Norwegian School of Economics
Bergen, Spring 2023

Behavioral Responses to Tax Salience: Evidence from Tanzania

An Empirical Analysis of the Effect of Tax Salience on Mobile Money Transactions

Sunniva Ingholm and Emil Løstegård

Supervisor: Vincent Somville

Master thesis, Economics and Business Administration

Major: Energy, Natural Resources & the Environment (ENE)

and Economics (ECN)

NORWEGIAN SCHOOL OF ECONOMICS

Acknowledgements

This thesis process has been a great experience which gave us the opportunity to collect primary data in Dar es Salaam, Tanzania, in February 2023. We want to express our sincere gratitude to our supervisor Vincent Somville. His knowledge, guidance, and availability throughout the process have been invaluable.

We also want to extend a special thanks to our supervisor at Chr. Michelsen Institue (CMI), Senior Researcher Ingrid Hoem Sjursen, for including us in their exciting research project. It has proved to be an inspiring process where we received firsthand experience conducting research, and we are happy that we got to gather data under the project. Without Ingrid, this thesis would not have been possible, and her guidance and enthusiasm have made this a rich and enjoyable learning experience. In addition, we would like to thank Ph.D. candidate Jasmin Vietz, for good collaboration during our time in Tanzania. At CMI, we want to extend our gratitude to Professor Odd-Helge Fjeldstad for good discussions and valuable input throughout this process. We have also valued being part of the academic environment at CMI where we have enjoyed interesting conversations spanning various fields of development studies. This has been very motivating.

Our time in Tanzania was a great experience, and we want to thank REPOA for assisting us in this project. The think tank has provided skillful enumerators, efficient recruitment of participants, and created suitable premises for a successful experiment.

Norwegian School of Economics

Bergen, May 2023

Your name here

Your name here

Abstract

A central assumption in public economics is that individuals optimize fully to tax policies, meaning they respond to tax changes similarly to price changes. However, a growing body of evidence in the field of tax salience suggests that individuals optimize imperfectly to the available information due to inattention. Our thesis contributes to the field of tax salience by empirically investigating whether individuals optimize fully to taxes.

We analyze the effect of salience on behavioral responses to taxes and fees on mobile money transactions in Tanzania. In February 2023, we collected primary data through a laboratory-in-the-field experiment in Dar es Salaam where we exploited the attention to the heavily disputed Electronic Money Transaction Levy. Using a willingness to pay elicitation method, we provide evidence that increasing the salience of taxes and fees of mobile money transactions reduces the willingness to pay for using mobile money services. The results provide support for the salience theory and suggest that the salience of a tax has policy implications. To our knowledge, we are the first to estimate a causal effect of salience on taxes levied on mobile transactions.

Keywords – Tax Salience, Multiple Price List, WTP, Tanzania, Mobile Money

Acronyms

BDM = Becker-deGroot-Marschak

COSTECH = Tanzania Commission for Science and Technology

GDPR = General Data Protection Regulation

MM = Mobile money

MPL = Multiple price list

MSB = Multiple-switching behavior

NSB = Never-switching behavior

NSD = Norwegian Social Science Data Services

OLS = Ordinary least squares

TZS = Tanzanian Shilling

USD = United States Dollar

VAT = Value-added tax

WTP = Willingness to pay

Contents

List of Figures

List of Tables

1 Introduction

1.1 Motivation and purpose

Mobilizing domestic revenues to finance public expenditures is crucial for economic and social development. Therefore, developing effective tax and benefit systems have become indispensable for countries in Sub-Saharan Africa. Many of these countries have recently imposed taxes on mobile money transactions as a new way of expanding the tax base (Matheson and Petit, 2021).

Mobile money (MM) has become a key driver of financial inclusion in Tanzania. Introduced in the late 2000s, it has steadily increased in volume and value of MM transactions, with 3.75 billion transactions in 2021 (GSMA, 2023). MM is a payment alternative to cash, where registered users can deposit cash in their digital MM account. It enables individuals to register and access bank services through their mobile operators, and the process is less extensive than registering for a traditional bank account. Hence, MM increases financial inclusion by making bank services accessible to more citizens.

In July 2021, Tanzania introduced the Electronic Money Transaction Levy, which taxes mobile transactions and is designed similarly to a sales tax (GSMA, 2021). Introducing such a tax caused public outcry and was vastly criticized by the citizens as well as tax stakeholders, ultimately resulting in the number of transactions declining heavily (GSMA, 2021). The main critique is that the tax hampers the progress of increasing financial inclusion in Tanzania, since the additional taxes will raise the barrier for registering and using MM services. However, the tax accounted for 1.7% of total domestic revenue for the government in the fiscal year of 2021/2022. In comparison, this is higher than the revenue generated from personal income taxes, which accounted for 1.1% of the domestic revenue (Tanzania Revenue Authority, 2023). Therefore, the levy has become an important source of tax revenue, which according to the government, was to be used for public goods like education, health, and infrastructure (Ministry of Finance and Planning, 2022).

A central assumption in public economics is that individuals optimize fully with respect to tax policies. Therefore, individuals respond to tax changes similarly to price changes (Ramsey, 1927). However, contrary to this assumption, a growing body of evidence suggests

that individuals optimize imperfectly to the available information due to inattention (Chetty et al., 2007; Finkelstein, 2009; Gamage and Shanske, 2011). Accordingly, new models have been introduced in order to reflect how people actually respond to taxes.

One such model is developed by Bordalo et al. (2022) where they introduced a general model of memory. They argue that salient attributes, or more visible features of a good, draw attention away from more goal-relevant non-salient attributes. For instance, Chetty et al. (2007) conducted a study in US grocery stores where they found that demand for goods was reduced when sales taxes were displayed in the price tag, as opposed to being added at the counter. Notably, the participants demonstrated knowledge about the sales tax when asked about it. This suggests that people tend to focus on the salient attributes of a good and ignore non-salient features.

Similarly to Chetty et al. (2007), we will investigate tax salience using the generalized model of memory introduced by Bordalo et al. (2022). Tax salience emphasizes that how taxes are displayed can affect how people respond to them (Varela, 2016). In particular, people are more likely to change their behavior in response to highly visible and salient taxes. Tax salience, therefore, implies that people respond to factors other than net tax liability. If reducing the salience of taxes results in citizens responding less to them, it can improve tax collection. In this scenario, a tax is considered efficient when minimizing disruptions to production and consumption decisions.

This thesis investigates empirically whether individuals optimize fully with respect to taxes by analyzing the effect of salience on behavioral responses to MM taxes and fees. We will elicit willingness-to-pay (WTP) by using a multiple price list (MPL) method to examine the effects of tax salience. Particularly, we will conduct a randomized laboratory-in-the-field experiment in Dar es Salaam, Tanzania, to investigate if increasing the salience reduces the WTP for taxes and fees related to MM. Additionally, for exploratory reasons, this thesis will explore attitudes, usage, and knowledge about MM in Tanzania and discuss potential policy implications.

1.2 Research question

Research on tax salience is relatively limited but has gained increasing interest in recent years. Tax salience will be examined by applying it to the current MM tax in Tanzania and investigating if the theory also applies to this setting. As the MM tax is a new phenomenon, research on the topic is also limited. To our knowledge, there are no similar studies that combine these two topics. This thesis seeks to provide a first step in addressing this gap. Thus, we aim to investigate the following research question:

Does the salience of the tax on mobile money transactions reduce the willingness to pay for using mobile money instead of cash in Tanzania?

This thesis is structured in the following sections: In section 2, we give a brief overview of the salience theory and related research, as well as explain the MPL method to elicit WTP. Section 3 gives background information about MM in Tanzania. Next, section 4 describes our data collection process, with a detailed description of the experimental design and sample, before introducing the empirical strategy in section 5. In section 6, we present the empirical results from the experiment, followed by a discussion of findings and limitations related to these results in section 7. Lastly, our conclusion is presented in section 8.

2 Literature Review

2.1 Salience theory

2.1.1 Top-down and bottom-up attention

Salience can be understood as the extent to which stimuli of a good draw attention. These stimuli could be any object, event, or piece of information that captures the attention of an individual (Bordalo et al., 2022). As previously mentioned, the salience theory challenges the assumptions of rationality and complete information, which are fundamental in standard economic theory. Therefore, one assumes individuals to be driven by top-down attention, which means that attention is unlimited and voluntarily allocated to reach the individual's goals and expectations (Treisman and Gelade, 1980). Accordingly, the individual would then consider all relevant attributes of a good and make accurate cost-benefit calculations of the information available. On the contrary, salience theory acknowledges that individuals do not have complete or precise knowledge of all relevant factors and may rely on salient cues or heuristics to make decisions. Their attention is then also driven bottom-up, which distracts individuals from their immediate goals or the relevant attributes of the good (Wolfe, 1994). This concept is crucial in understanding how individuals allocate their limited cognitive resources and make decisions (Bordalo et al., 2022).

Bordalo et al. (2022) introduce a general salience model, which is a model of memory that explains how the invisible attributes of a good are underweighted, and the visible attributes are overweighted. Furthermore, individuals tend to neglect the invisible attributes although they are familiar with them. Contrary to the assumed top-down attention in standard economic theory, the model acknowledges that salient stimuli of a good attract bottom-up attention. This is due to being in high contrast to the surroundings, being surprising relative to recalled experiences, or having prominent features. The focus of this study will be on how prominence stimuli drive bottom-up attention and how it is also affected by contrast and surprise stimuli. Unlike contrast or surprise, which come from comparing stimuli, prominence comes from factors exogenous to the stimulus itself.

Prominence stimuli refer to the idea that stimuli that are highly available to our senses or in memory are more salient. Typically, stimuli that have recently attracted attention continue to do so, even if they are no longer task-relevant. Bottom-up attention to prominent attributes of a good can explain why people neglect hidden attributes (Bordalo et al., 2022), e.g., sales taxes which was previously mentioned (Chetty et al., 2007)

Contrast stimuli capture the idea that a specific attribute of a good stands out when the good is compared to alternatives, e.g., a more creative design of a good contrasts the generic goods of competitors. Since it draws bottom-up attention, it distracts the consumer from paying enough attention to other attributes, such as price (Bordalo et al., 2022).

Surprise stimuli encapsulate that humans are highly sensitive to changes in sensory inputs, and it can therefore be viewed as the contrast between a stimulus and the memories it evokes (Bordalo et al., 2022). For example, when the consumer sees the price of a good, it also retrieves past prices from memory. In a situation where the retrieved price is lower than the current price, it immediately becomes a very salient feature of the good and thus comes across as surprising. The connection between the different types of stimuli is explained in detail in the next section.

2.1.2 The salience model

The salience model captures how the intrinsic valuation of a good is based on salient and non-salient attributes. Suppose an individual is given information $k = (a_k)_{k\epsilon P}$ which captures the value of the good's attributes a_k which is naturally seen and considered salient. What is not naturally seen, are the remaining attributes $(a_k)_{k\epsilon I}$, which then are considered non-salient.

The model captures the surprise aspect by incorporating memory in the model. The individual regains from its mind the norms $(a_k^n)_{k\epsilon P}$ and $(a_k^n)_{k\epsilon I}$ for the visible and invisible attributes, or normal values, which is based on previous experiences. The salience of attribute k is then given by $\sigma(a_k, a^{-n}_k)$. As the model assumes the attributes of a good to act as triggers for memory retrievals, these retrievals generate an average version of the good. Therefore, a^{-n}_k is the memory-based reference that arises from calculating the average of the retrieved norms for all goods in the choice set. The expression for the

salience of attribute k then captures that an individual overweighs the attribute that is considered most surprising by having the highest contrast from its normal value. However, the individual could also recall experiences regarding the hidden attributes $(a_k^n)_{k\epsilon I}$.

To model contrast, the model introduces a bounded salience function $\sigma(a_k; \overline{a}_k) \geq 0$, which measures the contrast of a good's attribute a_k relative to the average value $\overline{a}_k$ of the same attribute for the goods in the choice set. This function captures the ordering property of salience, which means that an attribute is more salient if the valuation of the attribute is more different from the average value in the choice set. Furthermore, a given difference in the attribute value is more salient at lower values because of diminishing sensitivity. $w_k \equiv w(\sigma_k; \sigma_{-_k}) \geq 0$ is the weighting function, where σ_k denotes the salience of attribute k, and the vector $\sigma_{-k} = \{\sigma_i\}_{i \neq k}$ denotes the salience of the good's other attributes. The weight w_k attached to an attribute of a good in a choice set increases in the salience σ_k of attribute k and decreases weakly in the salience of other attributes σ_{-_k}. This means that when an attribute of a good becomes more salient, it causes the individual to assign a higher weight to that attribute while simultaneously assigning a lower weight to the less salient attributes in the choice set.

π_k is the decision weight attached to attribute k, and one can think about it as the optimal weight based on the individual's preferences or objective probabilities. A salient attribute k in the choice set will then draw attention and gets overweighed, which then distorts the decision weights to $\hat{\pi}_k > \pi_k$, and the opposite if the attribute is not salient. Each of the three types of stimuli point to different mechanisms for distorting the weights, which differ in various situations. r_k is the probability to recall the attribute $k\epsilon I$. The intrinsic valuation of a good can then be understood with the following function:

$$\hat{V}_P = \sum_{k\epsilon P} w_k \pi_k a_k + \sum_{k\epsilon I} r_k \tilde{w}_k \pi_k a_k^n \tag{2.1}$$

The first term of equation (2.1) represents the valuation of the visible attributes $k\epsilon P$, where the weight w_k is distorted by contrast and surprise stimuli relative to the retrieved norm of the attributes, a_k^n. π_k is the decision weight attached to attribute k. The second term represents the valuation of the invisible attributes $k\epsilon I$, which tend to be underweighted on average because recalling them might fail, meaning that $r_k < 1$. Their

weights $\tilde{w}_k$ only depend on contrast and not surprise relative to a_k^n, since the actual values of the attributes are non-observed. Furthermore, π_k is the decision weight attached to the invisible attribute k. Therefore, equation (2.1) suggests that features that are prominent drive bottom-up attention, especially if they are contrasting or surprising, while features that are not prominent are non-salient, although they are goal-relevant. Prominence is therefore a foundation of framing effects, meaning that normatively irrelevant changes in the way a choice is presented alter the decision (Bordalo et al., 2022). This salience model thus shows that bottom-up attention often mentally competes with top-down attention. Importantly, both visible and invisible attributes are familiar to the individual, but the invisible attributes tend to be underweighted since they may not be recalled at the moment of decision.

2.1.3 Tax salience

The field of tax salience is the application of this theory related to tax and emphasizes that how taxes are displayed can affect how people respond to them (Varela, 2016).By the same mechanism, bottom-up attention sheds light on tax salience, an idea introduced by Chetty et al. (2007). They found that displaying prices that include sales taxes reduces demand in grocery stores in the US. Instead of adding the sales taxes at the counter, they were stated on the price tags. Participants of the study did report the sales taxes accurately when asked to state them, meaning they knew about the tax beforehand, which indicates selective memory. Furthermore, it shows that when valuing goods, consumers tend to neglect important attributes not immediately available to them in the purchasing decision, e.g., sales taxes are not seen and get neglected. In these cases, valuation is distorted due to misleading mental representations since some of the important features of the good are not observed. The model helps to explain this phenomenon by viewing mental representations as the product of selective memory. This creates bias in two ways: 1) since norms are based on past experiences, a biased database creates a biased valuation, and 2) recalling is associative, which means that norms are tilted toward past experiences most similar to the cue. This implies that valuation biases are individual-specific, because they are shaped by personal experiences (Bordalo et al., 2022).

A study by Finkelstein (2009) finds that after toll facilities introduced electronic toll collection, the drivers passing through the toll booths are significantly less aware of tolls

paid electronically. This shows that drivers underweight the cost of highway tolls when they are paid electronically. The effect reflects the prominence aspect of the salience model since attention is directed towards the convenience of electronic toll booths and away from having to pay the highway tolls, which is made less salient in this case. Our research is similar to Finkelstein (2009) since we are researching how individuals react to non-salient attributes such as taxes.

Similarly to Chetty et al. (2007) and Finkelstein (2009), our thesis will investigate how the salience of indirect taxes affects decision-making. However, our experimental treatment will be alternated to fit our context. We want to test the assumption that participants have knowledge of the MM taxes and fees. Therefore, the specific amount of taxes will not be displayed since we will have them estimate the taxes and fees. Our experiment is designed to encapsulate the prominence aspect of the model on consumer choice. It is expected that individuals make choices based on the salient attributes of the good at the time of decision, and do not consider the non-salient attributes. The treatment is to alternate r_k, i.e., the probability of recalling the less visible attribute $k \epsilon I$ by reminding the treatment group of the taxes and fees for sending MM. Inspired by Chetty et al. (2007), we assume that individuals already have somewhat accurate knowledge of the taxes and fees beforehand. Therefore, we expect to find that increasing the salience of the tax on MM transactions reduces the WTP for using MM instead of cash in Tanzania.

New literature shows some interesting implications of tax salience, specifically related to policy development (de la Cuesta et al., 2020; Varela, 2016; Brautigam et al., 2008). Digital financial service providers have seen increasing turnover and such taxes introduce an opportunity to tap into this revenue (Lees and Akol, 2021). Therefore, it can be argued that tapping into this income is a redistribution measure that benefits society as a whole. Another aspect is that since indirect taxes are less visible to citizens, they might provoke less willingness to punish the government politically and thus mute or eliminate the effect of taxation on citizens' demand for accountability (de la Cuesta et al., 2020; Brautigam et al., 2008). This could have serious implications for the country, where improving democratic accountability is important.

2.2 WTP elicitation using a multiple price list format

To investigate the effect of salience, an approach to elicit participants' WTP is necessary. In this case, WTP refers to an individual's relative preferences between two alternatives which is expressed in a monetary amount (Jack et al., 2022). To find this preference, a multiple price list (MPL) will be applied. WTP is a well-known economic concept and is a measurement that can be used to estimate demand for public or private goods (Wertenbroch and Skiera, 2002). Understanding people's WTP for MM taxes and fees is useful for governments since this knowledge allows them to implement better tax regimes.

The WTP is elicited by asking questions repeatedly to understand the individual's relative valuation of two goods. A commonly used WTP elicitation method in experiment settings is a standard Becker-deGroot-Marschak (BDM) random price mechanism which identifies a single WTP value (Burchardi et al., 2021). However, this thesis will use an MPL elicitation method that obtains WTP intervals instead. Similarly to the BDM method, it is incentive-compatible, meaning that participants are incentivized to act truly. Contrary, the MPL method differs by being easier to understand, which serves as an important advantage (Jack et al., 2022). Additionally, the format is relatively transparent and provides simple incentives for truthful revelation.

