Essays on Household Behavioral Responses to Economic Shocks and Social Constraints

Guillermo Cabanillas-Jiménez

A thesis presented for the degree of
Doctor of Philosophy in Economics

School of Economics
University of Kent
United Kingdom

December, 2021

Contents

List of Tables

List of Figures

Executive Summary

This PhD thesis intends to extend the economic research in the area of behavioral economics and go deeper in the understanding of individual and household behaviors under different scenarios that an economy or the World can present. More precisely, in the first two chapters we analyze household decisions on consumption allocation of different types of goods under windfall effects. In the third one, we focus on individual decisions in the labor market when specific religious events occur.

Going in more detail through all chapters, the first one analyzes how households that live in winning regions of the Spanish Christmas Lottery behave, in terms of consumption of different types of goods, compared with those households that live in non-winning regions of the lottery. Using the winning regions to construct an instrumental variable for total household expenditures, we estimate the Engel curves. The main findings in this chapter are that (i) the estimated total household expenditures elasticities on durable and non-durable goods are similar and, thus, households respond in a similar way to a shock to total household expenditures; and (ii) that the direct impact of the lottery winnings on household consumption of durable and non-durable goods imply a violation of the Permanent Income Hypothesis (PIH). Such findings lead to a contradiction with the theoretical predictions by Cerletti and Pijoan-Mas, 2014.

In the second chapter we proceed with a similar study as in the first chapter, where we aim to test, at the household level, how winning the Millions or the £30.000 Postcode Lottery in the United Kingdom affects household consumption behavior. On a first instance, we test the direct effect of the lottery winnings on household consumption of different types of goods. The main finding is that the PIH is violated, as consumption for durable and non-durable goods increase due to the lottery income shock. Later, we estimate the Engel curves, using the same approach as we did in the first chapter. We do find that the estimated elasticity of durable and non-durable goods to total household expenditures are similar, as happens in the first chapter. Therefore, these results lead again to a contradiction of the expected theoretical results, as we should expect non-durable goods to not react sensitively to the income shock.

Finally, in the third chapter we analyze individual behavior in the labor market under religious constraints, such as Ramadan, one of the big five pillars of Islam, where individuals that take part of this tradition have to fast for an entire lunar calendar (29-30 days) from sunrise to sunset. Specifically, we test, using a Heckman model with fixed effects, how individuals labor supply is affected under the Ramadan month. Using Malawian data, we find that women do reallocate their time significantly, from paid jobs to housework. On the other hand, we do not find any evidence of time reallocation of hours worked for men. These results prove that during Ramadan, individuals do not work less or decrease their productivity, contrary to the general beliefs.

After doing these studies, we find out that human behaviors do not always follow the predicted theories and we do not behave rationally or according to the theory. This is the case we find along this thesis, where classic theories like the Permanent Income Hypothesis are violated.

Declaration of Authenticity

This report contains genuine work conducted originally by the author. The work presented herein has not been submitted and/or accepted for the award of any other degree or diploma in any university. To the best of my knowledge and belief, this thesis contains no materials previously published or written by other person, except where due references has been made.

Place: Univerity od Kent (Canterbury, United Kingdom)

Date: 27/05/2020

Name: Guillermo Cabanillas-Jiménez

PhD in Economics

Acknowledgements

I would like to thank many people that took part of this big project, and that I had the chance to meet, along these four years. They encouraged me to pursue this PhD, something challenging but that gave me a chance to do all what I like: work on what I want, have a chance to teach and get involved in the world of Academia. This list of acknowledgements is long, therefore, I will try to be as short as possible, despite sometimes this is not possible.

First, I would like to acknowledge the University of Kent and the School of economics to believe in my project and give me the chance to become a doctor in Economics. Also, thanks to them for award me with the Vice-Chancellor Scholarship and provide me with all the resources needed to develop my research project.

Thanks to Dr. Maria Garcia-Alonso for all her support since I arrive to Canterbury -even from before starting my thesis. I cannot forget either about the convenors of the modules I have been teaching at the University of Kent: Dr. Olena Nizalova, Dr. Miltos Makris and Dr. Sylvain Barde who were always there supporting me and giving me advice any time I needed them. And of course, I need to mention my great students that were always willing to attend my sessions and making my job easier.

During these years I have also overseen organizing the student PhD seminars, the micro-forum and micro-reading group sessions. These last two seminars would not have been possible to organize without the help of Dr. Bansi Malde who came up with the idea and help us to organize the micro-research group.

To conclude this University block, I would like also to thank Kent Union. Thanks to them I discovered the university societies, specially the UKC Spanish Society who made my adaptation to Canterbury very easy and the UKC Salsa Society, the big pillar during my time in Canterbury, making me discover a new hobby and giving me the chance to also teach something different than economics in a university environment.

I want also to thank the IDEA program from the Universitat Autònoma de Barcelona (UAB), where I did my master studies and gave me most of the knowledge to develop my PhD. Especially, I would like to thank Dr. Joan Llull, who was my master thesis adviser, and Dr. Inés Macho, my best support during my two years at UAB and someone who pushed me to present in conferences afterwards. Both motivated me to continue my studies and to do what I wanted, jointly with my mates Niccolò Maffei, Pablo Amor and Lluís Puig, people that always were there and supported me during the master and after that.

Now I want to list all those friends that supported me along these four years, those that I know since primary school, undergraduate, *Woolf college* flat mates, etc. People that made the experience a great one and worthy to take: Albert, Gerard, Xavi, Montse, Agustí, Giulia, Aysu, Laura, Yasmine, Yannis, Pablo and Theo. Also to my village friends in *Segur de Calafell*: Marta, Laura, Sandra, Maria, Laia, Lola, Javito, Quilez brothers, David, Gonzalo, Albert, Yera, Mampel, Albarral, Dani, Chuso and Jordi.

Also, I would like to give a big thanks to a group of people who made my stay in Canterbury special and different and without them it would not have been the same and nothing could have been explained, my little family of Canterbury: Laura, Marta, Sara, Clara, Marina, Leire, Paula, Sofía, Diana, Carmen, David, Pablo and Luisín.

Another big thanks is for the best housemate one could have ever had: Giada, she became an important piece of this story, especially during lockdown periods where we both were offering support to other, making the experience a good one.

And I would not like to finish this section without mentioning the most important person along these four years: Antonio. The most important member of this family I built in Canterbury, the one who have been always supporting me and made the experience here a great one. Big thanks to this good friend.

And finally, I cannot finish this without acknowledging my supervisors: Dr. Guy Tchuente for all his help and making me improve a lot in my econometrics and programming skills and to my principal supervisor, the most important person and the one that really motivated me to pursue this PhD and made this a reality: Dr. Zaki Wahhaj.

Chapter 1

Introduction of the Thesis

This PhD thesis is structured in three chapters, focusing on the area of behavioral economics and intends to explore new areas of research in this topic. Along these three chapters, we are interested in analyzing how different individuals and households behave under different scenarios that the economy can present. More precisely, in the first two chapters we focus on the analysis of local windfall effects on household consumption behavior in Spain and the United Kingdom. The third chapter tests individual performance in the labor market when specific religious events occur. Specifically, we are interested in the effect of religious constraints on individuals' labor supply decisions when they face religious constraints that might alter their daily routine activities.

Thus, the aims and objectives of the research conducted along this PhD thesis are mainly three: the first two chapters aim to test, on first instance, if Spanish and British households behave rationally according to the Permanent Income Hypothesis (PIH) or not when they experience local windfall gains. The second main goal of these two chapters is to estimate the Engel curves to estimate the elasticities of total household expenditures, instrumented with the lottery income shock, on household consumption expenditures of specific goods, with special focus on durable and non-durable goods. The third goal of this thesis is to proof that when individuals experience religious events that can alter their daily commitments, their working hours and their productivity do not decrease.

The PIH - originally stated by Milton Friedman in 1957 - is a well-known hypothesis in consumer behavior analysis. The PIH states that individual's smooth consumption over their lifetime according to their permanent income (see Friedman, 1957). However, do individuals behave rationally according to this hypothesis? This is what we examine along the first two chapters of this thesis, using as local windfalls the winnings from the first prize of the Spanish Christmas Lottery in the first chapter, and the Millions and the £30.000 British Postcode Lottery in the second one. In this case, the first two chapters complement to each other, given the different size of the local average treatment effects coming from the different types of income shocks we analyze in each chapter. In the Spanish Christmas lottery scenario, the shock can be considered a universal one, as almost the whole country plays it (according to Garvía (2007), around 75% of the Spaniards play this specific lottery); therefore, the estimated elasticities will be considering a much larger population size compared to the Postcode Lottery scenario, as not everyone plays this lottery (only 14% of the UK population played it in 2019). Hence, in Chapter 2 we consider everyone that lives in the winning region and bought lottery as potential winners of the lottery; whereas, in Chapter 3, we consider only those households that live in winning postcodes (which on average there are 15 households per postcode in the UK) - therefore, the population treated in each scenario is significantly different. What is also relevant from Chapter 3, and makes a difference with Chapter 2, is that we are analyzing two different shocks: a small one from the £30,000 Postcode Lottery and a big one from the Millions one, which allows us to analyze and compare the effects of these two lottery prizes into household consumption behavior and test the PIH.

In the third chapter, we use Ramadan as our religious constraint to analyze how individuals that take part of this festivity behave in the labor market during the Holy month of Ramadan. The idea of this chapter is to analyze how culture and beliefs can affect human behaviors and economic outcomes, as religions impose some rules of behaviors and practices that might constraint individuals in different ways. However, are religion culture impositions orthogonal to the economic outcomes or not? If the second option holds, then we need to understand why this is happening, and this is what we will be analyzing in Chapter 4 as well. Based on that, the literature finds that Ramadan has negative implications on economic indicators, as it negatively correlates with economic growth, productivity and labor markets (see Campante and Yanagizawa-Drott, 2015 and Barro and McCleary, 2003). Specifically, Campante and Yanagizawa-Drott (2015) tests the impact of religious practices on macroeconomic indicators already mentioned and at the aggregate level for labor supply. However, the literature does not investigate the effects of cultural constraints on labor supply at the individual level. In Chapter 4, we test such effects on labor supply at the individual level and how households adjust the intensive margins of labor when they need to adjust to cultural constraints, by performing an intrahousehold analysis on labor supply indicators. Therefore, we want to test if religious constrains impose causality to individual labor supply decisions or not.

Exploring previous research in consumer behavior, we find that, from a theoretical point of view, windfall gains lead households to anticipate the purchases of durable goods before these get obsolete (see Grossman and Laroque, 1990). Those households that do not experience a windfall will renew their durable goods only when these get obsolete, or they sell the good to purchase a new one. Another important theoretical result in the literature is that households that face a one-off large income shock and are credit-constrained, use the money from the shock to increase their durable goods consumption and thus, buy items that before they were not able to do (see Cerletti and Pijoan-Mas, 2014). Moreover, the income shock allows households to alleviate their constraints by paying back any debts they may have. According to Cerletti and Pijoan-Mas (2014), non-durable goods smooth consumption under this scenario. These theoretical findings lead to a violation of the Permanent Income Hypothesis, as households do not smooth consumption for durable goods. These findings help us to explain, from a theoretical point of view, our empirical results and what we should expect from Spanish and British household's consumption behavior when they experience a windfall. Moreover, from the empirical point of view, Kuhn et al. (2011) find that households that win the Dutch Postcode Lottery do significantly increase their consumption allocation for durable goods, whereas they smooth consumption for non-durable ones; satisfying the theoretical predictions by Cerletti and Pijoan-Mas (2014).

Related to the third chapter, research in the area of religious constraints affecting economic outcomes finds that individual productivity significantly decreases in countries that celebrate the Holy month of Ramadan due to the low caloric intake during the day (see Schofield, 2014) and, as mentioned earlier, Ramadan also has negative implications in economic growth and labor markets (see Campante and Yanagizawa-Drott, 2015). Hence, there might be a general belief that during Ramadan people work less due to the productivity decrease during this period in countries that take part of this festivity. However, do individuals work less during Ramadan and are less productive? This might not be true, as during this period individuals spend time preparing activities for family and friends; therefore, they are still productive but working more for household activities. Thus, this is what we analyze in the third chapter.

The research questions we try to answer along the thesis are mainly two: (i) how lottery winnings affect household consumption expenditures of durable and non-durable goods in winning areas compared with non-winning areas, and (ii) how religious constraints affect individuals' time-allocation in the labor market in developing countries. The first question is answered along the first two chapters, where the winning area refers to winning regions of the Spanish Christmas Lottery in the first chapter, and winning British postcodes of the Millions and £30.000 Postcode Lottery in the second one. As mentioned above, these are two completely different types of shocks given the characteristics of each lottery and the amount of people playing it. The second question is the main focus of the third chapter. Thus, the underlying hypotheses of the thesis are that (i) households in winning areas do not alter their consumption and, thus, do not spend the lottery winnings immediately, compared with households living in non-winning areas, and (ii) individuals that celebrate Ramadan do not reduce their own productivity or work less hours during this period.

To answer these questions, we use data coming from different sources. For the first chapter, we use data from the Spanish Household Survey available in the Spanish Statistics Institute. To proceed with the estimations of the second chapter, we use the Understanding Society data, provided by the UK Data Service. Finally, for the third chapter, we use three different sources of data. The first one is the Malawian Integrated Household Panel Survey provided by the National Statistics Office (NSO), the other two sources are the Household Income and Expenditures Survey (HiES) and the Women's Life Choices and Attitudes Survey (WiLCAS) from Bangladesh. All datasets provide information about the heads of the households and household characteristics: age, gender, education, employment status, hours worked during the week, number of children and household size. Along the chapters we describe these datasets in detail.

To develop the analysis of our proposed research questions, we use two main econometric methodologies. In the first two chapters, we use the lottery income shock from each analyzed lottery as an instrumental variable for total household expenditures, to estimate the Engel curves for household consumption expenditures of different types of goods. In this case, given that the income shock provided by the lottery winnings is completely random, the exogeneity condition for instrumental variables is automatically satisfied. Later we need to test for the relevance condition. In this case, we observe that the relevance condition is also satisfied and, thus, the lottery income shock works as a good instrument for total household expenditures.

In the third chapter, we proceed with the estimation of the Heckman model with fixed effects following the method proposed by Professor Jeffrey Wooldridge (see Wooldridge, 1995). The reason for applying such estimation method is to solve for the problem of sample selection bias, as there might be individuals that might not work because of personal reasons. Therefore, we use labor force participation as our selecting variable, and the factors that may have an influence on it are: the area of land a household owns over the number of adults in that given household (see Julien et al., 2019), as well as its interaction with gender, age and marital status categories. We believe that in developing economies, the more adults that live in a household, the less likely an individual is to work in the household land and thus, he/she is more likely to look for a job outside home (see Nagler and Naudeé, 2014). Moreover, we also believe that individual characteristics together with the area of land owned over number of adults in a household might impact the decision of being part of the labor force as well. In this chapter we check whether our beliefs are true, by testing the relevance condition of the instrumental variables on labor force participation, and the exogeneity condition of the instruments.

The main finding for the first two chapters is that households living in winning areas of the lottery significantly increase consumption expenditures on durable and non-durable goods in comparison with non-winning areas. These results imply a violation of the Permanent Income Hypothesis in Spain and UK. Thus, when Spanish and British households experience a windfall effect, they spend money from the lottery winnings, instead of saving it for future periods. Moreover, we find that the proposed theoretical results by Cerletti and Pijoan-Mas (2014) are not satisfied either. However, the most interesting result in both chapters is that the estimated elasticities of household consumption of durable and non-durable goods to total household expenditures are similar. In the third chapter, we do not find any evidence at the aggregate level that Ramadan has an impact on labor supply indicators; in other words, we do not observe that individuals that celebrate the Ramadan festivity increase/decrease their productivity during this period. However, when we analyze the effects at the individual level, we find that women work less in their jobs, but they work more at home. Whereas men keep working more at their jobs. We discuss these findings in more detail along the results sections of these chapters.

Thus, the novelty of this thesis in the economics literature is that we use two different lottery income shocks to estimates income-elasticities on household consumption and analyze what people do with windfalls. We find similar reactions in household consumption expenditures on durable and non-durable goods to a shock to total expenditures, under the proposed scenario for the Spanish and the British households, something that, to our best knowledge, was never found before. Moreover, we also contribute with an individual analysis in the Malawian labor market during Ramadan, where contrary to general beliefs, our findings show that individuals do not work less: women redistribute their working hours by working more at home and less at their jobs, and men do not reduce their number of hours worked at their jobs during Ramadan.

As mentioned before, this thesis is structured in three chapters. The first one analyzes how household consumption expenditures of specific goods are affected in those regions awarded with the first prize of the Spanish Christmas Lottery. The second chapter focuses on the British People's Postcode Lottery, where we try to analyze how the lottery winnings affect consumption expenditures for those households that are located in winning postcodes of the Millions and the £30.000 Postcode Lottery. Finally, the third chapter focuses on how religious constraints, like Ramadan, affect individuals labor supply composition.

Chapter 2

Consumption Responses to Income Shocks: Evidence from the Spanish Income Lottery

Abstract

Using the earnings from the first prize of the Spanish Christmas Lottery (one of the most popular lottery games in Spain), we study how local windfall gains from the Christmas Lottery can affect household consumption behavior. After estimating the direct impact of the lottery winnings on household consumption for durable and non-durable goods, we find that households living in winning regions significantly increase consumption expenditures on these goods in the winning regions in comparison with the non-winning ones. These results lead also to a violation of the Permanent Income Hypothesis. Moreover, we estimate the Engel curves, and we find that the consumption of durable goods is sensitive to lottery winnings. But, contrary to our expectations, we find that the consumption of non-durable goods is unit-elastic to a shock to total expenditures, which means that non-durable goods consumption also reacts as a result of the lottery winnings. These findings are not in line with the theoretical predictions in Cerletti and Pijoan-Mas (2014), where non-durable goods should be inelastic to the lottery income shock.

Keywords: Consumption, Durable goods, Non-durable goods, Lottery, PIH, Winning Region
JEL Classification: C01, C23, C26, C55, C93, D12 and D14.

1 Introduction

Over recent years research into consumer behavior has focused on how different income shocks affect the behavior of individual agents (see Berniell, 2016; Hsieh, 2003; Kuhn et al., 2011). Based on the theoretical economic predictions, agents should smooth consumption by spending their average income in every period. This would imply that if agents experience a positive income shock, such as a windfall, it should act as a buffer-stock: they smooth consumption and save money for future periods, when income might be lower (see Adamopoulou and Zizza, 2017; Berger et al., 2015). This prediction is known as the Permanent Income Hypothesis (PIH), first proposed by Milton Friedman in 1957 (see Friedman, 1957). In 1978, Robert Hall tested this hypothesis; the main finding was that if agents' consumption is based on the information individuals have at the time of making the decision, past income and consumption decisions in previous periods should not have an influence on current consumption decisions. However, we should distinguish between positive and negative transitory shocks in income. Making such a distinction is important, as agents can behave differently under the varying scenarios that a given economy can present.

This chapter focuses on the effects of windfall gains in income in relation to household consumption expenditures. We consider the exogenous variation in income in local areas that result from payments made by the Spanish Christmas Lottery. More precisely, we focus on how the winnings from this lottery can affect household consumption behavior. However, there are some facts about the Spanish Christmas Lottery to be taken into consideration in this study. Firstly, the prize offers a large shock in income that creates a significant impact on the local economy of the winning region. On average, this shock increases the GDP of the winning province by 3.5%. This factor implies that winning regions are richer and, thus, have more disposable income with which to increase consumption. Secondly, the prize does not belong solely to one person - it is shared among all those individuals who bought the same lottery number,[1] therefore, the shock affects more than one household, making the analysis more heterogeneous. Thirdly, around 75% of the Spanish population enters this lottery, implying that ordinary citizens play it; this, therefore, alleviates potential disturbances created by gamblers' behavior (see Bermejo et al., 2019). Finally, winners are clustered. This is because the whole series of a lottery ticket is sold in (almost) one lottery outlet, making it easier to locate who the potential winners are and to analyze how the lottery winning regions behave after experiencing the income shock.

