian Journal of Agricultural and Resource Economics,* 59(4)**,** 479-498.

Rai, R. K., Shyamsundar, P., Nepal, M. & Bhatta, L. D. (2015). Differences in demand for watershed services: Understanding preferences through a choice experiment in the Koshi Basin of Nepal. *Ecological Economics,* 119**,** 274-283.

Samaj Pragati Sahayog (2016). Scoping Study on Infrastructure for Climate Resilient Growth through MGNREGA (DRAFT final report) for the Department for International Development (DFID).

Schiappacasse, I., Vásquez, F., Nahuelhual, L. & Echeverría, C. (2013). Labor as a welfare measure in contingent valuation: the value of a forest restoration project. *Ciencia e investigación agraria,* 40**,** 69-84.

Schläpfer, F. & Fischhoff, B. (2012). Task familiarity and contextual cues predict hypothetical bias in a meta-analysis of stated preference studies. *Ecological Economics,* 81**,** 44-47.

Shyamsundar, P. & Kramer, R. A. (1996). Tropical forest protection: An empirical analysis of the costs borne by local people. *Journal of Environmental Economics and Management,* 31(2)**,** 129-144.

Susilo, H., Takahashi, Y. & Yabe, M. (2017). The Opportunity Cost of Labor for Valuing Mangrove Restoration in Mahakam Delta, Indonesia. *Sustainability,* 9(12), 2169.

Sutton, W. R., Larson, D. M. & Jarvis, L. S. (2008). Assessing the costs of living with wildlife in developing countries using willingness to pay. *Environment and Development Economics,* 13(4), 475-495.

Swallow, B. M. & Woudyalew, M. (1994). Evaluating willingness to contribute to a local public good: application of contingent valuation to tsetse control in Ethiopia. *Ecological Economics,* 11(2), 153-161.

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review,* 34(4), 273-286.

Tilahun, M., Birner, R. & Ilukor, J. (2017). Household-level preferences for mitigation of Prosopis juliflora invasion in the Afar region of Ethiopia: a contingent valuation. *Journal of Environmental Planning and Management,* 60(2), 282-308.

Tilahun, M., Vranken, L., Muys, B., Deckers, J., Gebregziabher, K., Gebrehiwot, K., Bauer, H. & Mathijs, E. (2015). Rural Households' Demand for Frankincense Forest Conservation in Tigray, Ethiopia: A Contingent Valuation Analysis. *Land Degradation & Development,* 26(7), 642-653.

UNFAO. (2018). *India at a glance* [Online]. Available: https://bit.ly/3rtNhaS [Accessed 24 August 2018].

UNICEF. (2019). *Vitamin A* [Online]. Available: https://bit.ly/3phNa0a [Accessed 21 May 2019].

United Nations. (2015). *Transforming our World: The 2030 Agenda for Sustainable Development* [Online]. Available: https://bit.ly/2JriDOg [Accessed 17 November 2020].

Vasquez, W. F. (2014). Willingness to pay and willingness to work for improvements of municipal and community-managed water services. *Water Resources Research,* 50(10), 8002-8014.

Vondolia, G. K., Eggert, H., Navrud, S. & Stage, J. (2014). What do respondents bring to contingent valuation? A comparison of monetary and labour payment vehicles. *Journal of Environmental Economics and Policy,* 3(3), 253-267.

Vondolia, G. K. & Navrud, S. (2018). Are non-monetary payment modes more uncertain for stated preference elicitation in developing countries? *Journal of Choice Modelling,* 30, 73-87.

Whittington, D. & Cook, J. (2019). Valuing Changes in Time Use in Low- and Middle-Income Countries. *Journal of Benefit-Cost Analysis,* 10(S1), 51-72.

WHO. (2014). *Global Nutrition Targets 2025 - Anaemia Policy Brief* [Online]. Available: https://bit.ly/37P1WW6 [Accessed 17 November 2020].

WHO & FAO. (2006). *Guidelines on food fortification with micronutrients* [Online]. Available: https://bit.ly/38Auxhb [Accessed 17 November 2020].

WHO/UNICEF. (2017). *Progress on Drinking Water, Sanitation and Hygiene: 2017 Update and SDG Baselines* [Online]. Geneva. Available: https://www.unicef.org/publications/index_96611.html [Accessed 24 August 2018].

Wikimedia Commons. (2013). *File: India Orissa location map.svg by Milenioscuro* [Online]. Available: https://bit.ly/38FJZIQ [Accessed 15 December 2018].

Wikimedia Commons. (2016a). *File: India Chhattisgarh location map.svg by Milenioscuro* [Online]. Available: https://bit.ly/34KRgWV [Accessed 15 December 2018].
Wikimedia Commons. (2016b). *File: Indian-peninsula.png by P.Divya Reddy* [Online]. Available: https://bit.ly/3mR67VU [Accessed 15 December 2018].
World Bank. (2018a). *Poverty and Equity Database. Series: Number of poor at $1.90 a day (2011 PPP) (millions), Rural poverty headcount ratio at national poverty lines (% of rural population), Rural population.* [Online]. Available: https://bit.ly/3nPAHR6 [Accessed].
World Bank. (2018b). *Series: Labour force participation rate, female (% of female population ages 15-64) (ILO estimate) + Series: Labour force participation rate, male (% of male population ages 15-64) (ILO estimate)* [Online]. Available: https://bit.ly/34Nb4IW [Accessed 1 December 2018].
World Bank. (2018c). *World Development Indicators. Series: People using at least basic drinking water services, rural (% of rural population)* [Online]. Available: https://bit.ly/3rtNV8i [Accessed].