There are several challenges when using a MPL estimation method. One challenge is that many individuals show so-called "multiple-switching behavior" (MSB) or "never-switching behavior" (NSB) (Jack et al., 2022). Multiple switching means that participants make internally-inconsistent choices which makes it difficult to interpret their true WTP i.e., it requires the participants to switch between two alternatives at most once and the choices must be logical. Failing to do so could entail that participants are prone to choice error or to framing or order effects. To reduce MSB, it is important to make the incentive-compatible features clear, while also making the MPL easy to understand. Reducing the complexity of an experiment has proved to be more important when conducting experiments with samples where individuals have low numeracy (Dave et al., 2010). Given our limited knowledge about the sample while also conducting the experiment in a foreign country, such an approach proves beneficial considering these circumstances.

Participants who show NSB consistently choose the same alternative throughout the price list. Such behavior could express a strong preference for one of the alternatives. However, it could also be due to potential anchoring and order effects, or differences in default behavior (Andersen et al., 2006; Burchardi et al., 2021). When it comes to reducing NSB, Jack et al. (2022) argue that one of the most effective measures to reduce such behavior is to vary the order of alternatives within each binary choice. This minimizes the possible anchoring and order effects arising from repeatedly choosing the same option.

Another challenge to the MPL method is that it only elicits interval responses. However, in an experimental setting, interval responses could give more truthful estimations compared to point estimates (Andersen et al., 2006). In addition, by controlling for MSB and NSB behaviors, we believe a MPL elicitation method is an effective way of measuring WTP. The specific design of the MPL used to elicit WTP will be further explained in section 4.1.5.

3 Background

The context of MM and the Electronic Money Transaction Levy have several key aspects that are important to understand when conducting this tax salience experiment. As previously mentioned, the levy is designed similarly to a sales tax meaning it is taxed indirectly and naturally less salient. However, recent debates have drawn attention to this MM tax which has increased its salience. This section will provide more information about MM in Tanzania and why the levy is controversial.

3.1 Mobile money in Tanzania

MM has gained increased popularity in recent years. It was first introduced in 2008, and in 2021 it was used by 45% of the population above the age of 15 (Akyoo and Pallangyo, 2021). By March 2023, the total number of MM subscribers in Tanzania exceeded 44.35 million (Tanzania Communications Regulatory Authority, 2023).[1] As mentioned, MM is a payment alternative to cash in Tanzania. Similarly to Paypal, registered users can deposit cash in their digital MM account and make withdrawals or transfers. MM can additionally be used to pay government fees, apply for insurance and loans, and as a savings account (Suri, 2017). These transactions are managed by mobile communication providers like Vodacom, Tigo, and Airtel and their licensed agents. Registration is free, but transactions are fee-based (Economides and Jeziorski, 2017). This implies that individuals pay a fee to mobile operators for sending or withdrawing money. The MM account is connected to a SIM card and does not require WiFi or a smartphone. Therefore, anyone with a SIM-card, a mobile phone, and a national ID can use the services (Suri, 2017).

Mobile banking has proved to be a driving force for increasing financial inclusion in Tanzania. According to FinScope Tanzania (2006), millions of people in Tanzania, particularly those living in remote areas or on low incomes, had limited options for transferring money and accessing financial services before the introduction of MM. Therefore, digital payment services have shown to be a key driver of economic development since it allows more citizens to participate in the formal economy (GSMA, 2021).

[1] It is worth noting that the total number is likely lower since many individuals have several subscriptions.

There are many benefits to having a MM account. One benefit is that MM provides security when sending money across long distances, compared to sending cash by mail or bus (Suri, 2017). Another benefit of MM banking is that it facilitates savings by providing an easy and secure place to store money. Furthermore, it allows individuals to easily access savings instead of investing the savings in less fluid assets such as livestock (Mawejje and Lakuma, 2019). This helps to diversify risk and provide economic stability as poor people tend to be harder hit by accidents, climate change, and economic instability (Abiona and Koppensteiner, 2022). Another benefit is that MM accounts enable individuals to save and show a steady income. Being able to provide proof of savings in a MM account allows individuals to apply for loans and insurance, and then be able to invest in business and in society (Lotto, 2018).

From a government perspective, MM has the potential to help achieve the 2030 sustainable development goals by driving sustainable and inclusive growth and providing solutions to some of the world's most challenging development challenges (GSMA, 2019). According to the World Bank (2023), 64% of Tanzanians live in rural areas and the benefit of safe and easy transactions across long distances allows for easier financial redistribution. Furthermore, an argument can be made that more investments result in higher productivity which incentivizes business development (Goldin, 2016). However, there are some drawbacks to MM as well. The main drawbacks is potential security and surveillance risks, tech issues, and the extra charges for services, to mention a few (Kvartalnyi, 2023).

Despite significant efforts from the government and other actors in the financial market, there are still 46% without access to any type of banking (World Bank, 2021). Of these, 71% live in rural areas which indicates significant differences in access based on geographic location. Furthermore, 88% of the people who do not have access to the bank system, including mobile banking services, do not have education above primary school. 50% of the unbanked are estimated to lack the documentation needed to open an account. There is also a significant gender gap for account ownership overall, including both financial institutions and mobile money accounts. This means that men are more likely to have access to financial institutions than women (World Bank, 2021). These numbers indicate that there is still a way to go before reaching full financial inclusion in Tanzania.

3.2 Electronic Money Transaction Levy

In recent years, the levy, also known as Tozo, has become a word most people know in Tanzania. The levy on mobile transactions was first introduced in July 2021 through the National Payment Systems Act, section 46A (GSMA, 2021). It was introduced in an effort to increase the tax base, and the average transaction fee was increased from 3% to 369% depending on the transaction size (GSMA, 2023). In the fiscal year 2021/2022, the MM levy collected TZS 358 billion in total domestic tax revenue, which accounts for 1.7% of the total domestic revenues.[2] In comparison, the personal income taxes, only accounted for 1.1% of the total domestic revenues (Tanzania Revenue Authority, 2023). Therefore, it is evident that the MM tax is an important tax for the Tanzanian government.

The levy design

The design of the MM tax is similar to a sales tax as, the sender pays it at the point of transaction. However, it differs from a sales tax since it only applies to MM and it is not universal for all payment methods. The tax is regressive, which means that the tax burden decreases when the transaction amount increases (GSMA, 2023). Furthermore, the tax comes in addition to other taxes like VAT and sales taxes. Additionally, the MM tax compound itself, e.g., although one pays MM tax to send money to a family member, the family member still needs to pay the tax to use the same money later (Niesten and Abounabhan, 2023).

The levy prompted debates and critiques from citizens and stakeholders alike. After its introduction, the number of transfers and cash withdrawals fell heavily the two consecutive months before starting to recover but on a lower growth trajectory afterward (GSMA, 2021). Subsequently, this resulted in the downward revision of the MM levy by 30% with the maximum levy chargeable on MM transactions being revised from the previous TZS 10,000 to TZS 7,000 in September 2021. Furthermore, the Finance Act 2022 amended the National Payment Systems Act by further revising downwards the maximum levy chargeable on MM transactions from TZS 7,000 to TZS 4,000 and expanding the scope of such levy to cater for all electronic money transactions as opposed to the previous inclusion of electronic MM transactions only (Tanzania Revenue Authority, 2022; Ministry of Finance and Planning, 2022). In September 2022, the Government decided to remove

[2]Total net domestic revenue is used and all four departments in TRA are included.

the levy on electronic transactions from banks to MM accounts and vice versa, as well as electronic transactions within one bank (Niesten and Abounabhan, 2023). Due to the recent attention to the tax, we argue that most citizens above the age of 18 in Tanzania know approximately how much the tax is, which is an important assumption in our experiment.

Implications of the levy

The levy has several implications for the state and its citizens. The standard theory of optimal taxation states that a tax should be designed and implemented to maximize a social welfare function subject to a set of constraints (Ramsey, 1927). There are two criteria by which economists measure the outcomes of tax policy, efficiency and equity. Efficiency is primarily concerned with objective measures and does not involve ethical and normative judgments. Therefore, it only considers how resources are allocated (Mirrlees, 1971). For example, what differentiates the MM tax from a value-added tax (VAT) is that the VAT, taxes consumption, while the MM tax, taxes both input factors and consumption (Mpofu, 2022). This can be considered inefficient since the price per good becomes high when taxes are compounded. Therefore, the MM tax discourages individuals from using the service and it does not incentivize productivity. However, based on the theory, we also argue that the MM levy can be considered efficient since it is easy to monitor and easy to collect, thus making it attractive for tax authorities. If reducing the salience of a tax can reduce the amount that people negatively respond to the tax, it can increase the efficiency of tax collection.

The other criterion by which economists measure the outcome of tax policies is equity, which considers the distribution of resources, where social norms and judgments are used to estimate the effect (Bejaković, 2020). When the tax was implemented, the government argued that the revenues from the levy would be used to construct 10,000 classrooms, 4500 dispensaries, and 570 healthcare centers (Nyonzo, 2022). Thus, one can argue that the revenues would be distributed back to the people. However, little information was shared about who would benefit from these public goods. According to Niesten and Abounabhan (2023), only 23% of the adult population (age 15 or older) in Tanzania owned a traditional bank account in 2021. Isolating taxes on mobile transactions, 21% of the population with access to traditional bank accounts have an alternative way of transferring money without

paying the tax. Although traditional bank accounts come with their own fees, the absence of an equivalent tax calls into question the fair distribution of the tax burden. As people using a traditional bank account do not have to pay the tax, people without access are left with either using MM or cash. However, some research on distributive politics also looks into the input side (Persson and Tengs, 2023). A case can be made that people who are part of the formal economy already pay taxes in other ways (Anyidoho et al., 2022). Following this argument, one can argue that individuals who pay taxes should also receive more public goods. The tax is therefore meant to be skewed with the aim of taxing the informal sector, which is less taxed on average. This is an argument that was used by, e.g., the government of Ghana when introducing a similar tax (Anyidoho et al., 2022) .

There is currently limited data and research available on the implications of the MM tax on efficiency and equity, especially in the contexts of high informality and minimal access to traditional financial services. A recent meta-review found that there are no existing studies that provide robust insights as to how best to tax digital financial services fairly and transparently (Mader et al., 2022).

This study does not focus on MM but instead uses MM as an instrument in measuring tax salience. The tax salience effects on the MM tax has implications that will be discussed in section 7.

4 Data Collection

The data used in this thesis was collected in a laboratory-in-the-field experiment over two weeks in February 2023. The sample consists of vendors from 16 different markets in Dar es Salaam, Tanzania. A research permit to collect data in Tanzania was obtained through COSTECH[3] before collecting any data. The data collection was also approved prior to arrival by Norwegian Social Science Data Services (NSD).[4] We received ethical approval from the Ethics Committee of the Norwegian School of Economics.[5] The following sections will examine laboratory experiments as a data collection method, the experimental design and sample, and lastly, the context of the experiment will be explained.

4.1 Experimental design

4.1.1 Experiment as a research method

An experiment is a quantitative research method that aims to investigate a cause-and-effect relationship, and thus: "..useful for making predictions about the consequences of changing circumstances or policies" (Angrist and Pischke, 2009). Therefore, we explore if a change in an independent variable causes a change in a dependent variable. The main outcome has an explanatory purpose on tax salience theory. According to Angrist and Pischke (2009), this is the most credible and influential research design. Randomization eliminates selection bias, one of the most critical problems in empirical research. The design of our experiment is organized as a classic experiment where our sample is randomized into control and treatment groups. After the experiment, we will ask the participants questions related to their usage, knowledge, and attitudes related to MM.

Our experiment is a laboratory-in-the-field experiment. Gneezy and Imas (2017) formulates this type of experiment as being: "..conducted in a naturalistic environment targeting the theoretically relevant population but using a standardized, validated lab paradigm. Targeting the relevant population and setting increases the applicability of the results." Therefore, laboratory-in-the-field experiments should result in high internal validity,

[3]Tanzania Commission for Science and Technology
[4]NSD number: 267644
[5]Ethical approval number: NHH-IRB 42/22

meaning that it effectively establishes a cause-and-effect relationship between the given treatment and the observed outcome in the experiment, and other external factors are accounted for. Strong external validity, on the other hand, indicates that the results may be generalized to a broader sample (Saunders et al., 2009). However, this is not a priority in this study.

The experiment was conducted in a laboratory setting which allows for high control of the participants' surroundings. However, one drawback is that sample sizes are usually small due to practical barriers, e.g., the difficult and costly recruitment process and limited capacity in each laboratory session. Furthermore, people are aware that they are being observed in a lab, which may alter their behavior, e.g., they behave differently based on the expectations of how the experiment will inform future policy (Viceisza, 2016). Alternatively, social desirability bias could arise, where participants might answer the most socially desirable option instead of choosing the option they usually would do (Grimm, 2010). Therefore, it is difficult to know if actions are replicated in the real world.

The expected analysis approach should be outlined in a pre-analysis plan to increase the credibility of the findings and address concerns about specification searching (Duru et al., 2023). The plan should outline how the data will be analyzed and include information on primary outcomes, experimental design, randomization method, randomization unit, and sample size. Afterward, it should be registered publicly before data collection starts to a public database, e.g., asPredicted.

4.1.2 Practical implementation of the experiment

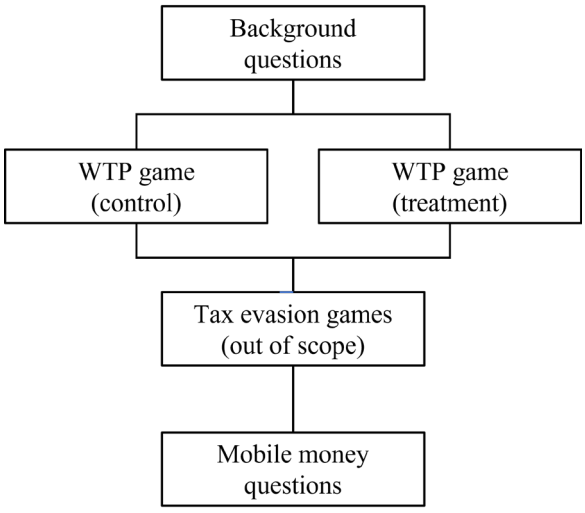
The survey started with information about the experiment and their rights, which was followed by a consent question. The participants could start the survey only if they agreed to participate. If they consented to participate in the experiment, participants were asked to write down their phone numbers on a paper slip with an ID number from 1 to 25. They had the option not to write their phone number as well. The phone numbers were used only to transfer the participation fee if they chose to receive it as MM. Due to GDPR concerns, the paper slips with phone numbers were shredded at the end of each day, following the routines outlined in the NSD notification form. The phone numbers could therefore not be connected to the individual afterward. During the day, the paper slips

were safely stored in a locked office.

The experiment had several parts, but not all are relevant to this thesis and will therefore not be elaborated on. Figure 4.1 shows a diagram of the experimental design. The first part consisted of background questions, which included questions related to age, gender, education, and occupation. In the second part, participants were asked in total seven questions about their preference for receiving the participation fee either in cash or in MM. Their answers to these questions allow us to elicit their WTP and place them in intervals. The participants were informed about how the game worked, and it was clearly communicated that they would randomly be paid one of the choices they made, highlighting the incentive-compatible features of the experiment.

The third part consisted of two different tax evasion games conducted by the other researchers involved in the project. These games are excluded from the analysis presented in this thesis. The fourth and final part included questions about participants' MM knowledge, attitudes, and usage. In this part, participants were asked about their money preferences, the frequency of using MM, their MM operator, and what they use MM for. Lastly, participants were asked to estimate the taxes and fees related to MM. When combining two different experiments, there is a chance of spillover effects, but this is not a concern as our experiment was the first one. The entire survey can be seen in its complete form in appendix A3.

Figure 4.1: Diagram of the experimental design



Each participant received a show-up fee of minimum TZS 30,000 (approx. USD 12.75 at

the time of data collection). This also included participants that showed up but did not consent to participate in the experiment. The participants could earn up to TZS 4,000 (approx. USD 1.7) extra by playing our WTP game. The exact amount they received varied depending on their choices. In total, each participant could receive a maximum of TZS 34,000 (approx. USD 18.71), which corresponds to approximately seven days worth of wage given daily minimum wage.[6] The experiment lasted for approximately one and a half hours.

Before registering the pre-analysis plan, a pilot and a focus group were conducted. However, no changes were made to the experimental design. The purpose was to test the planned design and to discover any deficiencies in the question formulations and possible sources of misinterpretations. The 18 participants who joined the pilot session also participated in the focus groups. They were recruited from a market similar to the sample participants. One main feedback we received was that many participants found the digital tablets difficult to use as they had little prior experience using them. Furthermore, they also commented that the text size should be larger. We also found that the survey needed to be shorter, as the slowest participants used almost two hours finishing it. Consequently, the text size was increased and the number of questions was reduced. We decided against including social desirability measures to reduce the length and because we concluded that this is not a major concern since MM is not a sensitive topic. Furthermore, we believe such quality controls increase the reliability and internal validity of the experiment.