The Spanish Christmas Lottery is a quasi-experiment conducted every Christmas in Spain, where the first prize awards the winner with a total of 400,000€ for each ticket bought. The random nature of the prize allocation means that winners experience an increase in income that they were likely not to be expecting, as the chance of winning the lottery is 1 over 100,000. Therefore, the exogeneity and the size of the first prize provide a good source of research material with which to investigate how households in winning regions react to a win: in other words, whether households increase savings and postpone consumption after experiencing the shock, satisfying the PIH, or, on the contrary, whether they spend the money from the lottery winnings. Hence, the research question we attempt to answer in this chapter is: how the Spanish Christmas Lottery income shock affects household consumption behavior in the Christmas Lottery winning regions, in comparison with the non-winning ones. Our underlying hypothesis is that given the average size of the shock,

[1]This is a syndicate game, where most of individuals share tickets with friends, colleagues or relatives.

the PIH holds, and households use the lottery prize to increase savings and use this money for future consumption.

Knowing the particularities of the Spanish Christmas Lottery and the randomness of the shock, we use information about both household consumption expenditures on different goods and the Christmas Lottery winning regions to identify the effect of a positive income shock on various categories of consumption expenditures. Using a two stage fixed effects estimator, we find that the windfall effect caused by receipt of the first prize in the Spanish Christmas Lottery has a significant impact on households' consumption behavior. In the first instance, we find that the effect of the lottery income shock has a positive and statistically significant impact on total household expenditures. This implies that it works as a good instrument for total household expenditures when we estimate the Engel curves in the second stage regression, as relevance and exogeneity property for instrumental variables is satisfied.

When examining the direct effect of the lottery winnings on household consumption, we observe that households increase their consumption in durable and non-durable goods due to the lottery winnings. In this case, these estimates capture the average lottery earnings per region on household consumption. In this instance, the average size of the income shock varies with the years, depending on the awarded region. Therefore, we need to consider the population in each region in every period of our analysis, as the awarded region changes every year and, thus, it will do the average earnings per household.[2] We observe in Table 11 that the average lottery earnings each year are 387€, the year with the lowest average and 2263€, the year with the highest average. Thus, we should expect the PIH to be feasible due to the lottery income shock.

Moreover, when analyzing the Engel curves, there is evidence that those households that live in the Christmas Lottery winning regions increase their consumption of non-durable and durable goods. However, the estimated effect for the aggregate of non-durable goods is unit-elastic and the effect for durable goods is elastic to total household expenditures. Specifically, we find that increasing total expenditures by 10% leads to an increase of 11.47% in household consumption of durable goods, whereas a 10% increase in total expenditures causes a 9.26% increase in household consumption of non-durable goods. Such findings imply that a positive shock to total expenditures (i.e., the lottery income shock) has a similar impact on household consumption expenditures for

[2]The average lottery earnings per household is equal to 160 series of each number times 10 tickets that each series contains, times the earnings from the first prize after tax: 320,000€, over the household size in the winning region:

Average earnings $= \frac{160 \times 10 \times 320{,}000}{\text{Total households per region}}$.

durable and non-durable goods. Thus, these results are not in line with the theoretical predictions made by Cerletti and Pijoan-Mas (2014) and Browning and Crossley (2009), where we should be expecting that households use the lottery winnings to increase their purchases of durable goods only.

This chapter belongs to a growing literature exploring consumer behavior and changes in household behavior when an income shock is experienced. Although this is not the first project to use the Spanish Christmas Lottery as an exogenous shock that affects agents' behavior (see Bagués and Esteve-Volart, 2016; Bermejo et al., 2019), it is the first one that uses such an income shock in order to analyze behavior in terms of household consumption expenditures in those regions that are awarded the first prize in the Christmas Lottery. As Garvía (2007) states, this is a particular type of lottery, as it is a social game rather than a gamblers' game. This means that the majority of the Spanish population participates in the game, and implies that most of the Spanish population is, therefore, eligible to be part of the treatment group and experience the lottery shock.

Theoretical findings in the literature about the behavior of agents regarding the consumption of durable and non-durable goods when they face a positive income shock, indicate a violation of the PIH. In this case, agents do spend money from the income shock on durable goods, meaning that these are sensitive to income shocks and respond significantly when there is an unexpected increase in income (see Cerletti and Pijoan-Mas, 2014). Moreover, the fact of receiving a one-time positive income shock leads agents to anticipate the purchase of durable goods rather than waiting until the item becomes completely obsolete (see Grossman and Laroque, 1990). The results from these two papers indicate that, on the other hand, the Life-Cycle Hypothesis (LCH) is satisfied. [3]

Related to this topic, other theoretical studies find that when income shocks are positive, individuals tend to become more impatient regarding their consumption and, rather than saving for future consumption, they prefer to consume more in the current period (see Haushofer et al., 2013). In contrast, if the shock happens to rich households, we should not expect to see many changes in their consumption expenditures (see Fagereng et al., 2016). In spite of these findings, changes in income lead to strong responses in consumption and play an important role in household decisions (see Krueger and Perri, 2010). Thus, when households experience positive income shocks, they have a larger propensity to consume; however, when the future is uncertain individuals tend to

[3] The Life-Cycle Hypothesis (LCH) concept was first introduced by Franco Modigliani in 1954. This hypothesis states that individuals smooth their consumption over their lifetime, planning their earnings along their life - borrowing during periods of low income and saving along times of high income (see Deaton, 2005). However, this chapter focuses on the PIH only.

save and postpone their consumption (see Kaplan and Violante, 2014).

Focusing on empirical research based on lottery prizes, very little has been investigated on how the impact of a lottery prize affects households' consumption behavior - only studied for countries like the Netherlands (see Kuhn et al., 2011), Norway (see Fagereng et al., 2016), United Kingdom (see Chapter 2) and Alaska (see Hsieh, 2003). As in this chapter, previous papers use a difference-in-difference method to compare household decisions in winning places as compared to households located in non-winning places. In countries such as the Netherlands, the main findings suggest that households spend the money from lottery winnings on durable goods, especially cars (see Kuhn et al., 2011).

This chapter performs a similar analysis to the papers described above. However, we differ from other studies by analyzing the effect of winning the Christmas lottery on household consumption expenditures at the region level, differentiating between households that live in winning regions compared with households that live in non-winning ones. Moreover, we use this income shock to estimate the Engel Curves, where given the characteristics of the Spanish Christmas Lottery explained above, the income shock can be considered a universal one since most of the population plays is and, thus, the estimated elasticities will be considering a large population size. Furthermore, the obtained results in this chapter are important, as contrary to what the literature finds, we encounter that household consumption expenditures for durable and non-durable goods react similarly to the lottery winnings.

The chapter is structured in six further sections: Section 2 and Section 3 are based on descriptive information about the lottery procedure and data description; Section 4 describes the identification strategy; Section 5 is based on the model we aim to study and the methods used to estimate it; in Section 6 we present the estimated results; and, finally, Section 7 concludes the chapter.

2 Background: The Spanish Christmas Lottery

The Spanish Christmas Lottery (*Lotería de Navidad*) is a national lottery game organized by the National Lottery Organization; the raffle takes place every year on December 22, and it has been played since 1812. This lottery is the biggest such event worldwide. It covers one fifth of the total lottery sales in Spain (see Bermejo et al., 2019). The Christmas Lottery is not a common type of lottery, such as where one buys a ticket, waits until the raffle occurs and only a few people participate in the game. In this case, around 75% (and increasing) of the population of Spain participate in it and one particular characteristic of this game is that ticket purchase is shared between friends, family or work-colleagues (see Bagués and Esteve-Volart, 2016). Therefore, we can assert that this lottery can be considered to be a social network event rather than a gamblers' event (see Garvía, 2007). In furtherance of this argument, we can note that most of the Christmas Lottery players buy tickets only for this lottery and do not play other lotteries held in Spain.

Each lottery ticket costs 20€, and the whole series (10 tickets) costs 200€. There are also shares and participations that cost between 2€ and 5€, and normally 1€ goes to charity. In recent years the amount of money spent by Spain's population on the Spanish Christmas Lottery is, on average, 64€ per person.[4] According to a survey run by the Center for Sociological Research, individuals planned to spend between 40€ and 60€ in 2004, and only around 8% of the sample population planned to spend more than 150€.

All lottery tickets have five-digit numbers. Since 2011 a total of 100,000 numbers are played in the Christmas Lottery draw,[5] ranging from 00,000 to 99,999. Each ticket number is split into 160 series, and each of these series consist of 10 fractions, known as décimos. Each fraction can be further divided into shares or minor units, known as participaciones. Out of all these numbers, a total of 1,807 will receive a prize; however, the probability of winning the lottery is very small: more precisely there is a 0.001% chance of winning the first prize, known as *el Gordo* - which in English would be translated as 'the fatty',[6] - this is 24 times lower than the likelihood of being hit by a car, as Professor E. Nualart[7] stated in *La Vanguardia*.. Table 1 shows how prizes are distributed, the amount of money associated with each ticket bought and the proportion one gets per each euro invested: Apart from these prizes, if someone's ticket contains the last numbers of 'the fatty', that individual also gets an amount of money per euro invested in return. Therefore, in total, 70% of the revenue is dedicated to prizes and the remaining 30% represents the commission paid to the outlets that sell the lottery tickets. In addition, those prizes that are above 2,500€ are taxed at the rate of 20% of the total amount, which goes to the Treasury. This tax was introduced in 2013. Before the tax was instituted, the amount won for the first prize was 300,000€ per ticket; therefore, for those who win the lottery after 2013, even if they have to pay tax, they still earn more than before: 320,000€ after tax per ticket, which is approximately 12 times higher than the

[4] Source: *El Economista*

[5] Until 2004, only 66,000 numbers were played, and between 2005 and 2010 this number increased to 85,000.

[6] Term that Bagués and Esteve-Volart (2016) used in their paper to refer to the first prize of the Christmas Lottery.

[7] Universitat Pompeu Fabra, Department of Economics.

Table 1: Distribution of the Lottery Prizes

Prize	**Numbers awarded**	**Amount won per ticket**	**Proportion**
First prize (*the "fatty"*)	1	400,000€	20,000€ per euro
Second prize	1	125,000€	6,250€ per euro
Third prize	1	50,000€	2,500€ per euro
Fourth prize	2	20,000€	1,000€ per euro
Fifth prize	8	6,000€	3,000€ per euro
Pedrea	1,974	100€	5€ per euro

Source: ABC - Lottery Prizes.

average household income in Spain (26,092€).[8]

Another characteristic of this lottery is that winners are clustered, with just a few towns winning the lottery every year; sometimes only one city is awarded as the winner. This is because each lottery outlet has been randomly assigned the numbers it has to sell. This means that the winner is more visible and it is easier to check whether those winning regions see a change in consumption behavior or not. One of the reasons that explains this factor is that the Christmas Lottery is a syndicate game, in other words, people who are in the same network want to play the same number (see Bagués and Esteve-Volart, 2016).

Lastly, there are two exceptional outlets where the Christmas Lottery is sold. The first one is a very famous outlet in Madrid, called *doña Manolita*. The second outlet is located in Sort (which means 'luck' in Catalan), a town that belongs to the province of Lleida; this outlet is known as *La Bruixa d'Or* (translated from Catalan as: 'the gold witch'). These two outlets are very famous for selling numbers that, over several years, have been awarded high prizes and, thus, for superstitious reasons, many people travel there from all over the country to buy tickets.

[8] *Source:* INE, 2014.

3 Data

The main data source we use in this chapter comes from the *Encuesta de Presupuestos Familiares* (EPF),[9] provided by the *Instituto Nacional de Estadística* (INE).[10] This is the Spanish Family Income Survey, in which households are randomly selected to take part. The sample is composed of 22,346 households covering the years from 1998 to 2016, where surveys in each wave can take place in any given period of the year. Data is presented in the form of panel data from the year 1998 up to 2005. After 2005 the INE changed the data collection process, and the institute began to present it in the form of cross-section data. The survey includes information about household income and expenditures, household characteristics, demographic variables for the head of the household (age, gender, education, marital status, etc.), employment status of the head of the household (whether he/she is employed, number of hours worked, type of contract, etc.), among other variables of interest. In this survey, household income is given monthly[11] and household expenditures are given yearly; therefore, expenditures are modified to a monthly variable. This implies that we need to assume that *households spend the same amount in each month of the year.*

This survey also provides information about the region that households live in - the sample takes into account individuals from all the 19 Spanish regions (including Ceuta and Melilla); however, the data presents an important drawback, as it does not provide information about which households won the lottery, and given that we do not know the city in which households live, it is hard to discern with high precision who may be the winners of the Christmas Lottery. Despite this handicap in the data, we can perform an analysis at the regional level on household consumption behavior and test how households that live in winning regions of the first prize of the lottery behave, in comparison with households living in non-winning regions in respect of the Spanish Christmas Lottery and, thus, have a general idea of the average consumption pattern between regions. In Table 12 in the appendix, we show the proportion of people interviewed in each region and we also provide an estimation of winners in each region - we find that all regions are represented in the sample, but those that are more populated have a higher household representation, like Andalucía, Cataluña or Madrid. When looking at the proportion of winners in each region, we observe that regions with larger amounts of population are more likely to include winners (also because such regions have been awarded more times with the first prize of the lottery); however, there are some regions that have zero chances of having winners, as these have not won the lottery in the years

[9] Family Income Survey

[10] Spanish National Statistics Institute

[11] The survey documents do not specify whether this is the household income in the previous month, or simply the household income in any given month.

included in the survey.

On the other hand, the EPF collects information about which households are lottery players and which are not. This source of information allows us to identify those households who are lottery participants and, thus, we can identify which households may potentially be part of our treatment group and which may be part of the control group. The treatment group is composed of those households that both live in the winning region and also bought lottery tickets. The remainder of the sample belongs to the control group. A condition of being a winner in the Christmas Lottery is that at least one member of the household needs to buy a lottery ticket. If this condition is not satisfied, the household automatically belongs to the control group.

Table 2: Summary statistics

Variable	**Mean**	**Std. Dev.**	**N**
Lottery $\times$ win_region	0.072	0.258	305550
win_region	0.152	0.36	305550
Lottery	0.433	0.496	305550
Lottery expenditure	91.869	36.16	273098
Age	31.165	28.394	239434
Marital_status	1.319	1.055	278490
Single	0.426	0.495	305550
Married	0.468	0.499	305550
Education	3.514	1.629	278490
Employed	0.424	0.494	278490
Retired	0.337	0.473	305550

Source: Instituto Nacional de Estadística (INE)

Table 2 provides a descriptive summary statistics of the sample, where 15% of the interviewed households live in a region that won 'the fatty' and, out of those, only 7% bought a lottery ticket and may potentially be winners of the Christmas Lottery. The average age of our sample is 31 years old, which implies that the heads of the households, on average, are relatively young, with a standard deviation of 28 years. In addition, the majority of our sample is either single or married - 43% are single and 47% are married; the remaining are widowed (6%) or separated/divorced (4%).

Only 42% of the sample is employed and we find that 33% of the heads of the households are retired. Moreover, we find that, regarding educational level, most of them lie between categories three and four, which means that they hold at least a secondary school degree. Finally, we observe that the average expenditure per capita in the lottery is 91.87€.

Table 3: Summary statistics: consumption levels for the different type of goods

	Winning Regions		Non-Winning Regions		Testing Differences	
Variable	Mean	Standard Deviation	Mean	Standard Deviation	$t-$test difference	$p-$value
Total Expenditures	6142.035	9116.764	7466.984	8062.632	-20.93	0.000
Durables	2600.31	3999.109	3130.787	3574.774	-19.09	0.000
Non_durables	3541.725	5186.04	4336.198	4608.617	-22.05	0.000
Savings	322.703	616.366	293.014	593.813	6.90	0.000
Food_in	494.232	682.966	756.257	705.810	-54.68	0.000
Alcohol and Tobacco	413.992	668.994	504.907	704.275	-19.34	0.000
Clothes	443.9	680.967	581.29	719.097	-28.71	0.000
House_rent	483.164	690.793	751.442	716.39	-55.33	0.000
House_durables	466.88	691.939	681.198	740.17	-44.02	0.000
Health	419.028	672.042	500.366	703.422	-17.23	0.000
Car Value	203.607	511.858	144.905	436.296	16.55	0.000
Transport	457.271	686.601	615.758	727.598	-32.83	0.000
Communication	463.125	693.414	675.36	747.225	-43.47	0.000
Gambling	471.313	695.78	221.525	529.395	52.08	0.000
Leisure	446.935	683.507	585.001	725.357	-28.73	0.000
Education	281.067	585.446	254.655	562.258	6.46	0.000
Food_out	459.041	686.111	608.818	721.418	-31.07	0.000
Holidays	315.778	611.247	292.49	591.395	5.46	0.000

Source: Instituto Nacional de Estadística (INE). Values presented in Euros.

Table 3 presents a summary statistics of household consumption expenditures by winning and non-winning regions. Performing such differentiation allows us to check for potential differences in household consumption behavior across treatment and control groups. To test for such differences, we perform a t-test under the null hypothesis that: *household consumption expenditures do not differ across winning and non-winning regions of the Spanish Christmas Lottery.* In this case we are testing the average effect of the lottery income shock, which is equal to the number of series and tickets in each series times the awarded prize, over the number of households in the winning region. In this case, Table 11 in the Append provides the average earnings per year. We observe that the average lottery earnings are between 387€, the year with the lowest average, and 2263€, the year with the highest average. Therefore, given the exogeneity of the lottery shock and the random selection of the interviewed households, we should not expect significant changes in consumption expenditures across regions under this scenario due to the lottery winnings.

In addition, we also perform an exogeneity test of the treatment group. In order to test this, we run a linear probability model, using as dependent variable a binary variable that indicates whether the household is located in a winning region or not, on all individual characteristics and control variables that will be included later in our regression analysis in Section 5. After performing the F-test, which gives the result 0.66, we fail to reject the null hypothesis that all estimated regressors are equal to zero, meaning that our identifying assumption (described later in Section 4) holds (see Table 13).

We observe from Table 3 that, in general, Christmas Lottery winning regions spend less than non-winning ones. When looking at the aggregates of durable and non-durable goods, households in winning regions spend 531€ less on durable goods, 794€ less on non-durable goods and 1325€ less on total household expenditures. These differences are statistically significant at the 1% significance level by looking at the performed t-tests, implying that there are differences in household consumption behavior across regions. On the other hand, households in winning regions save more than those that live in non-winning regions. Specifically, households in winning regions save 30€ more per month, and despite this being a small difference, we find it to be statistically significant at the 1% significance level. In general, households that live in winning regions do not consume more than those that live in non-winning regions. There are only a few categories in which there are positive differences toward households in the treatment group: namely, car value, education and holidays. Despite the differences being small, the t-test shows that such differences are statistically significant.

Table 4: Lottery Expenditures

Year	GDP	GDP pc	Lottery exp. per capita	% of lottery exp. to GDP pc
1998	118386400€	140236.8€	39.87€	0.29%
1999	121493500€	142041.3€	41.83€	0.30%
2000	125689700€	145465.5€	42.42€	0.29%
2001	130972800€	149243.1€	44.19€	0.30%
2002	136616500€	153629.1€	47.76€	0.31%
2003	142270900€	157438.8€	50.09€	0.32%
2004	147994900€	161861.8€	50.88€	0.32%
2005	154340900€	165057.9€	53.13€	0.33%
2006	160380700€	169275.2€	53.56€	0.32%
2007	165626800€	173197.8€	54.55€	0.32%
2008	171188800€	175453.3€	51.89€	0.30%
2009	162610600€	164495.8€	50.37€	0.31%
2010	162272700€	164185.9€	49.82€	0.31%
2011	162326500€	164724.6€	49.27€	0.30%

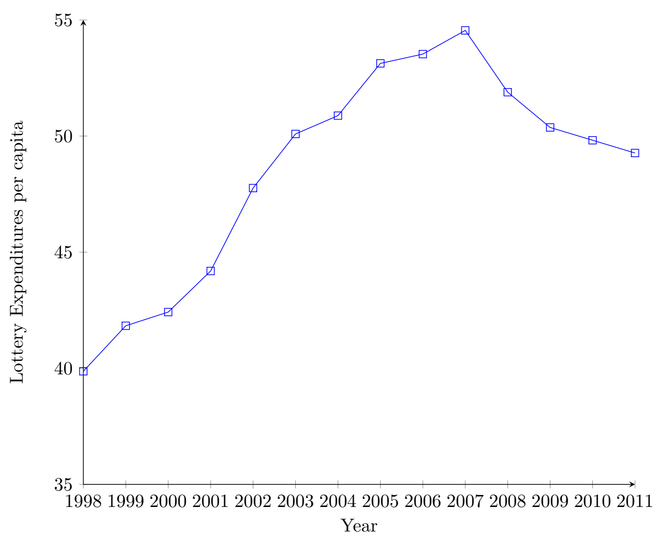
Source: Research Section - Prof. Manuel Bagues. *% of lottery exp. to GDP* shows the amount spent on the Christmas Lottery relative to GDP per capita and *Lottery exp. per capita* represents the average expenditures in the Christmas Lottery by the Spanish population.

Therefore, from Table 3, we observe that households that live in winning regions behave differently than those that live in non-winning regions. A reason that can explain these negative differences are that winning regions are likely to be poorer than non-winning regions. However, when we look at the standard deviation, we observe that for the aggregates (durable and non-durable goods and total expenditures), it is significantly higher in winning regions. This means that consumption variability is greater for those households that belong to the treatment group. However, the differences presented in Table 3 do not capture the elasticity of total household expenditures on the household consumption expenditures of specific goods, because only a small number of households that live in winning regions are likely to be potential winners of the Spanish Christmas Lottery. Therefore, we need to estimate a more sophisticated model of consumption, in which we evaluate the changes in household total expenditures on demand in different consumption categories. This is explained further and in more detail in Section 5.

Apart from the EPF, we also use regional and national data on Christmas Lottery expenditures, available in the *Sociedad Estatatal de Loterías y Apuestas del Estado* (SELAE). In addition, we use the national GDP and GDP per capita, both sources available in the INE, to measure the average lottery expenditure per year, relative to the Spanish GDP. Table 4 presents a summary statistics for all these variables of interest covering from 1998 to 2011. The reason our analysis extends only to 2011 is that after this year the data on Christmas Lottery average national expenditure differs across several sources and it is hard to find a general consensus.

From Table 4, we observe that the Christmas Lottery expenditure relative to GDP is equal to 0.3%. This fact has been stable not only throughout the years under analysis, but also in the previous two decades (see Bagués and Esteve-Volart, 2016). However, when we look at lottery expenditures in levels, this increased through the years until 2006, when it became stagnant and fell during the following years - coinciding with the economic recession in 2008. After these years lottery consumption was, on average, around 50€, but started to increase again. In 2018 consumption rose to 67.58€, according to the SELAE. We can observe this trend more clearly in Figure 1.

Figure 1: Average Spanish Christmas Lottery Expenditures



4 Identification Strategy

The identification strategy is based on the idea that winning the Spanish Christmas Lottery is akin to a random income shock. However, there are two caveats to be noted with this approach: (i) only households that participate in the Christmas Lottery can experience such a shock; and (ii) in our database, we do not observe winning households, but only whether, or not, in a given year, a particular household lived in a winning region - in other words, whether they lived in a region that had lottery winners.

We can assert that, in any given year, households in winning regions that had purchased a Christmas Lottery ticket have a non-zero probability of having won; all other households in that year have zero probability of having won the lottery and, thus, belong automatically to the control group. Therefore, we create an interaction term involving the binary variables *lottery* (whether a household had purchased a lottery ticket or not) and *win_region* (whether the household lives in a winning region of the Christmas Lottery or not) as an instrument for total household expenditures, as well as the *win_region* variable per se.

Households that purchase the Christmas Lottery ticket are likely to have different characteristics from those that do not, and winning regions may have systematically different characteristics from those regions that did not win (e.g. they may be more densely populated, have individuals who are more likely to purchase lottery tickets, etc.). Therefore, we need to control for region fixed-effects and year fixed-effects in our specifications. Thus, the interaction term is picking up, in a specific year, the difference in household consumption expenditures between households that play the Christmas Lottery and those that do not, differencing across regions that won the lottery versus regions that did not, after controlling for region fixed-effects and year fixed-effects.

Moreover, we need to keep the assumption stated by Bermejo et al. (2019), which applies to this chapter as well, that: *the winning province is randomly assigned conditional on expenditures on lottery tickets by province.*

Our identifying assumption is that this difference-in-difference effect on household consumption expenditures is due to lottery winnings rather than to region-year shocks correlated with the selection of the winning region in a given year. Because the selection of the winning region each year is random, there is no obvious reason why it would be correlated with other region-year shocks. Recall that the winning of the Spanish Christmas Lottery is an annual shock that takes place on every December 22.

5 Empirical Analysis

In this section, we investigate the existing relation between the lottery income shock and household consumption expenditures of specific categories of goods. We first run the reduced-form estimation, in which we analyze whether living in the winning region has a significant impact on households' consumption behavior. Secondly, we perform a robustness check on the reduced-form estimation, by including different head of the household characteristics as control variables into the regression, plus both region and year fixed-effects. We then proceed to estimate the Engel curves for the different categories of goods analyzed in this chapter. In order to do so, we proceed with an instrumental variable regression approach, to deal with potential endogeneity that may arise from total expenditures in the estimation of the Engel curves. In this case, we believe that total expenditures are endogenous because households can infer from the allocation of resources for their total consumption expenditures and, thus, they can decide how much to spend in each month. This implies that there may be some unobserved factors that will influence the expenditure decisions, and this can correlate with the error term of the regression.

In the first stage we test whether the fact of living in a winning region has an effect on the logarithm of total expenditures and whether the set of instruments is relevant or not. In the second stage, we test whether the logarithm of total household expenditures has an impact on household consumption behavior for specific categories of goods or not. The regressions used to estimate our outcomes of interest have been inspired by previous papers in the literature, in particular the Engel curves (see Banks et al, 1997; Browning and Collado, 2007).[12] We also extend our analysis by analyzing the effects of the lottery income shock on the labor market and on family composition.

[12]Some authors include past-time consumption in the analysis or other non-linearities, as the square logarithm of total expenditures (see Arellano et al., 2017; Blundell et al., 1993). However, in the case of our research, none of these effects are statistically significant, thus, we do not include them in the regression analysis.

5.1 Consumption Analysis

Starting with the simplest regression analysis, we test the effect of the random income shock caused by the Christmas Lottery on household consumption behavior. Our specification uses a difference-in-difference estimator that compares households' consumption expenditures for different types of goods in those regions awarded the first prize in the Christmas Lottery in comparison with other regions. The reduced-form regression is as follows:

$$\ln(c^g_{h,t}) = \beta_0 + \beta_1 win\_region_{h,t-1} + \beta_2 lottery_{h,t-1} + \beta_3 win_{h,t-1} \times lottery_{h,t-1} + u_{h,t} \quad (1)$$

where $c^g_{i,t}$ denotes household consumption expenditure of good g for household h in year t; $win\_region_{h,t-1}$ is a dummy variable that takes value one if the household lives in a winning region, and zero otherwise; $lottery_{h,t-1}$ is another dummy variable taking value one if household h participates in the Christmas Lottery in year $t-1$, and zero otherwise; $win_{h,t-1} \times lottery_{h,t-1}$ is the interaction term between the previous two dummies. The term $u_{h,t}$ represents the error term of the regression. Terms β_1 and β_3 are our coefficients of interest: the first represents the average difference in household consumption behavior between winning and non-winning regions; the second captures the average effect of those households that participate in the lottery and those that do not (considering those regions that won the lottery and those that did not) on household consumption expenditures of specific goods.

In addition to the reduced form in Equation (1), a robustness check is performed by adding household controls to the regression in order to check whether estimates are robust to the inclusion of household characteristics and fixed effects. The extended regression is as follows:

$$\begin{aligned} \ln(c^g_{h,t}) \quad = \quad & \beta_0 + \beta_1 win\_region_{h,t-1} + \beta_2 lottery_{h,t-1} + \beta_3 win_{h,t-1} \times lottery_{h,t-1} + X'_{h,t}\boldsymbol{\beta_4} \\ & + (gdp_{r,t}, \log(lot\_exp_{r,t-1}))'\boldsymbol{\beta_5} + \eta_{h,t} + \tau_t + u_{h,t} \end{aligned} \quad (2)$$

where, $X'_{h,t}$ is a vector of household characteristics including age, age square, education, marital status and employment status of the head of the household. We also include a vector of regional demographic characteristics: GDP per capita in each region represented by $gdp_{r,t}$, and the lottery expenditures per region, $\log(lot\_exp_{r,t-1})$. To ensure that unobserved effects are included despite the randomization of the treatment, we include $\eta_{h,t}$ as region fixed effect and τ_t as the year fixed effect.

This estimation process is itself interesting in investigating whether living in a winning region has an impact on household consumption behavior or not, as it allows us to test the PIH. However, this procedure does not capture the income-elasticity effect on household expenditures for different categories of goods. A more precise methodology for this is the estimation of the Engel curves. To achieve this estimate, we proceed with an instrumental variable estimation, in which we use the lottery income shock ($win_{h,t-1}$ and $win_{h,t-1} \times lottery_{h,t-1}$) as the set of instruments to estimate the effect of total expenditures on the consumption demand of different categories of goods. By taking this approach, we solve the endogeneity problem that arises from including total expenditures in the regression. Hence, the first stage regression is as follows:

$$\begin{aligned} \ln(exp_{h,t}) &= \beta_0 + (win\_region_{h,t-1}, lottery_{h,t-1}, win_{h,t-1} \times lottery_{h,t-1})'\boldsymbol{\beta_1} + X_{i,t}\boldsymbol{\beta_2} \\ &\quad +(gdp_{r,t}, lot\_exp_{r,t-1})'\boldsymbol{\beta_3} + \eta_{h,t} + \tau_t + \nu_{h,t} \end{aligned} \tag{3}$$

where the included coefficients are the same as those included in Equation (2). Once Equation (3) is estimated, we need to check that the relevance condition for the instruments holds. This can easily be checked by computing the F-test for instrumental variables.

Nonetheless, equations (1), (2) and (3) follow a two way fixed estimation (TWFE) with staggered treatment. This implies that the number of treated observations in each period of the survey is varying across years; in other words, the size of the treatment changes along the periods and some households that are treated in period t might not be treated in period $t+1$ but can be treated again in period $t+2$, and, thus, the treatment effect can be heterogeneous, given the size change in each period. This is because the assignation of the treatment (winning lottery regions) is random every year and any Spanish region that takes part of the lottery game can be a winner one, leading to a violation of the constant treatment effect assumption (see de Chaisemartin and D'Haultfoeuille, 2020). [13] Therefore, when this happens, the TWFE estimates might be biased and/or inconsistent, as there might be some households that are part of the treatment and control group along our analysis, which implies that the treatment group will be heterogeneous (see de Chaisemartin and D'Haultfoeuille, 2020; Callaway and Sant'Anna, 2021; and Goodman-Bacon, 2021). If this were the case in our study, we should expect to have a downward bias from the TWFE estimates, given the negative weights that the TWFE method assigns to periods with larger amounts of treated households, or to households that are treated for several years - this is because of the difference in treatment sizes across periods (see de Chaisemartin and D'Haultfoeuille, 2020). [14] This is something to take into account later in our results, as our treatment is staggered and it can lead

[13] The constant treatment effect assumption imposes that the treatment effect should be constant across groups and over the years. This implies that there is a pre-treatment period, where none of the observations in the sample are treated, and a post-treatment period, where some individuals in the sample will be treated after that given year.

[14] According to Goodman-Bacon (2021), the TWFE method assigns a weighted average treatment effect to the TWFE difference-in-difference (TWFEDD) estimators that compares timing groups to each other. If the constant treatment effect assumption holds, the TWFEDD estimates should not be biased. But if it is violated, the variance

to Type-I and Type-II errors (see Baker et al., 2021).

Next, we explore the effect of total expenditures, instrumented with the lottery income shock, on household consumption expenditures for the different categories of goods analyzed in this chapter. Hence, the second stage regression is as follows:

$$\ln(c^{g}_{h,t}) = \gamma_0 + \gamma_1 \ln(exp_{h,t}) + (X_{h,t})'\boldsymbol{\gamma_2} + (gdp_{r,t}, lot\_exp_{r,t-1})'\boldsymbol{\gamma_3} + \eta_{h,t} + \tau_t + u_{i,t} \tag{4}$$

where the $\ln(exp_{h,t})$ is the logarithm of total expenditures, estimated previously in Equation (3). Our coefficient of interest is γ_1, because it captures the elasticity effect of total household expenditures on household consumption behavior. According to the theory, we should expect the estimates of γ_1 to be between -1 and 1, as these would report inelastic effects and, thus, household consumption expenditures would not react to a shock to total household expenditures.

5.2 Labor Supply

In this subsection we analyze the effect of the random income shock caused by the Christmas Lottery on the labor market. More precisely, we want to observe whether living in a lottery winning region affects either the number of hours worked by households or their employment status (i.e., employed or not employed). The regression under this scenario is the following:

$$\begin{bmatrix} employed_{h,t} \\ num\_hours_{h,t} \end{bmatrix} = \beta_0 + \beta_1 win\_region_{h,t-1} + \beta_2 lottery_{h,t-1} + \beta_3 win_{h,t-1} * lottery_{h,t-1} + u_{h,t} \tag{5}$$

where $employed_{h,t}$ is a dummy variable that takes value one if the head of the household is employed, and zero otherwise; and $num\_hours_{h,t}$ represents the number of daily hours worked by the head of the household.

Equation (5) shows the reduced-form estimation for the labor market outcomes. As was the case for consumption, a robustness check is performed by adding household characteristics, and both region and time fixed effects as control variables to the regression. Thus, the extended regression

of the TWFEDD estimator might be incorrectly estimated and lead to inconsistent estimates, or the estimated treatment effect will be downward biased.

is as follows:

$$\begin{bmatrix} employed_{h,t} \\ num\_hours_{h,t} \end{bmatrix} = \beta_0 + \beta_1 win\_region_{h,t-1} + \beta_2 lottery_{h,t-1} + \beta_3 win_{h,t-1} * lottery_{h,t-1} + X'_{h,t}\boldsymbol{\beta_4}$$
$$+(gdp_{r,t}, lot\_exp_{r,t-1})'\boldsymbol{\beta_5} + \eta_{h,t} + \tau_t + u_{h,t} \tag{6}$$

In this case, the vector of individual controls, $X_{h,t}$, does not include the employment status in Equation (6), as this is one of our dependent variables.

5.3 Intergenerational analysis

When households experience a windfall, such as winning the lottery, their thinking about family composition tends to be affected and they increase the number of children, because they have more money or potential savings. In this subsection we analyze the effect of the income shock caused by the Christmas Lottery on family composition; in other words, we want to test whether living in a lottery winning region increases the number of children in the household or not. To test this idea, we estimate Equation (7) and Equation (8) using the forecast of two periods ahead for dependent children at home as our dependent variable:

$$child_{h,t+2} = \beta_0 + \beta_1 win\_region_{h,t-1} + \beta_2 lottery_{h,t-1} + \beta_3 win_{h,t-1} * lottery_{h,t-1} + u_{h,t} \tag{7}$$

where $child_{h,t+2}$ represents the number of children in the household, two periods after the lottery shock. Equation (7) presents the reduced-form estimation, whereas Equation (8) presents the extended regression, where the robustness check is performed by adding household characteristics, plus region and time fixed-effects as control variables to the regression, previously specified in subsection 5.1:

$$child_{i,t+2} = \beta_0 + \beta_1 win\_region_{h,t-1} + \beta_2 lottery_{h,t-1} + \beta_3 win_{h,t-1} * lottery_{h,t-1} + X'_{h,t}\boldsymbol{\beta_4}$$
$$+(gdp_{r,t}, lot\_exp_{r,t-1})'\boldsymbol{\beta_5} + \eta_{h,t} + \tau_t + u_{h,t} \tag{8}$$

6 Results

In this section, we present the estimated results for the different regressions introduced in Section 5. We are interested in examining the effect that living in the Christmas Lottery winning region has on household consumption expenditures of different types of goods; in other words, analyzing the effect of the lottery prize, not only on household consumption behavior, but also on labor supply and family composition.

6.1 Consumption Expenditures Estimation

Here we present the household consumption expenditure estimates for different types of goods. Beginning with presenting results for the reduced-form estimation, we later continue with the estimates of the Engel curves, under the instrumental variable approach.

Table 5 shows the results for the reduced-form regression in Equation (1). Observing, firstly, the interaction term estimates, $lottery \times win\_region$, we find a positive effect on the average consumption of durable and non-durable goods for those households that both live in a winning region and participated in the lottery, and this effect is statistically significant.

More precisely, performing a comparison across those households that potentially won the lottery and those that did not, we find that the estimated effect for these households implies an increase of 31.92% in the consumption of durable goods, and an increase of 18.53% in the consumption of non-durable goods, in comparison with those households that did not win the lottery.[15] Looking at specific goods in more detail, we find a statistically significant effect in all goods except for *alcohol and tobacco*, *clothes*, *transport* and *education*, where the effect of the interaction term is either close to zero or non-statistically significant. On the other hand, we find a negative and significant effect for *car value*, *food out*, *holidays* and *savings*. These results imply that households that live in winning regions and have entered the lottery save less, in comparison with those households that are not lottery participants and/or do not live in lottery winning regions. This means that lottery winnings are allocated to increasing consumption expenditures rather than to increasing savings. This might indicate that the PIH is not satisfied, as individuals do not keep their consumption allocation prior to the income shock. However, to test formally for that, we need to perform the F-test under Hypothesis 1.

[15] This effect comes from the log-linear regression model, which is computed by taking the exponential of the estimated coefficient of the interaction term. Thus, for households that live in Spanish Christmas Lottery winning regions and entered the lottery, the estimated effect is $(e^{0.277} - 1) \times 100 = 31.92\%$ for durable goods, and $(e^{0.170} - 1) \times 100 = 18.53\%$ for non-durable goods.

Table 5: Reduced-form estimation

PANEL A	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Food at home	Alcohol	Clothes	House Rent	House Investments	Health	Car Value	Transport
Win_region	-1.158***	-0.839***	-1.128***	-1.500***	-1.223***	-0.905***	-0.148***	-1.077***
	(-64.28)	(-38.42)	(-50.41)	(-67.63)	(-53.34)	(-41.81)	(-16.07)	(-46.92)
Lottery	-0.298***	0.378***	0.138***	-0.568***	-0.564***	0.296***	1.018***	0.0742***
	(-33.66)	(29.96)	(11.19)	(-49.87)	(-46.33)	(23.54)	(107.98)	(5.95)
Lottery×win_region	0.139***	-0.0393	0.0453	0.122***	0.311***	0.0990**	-0.182***	0.0399
	(5.60)	(-1.25)	(1.43)	(3.87)	(9.69)	(3.15)	(-8.68)	(1.24)
_cons	5.459***	2.931***	3.679***	5.032***	4.362***	2.942***	0.397***	3.789***
	(893.00)	(337.78)	(431.65)	(665.20)	(523.29)	(342.44)	(94.48)	(434.70)
F-test	3522.29	1493.00	2307.45	3817.68	1652.73	1249.17	306.78	2109.46
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	305550	305550	305550	305550	305550	305550	305550	305550
R-squared	0.0323	0.0120	0.0160	0.0396	0.0226	0.0113	0.0498	0.0140
PANEL B	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Communication	Leisure	Education	Food out Home	Holidays	Savings	Durables	Non-durables
win_region	-0.927***	-1.095***	-0.328***	-1.144***	-0.403***	-0.0317	-1.558***	-1.484***
	(-38.98)	(-48.99)	(-21.36)	(-50.26)	(-24.31)	(-1.70)	(-60.77)	(-66.24)
Lottery	-1.006***	0.164***	0.742***	0.291***	0.906***	0.856***	-0.486***	-0.275***
	(-72.56)	(13.27)	(66.10)	(23.28)	(77.36)	(74.85)	(-37.03)	(-24.39)
Lottery×win_region	0.569***	0.202***	0.0351	-0.0910**	-0.132***	-0.440***	0.277***	0.170***
	(16.83)	(6.45)	(1.35)	(-2.82)	(-4.83)	(-15.34)	(7.65)	(5.36)
_cons	3.879***	3.579***	1.091***	3.720***	1.343***	1.376***	5.771***	6.638***
	(417.97)	(413.48)	(164.83)	(426.32)	(191.85)	(199.13)	(653.33)	(880.27)
F-test	223.07	1643.64	196.87	2902.80	601.55	467.29	2501.13	3408.54
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	305550	305550	305550	305550	305550	305550	305550	305550
R-squared	0.0226	0.0136	0.0195	0.0188	0.0259	0.0201	0.0274	0.0314

Our dependent variables are the logarithm of consumption expenditures of each good category. The *win_region* coefficient reports the effect that living in a Spanish Christmas Lottery winning region has on household consumption expenditures for the different types of goods. *Lottery* estimates how the fact of participating (or not) in the Spanish Christmas Lottery affects household consumption behavior. Finally, *Lottery* × *win_region*, is the interaction term between the previous two variables. This coefficient captures the effect of a household that lives in a Christmas Lottery winning region and participates in it, on household consumption expenditures, in comparison with households that either live in other regions, or have not participated in the lottery, or both. In this specification, we only include the set of variables that belong to the lottery income shock; no other controls are included in this set. We compute robust standard errors. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The F-test performs a joint significant test of the lottery income shock variables (*win_region* and *lottery* × *win_region*), which later will be used as instrumental variables for total expenditures, on household consumption expenditures. The null hypothesis of the F-test is that: *living in the winning region of the Spanish Christmas Lottery has no effect on household consumption behavior.* In other words, this test is testing the validity of the PIH.