Our experiment was conducted in Dar es Salaam, Tanzania. It was coded in KoboToolbox and participants answered the questions on tablets individually. KoboToolbox is a free survey software with a specific focus on development studies and takes data protection seriously (KoBoToolbox, nd). The survey was first translated into Swahili and then translated back to English again by researchers in Tanzania who have extensive experience with such translations. Afterward, the formulations and translations were thoroughly checked by the enumerators. Extra time was spent on this to ensure that questions were interpreted the way it was intended. The topic of taxation is especially tricky to translate into the Tanzanian context correctly because of the many words and meanings on the topic (Jacobsen, 2022). For example, *kodi* and *ushuru* are both words that mean duties and

[6]The minimum daily wage for Tanzanians working in the trade, industries and commercial services was approx. TZS 4,423.40 (valid in May 2023) (AfricaPay, 2023)

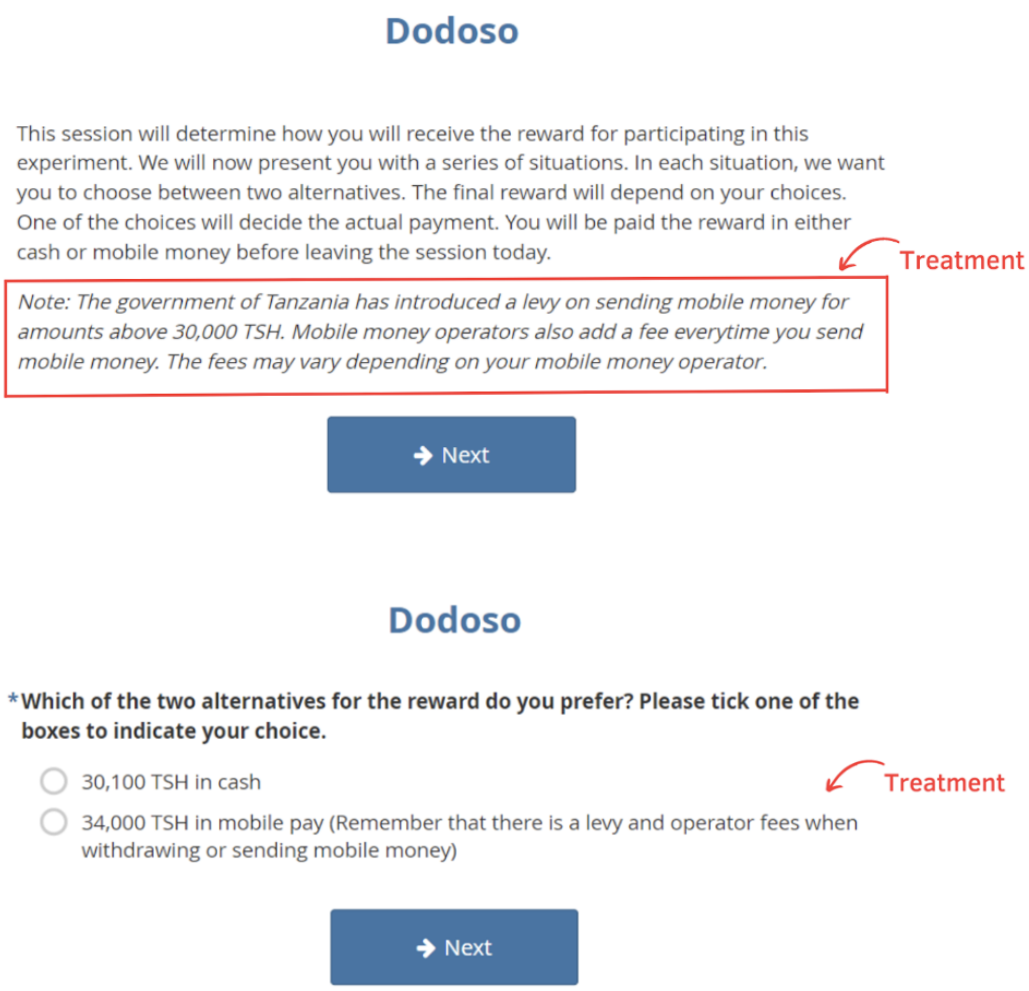
taxes, but their definition is not clear. In the treatment, the word for *tax* was therefore changed from *kodi* to *ushuru* after feedback from the focus group before starting the data collection. We have no indications of misspecification errors in the experiment due to the translation.

4.1.3 Experimental treatments

Participants were randomly assigned to either a control or a treatment group to study the salience effect of a reminder of MM taxes and fees. Recalling the salience model, the focus is on the prominence dimension, which implies that by increasing the salience of taxes and fees, participants that receive treatment will consider them in the decision-making. All participants were asked if they preferred receiving the monetary amount in cash or MM, while only the treatment group received a reminder about MM taxes and fees. In this way, the salience of the less visible attributes was increased, which is that consumers have to pay taxes and fees on MM usage. As the theory outlines, the treatment will ensure that this group recalls the non-salient attributes, which will then distort the weights related to these attributes. In the control group, participants were asked the same questions without any additional information. Figure 4.2 shows the treatment as presented in the English version of the survey (excluding the red parts).

Our treatment is considered strong, similar to the treatment in the study by Chetty et al. (2007). We did however consider several other treatment alternatives, both lower salience, e.g., just the first information note, and stronger salience, e.g., calculating the actual costs of MM usage or making the text in the information note red. Nevertheless, given the nature of a laboratory-in-the-field experiment, we saw no need for making the treatment more salient. At the same time, it seemed strong enough to capture a potential effect. Contrary to Chetty et al. (2007), who conducted a preliminary survey prior to their experiment where they asked about participants' knowledge of sales taxes, we will ask them after the WTP game, i.e., in the fourth part, to estimate the amount of MM taxes and fees they have to pay. An important feature of the salience theory is that they are aware of the taxes and fees but fail to recall that they exist. Furthermore, given the budgetary restrictions, one strong treatment effect was prioritized to receive strong results instead of dividing the sample into several treatment groups and receiving less powerful results. Power calculations will be presented in section 4.2.1.

Figure 4.2: Salience treatment as displayed in our experiment



In the recruitment process, the participants that were invited to join from each market were randomly chosen from a list of all the vendors. Then, the participants were randomized into treatment and control groups at the individual level using the built-in function in KoboToolbox, which accurately divided our sample into two equally-sized groups. Between-subject randomization increases the validity of the experiment compared to experiments where each session is given either treatment or control (Cuttler, 2023).

4.1.4 Randomization of the options in the choice sets

The options in the experiment were stacked vertically and we decided to randomize the order the options were displayed. This randomization was implemented to minimize the potential anchoring effects, which might arise by displaying similar options consecutively, resulting in always choosing the same options, which according to similar studies, proved to be quite prevalent Jack et al. (2022). This approach reduces NSB as it compels participants to stop and reflect on each choice set. Consequently, we believe this increases the validity of the results. However, randomization of choices can also lead to confusion if the participants are inattentive.

4.1.5 Constructing willingness to pay

The participants' WTP is elicited by using a MPL elicitation method, similar to Allcott and Kessler (2019). The participants are asked in total seven questions which allow for measuring their WTP. Participants first made a choice between receiving TZS 34,000 in MM or TZS 30,100 in cash. Then, the cash option continued to increase in the consecutive rounds until the cash and MM amounts were the same. Next, the MM amount was reduced while holding the cash amount constant until it reached the initial monetary value of the cash option in the first choice set. As stressed earlier, this setup was used as it offered an easy and understandable elicitation method. Table 4.1 illustrates the nature of the MPL.

Table 4.1: Overview of the MPL

Binary choices	**Mobile money**	**Cash**
Question 1	34,000	30,100
Question 2	34,000	31,600
Question 3	34,000	33,100
Question 4	34,000	34,000
Question 5	33,100	34,000
Question 6	31,600	34,000
Question 7	30,100	34,000

WTP for using MM was elicited by looking at the choices they made at each question. For example, if a participant chose TZS 33,100 in MM instead of TZS 34,000 in cash, the participant valued receiving it in MM with TZS 900 or less. If a participant chose TZS 34,000 in cash instead of TZS 31,600 in MM, the participant values receiving it in

MM with TZS 2,400 or less. A participant who answers like this, has a WTP between TZS 900 and TZS 2,400. The experiment also allowed participants to reveal negative WTP since participants might prefer cash to such an extent that they would be willing to give up a monetary amount to receive the participation fee in cash instead. Therefore, a negative WTP suggests a preference leaning towards cash, i.e., participants with a high willingness to give up money to receive it in cash. On the contrary, a positive WTP suggests a willingness towards paying to receive it in MM. When participants made all the seven choices, it allowed for placing the respondent's WTP into eight ranges, symmetric around zero ('= in TZS 1,000):

$$(-\infty, -3.9'], [-3.9', -2.4'], [-2.4', -0.9'], [-0.9', 0], [0, 0.9'], [0.9', 2.4'], [2.4', 3.9'], [3.9', \infty) \tag{4.1}$$

A MPL elicitation method creates interval values for WTP and some assumptions are needed to analyze the results. First, to simplify, we assign one unique WTP value for each interval. Next, we assign the mean of the endpoints of the WTP intervals, except for the highest and lowest interval ranges. For example, $[0, 0.9]$ have the mean WTP value of 0.45. Lastly, for the highest and lowest interval ranges, we assume a triangular distribution on the outer bounds, similarly to (Allcott and Kessler, 2019). This is because we expect most participants to end up on intervals closest to 0. For the outer interval ranges, mean WTP is calculated using a formula assuming the mass above TZS 3,900 to be distributed triangularly on the WTP range $[3.9', \infty)$, and similarly for the range $(-\infty, -3.9']$. To calculate mean WTPs, the assumption of triangular mass gives us the density for $[2.4, 3.9]$ to be 1.8%, and the mass in the range $[3.9', \infty)$ to be 23%. Similarly, we find the density on $[-3.9', -2.4']$ to be 3.2% and the mass in the range $(-\infty, -3.9']$ to be 44%.

Following Allcott and Kessler (2019) mode of calculating, we assume that this density is the maximum at TZS 3,900, and decreases to zero density above some upper bound. The upper bound is found using simple geometric calculations. First, we use the formula of the area of the triangle:

$$triangular\ area = \frac{base \times height}{2} \tag{4.2}$$

This area represents the mass in the range $[3.9', \infty)$ which we know is 23%, and then we find the upper bound:

$$23\% = \frac{(x - 3.9) \times 1.8\%}{2} \tag{4.3}$$

$$46\% = (x - 3.9) \times 1.8\%$$

$$25.55 = x - 3.9$$

$$x = 29.45$$

We find the mean by using the formula for the mean of a triangular distribution:

$$mean_{WTP} = \frac{a + b + c}{3} \tag{4.4}$$

where a is the lower bound, b is the upper bound, and c is the value of the mode, i.e., the value where the density is highest. Since we assume the highest density is at TZS 3,900, we will use this value. Therefore, we get the following:

$$mean_{WTP} = \frac{3.9 + 29.45 + 3.9}{3} \rightarrow mean_{[3.9', \infty)} = 12.4 \tag{4.5}$$

Similarly, for the lower bound, the formula of the area of the triangle is used since the area represents the mass in the range $(-\infty, -3.9']$ which we assume is 44%. We then find the upper bound:

$$44\% = \frac{(x - 3.9) \times 3.2\%}{2} \tag{4.6}$$

$$88\% = (x - 3.9) \times 3.2\%$$

$$27.5 = x - 3.9$$

$$x = 31.4$$

We find the mean by using the previously mentioned formula for the mean of a triangular distribution:

$$mean_{WTP} = \frac{3.9 + 31.4 + 3.9}{3} \rightarrow mean_{(-\infty, -3.9']} = -13.1 \tag{4.7}$$

This results in the mean on the lower bound being negative TZS 13,100, and the mean on the upper bound being TZS 12,400. Alternative ways of calculating the outer bounds are examined in robustness check 3 in section 5.

Our experiment is aimed at capturing individuals' valuation of MM, but it is important to reflect on where we would expect the participants to switch in the MPL. A profit-maximizing individual with neutral preferences, who fail to recall taxes and fees, will switch at question 4 in table 4.1. This is where the two options offer the same amount. Contrary, if the individual manages to recall that there are taxes and fees, the question in which they will shift presumably depends on their preference for either of the options and what they intend to use MM for. Table 4.2 shows a detailed overview of the total taxes and fees for different types of actions. Assuming an individual recalls the taxes and fees, with the purpose of sending them to friends or family, the individual will switch at choice 2 or 3, in table 4.1, to maximize the monetary amount. The reason for this is that the individual will incorporate the total amount of taxes and fees, which for sending inside a network is TZS 751, and outside is TZS 963, into their decision-making and they will maximize the monetary amounts by shifting at this point.

Table 4.2: Overview of the total amount of MM tax and fees for different use

Mobile money usage	**Tax(TZS)**	**Fees(TZS)**	**Total(TZS)**
Send to someone (within network)	351	400	751
Send to someone (outside of network)	351	612	963
Transfer to traditional bank account	351	2400	2751
Withdraw the money	351	1850	2201

Does the experiment correctly measure WTP?

Given the nature of this experiment, the effect is implicit since we, the researchers, are paying the taxes and fees when sending the participation fee. However, for individuals to either send or withdraw the money later, they would need to pay taxes and fees. Therefore, the treatment is still expected to measure WTP correctly since we remind them that there are taxes and fees involved, which they will eventually pay when using MM later.

4.2 Sample

In an effort to increase the validity of the experiment, we decided against using students in the sample, following Sears (1986) widely cited paper arguing that students provide a "narrow database", meaning: "College students are likely to have less-crystallized attitudes, less-formulated senses of self, stronger cognitive skills, stronger tendencies to comply with authority, and more unstable peer group relationships. The laboratory setting is likely to exaggerate all these differences". Instead, participants were recruited from markets located in different parts of Dar es Salaam. The city is the economic and administrative center of Tanzania (Peter and Yang, 2019), and the markets serve as the primary place for trading goods and services.

The recruitment was done by reaching out to the market leaders and each leader provided a list of all the vendors in their market. The market leaders were paid a reward of TZS 30,000 for this work. From each list, 50 participants were randomly selected to take part in the survey. There was a minimum age requirement of 18 years old to participate. It proved easy to have participants partaking in the experiment, which could be due to the high participation fee and the enumerators' close cooperation with the market leader. Some adjustments to the election process were made after the first day to exclude participants with bad eyesight or who could not read. The enumerators enforced these requirements by questioning the participants before boarding the bus. We conducted four sessions daily with 25 participants in each. A maximum of two groups were gathered from the same market. The groups arrived by bus and no information was exchanged since the second group was picked up from the market while the first group was completing the survey. In total, 32 sessions were conducted over eight days, which resulted in 796 consenting participants partaking in the experiment.

Prior to excluding outliers according to our registered pre-analysis plan, we examined the descriptive statistics of the entire sample. We placed these tables in the appendix since this is not the sample we use for analysis. Table A1.1 shows the descriptive statistics for the full sample, and we want to particularly address that the gender balance in the sample was significantly different ($p=0.086$) between the control and the treatment group. Furthermore, in table A1.2 in the appendix, we use a sample with only MSB participants and create a binary variable for MSB and regress the outcome variable on background

variables and salience treatment where we find that gender is significant at the 1 percent level. This indicates that women are more inclined to be MSB. However, after removing the MSB participants, the sample actually gets more balanced on the gender dimension. We will discuss this further in section 7.

Following our pre-analysis plan, 159 observations were removed due to multiple switching behavior (MSB). As mentioned earlier, a multiple-switcher is a participant that switches between MM and cash more than once during the experiment. This also includes the ones who switched from cash to MM since this is not logical given the nature of the MPL and thus indicates internally inconsistent choices. Descriptive statistics of our final sample are reported in table 4.3.

Table 4.3: Descriptive statistics of the sample

	Mean			**P-value of t-test**
	Control	Treatment	Total	Control vs. Treatment
Female	0.41	0.36	0.38	0.159
	(0.03)	(0.03)	(0.02)	
Above Median Age	0.45	0.54	0.49	0.036**
	(0.03)	(0.03)	(0.02)	
High School	0.26	0.29	0.28	0.430
	(0.02)	(0.03)	(0.02)	
Observations	322	315	637	

Mean (standard error) and p-value of t-test. $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Notes: The table reports descriptive statistics of the sample using the indicator variables used for the analysis. Column (1) reports means for control group, Column (2) reports means for the treatment group, Column (3) reports the means for the entire sample, and Column (4) reports p-values for a two-sided t-test for difference in means between control and treatment groups.

Table 4.3 shows that the share of participants above median age is significantly different (p=0.036) between treatment and control groups. It was challenging to find credible statistics for a sample comparison to citizens in Dar es Salaam. We will therefore compare the background characteristics to country-wide estimates in this section to provide an understanding of the Tanzanian context, although there are differences between urban and rural areas. The median age in the dataset is 42 years old. This is considerably higher than the country-wide median age of 17.8 years old, according to the last national census (UNDP, 2017). However, this is expected since the age requirement to participate was 18 years old. Another reason could be that becoming a part of an organized market requires

a very specific network and an understanding of the context, which is acquired over time (Kirumirah and Munishi, 2021).

38% of the participants in the dataset are female. Specifically, in the control group there are 41% female and 36% in the treatment group, as seen from table 4.3. There is no significant difference between the gender distribution in treatment and control groups. In Tanzania, the female proportion is estimated to be around 48.7% (Afrobarometer, 2021). Although there is limited research available on this matter, one plausible explanation could be that female vendors may have small children and therefore remain at the marketplace to manage the shop and attend childcare responsibilities, while the male counterpart participate in the experiment. There might also be a cultural explanation, e.g., that it is more common for men to own a shop (Idris, 2018).

An indicator for participants having completed high school or higher is included, as outlined in the registered pre-analysis plan. As seen in table 4.3, 28% of the entire sample has a degree from high school or higher. According to the Afrobarometer (2021), there are 24.3% who have high school education or higher in Tanzania. Furthermore, there is a 3% difference between the control and the treatment group, which is not significant.

All participants work in the marketplace and own their shops. Therefore, the occupation question in the survey indicates what type of products the participant is selling. As seen in table A1.3 in the appendix, the most chosen category was vegetables (25.9%). The "other" category was also frequently chosen, with specifications like fruits, small café (Swahili: mama lishe), and meat. This question was asked for exploratory reasons and indicates that our sample consists of a mixed group of people.

In the last part of our experiment, we asked which mobile operator the participants used. Table A1.4 in the appendix, shows that the sample uses Tigo with Tigo Pesa and Airtel Money the most. Worth noting is that many use several different mobile operators. A reason for this is that the operators have different promotion schemes to attract people to their platform, making it lucrative to create new MM accounts (Materu, 2023).

4.2.1 Power calculation

Prior to data collection, power calculations were conducted and included in our pre-analysis plan. Power calculations are a statistical tool to help compute sample size, statistical power, and minimum detectable effect size. According to the World Bank (2023), power calculations are an important process when conducting an experiment and should be calculated in the design stage. Statistical power is the probability that one or another statistical criterion can correctly reject the null hypothesis, in the case when the alternative hypothesis is true. The higher the power of the statistical test, the less likely it is to make a type II error and not reject the null hypothesis that is actually false (Dupont and Plummer, 1990).

For this experiment, power calculations were conducted to determine the required sample size for a two-sample means test. In line with common practice, a power of 0.8 was used. The power represents the probability of rejecting the null hypothesis when a specific alternative hypothesis is true, indicating a 20% chance of mistakenly failing to reject the null hypothesis when using a power of 0.8. The significance level, typically set at 5%, was also employed in this study (Jameel, 2023). The hypothesis for this study will be presented in section 5.

While exploring similar studies using MPL and salience treatments, we found limited information on power calculations. Therefore, we used generalized assumptions commonly employed in laboratory experiments. The minimum detectable effect, which is the smallest effect that, if true, has an 80% chance of producing an estimate that is statistically significant at the 5 percent level, was estimated to be 0.2 with a standard deviation of 1 (Bloom, 1995). Using a one-sided test, a total of 620 participants, 310 in each group, was estimated for a power of 0.8. The possibility of having several treatment groups was looked into but decided against after conducting these power calculations.

4.3 Context of the experiment

The experiment was organized by Chr. Michelsen Institute (CMI) and funded by the Research Council of Norway. The researchers associated with this particular experiment, alongside ourselves, were Senior Researcher Ingrid Hoem Sjursen at CMI and Ph.D. student Jasmin Vietz at the University of Hohenheim. We are master's students at NHH writing our MSc thesis at CMI. One of us also works as a research assistant for the Tax and Public Finance group at CMI. Our role in the project was to prepare research permits and ethical approvals, design our parts of the experiment, and plan and execute the experiment in Tanzania. We, the authors of this thesis, bore the full responsibility of designing and executing our part of the experiment. During the data collection, the authors ensured that the experiment ran smoothly, prepared participation payments, and ensured that the data was registered.