On the other hand, the coefficient of living in the winning region per se shows a negative estimate for the consumption of all goods, it being statistically significant for all goods as well. Contrary to the estimated effect for the interaction term, these estimates show that a household that lives in a Christmas Lottery winning region consumes less in comparison with a household that lives in a non-winning region, *irrespective of whether they bought a lottery ticket or not.* As explained earlier in Section 3, one of the potential reasons why these estimates are negative is because those households that live in poorer regions are more likely to purchase lottery tickets and, therefore, they are more likely to win. Thus, to capture whether there exists an effect from the lottery income shock on household consumption behavior, we need to look at the results of the F-tests in Table 5, performed under the following hypothesis:

Hypothesis 1:

H_0: Living in the Spanish Christmas Lottery winning region has no effect on household consumption behavior.
H_a: Living in the Spanish Christmas Lottery winning region has an effect on household consumption behavior.

The results from the F-test show that for those households that belong to the treatment group - in other words, that potentially win the lottery - there is a change in the household consumption expenditures of specific goods, in comparison with those households that live in non-winning regions or did not participate in the lottery. This effect is statistically significant at the level of 1% in all cases.

Therefore, such results might indicate a change in household consumer behavior in winning regions for those individuals that have entered the lottery; or that there, at least, exists a neighborhood-spread effect in consumption across neighbors in the same region, given that we do not know exactly which individuals have won the lottery and which have not.

In Table 6 we perform a robustness check from the reduced-form estimation, by adding household controls, jointly with both year and region fixed-effects, as proposed in Equation (2). We do not find many differences in comparison with the estimates of the reduced form in Table 5. The estimated effect of the interaction term is still positive and statistically significant on the aggregates of household consumption for durable and non-durable goods. However, when we estimate the pure effect of living in the Christmas Lottery winning region (i.e., adding the estimated coefficients for win_region and $lottery \times win\_region$), we still observe that households that live in winning regions spend less on household consumption of the goods analyzed in comparison with those that live in non-winning regions. Focusing on the aggregates of durable and non-durable goods, we observe that, on average, a household living in a winning region spends 14.13% less on durable goods and 11.26% less on non-durable goods than a household that lives in a non-winning region.

Table 6: Household consumption expenditures - Adding household controls

PANEL A	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Food at home	Alcohol	Clothes	House Rent	House Investments	Health	Car Value	Transport
win_region	–0.148***	0.0398	-0.0367	-0.168***	-0.166***	0.0448*	0.253***	-0.0316
	(-14.92)	(1.84)	(-1.88)	(-15.60)	(-11.99)	(2.05)	(13.66)	(-1.81)
Lottery	0.165***	1.052***	0.842***	0.0610***	0.303***	1.194***	1.449***	0.702***
	(27.75)	(80.55)	(71.66)	(9.40)	(36.32)	(90.63)	(129.78)	(66.75)
Lottery×win_region	0.0900***	-0.341***	-0.206***	0.0333*	0.000498	-0.332***	-0.460***	-0.159***
	(6.88)	(-11.90)	(-7.98)	(2.34)	(0.03)	(-11.48)	(-18.74)	(-6.90)
_cons	8.476***	10.03***	4.043***	10.80***	11.90***	1.982	1.711	6.927***
	(16.06)	(8.67)	(3.88)	(18.78)	(16.12)	(1.70)	(1.73)	(7.44)
F-test	31.26	178.29	142.62	144.28	131.88	159.10	144.01	110.38
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	211096	211096	211096	211096	211096	211096	211096	211096
R-squared	0.831	0.518	0.612	0.874	0.802	0.507	0.212	0.693
PANEL B	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Communication	Leisure	Education	Food out Home	Holidays	Savings	Durables	Non-durables
win_region	-0.172***	-0.116***	0.0903***	-0.0652***	0.102***	0.174***	-0.201***	-0.164***
	(-12.51)	(-6.16)	(3.45)	(-3.61)	(4.47)	(7.68)	(-15.62)	(-13.95)
Lottery	-1.006***	0.164***	0.742***	0.291***	0.906***	0.856***	-0.486***	-0.275***
	(23.22)	(83.49)	(109.84)	(74.72)	(98.07)	(73.22)	(49.70)	(54.87)
Lottery×win_region	-0.0114	-0.129***	-0.438***	-0.209***	-0.355***	-0.359***	0.0487**	0.0445**
	(-0.63)	(-5.18)	(-12.64)	(-8.74)	(-11.77)	(-11.99)	(2.87)	(2.87)
_cons	6.625***	3.009**	2.440***	-1.738	-10.16***	32.24***	12.54***	10.81***
	(9.03)	(3.00)	(6.15)	(-1.80)	(-8.35)	(26.73)	(18.31)	(17.28)
F-test	164.27	156.11	164.14	212.50	113.62	61.82	129.03	95.19
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	211096	211096	211096	211096	211096	211096	211096	211096
R-squared	0.814	0.633	0.308	0.675	0.347	0.391	0.867	0.855

Our dependent variables are the logarithm of consumption expenditures of each good category. The results in this table perform a robustness check of Table 5, by adding household and demographic characteristics, as well as both region and year fixed-effects. The coefficients presented in this table, *win_region*, *Lottery* and *lottery* × *win_region*, are as described in Table 5. In this specification, we also include as control variables the age of the head of the household and its square, the marital status of the head of the household and his/her educational level, employment status and whether he/she is retired or not. Moreover, we also include the logarithm of lottery expenditures per region and the regional log-GDP as demographic controls, and both regional and year fixed-effects. We compute robust standard errors. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of estimates available upon request.

The performance of the F-test is as described in Table 5 and it follows the null hypothesis described under Hypothesis 1. This test is examining the validity of the PIH.

Moreover, when we perform the F-test under Hypothesis 1, the null hypothesis is rejected, implying that the estimated effect of the lottery income shock is statistically significant and, thus, that it still has an impact on household consumption behavior when we control for household and demographic characteristics. Therefore, these findings lead to a violation of the PIH, under the conditions created by the Spanish Christmas Lottery winnings: households that live in winning regions do not smooth consumption, because it can be seen that households increase their consumption expenditures in durable and non-durable goods.

These findings are not in line with the theoretical results in Cerletti and PijoanMas (2014), which stated that households that experience a positive shock in income use it to increase their consumption of durable goods or to pay back their mortgages or other debts, and the consumption of non-durable goods should not show a reaction to the income shock. In the case of our research, a positive shock to household expenditures leads to an increase in both: the consumption of durable and non-durable goods. However, these estimates do not capture the elasticity effect of the income shock on consumption behavior. In order to estimate this, we need to estimate the Engel curves, as proposed in Equation (4).

Moreover, we also perform the test for the parallel trend assumption. In this case, we want to examine if households can somehow anticipate the lottery income shock and, thus, start increasing consumption of the different goods prior to the lottery draw celebration or, if otherwise, they do not and take the income shock as something completely unexpected. We observe from Figure 2 in the Appendix that the parallel trend assumption passed in all cases, as households do not increase consumption in the pre-treatment period and, thus, they do not anticipate the lottery winnings. Therefore, the parallel trend assumption is satisfied. Such results can be also confirmed in Table 16 in the Appendix.

Table 7 presents the results for the first stage regression presented in Equation (3). In Table 14 in the Appendix of this chapter, we report the full set of estimates for the first stage regression. The main finding is that for a household that lives in a winning region and participates in the Christmas Lottery, in comparison with a household that either does not participate in it, or does not live in a winning region, or both, total household expenditure increases by 5.61%, and is statistically significant. Thus, we have an indication that potential winners of the lottery may increase their total household expenditures. However, when looking at the *win_region* coefficient per se, we observe that the estimated coefficient is negative, and the aggregate estimate of both coefficients is negative. Nevertheless, as explained earlier, this might be caused by the fact that there are more people playing in regions with lower income. Furthermore, to check whether the lottery income shock has an impact on total household expenditure, we need to test the relevance condition for instrumental variables. To do so, we run an F-test under the following hypothesis:

Hypothesis 2:

$$H_0: \text{ The set of instruments for total expenditures is not relevant.}$$

$$H_a: \text{ The set of instruments for total expenditures is relevant.}$$

The results from the F-test for instrumental variables is 133.57. This shows that the relevance condition is satisfied and the set of instruments used is strong, because the resulting number is greater than 10. Thus, the relevance condition is fulfilled, as well as orthogonality or exogeneity, which is automatically satisfied because the winning regions are assigned completely randomly.

The results for the logarithm of total expenditures are only taken into account in the second stage, because the estimated results for the interaction term are in line with the theory and explain our expectations well: the effect of the lottery income shock for potential winners (i.e., those households that participate in the lottery and live in the winning region) has a positive impact on total household expenditures.

Table 8 shows the results for the second stage regression in Equation (4). It can be seen that durable goods are sensitive to total expenditures, adjusted to the lottery income shock in the first stage, because the estimated effect is above one. More precisely, a 10% increase in household total expenditures leads to an increase of 11.47% in household expenditures for durable goods. On the other hand, the adjusted total expenditures coefficient for non-durable goods is below one. This implies, from a theoretical point of view, that non-durable goods are not sensitive to a shock in total expenditures and, thus, the estimated effect is inelastic. Specifically, increasing total household expenditures by 10% implies an increase of 9.26% in household consumption of non-durable goods, and this effect is statistically significant as well. However, there are two things we need to test in this case: first we need to confirm the elasticity effects for durable and non-durable goods, and to check if the estimates are significantly different from each other.

In order to test whether the estimated effects for durable and non-durable goods are elastic or inelastic, we need to perform a t-test under the null hypothesis that the estimate for *log expenditures* is equal to one. In this case, we observe the elasticity test in Table 8 that we do reject the null hypothesis for durable goods at the 5% level of significance. However, we fail to reject the null hypothesis for non-durable goods at the level of 5%. This implies that durable goods are indeed elastic and sensitive to a positive shock to total household expenditures, but non-durable goods are unit elastic to total household expenditures.

Table 7: First stage regression - Total household expenditures

	(1)
	Total Expenditures
win_region	-0.179***
	(-14.54)
lottery	0.391***
	(52.71)
lottery×win_region	0.0546***
	(3.35)
_cons	12.17***
	(18.52)
Specification	Expenditures in logarithms
F-test for the IV	133.57
p-value	0.000
N	211096
R-squared	0.858

This table shows the results for the first stage estimation, using the lottery income shock variables, *win_region* and *lottery* × *win_region*, as instruments for total household expenditures. The coefficients *win_region*, *Lottery* and *lottery* × *win_region*, are as previously described in Table 5. We present the effect of the lottery income shock on total household expenditures in logarithms; we also did the estimations for total expenditures in levels, but the results were showing a negative impact of the lottery income shock on total expenditures, thus we avoid using expenditures in levels, as the estimates go against our expectations. Both specifications include as control variables the age of the head of the household and its square, the marital status of the head of the household and his/her educational level, employment status and whether he/she is retired or not. Moreover, we also include the logarithm of lottery expenditures per region and the regional log-GDP as demographic controls, and both regional and year fixed-effects. We compute robust standard errors. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of estimates available in Table 14 of the Appendix of this chapter.

The F-test performs the relevance condition test for the instrumental variables (*win_region* and *lottery*∗*win_region*), where the null hypothesis is that: *the set of instrumental variables for total household expenditures is not relevant.*

Table 8: Second stage estimation - Household consumption behavior

PANEL A	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Food at home	Alcohol	Clothes	House Rent	House Investments	Health	Car Value	Transport
log expenditures	0.725***	0.505***	0.684***	0.979***	1.037***	0.453***	-0.571***	0.549***
	(14.75)	(4.68)	(7.06)	(18.27)	(15.09)	(4.17)	(-6.19)	(6.34)
Household controls	✓	✓	✓	✓	✓	✓	✓	✓
Fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Elasticity test	31.42	21.14	10.62	0.16	0.29	25.31	290.31	26.98
p-value	0.0000	0.0000	0.0000	0.6893	0.5903	0.0000	0.0000	0.0000
N	211096	211096	211096	211096	211096	211096	211096	211096
R-squared	0.831	0.517	0.612	0.874	0.802	0.506	0.211	0.693
PANEL B	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Communication	Leisure	Education	Food out Home	Holidays	Savings	Durables	Non-durables
log expenditures	1.101***	1.009***	0.406***	0.870***	0.146	-0.294**	1.147***	0.926***
	(16.13)	(10.80)	(3.85)	(9.69)	(1.29)	(-2.62)	(17.99)	(15.89)
Household controls	✓	✓	✓	✓	✓	✓	✓	✓
Fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Elasticity test	2.21	0.01	31.66	2.11	56.87	132.75	5.31	1.62
p-value	0.1372	0.9249	0.0000	0.1464	0.0000	0.0000	0.0213	0.2025
N	211096	211096	211096	211096	211096	211096	211096	211096
R-squared	0.814	0.633	0.331	0.675	0.346	0.390	0.867	0.855

Our dependent variables are the logarithm of consumption expenditures of each good category. The *log expenditures* coefficient reports the estimates for total household expenditures, instrumented using *win_region* and *lottery* × *win_region* as instrumental variables in the first stage regression. This coefficient captures the elasticity effect of total household expenditures on household consumption expenditures for the different types of goods analyzed. All specifications include both year and region fixed-effects. Moreover, we also control for the age of the head of the household and its square, the marital status of the head of the household and his/her educational level, employment status and whether he/she is retired or not. In addition, we control for the logarithm of lottery expenditures per region and the regional log-GDP as demographic controls. We compute robust standard errors. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of estimates are available in Table 15 of the Appendix of this chapter. The reported Elasticity test examines the elasticity effect of expenditures towards household consumption, in words, whether the estimates are different from one.

For the second test, we perform a t-test under the null hypothesis that the estimated elasticities for durable and non-durable goods are the same. In this case we do reject the null hypothesis, as the estimated t-test is 9.08 with a p-value equals to zero. However, looking at the estimated elasticities, we observe that the difference between coefficients is of 0.22, which is small. Thus, economically speaking, the estimates of total household expenditures to durable and non-durable goods is similar are not far different from each other, something that we would not expect given the theoretical predictions in Cerletti and Pijoan-Mas (2014).

Analyzing the specific goods themselves, we observe that an increase in total household expenditures has a positive effect on household consumption for almost all the goods analyzed, except for car value and savings. More precisely, we observe that an increase of 10% in total household expenditures leads to a decrease of 2.94% in household savings. This result implies that the positive income shock to total household expenditures affects savings negatively. On the other hand, we find that when total expenditures increase by 10%, household consumption of food at home increases by 7.25%, alcohol consumption by 5.05%, monthly house rent by 9.79%, and transport costs (which can be understood as bus/metro tickets, petrol for the car, tolls, etc.) by 5.79%. What all these goods have in common is that they can be considered to be non-durable goods and, thus, we should expect them to be inelastic to total expenditures (i.e., the estimated effect of total expenditures should be below one), as indeed we find in our estimations. However, for leisure expenditures, we find that an increase of 10% in household expenditures leads to an increase in 10.1% in leisure activities. In this case we find that leisure is unit elastic to total expenditures.

On the other hand, when we estimate separately the effect of total household expenditures on durable goods, we find that an increase of 10% on household expenditures leads to an increase of 10.37% in the consumption of durable goods for the house (including house purchases); of 11.01% on communication goods (understood as phone, mobiles and related goods); of 4.06% on education; of 8.70% on eating out; of 6.84% in clothes, of 4.53% on health insurance, and to a decrease of 5.71% in car value. In this case, it can be seen that only household durables and communication goods are elastic to total expenditures; the remaining categories of goods report inelastic estimates to a shock to total expenditures.

However, when looking at the elasticity test in Table 8, we observe that monthly house rent, house investments, communication goods, leisure, eating out and non-durable goods are unit elastic to total household expenditures, as we fail to reject the null hypothesis that the estimated elasticity coefficient is different from one.

When comparing these results with previous empirical studies carried out on the topic, it can be seen that in the case of the Postcode Lottery in the United Kingdom, the effect of winning the Postcode Millions lottery leads also to a violation of the PIH, as household consumption of durable and non-durable goods increase due to the lottery income shock in winning postcodes (see Chapter 3). Moreover, when looking the elasticity estimates, Chapter 2 finds that an increase of 10% in total household expenditures leads to an increase of 13.80% in household durable goods consumption, and of 10.40% in non-durable goods consumption, and this effect implies that both goods are unit-elastic to total expenditures and the estimates are not significantly different from each other. Another empirical study carried out, using the Dutch Postcode Lottery, found that the lottery income shock led to an increase of 310€ on household durable goods for those households that won the lottery, in comparison with those households that did not (see Kuhn et al., 2011). However, that paper did not find statistically significant results of the lottery income shock on household consumption of non-durable goods.

Hence, given the results presented in this chapter, we conclude two important things: (i) that the PIH is violated under the scenario that the Spanish Christmas lottery presents, and (ii) economically speaking, the estimated elasticities of total household expenditures for durable and non-durable goods are not far different from each other and, thus, a positive income shock to total expenditures affects similarly household consumption expenditures for durable and non-durable goods.

6.2 Labor Supply Estimation

Table 9 presents the results for the estimated regression presented in Equations (5) and (6). In this case, columns (1) and (3) are presenting the results for the reduced-form estimates, whereas our preferred specifications are the ones presented in columns (2) and (4), as these are also controlling for household characteristics and fixed-effects.

In terms of the employment status of the head of the household, no effect from the lottery income shock is found. This implies that living in a lottery winning region does not affect the current employment status of households in comparison with those households that live in non-winning regions. Moreover, neither does the lottery income shock alter the employment status of the head of the household when we control for household and demographic characteristics, and year and region fixed-effects.

However, when estimating the reduced-form estimation on how the lottery income shock affects the number of hours worked, we find a positive and statistically significant effect for those households that live in Christmas Lottery winning regions (see column (3)). This means that the heads of households in the winning regions allocate more hours to work; specifically, the head of a household living in a winning region works 0.148 hours more than the head of a household in a non-winning region.[16] Looking at the F-test, we find this effect to be statistically significant. Although logic might consider that an increase in income should lead to a reduction in the number of hours worked.

[16] This effect is obtained by adding the estimated coefficients for win_region and $lottery \times win\_region$.

Table 9: Labor Supply analysis

	(1) Employed	(2) Employed	(3) Hours Worked	(4) Hours Worked
win_region	-0.00684*	-0.00639	-0.123***	-0.0605
	(-2.05)	(-1.73)	(-3.61)	(-1.43)
lottery	0.0124***	0.00678**	-0.329***	0.000878
	(6.43)	(3.05)	(-18.53)	(0.03)
lottery×win_region	0.00219	-0.00281	0.271***	0.0564
	(0.46)	(-0.58)	(5.50)	(1.00)
_cons	0.419***	-1.117***	4.206***	0.800
	(321.75)	(-5.68)	(348.30)	(0.36)
Household controls	No	Yes	No	Yes
Demographic controls	No	Yes	No	Yes
Region and time fixed effects	No	Yes	No	Yes
F-test	1.83	5.74	17.31	0.01
p-value	0.1766	0.0166	0.000	0.9265
N	278490	211096	174706	95840
R-Squared	0.000225	0.136	0.00197	0.000276

The coefficients presented in this Table, *win_region*, *Lottery* and *lottery* × *win_region*, are as described in Table 5. In this case, we are interested in how the lottery income shock affects household allocation of time in labor supply and to the probability of being employed. Specifications presented in column (1) and (3) are the reduced-form estimations, which only include the variables presented in this table, whereas specifications in columns (2) and (4) control as well for the age of the head of the household and its square, the marital status of the head of the household and his/her educational level. Moreover, in these two specifications, we also include the logarithm of lottery expenditures per region and the regional log-GDP as demographic controls, and regional and year fixed-effects. Thus, given the completeness of the estimations, we are interested in the results of these two specifications, instead of the reduced-form ones. We compute robust standard errors. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of estimates available upon request.

The F-test performs a joint significant test of the lottery income shock variables (*win_region* and *lottery*×*win_region*) on the probability of being employed and the number of hours worked per day. The null hypothesis of the F-test is that: *the lottery income shock has no effect on the employment status and daily hours worked.*

Nevertheless, when we control for household and demographic characteristics in column (4), we do not observe any significant effect of the lottery income shock on hours worked.[17] This is confirmed with the performance of the F-test, where we fail to reject the null hypothesis: *that the lottery income shock has no effect on the number of hours worked.*

6.3 Number of Children in the Household Estimation

Table 10 reports the estimated results for Equation (7) and Equation (8). We do not find any changes in family composition resulting from the lottery income shock. Therefore, living in the winning region does not affect the number of children, between 0 and 18 years old, in the household. When performing the robustness check by controlling for individual characteristics, demographic controls and fixed effects, the effect of the income shock is still not relevant for household composition in the periods after the income shock happened.

Finally, if we run a regression on the number of children in the household one year after winning the lottery or in the present time, we do not find significant results either. Therefore, there is no evidence that the lottery shock significantly affects family composition by resulting in an increase in the number of children.

[17] The aggregate effect of win_region and $lottery \times win\_region$ is close to zero.

Table 10: Intergenerational analysis

	(1)	(2)
	Children_{t+2}	Children_{t+2}
win_region	0.00830	0.0175
	(0.54)	(1.07)
lottery	0.0195	0.0202^{*}
	(1.92)	(1.99)
lottery×win_region	-0.0423	-0.0393
	(-1.88)	(-1.82)
_cons	4.509^{***}	4.192^{***}
	(682.16)	(4.52)
Household controls	No	Yes
Demographic controls	No	Yes
Region and time fixed effects	No	Yes
F-test	4.36	1.65
p-value	0.0369	0.1984
N	202738	188500
R-squared	0.0000332	0.000249

The coefficients presented in this Table, *win_region*, *Lottery* and *lottery* × *win_region*, are as described in Table 5. In this case, we are interested in how the lottery income shock affects family composition: whether it leads to households having more children, or not. The specification presented in column (1) represents the reduced-form estimation, which only includes the variables presented in this table, whereas the specification in column (2) controls as well for the age of the head of the household and its square, the marital status of the head of the household and his/her educational level, employment status and whether he/she is retired or not. Moreover, in these two specifications, we also include the logarithm of lottery expenditures per region and the regional log-GDP as demographic controls, and regional and year fixed-effects. Thus, given the completeness of this estimation, we are interested in the results of this last one, instead of the reduced-form one. We compute robust standard errors. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of estimates available upon request.

The F-test performs a joint significant test of the lottery income shock variables (*win_region* and *lottery* × *win_region*) on family composition. The null hypothesis of the F-test is that: *the lottery income shock has no effect on family composition* and, thus, households do not increase the number of newborns after experiencing the income shock.

7 Conclusion

There are several advantages of using the Spanish Christmas Lottery as a quasi-experiment with which to investigate household consumption behavior. Firstly, the economic impact of winning the lottery for the winning region is an increase of 3.5% on local GDP. Moreover, the population of Spain spends around 0.3% of national GDP on the Christmas Lottery. Secondly, the lottery ticket is not unique to one individual: the same number can be shared by groups of friends, colleagues or family members, and more than one person can get the same lottery ticket. This has two immediate consequences: (i) it makes the analysis more heterogeneous, as different individuals can be part of the treatment group, and (ii) it means that this lottery is a syndicate game, in which over 75% of the population participates (see Garvía, 2007). Finally, winners are clustered and easy to locate because each winning number is typically sold by one outlet.

This chapter takes advantage of the Spanish Christmas Lottery, a completely randomized and exogenous income shock, to study the causal effect of the lottery prize on household consumption behavior. After performing a fixed effect instrumental variable analysis, using the lottery income shock as an instrument for total household expenditures, we find that, when estimating the Engel curves in the second stage, the consumption of household durable goods is sensitive to a shock to total expenditures but the consumption of non-durable goods is not, because an increase of 10% in total household expenditures led to an increase of 11.47% in the consumption of household durable goods and an increase of 9.26% in the consumption of household non-durable goods. These effects imply that the consumption of durable goods is elastic to total expenditures, as the estimated elasticity is significantly different from one; while, in contrast, the consumption of non-durable goods is unit-elastic to total expenditures. From a statistical point of view, we find that these estimates are significantly different from each other, however, from an economic point of view the estimated elasticities are not far away one to another. This implies that households that experience a positive shock to total expenditures react likewise to consumption expenditures of durable and non-durable goods. And this is the novelty of this chapter, as these results lead to a contradiction of the theoretical results of Cerletti and Pijoan-Mas (2014), in which households that experience a positive income shock increase their consumption of durable goods, and smooth their consumption of non-durable goods.

Moreover, the results we found for the reduced-form estimation in Table 5 contradict the empirical outcomes in Kuhn et al. (2011), examining the Dutch Postcode Lottery, in which households that won the lottery significantly increased their consumption of durable goods only. Therefore, we also find a violation of the PIH under the scenario presented by the Spanish Christmas lottery.

However, when we analyze the implications of the lottery income shock on labor supply and household composition, we do not find any evidence that lottery winnings induce heads of households to change their employment status, the number of hours worked or alter their family composition.

Nonetheless, this chapter presents an important limitation. In the case of this research, we do not know which households are winners of the Spanish Christmas Lottery. However, we do know the region that they live in and whether they live in the capital of the province or not; thus, we can use this information as a proxy to locate the possible winners of the lottery. However, regions in Spain may be highly populated, such as Madrid, or may be large in area, such as Andalucía or Castilla; therefore, this introduces many limitations to the analysis of this chapter as it is hard to capture the real effect of the lottery income shock on household consumption. A potential solution to this is to proceed with a regional analysis and, thus, to see how Christmas Lottery winning regions behave in terms of consumption expenditures.

Finally, given the increase in consumption expenditures in almost all goods analyzed in response to the lottery income shock, we aim that the results obtained in this chapter may assist policy-makers in encouraging Spanish politicians to initiate new fiscal policy measures, such as tax rebates or reductions in personal income taxes (known in Spain as IRPF), in order to encourage household consumption, especially that of durable goods.

Appendix Chapter 2

Average Lottery Earnings

Table 11: Average lottery earnings per region

Year	Average
2016	387,72
2015	224,12
2014	1507,50
2013	753,81
2012	860,88
2011	2263,40
2010	1641,59
2009	387,72
2008	679,22
2007	918,29
2006	731,38
2005	240,20
2004	428,34
2003	914,65
2002	535,82
2001	847,62
2000	1157,63
1999	348,95
1998	667,95

Table 12: Proportion of people included in the sample and potential winners per region

Region	Proportion of interviwed people	Estimation of potential winners	Region	Proportion of interviwed people	Estimation of potential winners
Andalucía	11%	14%	C. Valenciana	7.97%	10%
Aragón	4.45%	10%	Extremadura	4.26%	1%
Asturias	3.96%	4%	Galicia	6.52%	2%
Baleares	3.67%	0%	Madrid	7.23%	18%
Canarias	4.65%	4%	Murcia	4.16%	6%
Cantabria	3.18%	0%	Navarra	4.08%	0%
Castilla-León	6.81%	3%	País Vasco	8.71%	14%
Castilla-La Mancha	5.30%	2%	La Rioja	3.14%	0%
Cataluña	9.68%	8%	Ceuta	0.88%	0%
			Melilla	0.25%	0%

This Table presents the proportion of households that take part of our sample by regions, as well as an expectation of the proportion of individuals in each reagion that can be potential winners of the lottery. In other words, of those individuals that were interviwed in a given year that the region became awarded with the first prize of the lottery, and they also bought lottery tickets. In general, we observe that larger regions (or more populated ones) take an important representation in the sample, like Andalucía, Cataluña or Madrid. When looking at the proportion of winners in each region, we observe that there are some regions that have zero chances of having winners, as these have not won the lottery in the years included in the survey.

Estimation Results

Table 13: Identification Strategy

Variable	Coefficient	(t-test)
age	0.000	(-1.58)
marital_status	0.000	(-0.37)
education	-0.005	(-1.60)
employed	-0.014	(-1.42)
retired	-0.009	(-1.54)
gender	0.005	(1.05)
_cons	0.190**	(2.91)
N	211096	
R-Squared	0.0014	
F-test	0.66	

Table 14: First stage regression - Total household expenditures

	(1)
	Total expenditures
Win_region	-0.179***
	(-14.54)
Lottery	0.391***
	(52.71)
Lottery*win_region	0.0546***
	(3.35)
Log lottery expenditure	-1.280***
	(-25.15)
log-GDP	-1.588***
	(-6.59)
Age	-0.0292***
	(-14.37)
Age^2	0.000243***
	(12.16)
Marital_status	0.0523***
	(14.45)
Education	0.0194***
	(8.59)
Employed	-0.000287
	(-0.04)
Retired	-0.153***
	(-7.51)
_cons	12.17***
	(18.52)
Specification	Expenditures in logarithms
N	211096
R-Squared	0.858

This Table presents the extended results of Table 7, where we described in detail the specifications and the estimation process. We compute robust standard errors. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

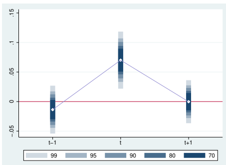
Table 15: Second stage estimation - Household consumption behavior

PANEL A	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Food at home	Alcohol	Clothes	House Rent	House Durable	Health	Car Value	Transport
Log expenditures	0.725***	0.505***	0.684***	0.979***	1.037***	0.453***	-0.571***	0.549***
	(14.75)	(4.68)	(7.06)	(18.27)	(15.09)	(4.17)	(-6.19)	(6.34)
Lottery	-0.108***	0.780***	0.525***	-0.326***	-0.114***	0.945***	1.586***	0.449***
	(14.75)	(4.68)	(7.06)	(18.27)	(15.09)	(4.17)	(-6.19)	(6.34)
log lottery expenditures	0.0968	-0.781***	-0.428**	-0.120	-0.0846	-0.687***	-1.199***	-0.539***
	(1.33)	(-4.90)	(-2.98)	(-1.52)	(-0.83)	(-4.27)	(-8.80)	(-4.21)
log gdp	0.104	-0.672	1.557***	-0.0302	-0.362	1.713***	-0.889*	0.134
	(0.51)	(-1.51)	(3.89)	(-0.14)	(-1.28)	(3.82)	(-2.34)	(0.37)
age	-0.0132***	0.0966***	0.0621***	-0.0148***	0.0112***	0.0527***	0.000892	0.138***
	(-6.04)	(20.21)	(14.44)	(-6.25)	(3.66)	(10.93)	(0.22)	(35.81)
age^2	0.000140***	-0.000890***	-0.000835***	0.000162***	-0.000106***	-0.000383***	-0.0000932*	-0.00162***
	(6.94)	(-20.18)	(-21.04)	(7.37)	(-3.78)	(-8.61)	(-2.47)	(-45.73)
marital	0.0138***	-0.0613***	-0.0600***	-0.00317	-0.0172**	-0.0724***	-0.0315***	-0.0920***
	(3.56)	(-7.21)	(-7.84)	(-0.75)	(-3.18)	(-8.44)	(-4.33)	(-13.43)
education	-0.00160	0.00410	0.0296***	-0.00856***	0.000268	0.0461***	0.0113**	0.0523***
	(-0.78)	(0.91)	(7.29)	(-3.81)	(0.09)	(10.11)	(2.92)	(14.37)
employed	-0.0142*	-0.0339*	0.0675***	-0.00483	-0.00389	0.0410**	0.0127	0.100***
	(-2.30)	(-2.50)	(5.53)	(-0.72)	(-0.45)	(2.99)	(1.10)	(9.16)
retired	-0.00144	-0.439***	0.122***	-0.00799	0.0339	-0.260***	-0.483***	0.151***
	(-0.08)	(-11.12)	(3.44)	(-0.41)	(1.35)	(-6.51)	(-14.30)	(4.76)
_cons	-0.302	3.573*	-4.492**	-1.126	-0.768	-3.836*	8.294***	0.0774
	(-0.40)	(2.17)	(-3.03)	(-1.37)	(-0.73)	(-2.31)	(5.88)	(0.06)
N	211096	211096	211096	211096	211096	211096	211096	211096
R-squared	0.831	0.517	0.612	0.874	0.802	0.506	0.211	0.693
PANEL B	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Communication	Leisure	Education	Food out Home	Holidays	Savings	Durables	Non-durables
Log expenditures	1.101***	1.009***	0.406***	0.870***	0.146	-0.294**	1.147***	0.926***
	(16.13)	(10.80)	(3.85)	(9.69)	(1.29)	(-2.62)	(17.99)	(15.89)
Lottery	-0.253***	0.514***	1.233***	0.421***	1.215***	1.042***	-0.0672*	0.0244
	(16.13)	(10.80)	(3.85)	(9.69)	(1.29)	(-2.62)	(17.99)	(15.89)
log lottery expenditures	0.0939	0.215	-0.505**	0.0705	-1.449***	-0.954***	-0.0403	-0.00506
	(0.93)	(1.56)	(-3.24)	(0.53)	(-8.66)	(-5.75)	(-0.43)	(-0.06)
log gdp	1.371***	2.024***	3.753***	3.418***	5.733***	-10.75***	-0.0751	0.171
	(4.87)	(5.25)	(8.62)	(9.23)	(12.28)	(-23.20)	(-0.29)	(0.71)
age	0.00656*	0.0910***	0.0983***	0.0922***	0.0856***	0.118***	0.00337	0.00139
	(2.16)	(21.95)	(20.99)	(23.15)	(17.02)	(23.61)	(1.19)	(0.54)
age^2	-0.0000240	-0.00108***	-0.00116***	-0.00113***	-0.00107***	-0.000521***	-0.00000599	-0.0000315
	(-0.86)	(-28.32)	(-26.88)	(-30.84)	(-23.16)	(-11.33)	(-0.23)	(-1.32)
marital	-0.0274***	-0.105***	-0.173***	-0.100***	-0.0609***	0.0334***	-0.0106*	-0.000621
	(-5.08)	(-14.29)	(-20.80)	(-14.12)	(-6.81)	(3.76)	(-2.11)	(-0.13)
education	0.0172***	0.0656***	0.0484***	0.0610***	0.0982***	-0.0923***	-0.00467	0.00351
	(6.00)	(16.76)	(10.95)	(16.20)	(20.69)	(-19.60)	(-1.75)	(1.44)
employed	0.0511***	0.0767***	-0.0197	0.169***	0.154***	-0.136***	-0.00105	0.00517
	(5.94)	(6.52)	(-1.48)	(14.95)	(10.83)	(-9.59)	(-0.13)	(0.70)
retired	-0.00470	0.0798*	-0.821***	0.0819*	0.316***	-0.384***	-0.0129	0.0163
	(-0.19)	(2.33)	(-21.21)	(2.49)	(7.62)	(-9.34)	(-0.55)	(0.76)
_cons	-6.844***	-9.427***	-10.62***	-12.54***	-12.24***	35.53***	-1.433	-0.460
	(-6.56)	(-6.60)	(-6.58)	(-9.14)	(-7.07)	(20.69)	(-1.47)	(-0.52)
N	211096	211096	211096	211096	211096	211096	211096	211096
R-squared	0.814	0.633	0.331	0.675	0.346	0.390	0.867	0.855

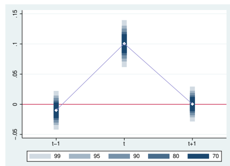
This Table presents the extended results of Table 8, where we described in detail the specifications and the estimation process. We compute robust standard errors. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Testing the Parallel Trend Assumption

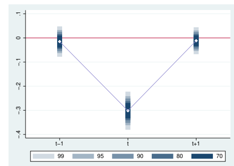
Figure 2: Parellel Trends - Treatment effect



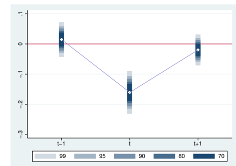
(a) Total Expenditures



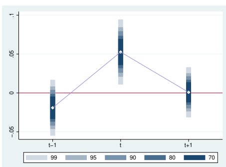
(b) Food at home



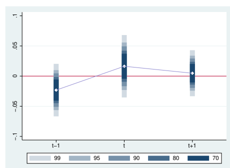
(c) Alcohol



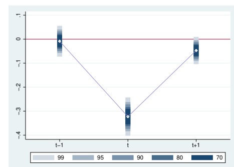
(d) Clothes



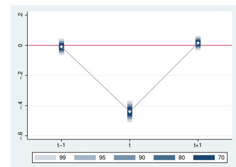
(e) House rent



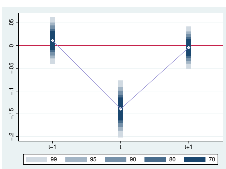
(f) House investments



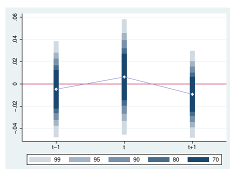
(g) Health



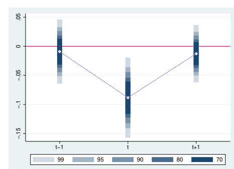
(h) Car value



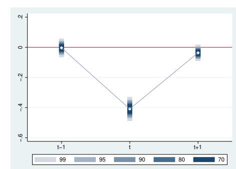
(i) Transport



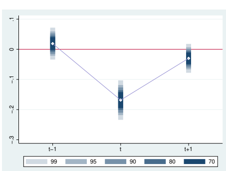
(j) Communication



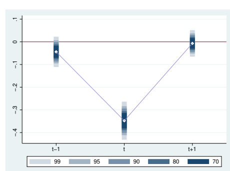
(k) Leisure



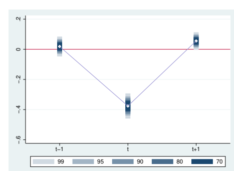
(l) Education



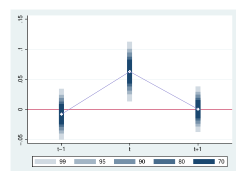
(m) Food out home



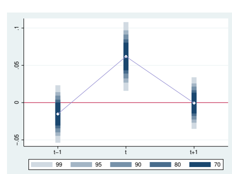
(n) Holidays



(o) Savings



(p) Durables



(q) Non-durables

Table 16: Testing the Parallel Trend Assumption

PANEL A	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Food at home	Alcohol	Clothes	House Rent	House Investments	Health	Car Value	Transport
F-test	0.61	0.35	0.44	1.81	1.87	0.12	0.20	0.32
p-value	0.4334	0.5558	0.5090	0.1781	0.1714	0.7266	0.6520	0.5725
PANEL B	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Communication	Leisure	Education	Food out Home	Holidays	Savings	Durables	Non-durables
F-test	0.08	0.18	0.02	0.85	2.85	0.53	0.20	1.01
p-value	0.7819	0.6720	0.8773	0.3568	0.0911	0.4684	0.6510	0.3155

In this table we present the results for the parallel trend assumption test, where we expect households to not anticipate the lottery income shock on their consumption of different goods. In this case, we report the F-test results and the p-values below. The main conclusions from this table are that households, who live in winning regions, do not alter their consumption, prior to the reception of the lottery earnings and, thus, the parallel trend assumption is satisfied. We use the *time varying* treatment command, following the analysis in Cerulli and Ventura, 2017.