The laboratory experiment was conducted at REPOA's facilities in Dar es Salaam. REPOA is a highly influential think tank ranked number 4 out of 679 listed think tanks in Sub-Saharan Africa (University of Pennsylvania, 2021). They are in charge of the Afrobarometer data collection in Tanzania and have extensive experience with data collection. They provided experienced enumerators to recruit and guide participants through the digital survey. Extra care was made to secure that participants were met with enumerators who spoke Swahili to ensure the effective execution of the lab sessions. The enumerators handled all dialogue and assistance with the participants and we were not visible to the participants. This was particularly important for the other part of the experiment due to the sensitivity of some of the questions.

5 Empirical Strategy

The empirical section of this thesis is built around the experiment in Tanzania and is in accordance with the pre-analysis plan.[7] As mentioned, the pre-analysis plan was registered prior to data collection for transparency and to ensure that *p-hacking* the analysis will not be possible, meaning that one makes decisions regarding analyses that are too specific for a study (P Simmons et al., 2021). This chapter will present the estimation method, heterogeneity analysis, and the robustness checks included in the empirical analysis.

5.1 The OLS model

An ordinary least squares (OLS) estimation is used to analyze the treatment effect on WTP. OLS is a simple linear regression model that estimates the relationship between two variables by minimizing the sum of squared residuals (Wooldridge, 2020). Several assumptions must be fulfilled to have an efficient and unbiased estimator, which we argue for having in Appendix A2. We estimate using the following empirical specification:

$$Y_i = \alpha + \delta Salience_i + \varepsilon_i \tag{5.1}$$

Y_i is the dependent variable and we will analyze it using two different specifications. The first specification is a binary dependent variable for each of the questions in the MPL which takes on the value of one if choosing cash, and zero if choosing MM.[8] We do this to investigate if the behavior is affected by the salience treatment regardless of how WTP is specified. The model is fit using a continuous dependent variable since the regression analysis yields more intuitive coefficient estimates while also being more flexible (Hellevik, 2009). Alternatively, a Probit regression model could be used, which is a non-linear regression model commonly used to model binary outcome variables. It estimates the probability of an event occurring. However, OLS models have been found to yield results of similar significance (Angrist and Pischke, 2009). Therefore, we analyze the dependent variable as continuous since it is easier to interpret. A robustness check will be conducted

[7] https://aspredicted.org/ws6hz.pdf

[8] This specification is not specified in our pre-analysis plan.

using a Probit regression model instead.

In the second specification, Y_i is a continuous dependent variable taking on one of the mean values of the WTP intervals previously specified. $Salience_i$ is an indicator variable that takes on the value of one if receiving salience treatment, and zero otherwise. This is the variable of main interest. ε_i is the error term.

Next, we conduct a multiple linear regression model using the second specification where we include several explanatory variables to analyze the WTP. This is the main model specification that was registered in our pre-analysis plan. We use the following specification:

$$Y_i = \alpha + \delta Salience_i + \beta X_i + \beta M_i + \varepsilon_i \tag{5.2}$$

Here, Y_i is the continuous dependent variable taking on the mean value of one of the WTP intervals. $Salience_i$, is the treatment variable. X_i is the background vector consisting of demographic characteristics. It includes an indicator variable for above median age taking on the value of one if over 42 years old and zero otherwise, an indicator variable for gender which takes on the value of one if female and zero if male, and an indicator variable for high school education which takes on the value of one if finished high school and zero otherwise. M_i is an indicator variable taking on the value of one if the participant is a high frequency user of MM (i.e., those choosing "Twice a week or more" and "Daily"), and zero otherwise. ε_i is the error term.

The coefficient δ is of particular interest as this is the treatment, and the main coefficient of interest to identify the effect of salience on WTP. In the analysis, the regression with no other variables except for $Salience_i$ will be estimated first. Then, controls will be added sequentially. All regressions are estimated with robust standard errors due to potential , as discussed in appendix A2. The results will be discussed based on the full specification.

The main hypothesis to be tested, by estimating equation (5.2), is formulated as follows:

Main hypothesis: Receiving the salience treatment reduces participants' willingness to pay for mobile money.

$$H_0 : \delta = 0 \ , \ H_A : \delta < 0 \tag{5.3}$$

5.2 Heterogeneity analysis

Heterogeneity in the treatment effect is investigated to shed light on potential underlying mechanisms affecting the dependent variable (Hu, 2023). This provides a more detailed view of treatment effects across the sample and may also reveal interesting sub-group characteristics, i.e., whether different sub-groups of the sample respond differently to the treatment (Angrist, 2004).

Heterogeneity analysis 1: Heterogeneity in background characteristics and those who switch once will be examined using interaction terms

A heterogeneity analysis will be conducted using interaction terms between treatment and the background indicator variables for *Female*, *Above Median Age*, and *High School*. Based on these regressions, the aim is to understand whether significant differences exist between sub-groups in the sample. Separate regressions for background indicator variables are created by using the following specification:

$$Y_i = \alpha + \delta Salience_i + \beta X_i + \theta Salience_i \times Var_i + \beta M_i + \varepsilon_i \tag{5.4}$$

In equation (.1), Y_i is the continuous dependent variable taking on the mean value of one of the WTP intervals. $Salience_i$ is the salience treatment, Var_i is an indicator variable for *Female*, *Above Median Age*, or *High School*. $Salience_i \times Var_i$ is an interaction term between the salience treatment and background indicator variables. θ is the estimated sub-group difference when receiving salience treatment. Conducting these regressions allows for studying whether there are significant differences in treatment effects between sub-groups in the experiment. ε_i is the error term.

Lastly, we will run a regression where we create an interaction term between treatment and an indicator variable *Switcher*, which are those individuals who switch at most one time in the MPL. We include an indicator variable for those individuals who switch to examine the part of the sample that falls in the WTP intervals between the outer bounds of the distribution. The rationale is to gain insights into how the salience treatment affects this particular sub-sample, as opposed to the part of the sample that exhibits a strong preference for MM or cash. Background characteristics is also included, and we use the following specification:

$$Y_i = \alpha + \delta Salience_i + \beta X_i + \beta M_i + \beta Switcher_i + \theta Salience_i \times Switcher_i + \varepsilon_i \quad (5.5)$$

In equation (5.5), Y_i is the continuous dependent variable taking on the mean value of one of the WTP intervals. $Salience_i$ is the salience treatment, X_i is a vector for background characteristics and M_i is the indicator for high usage of MM. $Switcher_i$ is an indicator variable taking on the value of one if switching at most one time and zero if the participant does not switch, and $Salience_i \times Switcher_i$ is the interaction term between this variable and the salience treatment variable. θ is the estimated sub-group difference of interest and ε_i is the error term.

5.3 Robustness checks

We do robustness checks to ensure structural validity of our findings, which is done by testing the sensitivity of the results to variations and assumptions in the model specification and estimation method (Lu and White, 2014). This section will explain the robustness checks conducted in this study.

Robustness check 1: Probit regression on each of the MPL questions

The MPL is designed as a set of binary questions, and it is therefore sensible to re-fit it to a Probit regression model (Noreen, 1988). Accordingly, we re-run equation (5.1) using this model specification. This is done to examine if salience treatment is statistically significant and has a consistent effect across all the questions.

Robustness check 2: Probit regression where WTP is a binary variable

Another regression analysis will be conducted modifying the WTP intervals into a binary variable. Specifically, the binary variable takes on a value of one for all the participants that exhibit a preference towards cash, meaning they fall on the negative side of the WTP distribution, and the value of zero if on the positive side, see equation (4.1). The latter indicates a preference towards MM, which include all the participants on the positive side of the WTP distribution. This modification is made because a substantial part of the sample falls on the outer bounds of the WTP distribution. Therefore, it would be interesting to investigate if we find a significant treatment effect when changing the dependent variable into a binary WTP variable indicating a preference for either cash or MM.

Robustness check 3: Alternative assumption of mean WTP

We assumed a triangular distribution on the mass above TZS 3,900 and below negative TZS 3,900 to estimate the mean WTP on the outer bounds, given the expectation that most people fall in the WTP intervals around zero. This led us to a mean WTP on the lower outer bound to be negative TZS 13,100, and the mean WTP on the upper outer bound to be positive TZS 12,400. However, this assumption has proved not to hold in practice, encouraging us to explore alternative methods for calculating the mean WTP values on the outer bounds. One approach is to use heuristic benchmarks like Allcott and Kessler (2019), which rely on the researchers' interpretation of plausible mean WTP values. It is plausible that the true WTP is somewhere between +/- TZS 3,900 and the estimated mean WTP values using a triangular distribution. Hence, regression models will be run assuming the mean WTP on the outer bounds to be half and one-third of the calculated mean WTP when using a triangular distribution.[9][10] We do not expect the WTP to be bigger than the calculated WTP using a triangular distribution assumption on the notion that the minimum wage is TZS 4,423.40 and, as our sample consists of vendors at informal markets, we believe their salary to be similar. Therefore, it seems unlikely that they are willing to pay that much in taxes and fees to use MM.

[9]Mean WTP (1/2): $(-\infty, -3.9'] = -7.8',\ [\ 3.9', \infty) = 7.5'$
[10]Mean WTP (1/3): $(-\infty, -3.9'] = -6.1',\ [\ 3.9', \infty) = 5.9'$

Additionally, we will use a Wald test to check if the salience treatment estimates in the model is significantly different from the actual total taxes and fees, i.e., if the effect size is statistically different from $\delta = -.963$ (Stata, 2023). This is of course an oversimplification of the model, and will only be used to look into the magnitude of the effect. The assumption entails that the taxes and fees are not considered in the control group and are perfectly calculated and incorporated in the treatment group. I.e., the prominence of the taxes and fees is non-existent in the control group and fully visible in the treatment group. We will test all three model specifications.

Robustness check 4: Testing different model specifications

Lastly, a check of the sensitivity to alternative model specifications will be conducted. According to Lu and White (2014), the *rcheck* Stata command can be used to check all possible model specifications. In this case, we ran the statistical model with 16 alternative specifications of WTP, using the main model specification given by equation (5.2). The idea is a "test-test-test" approach and the aim is to check if the results are robust to reasonable alternative specifications and not just one or a few versions of the model.

6 Empirical Analysis

In this section, we will first review the main results, which include the results from an OLS regression per question in the MPL, descriptive results of WTP, and main analysis results. Next, the results from other measures will be explored, which includes all other findings from the last part of our experiment related to MM usage, knowledge, and attitudes. Lastly, the results from the heterogeneity analysis and robustness checks will be examined.

6.1 Results: Salience effect on WTP

6.1.1 OLS regression per question in MPL

Table 6.1 uses the model given by equation (5.1) and shows an overview of all OLS regressions per question in the MPL, where each column number indicates the same order each question was asked. The results indicate a significant treatment effect at the 1 percent level and the effect is consistent across all questions in the MPL. This indicates that when receiving treatment, these participants are more inclined to choose the cash option in the choice set. Alternatively, we could have found various significant levels which then would indicate an inconsistent treatment effect.

Table 6.1: OLS regressions on all the MPL questions

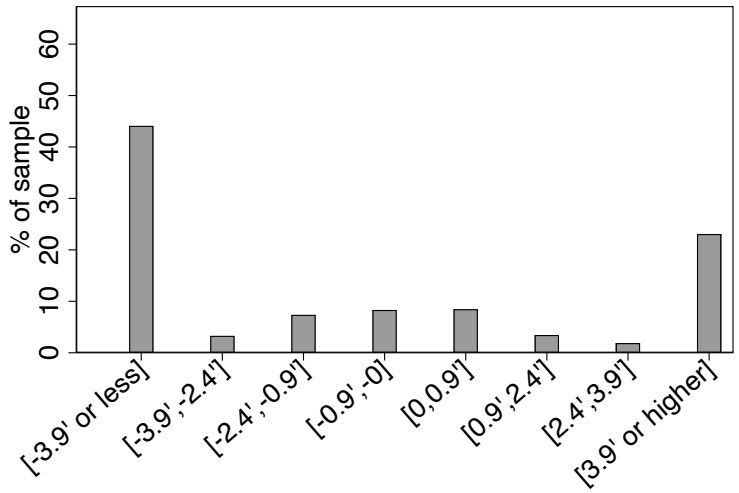
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Salience Treatment	0.104***	0.111***	0.113***	0.133***	0.107***	0.114***	0.114***
	(0.039)	(0.039)	(0.039)	(0.038)	(0.036)	(0.034)	(0.034)
Constant	0.404***	0.425***	0.497***	0.565***	0.655***	0.683***	0.699***
	(0.027)	(0.028)	(0.028)	(0.028)	(0.027)	(0.026)	(0.026)
Observations	637	637	637	637	637	637	637
R^2	0.011	0.012	0.013	0.019	0.014	0.017	0.018

Robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Notes: The outcome variable is a binary variable for the specific choice set taking on the value of one if choosing cash, and zero if choosing MM. *Salience Treatment* is an indicator variable taking the value of one if receiving treatment, and zero otherwise. Columns (1-7) represent the regressions for each of the questions in the same order as they were asked in the MPL.

6.1.2 Descriptive findings on WTP

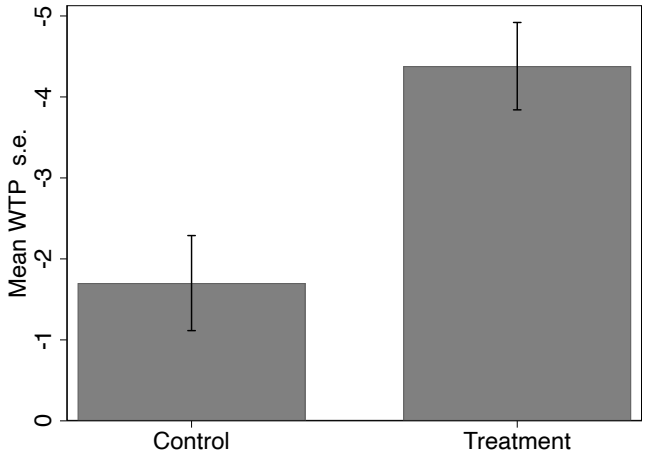
Figure 6.1 summarizes the participants' WTP for the entire sample. Most participants end up on the outer bounds, but many also end up in the interior WTP intervals. This indicates that many participants consistently chose cash or MM throughout the MPL.

Figure 6.1: WTP interval distribution for the sample ('=TZS 1,000)



The treatment effect can also be visualized by showing the mean WTP with standard error bars by treatment, as shown in figure 6.2. The results indicate that the mean WTP for the control and treatment groups are significantly different, where mean WTP for the treatment group is more negative, which indicates a preference for cash.

Figure 6.2: Mean WTP with standard errors by treatment

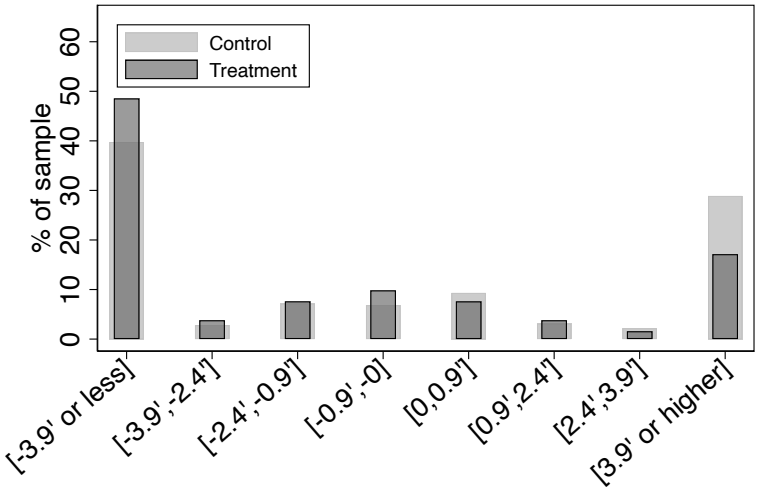


Notes: The figure shows the mean WTP values for the control and treatment groups. The mean WTP is in TZS 1,000. Y-axis is reversed for readability. The figure also indicates the estimated standard errors.

6.1.3 Main analysis: Salience effect on WTP

Figure 6.3 summarizes the participants' WTP per control and treatment groups. It shows that the distribution in the treatment group is more left-skewed, indicating that more participants have a negative WTP and a preference towards cash, i.e., they are willing to pay not to receive the reward in MM. Conversely, the distribution in the control group is more right-skewed, indicating a WTP for receiving the reward in MM.

Figure 6.3: WTP distribution by control and treatment groups ('=TZS 1,000)



Main result: The salience treatment has a strong, significant effect ($p < 0.01$) on participants' WTP for MM

In table 6.2, column (1) shows results estimated by equation (5.1), which is the simple OLS model where we regress the continuous variable of WTP on salience treatment ($p<0.01$). Column (5) shows the results estimated by equation (5.2), which shows the pre-registered empirical specification. In this column, we see that the salience treatment effect is estimated to be -2.763. This can be interpreted as WTP for using MM reduces with TZS 2,763 when receiving treatment. This amount is larger than the taxes and fees when transferring the participation fee [11], and the amount equals a bit over half a daily minimum wage.[12] This result is significant at the 1 percent level. Therefore, we reject $H_0 : \delta = 0$ that the salience treatment does not reduce the WTP for using MM.

[11] Approx. TZS 963 if sending to someone outside the mobile network, according to table 4.2

[12] Daily minimum wage citizens working in trade, industries, and commercial services was TZS 4,423.40 (AfricaPay, 2023)

Furthermore, in column (5), the variable *High Usage of MM* is significant at the 1 percent level, indicating that participants who use MM twice a week or more, generally have a TZS 2,815 higher WTP. Lastly, the high school indicator variable is significant at the 5 percent level, indicating a higher WTP for MM of TZS 2,815. The background characteristics variables will be investigated further in the heterogeneity analysis in section 6.3.