Chapter 3

Consumption responses to income shocks: Evidence from the British Postcode Lottery.

Abstract

The People's Postcode Lottery in the United Kingdom randomly awards a British postcode weekly with £30.000 and with 3 million pounds every month, distributed among all the households that play the lottery within the postcode. We analyze the effects of living in an awarded postcode of the lottery into household's consumption behavior and the time allocation of working hours. We find that the income shock from the Million's Lottery positively affects household consumption for durable goods (where the effects of the windfall are larger for long-lasting goods like cars and dwellings) and contrary to what theory predicts, households living in winning postcodes also increase their consumption for non-durable goods. Such findings are new in the literature, as these violate the Permanent Income Hypothesis. Moreover, when we estimate the Engel curves, we find that the estimated elasticity of total expenditures on durable and non-durable goods is similar, and households react similarly to a shock to total household expenditures. However, we do not find any effects on household consumption expenditures due to the £30.000 lottery. Neither do we find significant changes in hours worked due to both lottery income shocks.

Keywords: Consumption Behavior, Durable goods, Postcode Lottery, Winning Postcode
JEL Classification: C01, C23, C26, C55, C93, D12 and D14.

1 Introduction

Research in economic theory offers an interesting set of predictions about how income shocks can affect the economic behavior of individuals. Two of the most important theoretical predictions are *the Life-Cycle Hypothesis* (LCH), proposed by Franco Modigliani in 1954, and *the Permanent Income Hypothesis* (PIH), proposed by Milton Friedman in 1957. The first one, the LCH, predicts that individuals smooth their consumption over their lifetime, according to their earnings along their life - borrowing during periods of low income and saving in periods of high income (see Deaton, 2005). In contrast, the PIH states that individuals are expected to spend their money consistently with their long-term average income along their lifetime (see Friedman, 1957). Therefore, according to the PIH, when individuals experience a one-off positive income shock, they should smooth consumption and adjust it to the new scenario. The PIH is the hypothesis this chapter is focusing on. However, consistently with other theoretical findings, when individuals get a large enough cash transfer, they anticipate the purchase of durable goods before these get obsolete, but consumption of non-durable goods remain stable (see Grossman and Laroque, 1990 and Cerletti and Pijoan-Mas, 2014), violating the PIH proposed by Friedman (1957). Robert Hall empirically tested this hypothesis in 1978, where the main finding is that individual's consumption should not be inferred by past consumption; therefore, present consumption should not depend on previous decisions or any income shock and thus, the PIH should hold.

We analyze in this chapter how household consumption for durable and non-durable goods changes when a one-off, positive and unexpected income shock occurs. Specifically, we use winnings in the People's Postcode Lottery as an income shock and investigate its effect on household consumption behavior. Particularly, we are interested in two different type-size of income shocks: a big one coming from the Millions Postcode Lottery and a small one coming from the £30.000 Postcode lottery. Hence, the main research question we try to answer in this chapter is how two different size lottery income shocks, the Millions and the £30.000 Postcode Lottery winnings, affect the behavior of households that live in the winning postcodes towards household consumption of durable and non-durable goods. Moreover, we also estimate the Engel curves, where we obtain the elasticity of total household expenditures, instrumented with the lottery income shock, on durable and non-durable goods and other specific types of goods. Secondarily, we are also interested in analyzing how those individuals that live in winning postcodes of the Postcode Lottery reallocate their time from labor to leisure activities. Our underlying hypothesis is that those households that live in winning postcodes of the Millions Postcode Lottery do not spend more than those that live in non-winning postcodes and, thus, the PIH holds.

The amount that a winning postcode of the Millions lottery gets is £3 Million and £30.000 from the other lottery, which is shared across winning households in the same postcode - if more than one household in the winner postcode participates in the raffle. Due to the current legislation, the amount a household, who wins the Millions Postcode Lottery, can get is 10% of the total amount awarded, up to a maximum of £500,000, which is between around 8.5 and 14 times greater than the average income in UK. [18] Thus, the lottery earnings from both lotteries create a large

[18] In 2019 the average household income in UK was £35,300 - *Source*: ONS. Some researchers might find that the median is a better approximation for household income. According to the ONS, in 2019 it was £29,400, which makes the impact of the lottery shock around 10 times greater, in the worst-case scenario, and 17 times greater, in the best-case scenario, than the household income median.

enough income shock for households to allocate more resources to increase their consumption expenditures. Moreover, the fact that the prize is shared with more than one household in each winning postcode, makes the analysis more heterogeneous, as different households can be part of the treatment group. Besides, winners are clustered and located in a very specific zone, making them easier to be identified in our sample.

The Millions Postcode Lottery occurs every month in the UK by *Novamedia*, the company that owns the People's Postcode Lottery, whereas the £30.000 lottery occurs weekly. In 2019, around 14% of the British Households participated in the People's Postcode Lottery, and such ratio increases year by year - in 2012 only 1% of British Households played it. Thus, we use the Millions and the £30.000 Postcode Lottery as a quasi-experiment to estimate the effects of windfalls gains on household consumption behavior.

This chapter uses the Understanding Society secure data from years 2010 to 2019, where we have data about household consumption expenditures for different goods, their earnings and household characteristics, as well as individual characteristics of household members. On top of that, we also have precise information about where individuals live, making it easier to locate those that are the potential winners of the British Postcode Lottery. Given the innate randomness of the income shock and the absenteeism of externalities between winning and non-winning postcodes, households living in non-winning postcodes establish an accurate control group for those households living in the winning postcodes, our treatment group. Therefore, we can check for the effects of positive income shocks on household consumption behavior and the effects in leisure and time reallocation in their labor supply.

Our estimates, in first instance, are obtained by using a differences-in-differences estimation comparing the consumer behavior of winner households and non-winner ones in a developed economy with high presence in formal capital, financial and insurance markets. However, the estimates of this regression capture the average lottery earnings per postcode on household consumption. Hence, we should expect the PIH to be feasible due to the lottery income shock. However, we find that households that live in winning postcodes of the Millions Postcode Lottery, spend more on average, than those households that live in non-winning postcodes. This result implies that the PIH does not hold under this scenario.

Later, we perform a two-way fixed effects analysis to perform the analysis of the Engel curves of specific types of goods. In this case, we find that both, durable and non-durable goods, exhibit changes in consumption due to the million's lottery income shock only; implying that households that live in winning regions increase their consumption significantly. More precisely, we find that a 10% increase in total expenditures, previously instrumented with the million's lottery income shock, leads to an increase of 13.8% in household consumption for durable goods (observing an important increase in the value of cars) and of 10.4% in non-durable goods consumption. The estimates show that durable and non-durable have a unit-elastic effect towards the income shock, meaning that these two types of goods react proportionally to a shock to total household expenditures. Furthermore, we observe that the estimated elasticities are not significantly different from each other, meaning that household consumption for durable and non-durable goods react similarly to the lottery income shock. Hence, the estimated results for non-durable goods do not hold with the theoretical predictions by Cerletti and Pijoan-Mas (2014), where households use the income shock to increase their durable goods consumption only and non-durable goods consumption should not exhibit changes due to positive income shocks.

On the other hand, when we analyze the effects of the £30.000 lottery winnings we do not find significant changes in household consumption expenditures, neither on total household expenditures. Therefore, households living in winning postcodes of this lottery do not alter their expenditures, relative to households that live in non-winning postcodes. Therefore, the PIH holds under this scenario. Moreover, we observe that household savings do significantly increase by 37.79% in winning postcodes, due to the £30.000 lottery shock.

This chapter belongs to a growing literature on the response of household consumption decisions to exogenous income shocks, which tests the hypotheses presented earlier in this section. Focusing first on individual/household consumption behavior towards durable and non-durable goods,[19] the type of goods analyzed in this chapter, theory predicts that for non-durable goods there are not big changes in consumption patterns when a positive income shock occurs. Hence, most of the money is dedicated to savings and thus, consumer's behavior should not experience changes (see Michael and Becker, 1973 and Campbell and Cocco, 2007). Moreover, there exists relevant literature in this aspect that helps to understand the individual behavior when positive and transitory income shocks occur. Grossman and Laroque (1990) analyze a model of consumption based on a portfolio of durable goods. The authors assume that the value of durable goods start to depreciate after their purchase and households will buy a new one only when it becomes obsolete, or they sell it. However, if a transitory and positive income shock happens, agents will advance the purchase. Along similar lines, some other studies improved the model of Grossman and Laroque (1990) by

[19] Durable goods are those ones that have a long life period, like cars and non-durables are those ones that are consumed immediately after the purchase, like food.

adding non-durable goods into the utility function. This is the case of Martin (2003), who found a similar result to Grossman and Laroque (1990), where when such income shocks happen agents are prone to bring forward the purchase of a durable good and keep constant consumption of non-durable goods. Cerletti and Pijoan-Mas (2014) design a similar model in which durables and non-durables goods are assumed to be substitutes. Their main finding is that when an aggregate, transitory and positive income shock occurs, households spend more on durables and consumption of non-durable goods remain unchanged - in words, the PIH is violated in all cases.

On top of that, in recent years there have been empirical studies testing these theories and the hypotheses presented above. The most recent case is the first chapter of this thesis, where we test the PIH in Spain using the Spanish Christmas Lottery as income shock. Both chapters complement to each other, given the different size of the local average treatment effects coming from the different types of income shocks we analyze in each chapter. In the Spanish Christmas lottery scenario, the shock can be considered as a universal one, as 75% of the country plays it (see Garvía, 2007); therefore, we consider everyone that lives in the winning region and bought lottery as potential winners of the lottery and, thus, the estimated elasticities will be considering a much larger population size compared to the Postcode Lottery scenario, as not everyone plays this lottery. In this chapter, we consider only those households that live in winning postcodes (which on average there are 15 households per postcode in the UK) - therefore, the population treated in each scenario is significantly different. What is also relevant from Chapter 3, and makes a difference with Chapter 2, is that we are analyzing two different shocks-sizes, as described above, which allows us to analyze and compare the effects of these two lottery prizes into household consumption behavior and the PIH test.

Another closely related paper to this chapter is Kuhn et al. (2011). This paper uses the Dutch Postcode Lottery as income shock to analyze the effects on the PIH and the LCH, among other economic theories, including the social effects and happiness effects that winning the lottery can have into winner participants. In the Dutch case, authors run their own experiment, something that allows them to exactly identify the winners of the Dutch Postcode lottery and its neighbors to identify potential social effects and neighborhood spread effects of the lottery winnings. In both papers, the PIH does not hold, but the results of the Dutch postcode lottery paper are consistent with the LCH, in other words, households in The Netherlands use the lottery winnings to increase their allocation of resources into durable goods only.

These two papers are the most related ones to our study. There is other research based on the

Norwegian lottery winnings affecting household expenditures and savings (see Fagereng et al., 2016) and the payments of Alaska's Permanent Fund affecting the LCH (see Hsieh, 2003). In the first paper, authors find that lottery winnings lead to an increase of household expenditures in the year of winning. Whereas, in the second one, agents do smooth consumption and the PIH and the LCH are satisfied, as Alaskan Households do anticipate the payments from the Permanent Fund.

Despite these four papers mentioned above do a similar analysis to the one we are performing in this chapter; we differ from other studies by analyzing the effects of lottery winnings on consumption expenditures at the household level using a quasi-experiment designed from the British Millions Postcode Lottery. Therefore, we provide a new perspective in the UK, using new approaches of data analysis using a household survey with accurate information about where households live, allowing us to design sophisticated regressions where we control for individual, region and time fixed effects and in addition, we also include an interaction term between time and region fixed effects to control for any event occurred in any specific region at a specific period of time.

Therefore, the novelty of this chapter lies in the use of two different lottery income shocks to estimates income-elasticities on household consumption and also, in the relevance of the results obtained, which to our best knowledge, there is no other paper which finds that non-durable goods exhibit changes in consumption and do not smooth consumption under a positive, and transitory income shock,[20] like the Millions Postcode Lottery prize.

This chapter is structured in 7 further sections. The following section describes how the postcode lottery works. Section 3 focuses on the theoretical literature review. Section 4 does a description of the data set and its main variables of the study. In section 5 we describe the identification strategy of the project. Section 6 describes the regression approach we use. Section 7 presents the obtained results and finally Section 8 concludes the chapter.

2 The UK People's Postcode Lottery

The People's Postcode Lottery is one of the most popular lotteries played in the UK, with a share of 14% of its households playing it in 2019 according to the Annual Reports of the People's Postcode Lottery (see Table 20). Contrary to the National Lottery, which is the most popular lottery in

[20] If someone knows a paper where such thing occurs as well, I would really appreciate if you could send me the reference.

the UK (with 70% of the population playing it), the Postcode Lottery is a charity lottery game: 32% of its revenues are donated to charity. The Postcode Lottery was first launched in UK in 2005 as a trial scheme by the Dutch company *Novamedia BV*, which released its first postcode lottery in the Netherlands in 1989. It first started in the North of England and after its success, it was expanded through Scotland in 2007 and Wales in 2010, same year in which the lottery was distributed officially across the entire country. Since its start, the number of households playing the People's Postcode Lottery is increasing over the years, as we observe in Figure 3.

Each ticket number of the Postcode Lottery is the postcode number where the participant lives - each postcode in the UK is based on a 6-digit number. Hence, the likelihood of a household winning the Postcode Lottery is completely random and, conditional on participating in the raffle, is equal to one over the number of postcodes in the UK (approximately, 1.7 million). In addition, if more than one household in the same winning postcode participates in the draw, all players collect the prize. Therefore, the number of tickets purchased do not affect the likelihood of winning, simply the amount won conditional on winning (see Kuhn et al., 2011).

Households who want to participate in the People's Postcode Lottery need to pay in advance a monthly subscription of £10 to enter in every lottery draw. A household who pays the monthly subscription automatically participates in the raffle of all Postcode Lottery games that are drawn. The maximum amount of tickets/participations that participants can purchase is three.

There are three main types of Postcodes Lottery games played: the Millions Postcode Lottery, which takes place monthly since 2014,[21] in which each ticket holder living in the winning postcode wins an equal share of the total prize of £3 million.[22] This is the lottery this chapter focuses on, as it is the one presenting the largest shock in income, allowing households to increase their resources in consumption demand allocation. Specifically, a winner household receives an amount between £300,000 and £500,000 on average, which is between 8.5 and 14 times greater than the average income in UK.[23] There is also the £30,000 lottery which is played weekly, and the winner participants earn £30,000 per ticket purchased. We also analyze the effects of this lottery as a robustness check, to test whether it has any effect in consumption expenditures or not. The

[21]Before, in 2013 the draw took place 10 times spread along the year; in 2011 and 2012 it took place once every quarter and from 2008 to 2010 it was played once per year.

[22]Under the Gambling Act 2005 restrictions, a household who wins the Postcode Lottery can get a maximum prize of £500,000 or a 10% of the total draw proceeds. However, before July 2020, the maximum prize a household could get was £400,000.

[23]As explained in the introduction, the average household income in UK was £35,300 in 2019, according to the ONS source.

remaining lottery is the daily postcode lottery, where individuals can get a prize of £1,000 per ticket bought.

In addition to these lotteries, where the award is a monetary prize, there are also two extra raffles played monthly. One awards with a BWM plus £25,000 for each ticket holder who lives in the winning postcode, and the other one grants with £5,000 on holidays voucher plus £2,000 for expenses for each participant living in the winning postcode.

3 Theoretical Literature Review

Several theoretical studies have been done in the area of consumer behavior and in analyzing the effects of income shocks on household expenditures.

Research in the past 30 years have provided us some theoretical models of interest to help in explaining the behavior of agents when they face a transitory income shock like a windfall. In this case, given the empirical results we find in this chapter, the focus is on understanding why households tend to increase their consumption expenditure on durables and non-durables goods from a theoretical point of view and understand why savings do not increase.

One of the first papers to analyze this issue was Grossman and Laroque (1990). The authors analyze a model of optimal consumption and portfolio selection based on durable goods included in a *Constant Relative Risk Aversion* (CRRA) utility function. The portfolio of durable goods depreciates after the purchase (and over the time) of these goods and the full value of such goods can only be restored purchasing new ones after selling the old goods. However, the fact of purchasing/selling durable goods create transaction costs, due to market imperfections and potential asymmetries in information. Hence, the more imperfections the market presents, the harder it will be to sell the good; and vice-versa. Grossman and Laroque (1990) make also emphasis on the fact that households renew their durable goods if and only if the good has been depreciated enough to buy a new one, in the absence of a positive income shock. But, before buying a new durable good, the agent needs to sell the old good. In other words, the main idea is that individuals buy a durable (illiquid) good not in every period of time, only from time to time, as individuals incur a transaction cost every time they want to purchase such good. Nonetheless, the main finding in Grossman and Laroque (1990) is that when households experience a positive income shock (like winning the lottery), they wait less time in renewing their durable good of interest. However, this model presents a drawback for our purpose: authors only consider durable goods in the utility function and non-durables are excluded of the equation.

Martin (2003) improves the model proposed by Grossman and Laroque (1990) by adding non-durable goods into the CRRA utility function, where housing is the single durable good agents can get in this setup. In this case, Martin (2003) assumes that durable and non-durable goods are not separable in the utility function and are assumed to be complements. Under the absence of facing a positive income shock by households, the optimal solution from this problem is that whenever individuals want to buy a new durable good, they will save from buying less of the non-durable goods to recover from the expense. Martin (2003) and Grossman and Laroque (1990) agree on the optimal consumption path of durable goods: these are expensive and imply a transaction cost every time agents want to buy them; therefore, individuals are not buying them in every period. They only buy a new durable when it is depreciated enough and immediately after selling the one that they already own. However, when households face a positive income shock, agents are more prone to buy a new house.

Luengo-Prado (2006) designed a similar model to the previous two papers mentioned, where durable and non-durable goods are included in a CRRA utility function, such that durable goods are depreciated in each period after the purchase. In this case, Luengo-Prado (2006) assumes that in each period, households will have a different willingness for buying more or less of durable goods, relative to their desire of purchasing non-durable ones. In other words, if households do not have any willingness to purchase a durable good in a given period, they will only consume non-durable ones. The main result of this paper is that consumption for non-durable goods smooths when higher down payments happen; meaning that when households consume from durable goods, non-durable goods consumption will be adjusted to this payment. Hence, relating Luengo-Prado (2006) to our chapter, we should expect that if household consumption for durable goods increases, due to income shock, households should not alter their consumption for non-durable goods.

Finally, Cerletti and Pijoan-Mas (2014) propose a theoretical model which includes durables and non-durables in a CRRA utility function to analyze the household behavior when a transitory income shock occurs to them, as well as permanent ones.[24] In this setup, the principal assumption is that both goods are substitutes. The main finding in Cerletti and Pijoan-Mas (2014) is that when households experience a transitory positive income shock (i.e., a one-off large shock), durable goods expenditures react significantly compared to non-durable goods. However, when households

[24]In this case, we are focusing on the effect of transitory income shocks into household consumption, as this is the type of income shock we analyze in this chapter.

face such shock there are two situations that might arise if they are credit-constrained: (i) if the constraint is non-binding, the effect of the shock is the same for both types of goods and the size of the change in consumption expenditure is characterized by *the amount of insurance available.* (ii) if the constraint is binding, then, the effect in durable goods consumption is larger than that on non-durable goods, as consumers use the income shock to make the purchase that before they were not able to do, since they were credit constrained. Another important result in the paper is that when households experience a large positive income shock (like the one we are analyzing in this chapter - lottery winnings), it allows them to alleviate their constraints and dedicate the money from the shock to payback any debt they may have or increase their savings.

These papers provide important theoretical results about households' behavior relating to durable and non-durable goods optimal portfolio consumption when a positive income shock occurs. This is what is tested empirically in this chapter, using the Millions Postcode Lottery prize as income shock - a positive, transitory and idiosyncratic income shock. Hence, these proposed papers help us to understand how households such behave under the proposed scenario in this chapter and, thus, know what to expect from our analysis. Specially, we focus on Cerletti and Pijoan-Mas (2014), as it is the paper that describes the closest set-up to the one we are trying to test in this chapter. Therefore, following Cerletti and Pijoan-Mad (2014), we should be able to test empirically whether households decide to spend their income shock in durables instead of non-durable goods or if, in line with the Permanent Income Hypothesis, households smooth consumption and do save the money from the income shock for future periods.

4 Data

The main source of data used in this chapter is the Understanding Society (UKHLS) secure survey database, provided by the UK Data Service (see University of Essex, *Institute for Social and Economic Research*, 2020). Such data is based on survey of a random sample of UK households, which can take place in any given moment of the year, all around United Kingdom, composed by a total of 86,094 households covering the period from 2010 to 2019. Data is presented in the form of panel data, which means that we can observe households' behavior across different years covered by the dataset. The survey contains information about household composition, demographic variables (age, gender, ethnic group, education, marital status, employment status, among others), labor supply, household income and consumption expenditures, car ownership, current house value, among other variables of interest. About monetary variables, households are asked about their income and expenditures in the previous month to the survey, except for utility bills, which are given yearly. To solve for that, we convert all year variables in monthly ones, dividing these over 12, assuming that *households spend the same amount in utility bills in each given month of the year.*

The Understanding Society survey also provides information about where households live - at the city level, in the publicly available version of the dataset. However, the UK Data Service also offers a restricted and secure access data to the Understanding Society, where we can locate where households live. Specifically, there is information about the geographical (x, y) coordinates that help us to identify those coordinates that belong to winning postcodes. In the UK, there are approximately 15 households per postcode on average, with a maximum of 100 households allowed per postcode unit. For the purpose of this chapter, we make use of the restricted data, as it provides an accurate approximation of where households live, making the treatment assignation more precise.

Nonetheless, we need to emphasize that this survey is not a randomized control trial where the surveyors look specifically for those households that live in winning postcodes of the Postcode Lottery (any prize) and try to form a control group with those households living in non-winning ones. Therefore, we cannot ensure that households that have won (or are allocated winning postcodes) are interviewed during the survey. Besides, a postcode can result a winning one if and only if at least one household living there bought a Postcode Lottery ticket. However, we do not observe whether households participate or not in the Postcode Lottery. We neither know the participation rate in the Postcode Lottery by cities or regions. Unfortunately, this is an information that is not provided in the Annual Reports of the People's Postcode Lottery, and they were not able to provide us with more detailed information about disaggregated participation per regions or cities. Therefore, the Annual Reports only contain information about total participation per year in the Postcode Lottery and the winner postcodes. This implies that the treatment and the control group allocation is based exclusively on the postcode where households live, as we cannot observe which households bought lottery.

Table 17 provides a descriptive summary statistics of our sample, where we observe that 0.01% of our sample is located in a postcode that won the Postcode Lottery in the year of the survey; i.e., only 0.01% of the sample belongs to the treatment group. The average age of the heads of the households in our sample is 52 years old and 54% of them are employed. On the other hand, only 10% of the heads of the households in the sample have a second job and not many of them have children, where household size in our sample is less than 3. Finally, the 49% of the heads of the households in our sample are males.

Table 17: Summary statistics

Variable	**Mean**	**Std. Dev.**	**N**
Win_postcode	0.0001	0.04	222,129
Prob. playing lottery	0.033	0.034	222,129
Age	52.18	17.051	222,102
Married	0.37	1.09	222,129
Job_status	0.536	0.499	222,129
Has_sec_job	0.1	0.3	222,129
Num_kids	0.492	0.896	222,129
Ethnic_group	5.179	13.21	220,922
Household size	2.519	1.41	220,922
Male	0.495835	0.4355892	222,129

Source: Understanding Society: Waves 1-10, 2009-2019: Secure Access, UK Data Service. SN: 6676.

In Table 18 and Table 19, we present a summary statistics of our dependent variables differentiating across winning and non-winning postcodes of the Millions and the £30.000 Postcode Lottery, respectively. By doing such differentiation, we want to observe whether there exist significant differences in household consumption expenditures between treatment and control groups. To test for such differences in average consumption across groups, we perform a *t*-test to check for such changes. Hence, given the randomness of the shock and that households interviewed in the survey are randomly selected, we should not expect statistically significant differences in household consumption expenditures across groups.[25]

[25] It needs to be mentioned that for security reasons, some goods we analyze in this chapter do not appear in the summary statistics. More precisely, we do not include consumption expenditures in fuel, house value or rent, mortgage and insurance. The reason is because either the mean or the standard deviation for some goods can lead to inadvertent disclosure of personal information.

Table 18: Summary statistics: consumption levels for the different type of goods - Millions Postcode Lottery

	Winning Postcodes		Non-Winning Postcodes		Testing Differences	
Variable	Mean	Standard Deviation	Mean	Standard Deviation	$t-$test difference	$p-$value
Total expenditures	260837.5	6025.95	190112.5	441025.9	25.51	0.000
Durables	259250	6020.80	169440.20	423764	32.57	0.000
Non-durables	1587.5	400.45	1514.19	882.65	0.42	0.672
Utility bills	1125	340.343	1001.38	677.42	0.84	0.402
Food out of home	75	57.45	78.67	159.79	-0.15	0.883
Food at home	375	95.74	284.62	200.80	2.18	0.029
Car value	9250	6020.80	4257.44	9828.15	1.91	0.056
Alcohol and tobacco	12.5	9.57	44.31	81.50	-7.67	0.000
Savings	3350.56	1163.34	78.88	7120.37	6.49	0.000

Source: Understanding Society: Waves 1-10, 2009-2019: Secure Access, UK Data Service. SN: 6676. The summary statistics presents the monthly expenditures values, differentiating between those households who live in winning postcodes of the Millions Postcode Lottery and those who do not, in Pounds. For safety reasons and to avoid inadvertent disclosure of personal information, some goods have been omitted from the summary statistics table.

Moreover, we also performed an exogeneity test of the treatment by running a linear probability model, regressing a binary variable indicating whether the household is located in a winning postcode on all individual controls included in Table 17 - which are the control variables included in our regression analysis later. In this case we fail to reject the null hypothesis that all estimated coefficients are equal to zero; specifically, the resulting F-tests for both treatments lotteries are 0.80, implying that our identifying assumption, described in Section 5, holds (see Table 29 in the Appendix of this chapter).

Table 19: Summary statistics: consumption levels for the different type of goods - £30.000 Postcode Lottery

	Winning Postcodes		Non-Winning Postcodes		Testing Differences	
Variable	Mean	Standard Deviation	Mean	Standard Deviation	$t-$test difference	$p-$value
Total expenditures	57289.5	89778.51	206210	462094	-5.01	0.000
Durables	54447.5	89193.84	183514.8	444541.4	-4.37	0.000
Non-Durables	2842	1090.23	1639.21	925.34	3.34	0.001
Utility bills	2335	860.28	1073.401	712.86	4.43	0.000
Food out of home	117	88.77	86.96	144.89	1.02	0.306
Food at home	390	175.66	316.96	215.99	1.26	0.209
Car value	3500	3338.10	4840.09	10243.85	-1.21	0.225
Monthly rent	3127.5	4316.36	1168.9	6279.94	1.37	0.170
Savings	3662.17	1282.611	3136.45	4406.47	1.24	0.215

Source: Understanding Society: Waves 1-10, 2009-2019: Secure Access, UK Data Service. SN: 6676. The summary statistics presents the monthly expenditures values, differentiating between those households who live in winning postcodes of the Millions Postcode Lottery and those who do not, in Pounds. For safety reasons and to avoid inadvertent disclosure of personal information, some goods have been omitted from the summary statistics table.

We observe from Table 18 that there exists differences across postcodes in household consumption expenditures. More precisely, the average of total expenditures for those households that live in winning postcodes is £70,725 greater than for those that live in non-winning postcodes. Such effect is statistically significant at the 1% significance level. Analyzing the aggregate of durable and non-durable goods, Table 18 shows that households that live in winning postcodes spend £89,810 more in durable goods than those living in non-winning postcodes, being this effect statistically significant. On the other hand, we do not observe significant differences in non-durable goods consumption. Going in further detail, the car value's is greater in winning postcodes by £4,993, but we do not find this effect to be statistically significant. In the case of savings, these are £3,272 higher for those households living in winning postcodes relative to households living in other postcodes, where such effect is statistically significant at the 1% significance level.

In Table 19 we present the summary statistics for household consumption expenditures of different types of goods, but in this case we differentiate across winning and non-winning postcodes of the £30.000 Postcode Lottery. Contrary to what happened with the Millions Postcode Lottery, expenditures do significantly decrease in winning postcodes (more precisely households in non-winning postcodes spend £148,921 more than in winning ones, on average), as well as durable goods, where households spend £129.094 more in non-winning postcodes. Observing at other goods, we observe a significant increase in household utility bills by £1261 and non-durable goods consumption by £1203 in winning postcodes, relative to households living in non-winning postcodes. For the remaining type of goods, we do not observe significant changes in household consumption behavior across groups. On the other hand, we do observe an increase in household savings in winning postcodes, which might help to explain why expenditures do not increase in those households belonging to the treatment group; but such difference is not statistically significant.

In general, differences in household consumption expenditures are positive towards winning postcodes of the Millions Postcode Lottery and thus, we believe that there is an effect in household consumption behavior due to the lottery income shock. Therefore, these patterns suggest that the survey includes households that, indeed, played and won the Millions Postcode Lottery. However, this cannot be said from the observed differences from the £30.000 Postcode Lottery, where changes in household consumption across groups are not statistically significant for the majority of goods and total household expenditures are significantly lower in winning postcodes than in non-winning ones.

Nevertheless, these differences do not tell us what the expenditures elasticity of specific consumption categories is, because only a small fraction of households in the winning postcodes are likely to have played and won the lottery. For this reason, we estimate a structural model of household consumption to estimate the effect of changes in household expenditures on spending in different consumption categories, as explained later in Section 6.

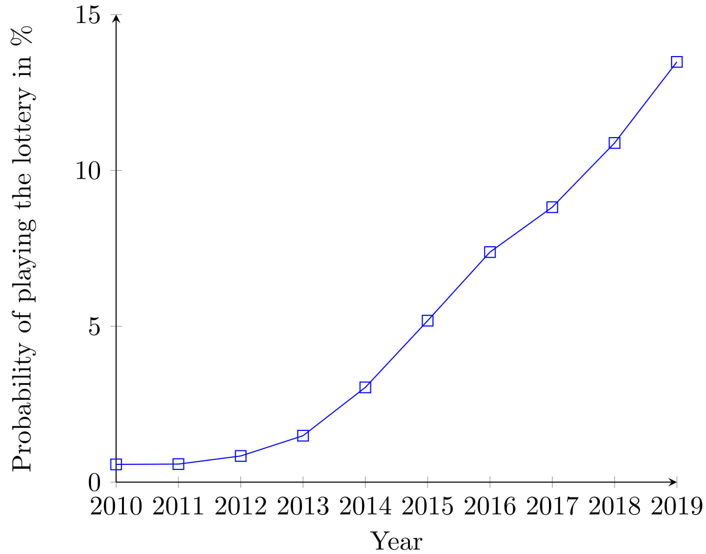
Table 20: Lottery Participation Across Years

Year	**Participation**	**Households**	**% of people playing th PCL**
2010	153899	26240000	0.57%
2011	171277	26409000	0.58%
2012	225179	26620000	0.84%
2013	397896	26663000	1.49%
2014	813141	26734000	3.04%
2015	1400000	27046000	5.18%
2016	2000905	27109000	7.38%
2017	2400000	27226000	8.82%
2018	3000000	27576000	10.88%
2019	3750000	27824000	13.48%

Source: Annual Records, Postcode Lottery.
Participation refers to the amount of households who played the PCL in each given year.
All these data is according to the Labor Force Survey (LFS) and the Annual Reports provided by the People's Postcode Lottery website.

Table 20 provides detailed information about the number of households that play the Postcode Lottery every year. It also shows the total households in the United Kingdom. We can observe the increasing participation trend in Figure 3, where the fraction of UK households that play the Postcode Lottery during the first years it was launched is very low, but it increases over the years, reaching a participation rate of 13% in the entire country in 2019. However, and as mentioned before, we do not know with certainty which households play the Postcode lottery and which ones do not. This creates an important limitation to this chapter, as it does not allow us to exactly identify potential winners of the Postcode Lottery. Thus, we cannot assert whether a household that live in a winning postcode is a real winner of the lottery or just a neighbor of a winner household. Then, to adjust for this drawback in the treatment group formation, we use the probability that a household has of playing the Postcode Lottery each year as it is a useful instrument. This is because the likelihood that any single household played the lottery in a particular year is, arguably, increasing in the proportion of UK households that played the lottery a given year.

Figure 3: Probability of playing the People's Postcode Lottery in UK per year



5 Empirical Strategy

The identification strategy of this chapter is based on the idea that winning the Millions or the £30.000 Postcode Lottery is akin to a random income shock. Therefore, by assuming this as a complete exogenous shock into income, we can study the consumption behavior of households in winning postcodes.

Nevertheless, there are few drawbacks that arise from this project given the available data set we have:

1. We do not know the exact individuals that bought lottery tickets. However, the annual reports of the UK Postcode Lottery [26] tell us the number of households who plays the UK Postcode Lottery in each year. Therefore, by dividing it over the total amount of households in the UK, we can get the probability that each household plays the lottery in a given year.

2. Neither do we know who are the individuals that won the lottery. However, we know the postcode where they live. Hence, we can say that those that live in a winning postcode have a positive probability of being lottery winners while residents in other postcodes have

[26] See PCL website

a zero probability of being lottery winners. Moreover, given the significant differences in expenditures provided in Table 18, we should suspect that those identified in the winning group of the Millions Postcode Lottery are, indeed, real winners.

Given these circumstances, we create two dummy variables indicating whether a household lives in a winning postcode of the Millions Lottery or the £30.000 Lottery, respectively. These two dummies can work as potential instrumental variables for the *logarithm of total expenditures* -our endogenous variable- when estimating the Engel curves for household consumption for different types of goods. However, the relevance condition is fulfilled if and only if the lottery shock significantly affects total expenditures and significantly changes consumption expenditures in the reduced form estimation. The exogeneity condition is automatically satisfied, given the randomness of the shock.

We also generate a variable which computes the probability that a household has of participating in the postcode lottery each year. We interact this variable with the lottery shock, described above. The interaction term captures the effect of the lottery income shock, given the likelihood that a household has of being a lottery participant, into consumption expenditures. Thus, we include the interaction term in the set of instruments for the logarithm of total expenditures. Hence, we regress total expenditures on the interaction term generated: the likelihood of playing the lottery and the dummy variable indicating whether a household lives in a winning lottery postcode, and each variable per se. If the set of instrumental variables significantly affects total expenditures, then the lottery shock is a valid instrument, as the relevance condition is fulfilled. [27]

Moreover, the parameter we want to identify is how the logarithm of total expenditures instrumented with the set of instruments described above affects consumption demands, i.e., see how sensitive demand for goods is with respect to the income shock.

Our dataset is a panel data, containing household information for several waves across years. This allows us to introduce household and time fixed effects in our estimations, as well as region fixed effects.

Therefore, our identifying assumption is that in a given region and year, the change over time in consumption expenditures for those households located in winning postcodes are, on average,

[27] We use only that lottery shock that significantly affects consumption expenditures in its reduced form and total expenditures. If both lotteries -Millions Postcode Lottery and £30.000 Postcode Lottery- have a significant effect, both will be included in the set of instruments. If not, we only include that lottery shock that has a significant impact on expenditures.

comparable to the change in consumption expenditures over time for those households located in non-winning postcodes. The choice of the winning postcode in any given year is exogenous and random, implying no obvious reason to say that it is correlated with other individual, city and year shocks. Recall that the Millions Postcode Lottery draws take place every month since 2014, and the £30.000 Postcode Lottery takes place weekly; meaning that along the year we have several postcodes that win the postcode lottery and can be part of our treatment group.

6 The econometric model

In this section, we specify the regression models of consumption demand in the presence of the lottery income shock. We first investigate, using a reduced-form model, whether living in a winning postcode of the Millions or the £30.000 Postcode Lottery has an immediate effect on household consumption behavior for different types of goods. Then, we proceed with an instrumental variable regression analysis to estimate the Engel curves for specific types of goods.

In the first stage, we test whether the fact of living in a winning postcode of the Millions or the £30.000 Postcode Lottery influences the logarithm of total expenditures, our endogenous variable. However, we only use the lottery shock if it significantly affects consumption in the reduced form. Otherwise, it will not have any effect on total expenditures and thus, the relevance condition will not be satisfied. If such condition is satisfied, we will include both shocks as a set of instrumental variables, if not, we only include that one that satisfies this condition, as the exogeneity condition is automatically fulfilled given the randomness of the income shock. The second step consists in analyzing whether the logarithm of total expenditures has an impact on household consumption behavior of specific good categories.

Finally, we want to test whether the lottery income shocks have implications in the number of hours worked by individuals that live in winning postcodes, compared to individuals living in non-winning ones, and see if they keep them unchanged, or per contrary, individuals do decrease their labor hours and dedicate more time to leisure. In this case, data provides us this information at the individual level.

6.1 Household consumption analysis

Testing first the pure effect of the lottery shock on consumption, we assume that the reduced form equation for household consumption is as follows:

$$\log(c^g_{h,t}) = beta_0 + \beta_1 prob\_lottery_{h,t} + \beta_2 win_{h,t} + \beta_3 win_{h,t} \times prob\_lottery_{h,t} + u_{h,t} \tag{9}$$

where $c^g_{h,t}$ represents the consumption expenditure for good g of household h at year t. $prob\_lottery_{h,t}$ denotes the probability that each British Household has of playing the Postcode Lottery each year; $win_{h,t}$ is a dummy variable taking value one if the household lives in the winning postcode of the Postcode Lottery and zero otherwise; and $win_{h,t} \times prob\_lottery_{h,t}$ is the interaction term between these two variables. Finally, $u_{h,t}$ represents the error term. β_2 and β_3 are our coefficients of interest. The first one measures the own lottery income effect, i.e., the increase/decrease in household consumption in winning postcodes, relative to those households that live in non-winning postcodes. The second one, represents the expected effect of the lottery shock in winning postcodes (compared with non-winning ones) into household consumption, given the probability a household h has of playing the lottery in year t. We run this regression for the Millions and the £30.000 Postcode Lottery separately.

Equation (9) test the direct average effect of the lottery shock on household expenditures on specific consumption goods. Thus, this regression analysis allows us to test the PIH. However, these estimates do not capture how sensitive consumption of the different goods analyzed is regarding the lottery income shock. To estimate the elasticity of each good with respect to the income shock, we regress the Engel curves on total expenditures.

We use the lottery income shock ($win_{h,t}$ and $win_{h,t} \times prob\_lottery_{h,t}$) as the set of instrumental variables to estimate the effect of total expenditures on consumption of specific good categories. This allows solving the endogeneity issue arising from the total expenditures. The endogeneity results as total expenditures are defined as the sum of expenditures on individual good components, which are endogenous to the decision making (see Pitarakis and Tridimas, 1999). To solve for this issue, we run a panel data instrumental variable regression, in where the first stage regression is:

$$\log(exp_{h,t}) = \beta_0 + \beta_1 win_{h,t} + X'_{h,t}\boldsymbol{\beta_2} + \rho_{h,c} + \eta_h + \tau_t + \rho_{h,c} \times \tau_t + \nu_{h,t} \tag{10}$$

where $X_{h,t}$ is a vector of household characteristics including age, age squared, marital status, employment status, number of children in the household, race of the head of the household and household size. $\rho_{h,c}$ represents region fixed effects which controls for unobserved time-invariant region heterogeneity; η_h is a household fixed, and τ_t is a year fixed effect. Finally, $\nu_{h,t}$ is the error term of the first stage regression. In this case we only include the Millions Postcode Lottery shock, as it is the only one that significantly affects total expenditures and thus, it satisfies the relevance condition, as we describe later in Section 7.