Table 6.2: Main regression analysis

	(1)	(2)	(3)	(4)	(5)
Salience Treatment	-2.679***	-2.670***	-2.688***	-2.814***	-2.763***
	(0.797)	(0.796)	(0.800)	(0.799)	(0.792)
Female		0.156	0.185	0.212	0.481
		(0.829)	(0.839)	(0.837)	(0.834)
Above Median Age			0.234	0.909	0.701
			(0.812)	(0.854)	(0.848)
High School				2.535***	2.262**
				(0.924)	(0.918)
High Usage of MM					2.815***
					(0.801)
Constant	-1.701***	-1.765***	-1.883**	-2.869***	-4.300***
	(0.587)	(0.657)	(0.768)	(0.850)	(0.911)
Observations	637	637	637	637	637
R^2	0.017	0.017	0.018	0.029	0.048

Robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Notes: The outcome variable is WTP taking on the mean value of the WTP intervals. Column (1) shows the simple specification with only WTP on *Salience Treatment*, where *Salience Treatment* is an indicator variable taking on the value of one if receiving treatment and zero otherwise. Column (2) shows the estimates including a gender indicator variable *Female*, which takes on the value of one if female, and zero if male. Column (3) shows the estimates including an indicator variable *Above Median Age*, which takes on the value of one if above 42 years old, and the value of zero if below. Column (4) shows the estimates containing the indicator variable *High School*, which takes on the value of one if obtained high school diploma, and zero otherwise. Lastly, Column (5) shows the main model specification where also the indicator variable *High Usage of MM* is included, which takes on the value of one if using MM twice a week or more, and zero otherwise.

6.2 Results: Other measures

In the fourth and last part of our experiment, we asked the participants about their MM usage, attitudes, and knowledge for exploratory reasons. In this section, we will look at the results from this part. Since the questions were asked after receiving salience treatment and the information treatment from the other experiment, we have checked for spillover effects. The tests show no indications that these treatments have affected the answers given in this part.

Result 1: The salience assumption is validated since participants in both groups come close to estimating the exact amount of taxes and fees

In the final part of our experiment, we asked the participants to estimate how much they have to pay in taxes and fees. As previously mentioned, a key feature of tax salience is that people are aware of the taxes and fees, but they do not incorporate this information in the decision-making since the information is not salient. Therefore, we asked them how much in total they would need to pay in taxes and fees when sending TZS 40,000. Afterward, we asked them to estimate the amount they pay in taxes, which ultimately showed them how much of the total goes to pay the fees.

From table 6.3, we found that the mean estimation of taxes and fees was TZS 912.4, and the participants generally estimated 36% to be the tax cost and 64% to be fees. This is similar to the actual proportions of 38.3% being taxes and 61.7% being fees, which is the average amount if transactions are made across different mobile operators. However, the actual total amount of taxes and fees is TZS 963, which is higher than the mean guessed amount of TZS 912.4. This indicates that our sample believes that taxes are lower and that fees are proportionally higher, than in reality. However, it is rather similar to actual taxes and fees, which validates that they indeed have knowledge about the amount of taxes and fees.

Table 6.3: Estimation of taxes and fees

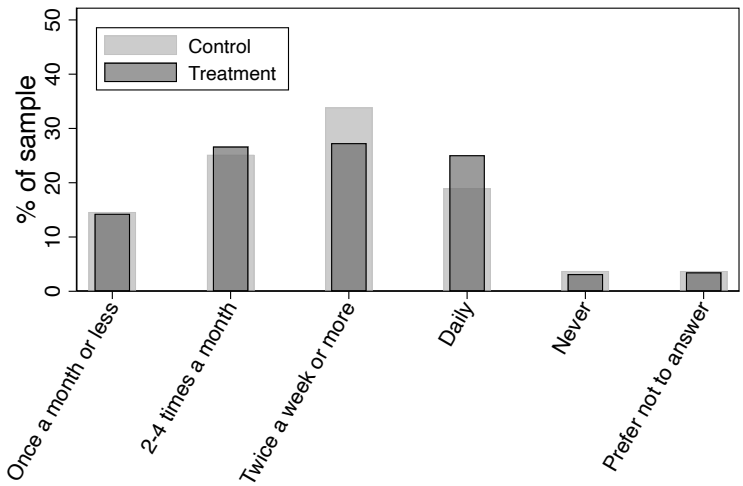
Control	**Mean**
MM levy	336.68
Fees	571.39
Total MM levy and fees	908.07
Treatment	
MM levy	312.02
Fees	604.81
Total MM levy and fees	916.83
Total	
MM levy	324.48
Fees	587.92
Total MM levy and fees	912.40

Notes: The question they were asked was the following: "If you send TZS 40,000 mobile money, how much would you need to pay in total in taxes and fees when sending money (in TZS)? If you do not know, please make a guess."

Result 2: Mobile money services are widely used

Looking further into people's use of MM, we asked them how often they used mobile payments. In figure 6.4, we see that 52.5% of our sample uses mobile payments twice a week or more often. 3.5% answered that they had never used it.

Figure 6.4: Distribution of how often MM is used



Notes: The question they were asked was the following: "How often do you use mobile pay?"

Result 3: Most participants tend to use MM services for transferring to family members, and many also use it for government payments

We asked the participants what they use MM services for. As seen in table 6.4, 58% answered that they use it to transfer money to family members and 33.4% use it to buy goods and services. Interestingly, the third most chosen option was government payments which approximately 21% of the sample answered.

Table 6.4: Overview of what the participants use MM services for

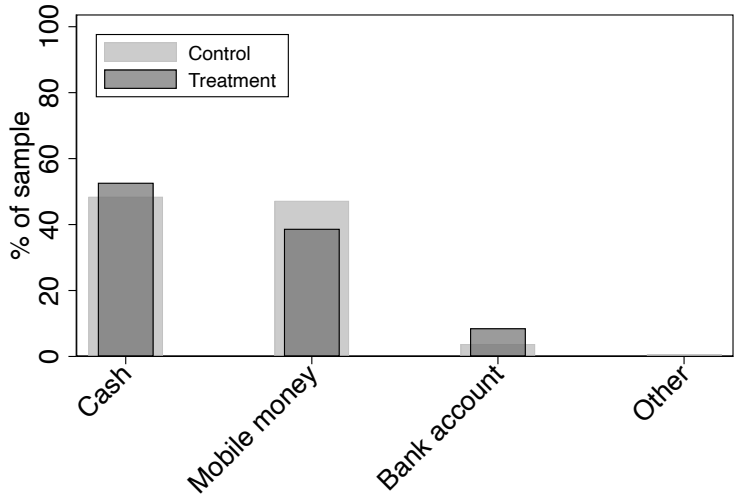
MM usage areas	
Transferring to family members	58.3%
Transferring to friends	11.1%
Government payments	20.6%
Buying goods and services	33.4%
Selling goods and services	16.6%
Other	3.7%
Prefer not to say	2.9%

Notes: The question they were asked was the following: "What do you use mobile pay for? Please select several options if that is correct for you."

Result 4: Cash is the preferred option when asked directly about it

To gain insights about our sample, we asked everyone their preferred method of receiving money. It was only allowed to choose one of the options since we wanted to explore their preferred option. Figure 6.5 shows an overview of participants' preference for either cash, MM, or receiving on a traditional bank account. For the entire sample, we find that 50% prefer cash, while 43% prefer mobile money, and the remaining 7% prefer receiving it in their bank account. This question was asked in the last part, and we can therefore not be certain about their true preference since they might answer based on their choices in the MPL and thus be affected by the information treatment in the part of the experiment that is out of our scope. From figure 6.5, we see that more people in the treatment group choose cash, which aligns with the results in figure 6.3. Furthermore, as mentioned, we should approach these findings with caution since their response to previous questions might influence the answer to this question.

Figure 6.5: Distribution of preferred method of receiving and sending money



Notes: The question they were asked was the following: "How do you generally prefer to send and receive money?"

6.3 Results: Heterogeneity analysis

Heterogeneity analysis 1: No strong indication that sub-samples related to the background characteristics behave differently to the salience treatment (except for high school education ($p<0.1$)), but participants who switch once react significantly different to the salience treatment

We ran the regressions with interaction terms between the salience treatment and each of the background indicator variables. Additionally, we created an interaction term between the treatment and the indicator variable for participants who switch once. In columns (2) and (3), we see that $Salience\ Treatment \times Female$ and $Salience\ Treatment \times Above\ Median\ Age$ are not significant. This indicates no sub-group differences, which is favorable given that table 4.3 shows that our sample is not balanced when it comes to *Above Median Age* ($p<0.05$). In column (4), we find that the interaction term $Salience\ Treatment \times High\ School$ is significant at the 10 percent level. This indicates that the participants with high school education are considerably less willing to pay for MM when given salience treatment. Furthermore, we find that the sub-group $Salience\ Treatment \times Switcher$ is significant at the 1 percent level. This indicates that when this sub-sample receives salience treatment, their WTP increases, ultimately having a WTP of negative TZS 439.

Table 6.5: Heterogeneity analysis

	(1)	(2)	(3)	(4)	(5)
Salience Treatment	-2.763***	-2.658***	-3.393***	-1.920**	-4.057***
	(0.792)	(0.997)	(1.093)	(0.939)	(1.139)
Female	0.481	0.611	0.504	0.437	0.544
	(0.834)	(1.202)	(0.834)	(0.831)	(0.824)
Above Median Age	0.701	0.694	0.075	0.690	0.722
	(0.848)	(0.850)	(1.191)	(0.848)	(0.835)
High School	2.262**	2.254**	2.269**	3.803***	1.932**
	(0.918)	(0.919)	(0.918)	(1.316)	(0.913)
High Usage of MM	2.815***	2.824***	2.818***	2.893***	2.827***
	(0.801)	(0.802)	(0.800)	(0.800)	(0.783)
Salience Treatment × Female		-0.271			
		(1.634)			
Salience Treatment × Above Median Age			1.273		
			(1.584)		
Salience Treatment × High School				-3.037*	
				(1.730)	
Switcher					2.210**
					(0.869)
Salience Treatment × Switcher					3.618***
					(1.192)
Constant	-4.300***	-4.352***	-4.029***	-4.725***	-4.947***
	(0.911)	(0.961)	(0.981)	(0.943)	(1.107)
Observations	637	637	637	637	637
R^2	0.048	0.048	0.049	0.052	0.090

Robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Notes: The table shows all the regressions in the heterogeneity analysis. The outcome variable is WTP. Column (1) shows the main regression model. Column (2) shows the regression where *Salience Treatment* × *Female* is the interaction term. Column (3) shows the regression where *Salience Treatment* × *Above Median Age* is the interaction term. Column (4) shows the regression where *Salience Treatment*×*High School* is the interaction term. Lastly, column (5) shows the regression where the indicator variable takes on the value of one if switching in the MPL once and zero if no switches occur. *Salience Treatment* × *Switcher* is the interaction term.

6.4 Results: Robustness checks

Robustness check 1: Probit regression on each question in the MPL show significant estimates (p<0.01)

Table 6.6 shows an overview of all the Probit regressions, where the column number indicates the choice set. We see that the treatment coefficient is significant at the 1 percent level for all the questions, indicating that the treatment has a consistent effect on all questions in our experiment. From table 6.6, we see that the constant is negative for columns (1-3) since more people choose MM at the beginning of the game, likely because the money offered in MM is higher than in cash, before switching and choosing cash, which indicates a negative WTP. In column (3), the constant is not significant and one

plausible reason could be that equal amounts of participants choose cash and MM, since this is where many shifts from MM to cash in the MPL. Nevertheless, this check validates our estimates.

Table 6.6: Probit regression on all the MPL questions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Salience Treatment	0.264***	0.280***	0.286***	0.356***	0.313***	0.354***	0.367***
	(0.100)	(0.100)	(0.100)	(0.102)	(0.106)	(0.108)	(0.110)
Constant	-0.244***	-0.188***	-0.008	0.164**	0.400***	0.477***	0.521***
	(0.071)	(0.070)	(0.070)	(0.070)	(0.072)	(0.073)	(0.073)
Observations	637	637	637	637	637	637	637

Robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Notes: The table reports Probit regressions where the outcome variable is a binary variable taking on the value of one if choosing cash and zero if choosing MM. Columns (1-7) shows the regression of outcome variable on *Salience Treatment*.

Robustness check 2: Probit regression where WTP is a binary variable validates main regression results

As mentioned in the empirical strategy, we conduct a regression analysis modifying the WTP intervals into a binary variable taking on the value of one if negative WTP and zero if positive WTP. As seen in table 6.7, the salience treatment is significant at the 1 percent level and thus similar to the main results in table 6.2. This indicates that the results are robust to an alternative specification.

Table 6.7: Probit regression with binary outcome variable for WTP

	(1)
Salience Treatment	0.355***
	(0.104)
Female	-0.085
	(0.107)
Above Median Age	0.059
	(0.109)
High Usage of MM	-0.296***
	(0.105)
High School	-0.214*
	(0.118)
Constant	0.388***
	(0.122)
Observations	637

Robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table shows a Probit regression where the outcome variable is WTP taking on the value of one if negative WTP and zero if positive WTP interval. Column (1) shows the regression using the main model specification with a binary outcome variable of WTP.

Robustness check 3: Estimates are sensitive to changes in alternative assumptions about mean WTP on the outer bounds

Table 6.8 shows an overview of the regression results using alternative assumptions for mean WTP on the outer bounds. From the table, we see that *Salience Treatment* coefficients change in value when changing the mean WTP on the outer bounds. Since the triangular distribution assumption likely does not hold in reality, alternative heuristic calculations of half and one-third of the triangular distribution were used.[13][14] From table 6.8, all *Salience Treatment* coefficients are significant but they decrease in value when using alternative assumptions. Column (2) is the heuristic benchmark assuming half the mean value calculated using a triangular distribution. From this column, we can see that WTP decreases with TZS 1,681 when receiving treatment. Column (3) is the heuristic benchmark assuming one-third of the initial mean values, and when receiving salience treatment, WTP decreases with TZS 1,331. Since most participants fall into these intervals, the estimated mean WTP is highly sensitive to the mean on the outer bounds.

[13]Mean WTP (1/2): $(-\infty, -3.9'] = -7.8,\ [\ 3.9', \infty) = 7.5$
[14]Mean WTP (1/3): $(-\infty, -3.9'] = -6.1,\ [\ 3.9', \infty) = 5.9$

Table 6.8: Regression analysis with alternative WTP specifications

	(1)	(2)	(3)
Salience Treatment	-2.763***	-1.681***	-1.331***
	(0.792)	(0.477)	(0.376)
Female	0.481	0.308	0.252
	(0.834)	(0.502)	(0.395)
Above Median Age	0.701	0.379	0.274
	(0.848)	(0.511)	(0.403)
High Usage of MM	2.815***	1.734***	1.385***
	(0.801)	(0.482)	(0.380)
High School	2.262**	1.338**	1.040**
	(0.918)	(0.554)	(0.437)
Constant	-4.300***	-2.577***	-2.029***
	(0.911)	(0.550)	(0.434)
Observations	637	637	637
R^2	0.048	0.049	0.049

Robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Notes: The table shows OLS regressions where the outcome variable is WTP taking on the mean value of the WTP intervals. Column (1) shows the regression using a triangular distribution with mean WTP values at the negative and positive outer bounds, which are -13.1 and 12.4, respectively. Column (2) shows the regression using a heuristic benchmark where we take half the values of the mean WTP values using a triangular distribution, which provides mean WTP on the negative and positive outer bounds to be -7.8 and 7.5, respectively. Column (3) shows the regression where we use one-third of the mean WTP values when using a triangular distribution, which gives mean WTP on the negative and positive outer bounds to be -6.1 and 5.9, respectively.

Using the Wald test, we check if the δ estimate in the main regression in column (1), is statistically different from the total amount of taxes and fees. The test returns a p-value of 0.0233, which means that the effect size is statistically different from total taxes and fees at the 5 percent level, and this implies that the magnitude of the treatment effect is large.

However, we also checked the alternative WTP specifications in columns (2) and (3). Running the same test, we found (p=0.1324) and (p=0.3278) for column (2) and column (3), respectively. This indicates that the effect is not statistically different and implies that the magnitude of the effect depends on the assumptions of the specification of WTP outer bounds. Implications for the magnitude of the effect will be discussed in section 7.

Robustness check 4: The results are validated in all 16 model specifications The command checked statistical analyses with all combinations of the main model specification given by equation (5.2), and it checked whether the treatment variable was significant at the 1 percent level. It found that all the 16 model specifications were significant at the 1 percent level, which indicates that the treatment effect is robust under these alternative model specifications.

7 Discussion

7.1 Discussion of the results

7.1.1 Main findings support the salience theory

Taking a closer look at the results, we see that when choosing between receiving the participation fee in MM or cash, the participants calculate the valuation of each alternative, similar to equation (2.1). The monetary amount displayed is a visible attribute they base their valuation on, as well as recent experiences with MM. The weight they assign to each visible attribute is determined by whether they are contrasting or surprising relative to their prior experience with the payment options. When it comes to the invisible attributes, we have taxes and fees related to MM usage. As equation (2.1) explains, there is a probability (r_k) of recalling the taxes and fees. Therefore, when increasing the salience of taxes and fees for the treatment group, which is accomplished by including reminders in the experiment, we ensure that they recall the non-salient attributes. This distorts the weights related to the non-salient attributes and makes those features more prominent, driving bottom-up attention by being considered contrasting to the cash alternative.

Both control and treatment groups come close to estimating the taxes and fees when explicitly asked about it, thereby providing evidence for the salience theory. Furthermore, our findings indicate that bottom-up attention is driven by prominence stimuli and not by surprise. If it were driven by the latter, the participants would not have managed to calculate the taxes and fees correctly. However, it is important to note that the two alternatives are not perfect substitutes, and individual preferences for each option can influence their valuation tremendously. Nevertheless, we can still explain this within the framework of the salience model. They might regard other attributes higher, e.g., the convenience feature of MM, thus making it the preferred option. We also see in table 6.1 that the salience effect is strong without using the WTP calculations. Therefore, we are confident that the salience effect is identified, which is consistent with previous studies (Finkelstein, 2009; Chetty et al., 2007). By applying the salience model to a new setting, we contribute to the existing body of literature on this topic.

7.1.2 Magnitude of the measured salience effect

The measured effect of TZS 2,763 (approx. USD 1.17) is around three times as high as the actual amount of taxes and fees. The estimated differences in mean outcomes between the treatment and the control group show a statistical effect size of 0.2657 (using Hedges' g), which is higher than the minimum detectable effect size we assumed in our pre-analysis plan. To evaluate the mechanisms affecting the magnitude of the salience effect, we offered an oversimplified example presented in robustness check 3. In this case, we assume that everyone in the control group is inattentive to the taxes and fees since these are non-salient attributes ($r_k = 0$), and that everyone who receives the treatment calculate taxes and fees perfectly and incorporate them in their decision-making ($r_k = 1$). Given that cash and MM are perfect substitutes, we would expect the treated group on average to have a lower WTP for MM compared to control group since they incorporate the knowledge about them. More specifically, their WTP would equal the total amount of taxes and fees (TZS 963). However, the effect size seems to be statistically significantly higher than the taxes and fees at the 5 percent level. Even with this simplification, it is evident that individuals, on average, seem to overreact to the salience treatment.

A plausible reason for these results can be that the assumptions related to WTP interval calculations are inaccurate. Initially, we assumed a triangular distribution of the mass +/- TZS 3,900 before collecting data. However, this assumption is premised on the belief that most participants fall in the intervals in the middle of the WTP range. Using heuristic benchmarks, as shown in robustness check 3, we see that when using the oversimplified example where taxes and fees are fully non-salient in the control group and fully salient in the treatment, assuming that cash and MM are perfect substitutes, the effect is not statistically different from total taxes and fees. By using alternative WTP specifications for the outer bounds, the results seem to align with the salience theory, and we can not statistically say there is an overreaction to treatment in this specific scenario. Consequently, our estimates are sensitive to the specific assumptions regarding the mean value of the outer bounds. Nonetheless, the salience effect itself remains robust even when altering the underlying assumptions, which is enhanced by the robust findings when conducting OLS regressions for each of the questions in the MPL.

It is also essential to acknowledge that the MM levy has recently been debated in Tanzanian news and consequently been reduced. Based on the salience model, it assumes that the attributes of a good act as triggers for memory retrievals, which generate an average version of the good. As the tax has been reduced several times since 2021, an individual's average memory of what it is, might be higher than the actual amount. This mechanism might result in participants favoring the cash option. The fact that the tax is already quite prominent could explain why many participants generally managed to estimate the taxes and fees when specifically asked about them. Interestingly, despite being familiar with the taxes and fees, we still identified a significant effect, indicating that many individuals still respond to the salience treatment. However, it is important not to overstate the magnitude of this effect, as it suggests that participants were already familiar with the taxes and fees prior to the experiment.

7.1.3 Exploratory findings

Our experiment also explored various aspects regarding MM usage in the sample. One notable finding is that 52.5% of the participants in our sample use MM services twice a week or more often. Conversely, 3.5% of our sample have never used it. Although many participants use MM frequently, we see from the results of the experiment that most participants can be found on the negative side of the WTP range, indicating a preference towards cash. Although, MM is frequently used this indicates that cash is the preferred option for most participants. Our results also indicate that participants who use MM frequently have a higher WTP for using it. Furthermore, when the participants were asked what they use MM for, 58.3 % answered that they use it for transferring money to family members, while 20.6% of the respondents reported using MM services to pay government fees. This it is of interest to explore further, since being able to pay these fees digitally might improve the government's collection of them. Since our sample consists of vendors, it is also interesting that only 16.6% answered that they use MM services to sell goods and services, which indicates that most of them mainly use cash in their business.

We conducted a heterogeneity analysis to explore whether sub-groups within our sample behave differently when receiving salience treatment. From table 6.5, we see that participants who have completed high school and receive salience treatment have a lower WTP ($p<0.1$) compared to those who do not receive treatment, which indicates that

this sub-sample are more inclined to choose cash. However, we should be cautious about placing extensive emphasis on these results. Another result of the heterogeneity analysis is that when those participants who switch in the MPL once, and receive treatment, they have a negative WTP ($p<0.01$) but a less negative WTP for MM compared to the rest of the sample, which still indicates a preference towards cash. These results seem to make sense given the WTP distribution presented in figure 6.3. Lastly, we excluded multiple switchers from our sample following our pre-analysis plan, and interestingly, we found that women are more likely to be multiple switchers in this sub-sample. Therefore, such behavior might be driven by gender differences.

7.2 Limitations

Several limitations to our study need to be addressed. First, looking at the distribution of observations on the WTP intervals, we find that most participants can be found on the outer bounds, meaning that many consequently chose cash or MM throughout the MPL. This is a limitation because we do not manage to pinpoint precise WTP intervals and, in extension, a more precise effect size since the participants on the outer bounds drive up the estimates in this particular experiment design, as discussed in robustness check 3. One reason for this distribution could be that the monetary amounts offered in each choice set were not considered sufficiently different, as the difference between the maximum and minimum amounts was TZS 4,000 (approx. 1.70 USD). As indicated by Jack et al. (2022), if participants become indifferent to the alternatives, they might stick to their initial strategy since calculating the valuation of each good becomes too "mentally costly" compared to how much more they receive. However, increasing the monetary amounts was not possible due to budgetary restrictions and it was necessary to be above the minimum amount of TZS 30,000 for the tax to apply, and the budget did not allow for more than TZS 34,000 per person.

A plausible reason why the magnitude of the effect is much larger than expected could be because of factors not accounted for in our main analysis. Due to the limited length of the survey, questions related to income, trust in government, and other reasons behind an individual's preferences for MM or cash, were not included. However, several studies have indicated that these factors are relevant in the Sub-Saharan African context (Tetteh et al.,

2023; Niesten and Abounabhan, 2023). For example, since the tax itself is regressive, low-income individuals must pay more relative to their income and are likely to be more inclined to choose cash (Niesten and Abounabhan, 2023). This is a common concern in experiments, and there is always a trade-off between survey length and insight. However, we did include a question about preferences in the final part of the survey, though the location of this question is prone to be affected by previous answers.

Another reason could be that since people are aware they are being observed, they might behave differently based on the expectations of how the experiment will inform future policy decisions or choose the option considered more socially acceptable, i.e., social desirability bias, as previously mentioned. However, since the experiment was related to their actual reward in the experiment, we do not consider this to be prominent, although it would be valuable to include measures that allows us to check for such behavior. Lastly, one other plausible reason is that the laboratory design makes participants aware that they are being observed, which may alter their behavior (Viceisza, 2016). This is however not a major concern since our questions were not sensitive, but it should be addressed since it is a general concern when conducting an experiment.

The next limitation is the fact that we excluded 159 multiple switchers from the sample. This is a limitation since we ended up with much fewer observations. It could have potentially led to an unbalanced sample and weakened statistical power. We did account for potentially removing MSB participants in our planning phase, but it turned out to be more prominent in our study compared to similar studies using such a design (Allcott and Kessler, 2019; Jack et al., 2022). However, removing these actually made the sample more balanced, particularly on the gender dimension, according to tables 4.3 and A1.1. Conversely, one can argue that the design allows us to detect the ones that do not understand or are inattentive, ultimately resulting in more truthful estimates.

One plausible reason is that MSB participants did not understand or pay sufficient attention when participating in the experiment. In order to combat inattention from participants, we considered including an additional check to ensure that participants actually wanted to choose the option indicated. However, this was decided against since the experiment was already lengthy, and increasing the length could have increased the inattention instead. We also looked into the possibility of designing the experiment more

efficiently, similar to the alternative phone survey in the study by Allcott and Kessler (2019). If we assume consistent choices, the fourth question could have been asked first, where the MM and cash monetary amount are equivalent, and then the next question would be based on their answer to this question. Such a design would have reduced the number of questions to three, and MSB would occur less. As mentioned in section 4.2, we decided against recruiting students since they have many similar characteristics to WIERD participants, i.e., Western, Industrialized, Educated, Rich, and Democratic (Friedman, 2012). Therefore, when comparing the number of MSB in studies that use such samples, they might not be transferable to our context. For example, according to our descriptive statistics in table 4.3, only 27% of our sample have finished high school education.

Lastly, there is a limitation in relation to the recruitment process since we introduced an exclusion criterion of participants who could not read after the recruitment was already implemented. This is a limitation of external validity since it affects the representation of the sample. After the first day of sessions, it became evident that illiteracy might be more prominent than expected. Since the experiment was designed to be answered individually on tablets and due to the sensitivity of some of the questions, this requirement was added. However, this only affects the degree to which the findings can be generalized to other contexts and does not impact the findings and internal validity of our research.

7.2.1 Discussion of validity

An important aspect to consider is the internal validity of the results. This is important to be able to evaluate the causal effect of a treatment (Slack and Draugalis Jr, 2001). A benefit of laboratory experiments is high control of the participants' surroundings. In the experiment, the only difference was the extra information about MM taxes and fees provided to participants in the salience treatment group. Participants only met the enumerators from Tanzania that were briefed on the experiment. Notably, the participants were not informed that the experiment was about taxation, and the causal interpretation of treatments is likely to be valid. Another benefit of laboratory experiments is the opportunity for randomizing who receives treatment. This eliminates selection bias, allowing us to find a causal effect.

Extensive checks have been conducted to ensure that the findings are valid. One major concern was the gender and age imbalance between the control and treatment groups within the full sample. After removing MSB participants, the sample became more balanced on the gender dimension making it less concerning but the age dimension was still unbalanced. However, in our heterogeneity analysis, we found that gender and age sub-groups do not behave significantly differently when receiving treatment. Contrarily, we find that the participants with high school education who receive treatment behave differently. This indicates that the treatment effect might be stronger because of the 2.8% higher amount of high school graduates in the treatment group. However, since this estimate is only significant at the 10 percent level, we attach less importance to it. Extensive robustness checks have also been conducted to ensure the findings are valid, which have all deemed strong results. Therefore, our results are considered reliable and internally valid.

External validity is not a priority in this study, but it is worth commenting on why this is not the case. The sample for this experiment is quite specific, which is vendors at informal markets in Dar es Salaam. From Afrobarometer (2021), we find that the use of MM is higher in urban areas compared to rural areas. Additionally, rural areas have less access to places where withdrawals and deposits can be made, and connectivity is more limited. Furthermore, people in urban areas generally have higher education than rural areas. Additionally, they might have better access to news and thus have followed the debate more closely. Therefore, our sample is not representative for the entire population in Tanzania.

7.3 Implications

The results of this thesis provide important implications for understanding citizens' behavior and policy. First, the salience of the tax affects how individuals react to it. According to the findings presented in this thesis, a highly salient tax provokes a stronger behavioral reaction than a less salient tax. Therefore, less salient MM taxes is favorable for a government trying to generate higher domestic revenue with as little reaction from citizens as possible. On the other hand, if a government aims at nudging citizens into behaving a certain way, the tax should be salient enough to provoke a reaction (Varela,

2016). For example, given the goal is to nudge citizens to make more healthy choices when buying groceries, taxing products containing refined sugar or alcohol would make these choices less appealing if the tax policy was designed in a highly salient way. However, if the tax is designed to be less salient, the desired reaction would not occur, at least not to the same degree.

Finally, some ethical implications of designing less salient tax policies should be addressed. According to the fiscal contract theory, people assent to pay taxes because they value the public services they get access to (World Bank, 2023). When decreasing the salience of taxes, one can argue that a government to a certain degree manipulates its citizens for tax money. This thesis finds that the salience of the tax affects individuals' willingness to pay taxes (to use MM). Therefore, complementing de la Cuesta et al. (2020) who finds that indirect, or less salient, taxes provoke less willingness to punish the government politically, we argue that these types of taxes result in a higher willingness to accept the tax than if the tax were more salient. Therefore, our results complement those of de la Cuesta et al. (2020), that less salient taxes are favorable for the government to implement since indirect taxes generate more domestic revenues. Thus, governments have to critically reflect upon introducing such taxes as they can have long-term effects.

8 Conclusion

The main objective of this thesis was to investigate empirically whether individuals optimize fully with respect to taxes. We did this by analyzing the effect of salience on behavioral responses to MM taxes and fees using a MPL method to elicit WTP. The data was collected through a randomized laboratory-in-the-field experiment in Dar es Salaam, Tanzania.

We find that increasing the salience of MM taxes and fees reduces participants' WTP for using MM, which provides support for the salience theory. Our finding suggests that individuals consistently underweigh non-salient features in daily decision-making. The measured effect size of TZS 2,657 (or USD 1.17) is around three times as high as the actual amount of taxes and fees. We believe our estimates are sensitive to assumptions made regarding the mean value at the outer bounds of the WTP range. Furthermore, the substantial effect might be explained by the fact that the tax was already quite salient due to being recently debated in the news. Nevertheless, the results prove to be robust.

In addition to identifying the salience effect, our study shows other interesting aspects. We find that high school graduates are more inclined to choose MM to cash on average. However, when receiving the salience treatment, this group becomes less willing to pay for MM, indicating an overreaction. Furthermore, we find that half the sample uses MM services twice a week or more often, but more participants prefer cash to MM in the experiment. This highlights that cash is still the preferred option within this sample. Regarding the limitations of our study, the main consideration is the fact that most participants can be found on the outer bounds of the WTP interval range, and therefore we do not manage to calculate precise WTP intervals.

This thesis has investigated if tax salience theory applies to MM tax in the context of vendors working in informal markets in Dar es Salaam, Tanzania. The fact that the salience of the tax affects how individuals react to it has some practical policy implications. First, the salience of a tax needs to reflect the aim of the tax. Second, less salient taxes are favorable for a government to implement since our findings imply that such taxes can improve the efficiency of tax collection and generate more domestic revenue.

With the limitations of this study in mind, further research on this topic is highly recommended. Repeating the study with a larger sample would allow several treatment groups and different sample characteristics, i.e., not just looking at vendors from informal markets in Dar es Salaam. A larger sample would also make it possible to investigate different treatment strengths and how these affect different respondents. Furthermore, we were unable to examine aspects of preference that we believe drive the high willingness to pay for using MM, e.g., the effects of income and trust in government. Therefore, this serves as an interesting avenue to investigate further.

We provide evidence that the salience of the taxes and fees on MM transactions in fact reduces willingness to pay for using MM instead of cash. Furthermore, we have applied the tax salience theory to a new setting. To our knowledge, no similar studies combine the two topics of salience and MM tax, and we believe this thesis provides a first step in addressing this gap.

References

Abiona, O. and Koppensteiner, M. F. (2022). Financial inclusion, shocks, and poverty evidence from the expansion of mobile money in tanzania. *Journal of Human Resources*, 57(2):435–464.

AfricaPay (2023). Minimum Wages: Mainland Trade, Industries, and Commercial Services. Last accessed 27th May 2023. https://africapay.org/tanzania/salary/minimum-wages/6182-mainland-trade-industries-and-commercial-services.

Afrobarometer (2021). Afrobarometer round 8 survey in tanzania. Last accessed 27th May 2023. https://www.afrobarometer.org.

Akyoo, S. E. and Pallangyo, B. T. (2021). Access and use of mobile money by microentrepreneurs in moshi district, tanzania.

Allcott, H. and Kessler, J. B. (2019). The welfare effects of nudges: A case study of energy use social comparisons. *American Economic Journal: Applied Economics*, 11(1):236–276.

Andersen, S., Harrison, G. W., Lau, M. I., and Rutström, E. E. (2006). Elicitation using multiple price list formats. *Experimental Economics*, 9:383–405.

Angrist, J. D. (2004). Treatment effect heterogeneity in theory and practice. *The economic journal*, 114(494):C52–C83.

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Anyidoho, N. A., Gallien, M., Rogan, M., and van den Boogaard, V. (2022). Mobile money taxation and informal workers: Evidence from ghana's e-levy. *Development Policy Review*, page e12704.

Bejaković, P. (2020). How to achieve efficiency and equity in the tax system? *Revija za socijalnu politiku*, 27(2):137–150.

Bloom, H. S. (1995). Minimum detectable effects: A simple way to report the statistical power of experimental designs. *Evaluation review*, 19(5):547–556.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2022). Salience. *Annual Review of Economics*, 14:521–544.

Brautigam, D., Fjeldstad, O.-H., and Moore, M. (2008). *Taxation and state-building in developing countries: Capacity and consent*. Cambridge University Press.

Burchardi, K. B., De Quidt, J., Gulesci, S., Lerva, B., and Tripodi, S. (2021). Testing willingness to pay elicitation mechanisms in the field: Evidence from uganda. *Journal of Development Economics*, 152:102701.

Chetty, R., Looney, A., and Kroft, K. (2007). Salience and taxation. *NBER Working Paper*, 13330.

Cuttler, C. (2023). Experimental Design. Last accessed 24th May 2023. https://opentext.wsu.edu/carriecuttler/chapter/experimental-design/.

Dave, C., Eckel, C. C., Johnson, C. A., and Rojas, C. (2010). Eliciting risk preferences: When is simple better? *Journal of Risk and Uncertainty*, 41:219–243.

de la Cuesta, B., Martin, L., Milner, H. V., and Nelson, D. L. (2020). Do indirect taxes bite? how hiding taxes erases accountability demands from citizens.

Dupont, W. D. and Plummer, W. D. (1990). Power and sample size calculations: a review and computer program. *Controlled clinical trials*, 11(2):116–128.

Duru, M., Sautmann, A., Turitto, J., and Welch, K. (2023). Pre-Analysis Plans. Last accessed 27th May 2023. https://www.povertyactionlab.org/resource/pre-analysis-plans.

Economides, N. and Jeziorski, P. (2017). Mobile money in tanzania. *Marketing Science*, 36(6):815–837.

Finkelstein, A. (2009). E-ztax: Tax salience and tax rates. *The Quarterly Journal of Economics*, 124(3):969–1010.

FinScope Tanzania, S. (2006). Findings report for demand for financial sevices and barriers. Last accessed 26th May 2023. http://www.fsdt.or.tz/.

Friedman, J. (2012). Weird samples and external validity. Last accessed 30th May 2023. https://blogs.worldbank.org/impactevaluations/weird-samples-and-external-validity.

Gamage, D. and Shanske, D. (2011). Three essays on tax salience: Market salience and political salience. *Tax L. Rev.*, 65:19.

Gneezy, U. and Imas, A. (2017). Chapter 10 - lab in the field: Measuring preferences in the wild. In Banerjee, A. V. and Duflo, E., editors, *Handbook of Field Experiments*, volume 1 of *Handbook of Economic Field Experiments*, pages 439–464. North-Holland.

Goldin, C. D. (2016). Human capital. Retrieved from: https://dash.harvard.edu/handle/1/34309590.

Grimm, P. (2010). Social desirability bias. *Wiley international encyclopedia of marketing*.

GSMA (2019). The state of mobile internet connectivity report 2019. Last accessed 25 May 2023. https://www.gsma.com/mobilefordevelopment/resources/the-state-of-mobile-internet-connectivity-report-2019/.

GSMA (2021). The mobile economy 2021. Last accessed 25 May 2023 . https://www.gsma.com/mobileeconomy/wp-content/uploads/2021/07/GSMA_MobileEconomy2021_3.pdf.

GSMA (2023). State of the Industry Report on Mobile Money 2023. Last accessed 24th May 2023. https://www.gsma.com/mobilefordevelopment/resources/state-of-the-industry-report-on-mobile-money-2023-2/.

Hellevik, O. (2009). Linear versus logistic regression when the dependent variable is a dichotomy. *Quality & Quantity*, 43:59–74.

Hu, A. (2023). Heterogeneous treatment effects analysis for social scientists: A review. *Social Science Research*, 109:102810.

Idris, I. (2018). Barriers to women's economic inclusion in tanzania. Last accessed 27th May 2023. https://opendocs.ids.ac.uk/opendocs/handle/20.500.12413/13864.

Jack, B. K., McDermott, K., and Sautmann, A. (2022). Multiple price lists for willingness to pay elicitation. *Journal of Development Economics*, 159:102977.

Jacobsen, A. E. (2022). Taxation in the informal economy: Towards a meaning-oriented conceptualization. Presented at the TaxCapDev-conference September 1st, 2022, Bergen, Norway. https://www.taxcapdev.no/event/test-power-and-politics-perspectives-on-taxation-and-state-building-in-africa/.

Jameel, A. L. (2023). Quick Guide to Power Calculations. Last accessed 24th May 2023. https://www.povertyactionlab.org/resource/quick-guide-power-calculations.

Kirumirah, M. and Munishi, E. (2021). Characterizing street vendors in the urban settings of tanzania: Towards sustainable solutions to vendors' challenges. In *Sustainable Education and Development 9*, pages 245–261. Springer.

KoBoToolbox (n.d.). Kobotoolbox. Last accessed 28th May 2023. https://www.kobotoolbox.org/about-us/software/.

Kvartalnyi, N. (2023). What is mobile banking: Advantages and disadvantages of mobile banking. Last accessed 24th May 2023. https://inoxoft.com/blog/what-is-mobile-banking-advantages-and-disadvantages-of-mobile-banking/#:~:text=The%20advantages%20of%20mobile%20banking%20include%2024%2F7%20access%20to,and%20extra%20charges%20for%20services.

Lees, A. and Akol, D. (2021). There and back again: The making of uganda's mobile money tax. *ICTD Working Paper*. Last accessed 25 May 2023. Retrieved from https://www.ictd.ac/publication/there-and-back-again-the-making-of-ugandas-mobile-money-tax-rib/.

Lotto, J. (2018). Examination of the status of financial inclusion and its determinants in tanzania. *Sustainability*, 10(8):2873.

Lu, X. and White, H. (2014). Robustness checks and robustness tests in applied economics. *Journal of econometrics*, 178:194–206.

Mader, P., Duvendack, M., Lees, A., Larquemin, A., and Macdonald, K. (2022). Enablers, barriers and impacts of digital financial services: Insights from an evidence gap map and implications for taxation.

Materu, B. (2023). Tanzania to Limit SIM Cards to One per Network. Last accessed 27th May 2023. https://www.theeastafrican.co.ke/tea/business/tanzania-to-limit-sim-cards-to-one-per-network--1416840.

Matheson, T. and Petit, P. (2021). Taxing telecommunications in developing countries. *International Tax and Public Finance*, 28:248–280. = https://EconPapers.repec.org/RePEc:kap:itaxpf:v:28:y:2021:i:1:d:10.1007$_s$10797 − 020 − 09621 − 6.

Mawejje, J. and Lakuma, P. (2019). Macroeconomic effects of mobile money: evidence from uganda. *Financial innovation*, 5(1):1–20.

Ministry of Finance and Planning (2022). KAULI YA SERIKALI. Last accessed 24th May

2023. https://www.mof.go.tz/uploads/text-editor/files/KAULI%20YA%20SERIKALI%20SERIKALI%20SEPT%202022_1663668606.pdf.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *The review of economic studies*, 38(2):175–208.

Mpofu, F. Y. (2022). Taxing the digital economy through consumption taxes (vat) in african countries: Possibilities, constraints and implications. *International Journal of Financial Studies*, 10(3):65.

Niesten, H. and Abounabhan, M. (2023). Taxation of Digital Financial Services in Tanzania. Last accessed 24th May 2023. https://www.ictd.ac/publication/taxation-of-digital-financial-services-in-tanzania/.

Noreen, E. (1988). An empirical comparison of probit and ols regression hypothesis tests. *Journal of Accounting Research*, 26(1):119–133.

Nyonzo, F. (2022). Tanzania's solidarity tax. *Review of African Political Economy*, 49(174):643–651.

P Simmons, J., D Nelson, L., and Simonsohn, U. (2021). Pre-registration: Why and how. *Journal of Consumer Psychology*, 31(1):151–162.

Persson, A. and Tengs, E. (2023). Who pays for distributive politics? reassessing the patronage state in sub-saharan africa. *Working paper presented at CMI, Bergen, March 14th, 2023*.

Peter, L. L. and Yang, Y. (2019). Urban planning historical review of master plans and the way towards a sustainable city: Dar es salaam, tanzania. *Frontiers of Architectural Research*, 8(3):359–377.

Ramsey, F. P. (1927). A contribution to the theory of taxation. *The economic journal*, 37(145):47–61.

Saunders, M., Lewis, P., and Thornhill, A. (2009). *Research methods for business students*. Pearson education.

Sears, D. O. (1986). College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *Journal of personality and social psychology*, 51(3):515.

Slack, M. K. and Draugalis Jr, J. R. (2001). Establishing the internal and external validity of experimental studies. *American journal of health-system pharmacy*, 58(22):2173–2181.

Stata (2023). rtest. Last accessed 24th May 2023.https://www.stata.com/manuals/rtest.pdf.

Suri, T. (2017). Mobile money. *Annual Review of Economics*, 9:pp. 497–520. Last accessed 25 May 2023. https://www.jstor.org/stable/26774443.

Tanzania Communications Regulatory Authority (2023). Communication statistics: Quarter ending march 2023. Last accessed 25th May 2023. https://www.tcra.go.tz/uploads/text-editor/files/Communication%20Statistics%20for%20Q3%202023_1682344173.pdf.

Tanzania Revenue Authority (2022). THE FINANCE ACT 2022. Last accessed 24th May 2023. https://www.tra.go.tz/Images/THEFINANCEACT2022.pdf.

Tanzania Revenue Authority (2023). Tax Collection Statistics. Last accessed 24 May 2023. https://www.tra.go.tz/index.php/tax-collection-statistics.

Tetteh, C. K., Amoah, A., Kwablah, E., Asiama, R. K., and Ahiabor, G. (2023). A test of behavioural changes to electronic levy: Evidence from mobile money transactions in a developing country. *Cogent Economics & Finance*, 11(1):2202963.

Treisman, A. M. and Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology*, 12(1):97–136.

UNDP (2017). National human development report 2017: Tanzania. *UNDP (United Nations Development Programme)*. Last accessed 27th May 2023.http://www.esrf.or.tz/docs/thdr2017launch.pdf.

University of Pennsylvania (2021). Think Tanks and Civil Societies Program. Last accessed 24th May 2023. https://repository.upenn.edu/think_tanks/18/.

Varela, P. (2016). What is tax salience? Last accessed 26th May 2023. https://crawford.anu.edu.au/publication/12581/what-tax-salience.

Viceisza, A. C. (2016). Creating a lab in the field: Economics experiments for policymaking. *Journal of Economic Surveys*, 30(5):835–854.

Wertenbroch, K. and Skiera, B. (2002). Measuring consumers' willingness to pay at the point of purchase. *Journal of marketing research*, 39(2):228–241.

Wolfe, J. M. (1994). Guided search 2.0 a revised model of visual search. *Psychonomic bulletin & review*, 1:202–238.

Wooldridge, J. M. (2013). Introductory econometrics: A modern approach. Last accessed 27 May 2023. https://economics.ut.ac.ir/documents/3030266/14100645/Jeffrey_M._Wooldridge_Introductory_Econometrics_A_Modern_Approach__2012.pdf.

Wooldridge, J. M. (2020). *Introductory econometrics: A modern approach*. Cengage learning.

World Bank (2021). Global findex database. Last accessed 25 May 2023. https://www.worldbank.org/en/publication/globalfindex.

World Bank (2023). Power Calculations. Last accessed 27th May 2023. https://dimewiki.worldbank.org/Power_Calculations.

World Bank (2023). The state of mobile internet connectivity report 2019. Last accessed 25 May 2023. https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?locations=TZ.

World Bank (2023). Why Paying Taxes Matters. Last accessed 29th May 2023. https://subnational.doingbusiness.org/en/data/exploretopics/paying-taxes/why-matters.

Appendix

A1 Descriptive statistics

Table A1.1: Descriptive statistics of sample including MSB

	Mean			**P-value of t-test**
	Control	Treatment	Total	Control vs. Treatment
Female	0.44	0.38	0.41	0.086*
	(0.02)	(0.02)	(0.02)	
Above Median Age	0.42	0.50	0.46	0.027**
	(0.02)	(0.03)	(0.02)	
High School	0.26	0.27	0.27	0.781
	(0.02)	(0.02)	(0.02)	
Observations	397	399	796	

Mean (standard error) and p-value of t-test. $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Notes: The table reports descriptive statistics of the entire sample using the indicator variables used for the analysis. Column (1) reports means for control group, Column (2) reports means for the treatment group, Column (3) reports the means for the entire sample, and Column (4) reports p-values for a two-sided t-test for difference in means between control and treatment groups.

Table A1.2: MSB regression analysis

	(1)	(2)	(3)	(4)	(5)
Female	0.084***				0.090***
	(0.030)				(0.030)
Above Median Age		0.033			0.031
		(0.029)			(0.030)
High School			-0.047		-0.043
			(0.031)		(0.032)
Salience Treatment				0.022	0.025
				(0.028)	(0.028)
Constant	0.166***	0.185***	0.212***	0.189***	0.147***
	(0.017)	(0.019)	(0.017)	(0.020)	(0.028)
Observations	796	796	796	796	796
R^2	0.011	0.002	0.003	0.001	0.016

Robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Notes: The table shows OLS regressions where the outcome variable "MSB" is an indicator variable taking on the value of one if the observation is a multiple switcher, and zero otherwise. Column (1) shows regression of outcome variable on *Female* indicator variable. Column (2) shows regression of outcome variable on *Above Median Age* variable. Column (3) shows regression of outcome variable on *High School* indicator variable. Column (4) shows the regression of outcome variable on *Salience Treatment*. Lastly, Column (5) shows the outcome variable on all the independent variables in the main model specification.

Table A1.3: Summary of occupation (type of vendor)

Occupation	
Vegetables	25.9%
Spices	10.2%
Grains	13.9%
Clothes	14.9%
Electronics	1.3%
Other	32.7%
Don't know	0.3%
Prefer not to answer	0.8%

Notes: The overview of occupation was created by the answer to the background question "What kind of product are you selling?"

Table A1.4: Overview of the MM operators our sample uses

MM operators	
Vodacom with M-Pesa	16%
Tigo with Tigo Pesa	49.1%
Airtel with Airtel Money	39.4%
Halotel with Halopesa	1.7%
TTCL	0.5%
Zantel with Ezy Pesa	0%
Other	2.6%

Notes: The overview of the MM operators was created by asking the question "What mobile pay operator(s) do you use?". It allowed for choosing more than one alternative.

A2 OLS assumptions

We argue that our regression model fulfills the assumptions for an OLS regression model. We estimate the WTP when receiving salience treatment using STATA/SE 17.0. The first four assumptions must hold to get unbiased results (Wooldridge, 2013). Furthermore, there is an additional assumption to obtain efficient estimators. We will briefly review the assumptions and explain why our regression fulfills them.

1st assumption: Linear in parameters

To achieve linearity in parameters, our model is modeled in the following format (Wooldridge, 2013):

$$Y = \beta_0 + \beta_1 X_1 + ... + \beta_k X_k + u \tag{.1}$$

The relationship between the predictors and the outcome variable should be linear. Our main model specification in equation (5.2) follows this format.

2nd assumption: Random sampling

Random sampling is essential to ensure the representativeness of the population and must be drawn randomly from the population (Wooldridge, 2013). We randomly assigning treatment to the participants. Therefore, this assumption is fulfilled. More information about the procedure is in section 4.

3rd assumption: No perfect collinearity

The sample must not have constant independent variables and no exact linear relationship among them (Wooldridge, 2013). However, it is important to note that this assumption allows correlated independent variables, but importantly, they cannot be perfectly correlated. We can check for multicollinearity (since we have more than two independent variables) by using a variance inflation factor(VIF) metric. It measures the correlation strength between the independent variables in a regression model. A value less than 10 indicates no trouble with multicollinearity and there is no need to make any adjustments. However, Wooldridge (2013) argues that one should be careful with interpreting the results based on this test since many factors could explain a high VIF value. In our case, by

running a VIF test, we find that all independent variables are slightly over the value of 1, where the mean VIF is 1.07. Alternatively, one can create a correlation matrix to see if there is a correlation among the variables. From table A2.1, the variables have a very low correlation. Therefore, we do not consider multicollinearity an issue and use this test as a mere support for no perfect collinearity.

Table A2.1: Correlation matrix

Variables	(1)	(2)	(3)	(4)	(5)	(6)
(1)WTP	1.0000					
(2)Salience Treatment	-0.1321	1.0000				
(3)Female	0.0148	-0.0559	1.0000			
(4)Above Median Age	-0.0006	0.0831	-0.1270	1.0000		
(5)High School	0.0948	0.0314	0.0231	-0.2911	1.0000	
(6)High Usage of MM	0.1451	-0.0041	-0.0991	0.0587	0.0626	1.0000

Note: The table show the correlation matrix.

4th assumption: Zero conditional mean

The assumption tells us that the error term must have an expected value of zero given any values of the independent variables, as this creates an endogeneity problem, which is caused by omitted variable bias, measurement errors or simultaneity (Wooldridge, 2013). We believe the most important factors are controlled for in our analysis.

There is no reason to expect omitted variable bias since it is not likely that variables not included correlate with the salience treatment and are a determinant of the dependent variable. On the contrary, measurement errors could occur due to poor quality of the survey questions or the design of the experiment. However, the obtained results do not indicate that this is an issue in our study. Furthermore, reverse causality is not present since the dependent variable does not affect the treatment status since the treatment is assigned randomly in the experiment.

5th assumption: Homoskedasticity

The last assumption is important to have an efficient regression model, and it means that the error term should have a constant variance given any value of the independent variables, and if not, the model exhibits heteroskedasticity (Wooldridge, 2013). Our regression model suffers from heteroskedasticity since the dependent variable takes on eight different values and most participants end up in the two outer bounds. We mitigate

this by using robust standard errors so that we obtain unbiased standard errors.

A3 The survey

In the digital survey, pages 75-88 is out of our scope as it is not part of our analysis. On pages 72-74, we have the experiment questions for the treatment group and control group, respectively.

Dodoso

Please put the number at the table into the field below and then hand the tablet to the participant.

Welcome to this research study!

This is the consent form for participating in a study conducted by REPOA – an independent Tanzanian research institution and CMI – an independent Norwegian research institution. The purpose of the study is to learn about your views on social issues and economic decisions.

If you choose to take part in the project, this will involve that you fill in an electronic survey. In the survey, we will ask you about your views, and we will ask you to make some economic decisions. All the information you submit in the survey will be treated with strict confidentiality, and in compliance with privacy regulations. Your participation in the study will take about 1 hour and your answers will provide important inputs for policies to improve the Tanzanian society.

As a compensation for your time, you will receive a payment of TZS 30 000 at the end of this session. Depending on the choices you make, you may make additional earnings during the experiment. Depending on your choices, we may need your phone number to pay you. We will therefore ask you to provide your phone number as part of this study. Your phone number will only be used to pay you. It will not be used for any other purposes. Your phone number will be recorded separately on the sheet of paper on your desk. The sheet of paper will be shredded at the end of today. During today, your phone number will be stored safely in a safe place that only the researchers involved in this project have access to. After today all your responses are completely anonymous, and it will not be possible to pick you out from what you say or do in the survey.

Participation is voluntary

Participation in this study is voluntary. If you choose to participate, you can stop the survey at any time without giving a reason and without penalty. We will process your phone number during today based on your consent. During today, you have the right to request that the phone number is deleted. This afternoon, your phone number will be deleted from our file in any case.

By chance, you have been selected for participation in this study as one of 50 traders from the market you trade in.

If you have questions about the project, or wish to exercise your rights, please contact: Dr. Lucas Katera at katera@repoa.or.tz.

Please confirm that you have received and understand the information about the research study and whom to contact in case of questions and that you consent to take part in the study below. Do you agree to participate in the study?

- (●) Yes
- ( ) No

Thank you for participating in this study. From now on until the end of the session, please refrain from communicating with other participants. If you do not abide by this rule, we will have to exclude you from the study.

We kindly ask you to read the instructions thoroughly. If you have any questions after reading the instructions or during the study, please raise your hand and one of the instructors will come to you and answer your question in person. Your payment and your decisions throughout the study will be treated confidentially. None of the other participants is informed, neither during nor after the study, about your decisions in the study or your payment.

Phone number

First, we ask you to write your phone number that you use for mobile money on the sheet of paper on your desk. Your phone number may only be used to make the payment to you. It will not be used for any other purposes. The paper slip with your phone number will be shredded at the end of today.

When you have provided your phone number, please confirm that you have done so by ticking the box below.

- ◯ I confirm that I have provided my phone number
- ◯ I do not want to provide my phone number

To start with, please answer the following questions about yourself.

How old are you?

What is your gender?

- ◯ Female
- ◯ Male
- ◯ Prefer not to answer

What is your highest level of education?

- ◯ No schooling
- ◯ Informal schooling (including Koranic schooling)
- ◯ Some primary schooling
- ◯ Primary school completed
- ◯ Intermediate school or some secondary school/high school
- ◯ Secondary school/high school completed
- ◯ Post-secondary qualifications other than university, e.g. diploma or degree from polytechnic or college
- ◯ Some university
- ◯ University completed (Bachelor)
- ◯ Post-graduate (Master)
- ◯ Don't know
- ◯ Prefer not to answer

What kind of product are you selling?

- ◯ Vegtables
- ◯ Spices
- ◯ Grains
- ◯ Clothes
- ◯ Electronics
- ◯ Other (please specify)
- ◯ Don't know
- ◯ Prefer not to answer

Please specify

_____

What is your religion, if any?

- ◯ Christian
- ◯ Muslim
- ◯ None
- ◯ Other (please specify)
- ◯ Prefer not to answer

Please specify if you would like to

_____

This session will determine how you will receive the reward for participating in this experiment. We will now present you with a series of situations. In each situation, we want you to choose between two alternatives. The final reward will depend on your choices. One of the choices will decide the actual payment. You will be paid the reward in either cash or mobile money before leaving the session today.

Note: The government of Tanzania has introduced a levy on sending mobile money for amounts above 30,000 TSH. Mobile money operators also add a fee everytime you send mobile money. The fees may vary depending on your mobile money operator.

_____

Which of the two alternatives for the reward do you prefer? Please tick one of the boxes to indicate your choice.

- ◯ 30,100 TSH in cash
- ◯ 34,000 TSH in mobile pay (Remember that there is a levy and operator fees when withdrawing or sending mobile money)

Which of the two alternatives for the reward do you prefer? Please tick one of the boxes to indicate your choice.

- ◯ 31,600 TSH in cash
- ◯ 34,000 TSH in mobile pay (Remember that there is a levy and operator fees when withdrawing or sending mobile money)

Which of the two alternatives for the reward do you prefer? Please tick one of the boxes to indicate your choice.

- 34,000 TSH in mobile pay (Remember that there is a levy and operator fees when withdrawing or sending mobile money)
- 33,100 TSH in cash

Which of the two alternatives for the reward do you prefer? Please tick one of the boxes to indicate your choice.

- 34,000 TSH in cash
- 34,000 TSH in mobile pay (Remember that there is a levy and operator fees when withdrawing or sending mobile money)

Which of the two alternatives for the reward do you prefer? Please tick one of the boxes to indicate your choice.

- 34,000 TSH in cash
- 33,100 TSH in mobile pay (Remember that there is a levy and operator fees when withdrawing or sending mobile money)

Which of the two alternatives for the reward do you prefer? Please tick one of the boxes to indicate your choice.

- 31,600 TSH in mobile pay (Remember that there is a levy and operator fees when withdrawing or sending mobile money)
- 34,000 TSH in cash

Which of the two alternatives for the reward do you prefer? Please tick one of the boxes to indicate your choice.

- 34,000 TSH in cash
- 30,100 TSH in mobile pay (Remember that there is a levy and operator fees when withdrawing or sending mobile money)

This session will determine how you will receive the reward for participating in this experiment. We will now present you with a series of situations. In each situation, we want you to choose between two alternatives. The final reward will depend on your choices. One of the situations will decide the actual outcomes. You will be paid the reward in either cash or mobile money before leaving the session today.

Which of the two alternatives for the reward do you prefer? Please tick one of the boxes to indicate your choice.

- 34,000 TSH in mobile pay
- 30,100 TSH in cash

Which of the two alternatives for the reward do you prefer? Please tick one of the boxes to indicate your choice.

- 34,000 TSH in mobile pay
- 31,600 TSH in cash

Which of the two alternatives for the reward do you prefer? Please tick one of the boxes to indicate your choice.

- 33,100 TSH in cash
- 34,000 TSH in mobile pay

Which of the two alternatives for the reward do you prefer? Please tick one of the boxes to indicate your choice.

- 34,000 TSH in mobile pay
- 34,000 TSH in cash

Which of the two alternatives for the reward do you prefer? Please tick one of the boxes to indicate your choice.

- 33,100 TSH in mobile pay
- 34,000 TSH in cash

Which of the two alternatives for the reward do you prefer? Please tick one of the boxes to indicate your choice.

- 31,600 TSH in mobile pay
- 34,000 TSH in cash

Which of the two alternatives for the reward do you prefer? Please tick one of the boxes to indicate your choice.

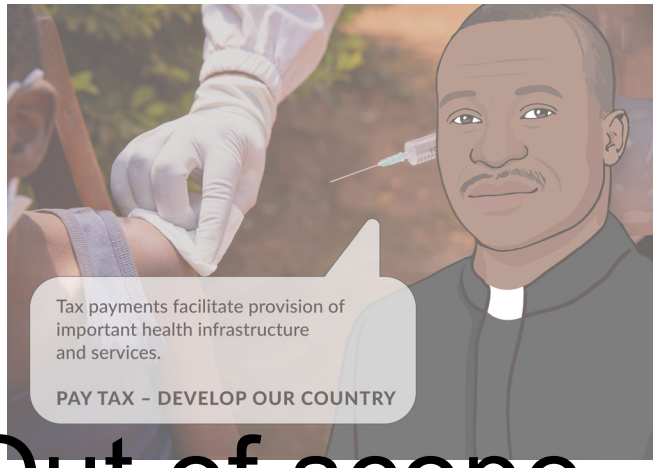
- 30,100 TSH in mobile pay
- 34,000 TSH in cash

Thank you. We will now continue with the next part of the survey.

Please take a careful look at the picture below.

01.06.2023, 11:05 Dodoso

Please take a careful look at the picture below.



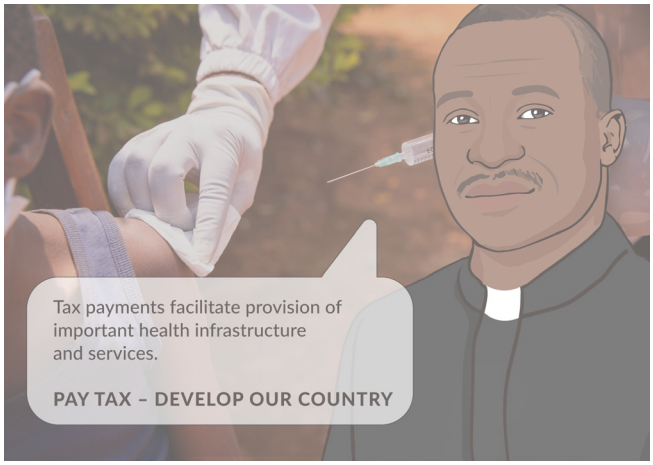
Out of scope

Please take a careful look at the picture below.

Please take a careful look at the picture below.

Out of scope

Please take a careful look at the picture below.

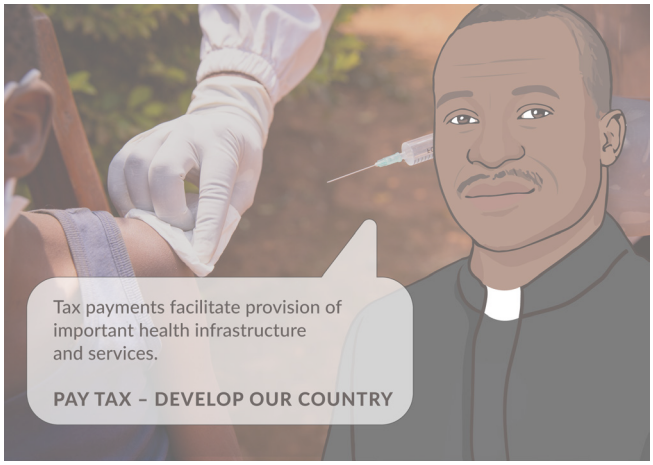


01.06.2023, 11:05 Dodoso

Please take a careful look at the picture below.

Out of scope

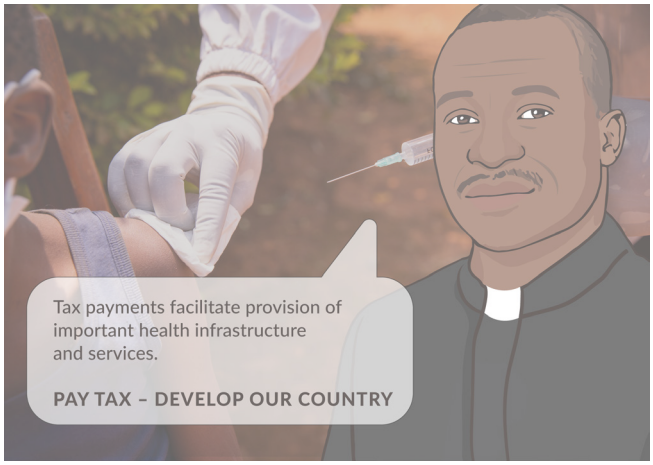
Please take a careful look at the picture below.



Please take a careful look at the picture below.

Out of scope

Please take a careful look at the picture below.

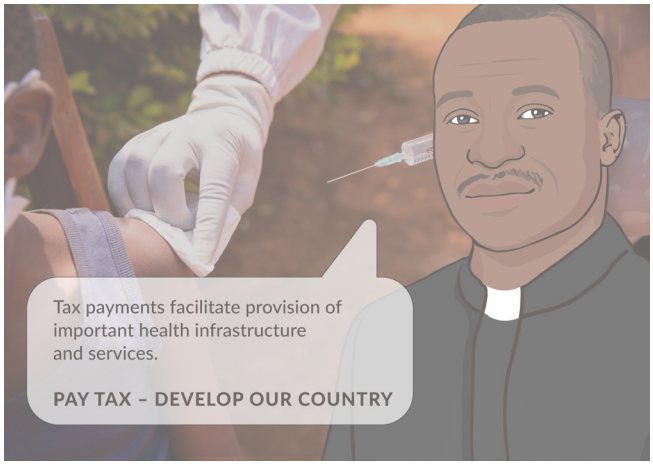


01.06.2023, 11:05 Dodoso

Please take a careful look at the picture below.

Out of scope

Please take a careful look at the picture below.

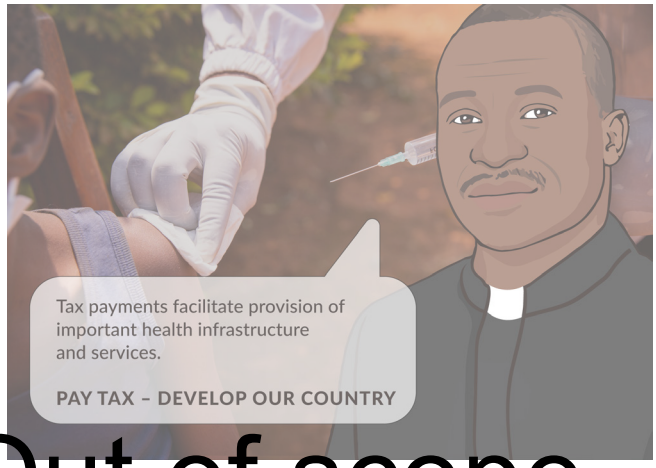


Please take a careful look at the picture below.

Out of scope

Please take a careful look at the picture below.

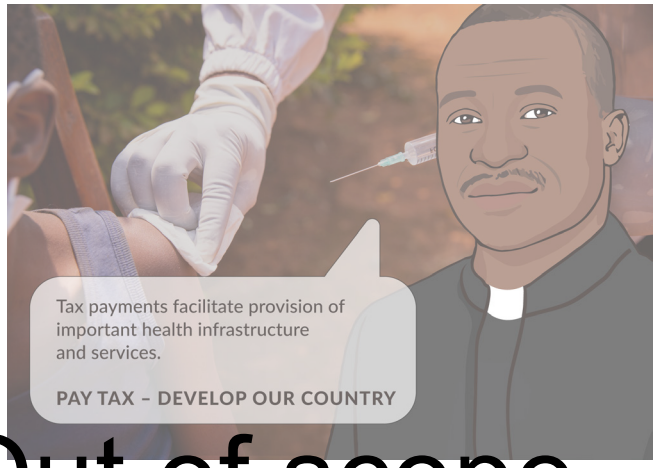
Please take a careful look at the picture below.



Out of scope

Please take a careful look at the picture below.

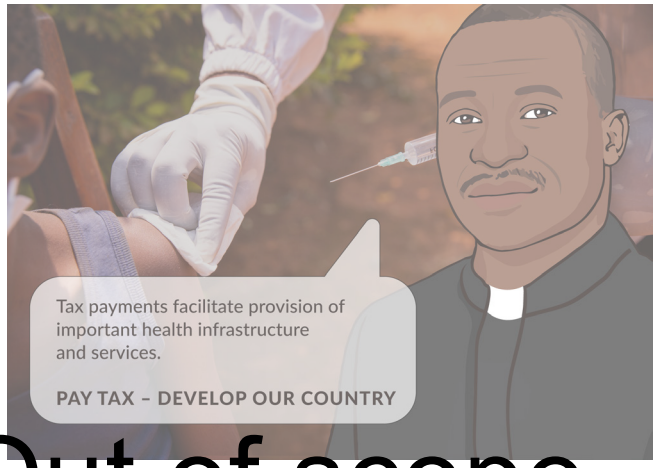
Please take a careful look at the picture below.



Out of scope

Please take a careful look at the picture below.

Please take a careful look at the picture below.



Out of scope

How credible do you find the information in the picture?

- ◯ Very
- ◯ Somewhat
- ◯ Just a little
- ◯ Not at all

How much do you trust the information in the picture?

- ◯ Very
- ◯ Somewhat
- ◯ Just a little
- ◯ Not at all

How much of the information in the picture was new to you?

- ◯ All of it
- ◯ Most of it
- ◯ Some of it
- ◯ None of ot

In this part of the study, you will receive an extra payment. The size of this payment depends on a decision you make.

You will make two decisions. One of them will be picked by chance and determine your payment.

You will receive the payment in cash at the end of the session. To keep your decisions anonymous, you will receive the money in a sealed envelope.

In each decision, you receive an income of 10,000 TSH. This income is subject to a tax of 20%. Your tax payment depends on how much income you report:

Tax Payment = 20% of Reported Income

You can decide how much income you report. Your final payment is your income minus the tax payment:

Final Income = 10 000 – (Tax Payment)

Your tax payment will be transferred to the recipient stated in the instructions for each decision.

The tax collected in this decision will be added to the tax revenue of Tanzania. How much of the 10 000 TSH would you like to report?

- ◯ 10,000 TSH
- ◯ 9,000 TSH
- ◯ 8,000 TSH
- ◯ 7,000 TSH
- ◯ 6,000 TSH
- ◯ 5,000 TSH
- ◯ 4,000 TSH
- ◯ 3,000 TSH
- ◯ 2,000 TSH
- ◯ 1,000 TSH
- ◯ 0 TSH

Out of scope

Your final income is **TSH**.
You reported an income of TSH.
Your tax payment is TSH.

The tax collected in this decision will be added to the budget of a community health initiative in Tanzania. How much of the 10 000 TSH would you like to report?

- ◯ 10,000 TSH
- ◯ 9,000 TSH
- ◯ 8,000 TSH
- ◯ 7,000 TSH
- ◯ 6,000 TSH
- ◯ 5,000 TSH
- ◯ 4,000 TSH
- ◯ 3,000 TSH
- ◯ 2,000 TSH
- ◯ 1,000 TSH
- ◯ 0 TSH

Your final income is **NaN TSH**.
You reported an income of TSH.
Your tax payment is NaN TSH.

Out of scope

The tax collected in this decision will be added to the tax revenue of Tanzania. How much of the 10 000 TSH would you like to report?

- ◯ 10,000 TSH
- ◯ 9,000 TSH
- ◯ 8,000 TSH
- ◯ 7,000 TSH
- ◯ 6,000 TSH
- ◯ 5,000 TSH
- ◯ 4,000 TSH
- ◯ 3,000 TSH
- ◯ 2,000 TSH
- ◯ 1,000 TSH
- ◯ 0 TSH

Your final income is **NaN TSH**.
You reported an income of TSH.
Your tax payment is NaN TSH.

We now ask you to complete a short questionnaire while we prepare your payments.

Who finances provision of health infrastructure and services in Tanzania?

How satisfied or dissatisfied are you with the following public services provided by the government?

Provision of roads and bridges

- Very satisfied
- Satisfied
- Neither satisfied nor dissatisfied
- Dissatisfied
- Very dissatisfied

Provision of public transport

- Very satisfied
- Satisfied
- Neither satisfied nor dissatisfied
- Dissatisfied
- Very dissatisfied

Out of scope

Provision of health infrastructure and services

- Very satisfied
- Satisfied
- Neither satisfied nor dissatisfied
- Dissatisfied
- Very dissatisfied

Provision of education services

- Very satisfied
- Satisfied
- Neither satisfied nor dissatisfied
- Dissatisfied
- Very dissatisfied

Please consider the following statement:
Taxpayers must always pay the taxes that they owe to the tax authority.

How much do you agree with the statement?

- ◯ Stongly Agree
- ◯ Agree
- ◯ Neither agree nor disagree
- ◯ Disagree
- ◯ Strongly Disagree

Out of the 24 other participants in the room, how many do you think Agree or Strongly Agree with the statement?

_____

Please consider the following statement:
Taxpayers could refuse to pay taxes if they are not receiving public services of adequate quality.

_____

How much do you agree with the statement?

- ◯ Stongly Agree
- ◯ Agree
- ◯ Neither agree nor disagree
- ◯ Disagree
- ◯ Strongly Disagree

Out of scope

Out of the 24 other participants in the room, how many do you think Agree or Strongly Agree with the statement?

_____

Out of ten taxpayers, how many do you think cheat on their taxes

_____

Out of ten taxpayers who cheat on their taxes, how many do you think are detected by the Tanzanian Revenue Authority?

_____

How much do you trust the national government?

- ◯ A lot
- ◯ Somewhat
- ◯ Just a little
- ◯ Not at all
- ◯ Don't know / Haven't heard enough

How much do you trust your local government authority?

- ◯ A lot
- ◯ Somewhat
- ◯ Just a little
- ◯ Not at all
- ◯ Don't know / Haven't heard enough

How much do you trust the Tanzania Revenue Authority?

- ◯ A lot
- ◯ Somewhat
- ◯ Just a little
- ◯ Not at all
- ◯ Don't know / Haven't heard enough

How much do you trust Religous Christian leaders?

- ◯ A lot
- ◯ Somewhat
- ◯ Just a little
- ◯ Not at all
- ◯ Don't know / Haven't heard enough

Out of scope

How much do you trust Religous Muslim leaders?

- ◯ A lot
- ◯ Somewhat
- ◯ Just a little
- ◯ Not at all
- ◯ Don't know / Haven't heard enough

How willing are you to give to good causes without expecting anything in return?

- ◯ Very
- ◯ Somewhat
- ◯ Just a little
- ◯ Not at all

How important would you say religion is in your life?

- ◯ Very important
- ◯ Rather important
- ◯ Not very important
- ◯ Not at all important
- ◯ Prefer not to answer

Finally, we want to know more about how you send and receive money. How do you generally prefer to send and receive money?

- ◯ Cash
- ◯ Mobile pay
- ◯ Bank account
- ◯ Other

What mobile pay operator(s) do you use?

- ☐ Vodacom with M-Pesa
- ☐ Tigo with Tigo Pesa
- ☐ Airtel with Airtel Money
- ☐ Halotel with Halopesa
- ☐ TTCL
- ☐ Zantel with Ezy Pesa
- ☐ Other

How often do you use mobile pay?

- ◯ Once a month or less
- ◯ 2-4 times a month
- ◯ Twice a week or more
- ◯ Daily
- ◯ Never
- ◯ Prefer not to answer

What do you use mobile pay for? Please select several options if that is correct for you.

- ☐ Transferring to family members
- ☐ Transferring to friends
- ☐ Goverment payments
- ☐ Buying goods and services
- ☐ Selling goods and services
- ☐ Other
- ☐ Prefer not to answer

Please specify if you would like to

_____

If you send (not withdraw) 40,000 TSH mobile money, approximatly how much extra would you need to pay for levy and fees when sending money (in TSH)? If you do not know, please make a guess.

- ◯ 0 TSH
- ◯ 200 TSH
- ◯ 400 TSH
- ◯ 600 TSH
- ◯ 800 TSH
- ◯ 1,000 TSH
- ◯ 1,200 TSH
- ◯ 1,400 TSH

Of the TSH, how much is the government levy?

_____

You have not consented to the survey. Please contact the enumerators.

_____