We also include an interaction term between time and region fixed effects in Equation (10), $\rho_{h,c} \times \tau_t$. This is done to control for shocks that are common to households in a specific region in a specific year. Therefore, we make sure that unobserved factors are also included in the regression, despite the randomization of the treatment effect.

In an alternative specification, we expand the set of instruments by including the interaction term between $win_{h,t}$ and $lottery_{h,t}$. Therefore, the second specification for the first stage regression is as follows:

$$\begin{aligned} \log(exp_{h,t}) \quad = \quad & \beta_0 + \beta_1 prob\_lottery_{h,t} + \beta_2 win_{h,t} + \beta_3 win_{h,t} \times prob\_lottery_{h,t} + X'_{h,t}\boldsymbol{\beta_4} + \\ & + \rho_{h,c} + \eta_h + \tau_t + \rho_{h,c} \times \tau_t + \nu_{h,t} \end{aligned} \tag{11}$$

Equation (11) presents a similar specification as in Equation (10), except that we expanded the set of instrumental variables, by including the interaction term $win_{h,t} \times prob\_lottery_{h,t}$.[28]

Next, we test the effect of total expenditures, instrumented with the lottery income shock, on household expenditures in different consumption categories. We use the logarithm of consumption expenditure for good g as dependent variable in the second stage. More precisely, we consider that household consumption for different good categories is given by:

$$\log(c^g_{h,t}) = \gamma_0 + \gamma_1 \log(exp_{h,t}) + \gamma_2 lottery_{h,t} + X'_{h,t}\boldsymbol{\gamma_3} + \rho_{h,c} + \eta_h + \tau_t + \rho_{h,c} \times \tau_t + u_{h,t} \tag{12}$$

where $\log(exp_{h,t})$ is the logarithm of total expenditures defined in Equation (11). As in the first stage, Equation (12) also includes region, household and time fixed effects, as well as the interaction between region and time fixed effects as explained above. Such approach represents a non-linear function, as we use the logarithm of total expenditures as our non-linear variable (see Blundell et al., 1993 [29]). The literature also suggests that age square is another non-linear control variable which plays an important role in consumption behavior analysis, as it tests how preferences of

[28] We also analyze the first-stage regressions in levels, under the same specifications as in Equation (10) and Equation (11). However, the relevance condition for instrumental variables is not satisfied, as the performed F-tests are not statistically significant.

[29] The authors in this paper also introduce the log-expenditures squared variable, but in our case it is not statistically significant, hence we do not include it in our analysis.

agents (i.e., consumption expenditures of different goods) change along the years (see Deaton, 1992 and Disney et al., 2010).

6.2 Labor supply analysis

The final thing we analyze in this chapter is whether the lottery income shock, either the Millions or the £30.000 Postcode Lottery, has an impact on labor supply indicators. In words, we want to check if lottery winnings affect the number of regular hours worked by individuals that live in winning postcodes of the lottery, relative to those individuals that live in non-winning ones. But we also test for the number of extra hours worked and the likelihood of being employed. To do so, we analyze the reduced-form estimation of the lottery income shock on labor supply. Thus, the regression is as follows:

$$\vec{labor}_{i,t} = \beta_0 + \beta_1 lottery_{i,t} + \beta_2 win_{i,t} + \beta_3 win\_lottery_{i,t} + u_{i,t} \tag{13}$$

where $\vec{labor}_{i,t}$ represents the set of labor supply indicators, which are the total number of hours worked and the number of extra hours worked per week by individual i, respectively and a dummy variable indicating whether individual i is employed or not. In this case, we can do the analysis at the individual level, as we have this information available for everyone interviewed in the sample.

We do not proceed with the extended-form analysis by controlling for individual characteristics, because the lottery income shock has no effect on labor supply indicators, as we discuss in the next section, from the estimates in Table 27.

7 Results

Our main interest is to estimate how the Postcode Lottery prize affects consumption behavior of winner households compared to non-winner ones. We start presenting the direct effect of the lottery income shock into household consumption expenditures of different goods, by estimating the reduced-form estimation and later, we present the results for the analysis of the Engel curves. Moreover, we are also interested in testing the effect of the lottery prize in time reallocation, i.e., observe if individuals allocate more time to potential leisure activities by reducing their working hours or not.

7.1 Household consumption behavior

Table 21 shows the results for the reduced-form estimation presented in Equation (9). In this case, we are interested in the estimated effect of winning the Millions Postcode Lottery, win and $win \times prob\_lottery$ variables, on household consumption expenditures of specific goods. Furthermore, we assess whether household consumption behavior significantly responds to the windfall gain or not, on average, by performing a joint significance test between win and $win \times prob\_lottery$ variables - the ones that capture the effect of the lottery income shock. Starting with the interaction term estimates, which capture the expected postcode effect of the lottery prize, we find evidence that the interaction term significantly affects household consumption behavior, causing an increase in household consumption expenditures for almost all goods analyzed to those households that live in winning postcodes and have a non-zero probability for participating in the Postcode Lottery, compared with those that live in non-winning postcodes. This result is true for all goods except for house insurance, house rent, household savings and the aggregate of durable goods expenditures, where the estimated effect is negative.

Focusing first on the aggregates of durable and non-durable goods, we observe that the effect of the interaction term leads to a decrease on household consumption for durable goods, and this effect is statistically significant, whereas we do not find a significant effect of the interaction term estimate for non-durable goods. Looking at household consumption expenditures of specific goods, we observe that the estimated effect of the interaction term leads to an increase in the owned number of cars, as well as the value of the car, for those households that live in winning regions. Besides, compared to households in non-winning postcodes, those living in winning postcodes have a higher house value which might be due to the purchase of house(s) or investment in their existing house(s), and also spend more in eating out. Comparing that with specific non-durable goods, we observe that utility bills or food at home do not exhibit significant changes across households in different postcodes.

Observing the effect of living in the winning postcode (win) itself, we detect that households living in postcodes that won the lottery significantly change their consumption behavior too (relative with those that live in non-winning postcodes), for all analyzed goods except for food outside the house and alcohol consumption. We observe a significant increase in durable and non-durable goods expenditures due to the lottery income shock, as well as for most of the analyzed goods. Combining the estimates of the win coefficient with the interaction term (described above), there is evidence that living in the winning postcode of the Millions Postcode Lottery leads to an increase in household consumption expenditures. However, to check whether this change in household consumption is statistically significant, we need to perform an F-test to check the pure effect of lottery winnings on household consumption expenditures for the different types of goods analyzed. The null hypothesis is as follows:

Table 21: Reduced form estimation - Millions Postcode Lottery

PANEL A	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Utility bills	Fuel	Food out of home	Food at home	Car value	Num. of cars	House value
win	0.884***	-0.874***	0.0579	0.588***	3.534***	1.799***	4.842***
	(6.90)	(-133.80)	(0.30)	(4.35)	(12.92)	(645.72)	(276.26)
prob_lottery	-3.334***	-2.405***	2.304***	-2.558***	12.51***	-0.488***	-0.815*
	(-22.76)	(-18.61)	(19.92)	(-31.33)	(47.87)	(-7.84)	(-2.16)
win× prob_lottery	0.630	2.405***	16.77**	0.848	11.77**	0.488***	0.815*
	(0.15)	(18.61)	(2.70)	(0.27)	(2.92)	(7.84)	(2.16)
F-test	53.02	1555.00	9.07	19.77	2305.33	3900.00	7407.42
p-value	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0002
N	222129	222129	222129	222129	222129	222129	222129
R-squared	0.00258	0.00147	0.00178	0.00497	0.0105	0.000357	0.0000335
PANEL B	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Insurance	Mortgage	Alcohol	Rent	Savings	Durables	Non-durables
win	-1.227***	-1.957***	-0.707	-1.854***	1.741***	2.587***	0.278**
	(-113.11)	(-230.83)	(-0.49)	(-186.17)	(7.16)	(185.79)	(2.59)
prob_lottery	23.40***	-1.505***	-0.839***	1.727***	3.635***	1.687***	-2.638***
	(81.95)	(-8.38)	(-6.72)	(7.92)	(39.09)	(6.93)	(-31.11)
win× prob_lottery	-23.40***	1.505***	9.654	-1.727***	-16.32***	-0.831**	1.745
	(-81.95)	(8.38)	(0.46)	(-7.92)	(-3.36)	(-3.01)	(0.53)
F-test	2487.98	4898.53	0.12	3465.03	49.01	4091.09	8.54
p-value	0.0000	0.0000	0.8852	0.0000	0.0000	0.0000	0.0000
N	222129	222129	222129	222129	139177	222129	222129
R-squared	0.0341	0.000323	0.000209	0.000298	0.0106	0.000236	0.00543

The *win* coefficient reports the effect that living in a winning postcode of the Millions Postcode lottery has on household consumption expenditures. *prob_lottery* estimates how the probability that a given household has of playing the lottery in a given year affects household consumption behavior. Finally, $win \times prob\_lottery$, is the interaction term between the previous two variables. The estimate of this coefficient provides the effect on household consumption expenditures for a given household who has non-zero probability of playing the lottery and lives in a winning postcode, compared to households who live in non-winning postcodes, have zero probability of playing the lottery, or both. In this specification, we only include the set of variables that belong to the lottery income shock, no other controls are included in this set. The constant term is not reported to avoid inadvertent disclosure of personal information. We compute robust standard errors, clustered at the household level. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The F-test performs a joint significant test of the lottery income shock variables (win and $win \times prob\_lottery$), which later will be used as instrumental variables for total expenditures, under the null hypothesis that the lottery income shock has no effect on household consumption behavior. In other words, this test is testing the validity of the PIH.

Hypothesis 1:

H_0 : The lottery income shock has no effect on household consumption behavior.

H_a : The lottery income shock has an effect on household consumption behavior.

We find that the joint significance test leads to a rejection of the null hypothesis, stated under *Hypothesis 1*, for all the analyzed goods, except for alcohol consumption where we fail to reject the null hypothesis. Hence, there is evidence from the reduced-form analysis that there exists an effect on household consumption expenditures for those households that live in the winning postcode of the Millions Postcode Lottery compared with those that live in non-winning ones. Moreover, this effect gains relevance when we estimate the reduced-form on total household expenditures (see column (2) in Table 25). In this case, we find that the lottery income shock significantly increases total household expenditures, being this effect statistically significant when we look at the performed F-test, rejecting the null hypothesis proposed in *Hypothesis 1.* Therefore, under the presented scenario by the Millions Postcode Lottery, we can ensure that the PIH is violated as, on average, households use the lottery winnings to increase their consumption for durable and non-durable goods. However, to confirm the failure of the PIH, in Table 22 we test also for the average effect of the lottery income shock on household consumption for durable and non-durable goods. In other words, we want to see if the average household consumption of durable and non-durable goods in winning postcodes significantly increases compared with non-winning postcodes.

Table 22: Testing the average effect of the Million lottery income shock on durable and non-durable goods

Year	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Durable goods	122.95	122.95	122.86	122.62	122.00	121.14	120.14	119.71	118.90	117.86
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Non-durable goods	5.14	5.14	5.23	5.43	5.91	6.57	7.27	7.71	8.36	9.16
	(0.0004)	(0.0004)	(0.0003)	(0.0002)	(0.0001)	(0.0001)	(0.000)	(0.000)	(0.000)	(0.000)

The aim of this table is to present a formal t-test for the average income effect of the Millions postcode lottery on household consumption for durable and non-durable goods, in each analyzed year. This tests takes into account the estimated effects of win and $win \times prob\_lottery$, and adds them up to get the expected effect of the lottery winnings on household consumption: $\mathbb{E}(c^g_{h,t}) = \beta_{win} \times 1 + \beta_{win \times prob\_lottery} \times 1 \times prob\_lottery$. In this case, we focus on the aggregate effects of durable and non-durable goods and we try to confirm the F-test results in Table 21 about the PIH. In parenthesis we present the p-values of the t-tests.

Table 22 shows the resulting t-test under the null hypothesis that the average effect of the Millions lottery income shock has no effect on durable and non-durable goods and, thus, the PIH is satisfied.

In this case, we observe that in all years, the average effect of the Millions Postcode Lottery winnings on household durable and non-durable goods consumption is statistically significant and, thus, the violation of the PIH is confirmed.

In addition to all the estimations, we also perform the test for the parallel trend assumption. In this case, we want to examine if households can somehow anticipate the lottery income shock and, thus, start increasing consumption of the different goods prior to the lottery draw celebration or, if otherwise, they do not and take the income shock as something completely unexpected. We perform it for the two types of lottery shocks analyzed in this chapter. In Table 32, we observe the results of the test for the Millions Postcode Lottery. We conclude that that there is no increase in consumption in the pre-treatment period due to the lottery winnings; therefore, for the Millions Postcode lottery, the parallel trend assumption is satisfied at the 5% significance level for all goods analyzed. Therefore, we conclude that, in general, households that live in winning postcodes do not increase consumption in the pre-treatment period and, thus, they do not anticipate the lottery winnings.

Table 23 shows the results for the £30.000 Postcode Lottery. In this case, we do not observe many significant changes in household consumption expenditures for the different goods we analyze in this chapter. We observe that the overall effect of the lottery income shock (i.e., the addition of *win* and $win \times prob\_lottery$ variables) is only significant for household consumption goods like food at home and eating out, insurance, mortgage and rent payments. For these specific goods, we observe that, on average, those households that live in winning postcodes of the £30.000 Postcode Lottery significantly decrease their expenditures, relative with households that live in other postcodes, except for rent payments, which, on average, the lottery income shock increases household consumption on this good by 1403%.[30] The remaining goods do not present statistically significant alterations on household consumption due to the lottery winnings. Neither do we observe significant changes in total household expenditures (see column (1) in Table 25). More precisely, winning postcodes spend on average 88.5% less than non-winning ones in total,[31] but this effect is not statistically significant.

However, for those households living in winning postcodes, we observe that savings increase compared with those households that live in non-winning ones, and the effect is statistically significant.

[30]The total effect of lottery winnings is computed by taking the exponential, as our dependent variable is presented in logarithmic terms: $(e^{4.153-38.99\times 0.037} - 1) \times 100) = 1403.48\%$

[31]Taking the exponential to compute the average lottery effect on total household expenditures: $(e^{-2.89+19.67\times 0.037} - 1) \times 100) = -88.5\%$

Table 23: Reduced form estimation - £30.000 Postcode Lottery

PANEL A	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Utility bills	Fuel	Food out of home	Food at home	Car value	Num. of cars	House value
win	0.258	-0.154	-3.727***	-0.400*	2.209	-1.163*	-6.519
	(1.63)	(-1.13)	(-10.60)	(-2.22)	(0.97)	(-2.27)	(-1.86)
prob_lottery	-4.623***	-0.485	-2.208***	-3.039***	15.32***	5.119***	0.852
	(-4.80)	(-0.89)	(-3.37)	(-6.48)	(12.61)	(18.13)	(0.51)
win× prob_lottery	2.467	1.422	50.28***	7.969***	-38.13	6.387	40.69
	(1.52)	(1.23)	(12.58)	(3.54)	(-1.43)	(0.98)	(0.88)
F-test	3.41	1.52	162.92	13.33	2.13	0.75	0.64
p-value	0.0647	0.2181	0.0000	0.0003	0.1440	0.3870	0.4227
N	220817	220817	220817	220817	220817	220817	220817
R-squared	0.0120	0.0118	0.0253	0.0761	0.450	0.132	0.0133
PANEL B	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Insurance	Mortgage	Alcohol	Rent	Savings	Durables	Non-durables
win	-17.76***	-3.110***	-2.024	4.153**	0.823***	-0.956	0.151
	(-3.58)	(-4.75)	(-1.84)	(2.63)	(5.97)	(-0.64)	(1.64)
prob_lottery	-5.425**	-0.379	-0.00983	0.574	0.777***	3.356**	-2.836***
	(-2.71)	(-0.41)	(-0.02)	(0.52)	(3.34)	(2.88)	(-6.27)
win× prob_lottery	147.7***	27.42***	15.19	-38.99**	-13.58***	-3.251	2.531*
	(3.89)	(4.94)	(1.82)	(-3.17)	(-12.56)	(-0.16)	(2.36)
F-test	15.40	24.65	3.29	10.55	180.88	0.05	7.44
p-value	0.0001	0.0000	0.0699	0.0012	0.0000	0.8204	0.0064
N	220817	220817	220817	220817	217943	220817	220817
R-squared	0.264	0.0176	0.0110	0.00600	0.131	0.0224	0.0204

The *win* coefficient reports the effect that living in a winning postcode of the £30.000 Postcode lottery has on household consumption behavior. *prob_lottery* and the interaction term, $win \times prob\_lottery$, variables are as described in Table 21. In this specification, we only include the set of variables that belong to the lottery income shock, no other controls are included in this set. The constant term is not reported to avoid inadvertent disclosure of personal information. We compute robust standard errors, clustered at the household level. *t*-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
The *F*-test performs a joint significant test of the lottery income shock variables (*win* and $win \times prob\_lottery$), which later will be used as instrumental variables for total expenditures, under the null hypothesis that the lottery income shock has no effect on household consumption behavior. In other words, this test is testing the validity of the PIH.

Specifically, on average, savings increase by 37.79% for those households that live in winning postcodes. [32] When performing the *F*-test, under *Hypothesis 1*, we do fail to reject the null hypothesis for the majority of the specific goods analyzed, including the aggregates of durable

[32] Taking the exponential to compute the average lottery effect on total household savings: $(e^{0.823-13.58\times0.037} - 1) \times 100) = 37.79\%$

goods and total expenditures. Therefore, we can assert that those households that live in winning postcodes of the £30.000 Postcode Lottery dedicate the lottery winnings to increase their savings and not to increase their consumption. Given these results, we do not include this prize in our instrumental variable analysis of the elasticity of expenditures on different consumption goods to total expenditures, as the relevance condition is not satisfied. Moreover, we suspect that given these results, the PIH holds under the proposed scenario of the £30.000 Postcode Lottery. This is confirmed when looking at the performed t-test in Table 24 of the average effect of the lottery winnings on durable and non-durable goods. In this case, we fail to reject the null hypothesis that the average winnings of the £30,000 lottery have no impact on household consumption for durable and non-durable goods.

Table 24: Testing the average effect of the £30,000 lottery income shock on durable and non-durable goods

Year	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Durable goods	-0.08	-0.07	-0.07	-0.06	-0.06	-0.06	-0.05	-0.05	-0.05	-0.05
	(0.9362)	(0.9442)	(0.9442)	(0.9522)	(0.9523)	(0.9526)	(0.9601)	(0.9610)	(0.9612)	(0.9612)
Non-durable goods	0.50	0.44	0.38	0.35	0.28	0.23	0.19	0.17	0.17	0.17
	(0.6171)	(0.6599)	(0.7039)	(0.7263)	(0.7795)	(0.8181)	(0.8493)	(0.8650)	(0.8651)	(0.8651)

The aim of this table is to present a formal t-test for the average income effect of the £30,000 postcode lottery on household consumption for durable and non-durable goods, in each analyzed year. This tests takes into account the estimated effects of win and $win \times prob\_lottery$, and adds them up to get the expected effect of the lottery winnings on household consumption: $\mathbb{E}(c^{g}_{h,t}) = \beta_{win} \times 1 + \beta_{win \times prob\_lottery} \times 1 \times prob\_lottery$. In this case, we focus on the aggregate effects of durable and non-durable goods and we try to confirm the F-test results in Table 23 about the PIH. In parenthesis we present the p-values of the t-tests.

Table 25 reports the results for the first stage regression. The full set of estimated coefficients are reported in Table 30 of the Appendix to this chapter. From Equation (10), we present the estimated results in levels (see column (3)) and logarithms (see column (4)). We observe that in both cases, households living in winning postcodes increase their total expenditures compared with households living in non-winning ones. This effect is statistically significant for both specifications, meaning that the relevance condition for instrumental variables is satisfied. However, when performing the F-test for instrumental variables in the first stage, we find that living in the winning postcode is a weak instrument for total expenditures in levels, as the performed F-test is lower than 10. On the other hand, when we estimate total expenditures in logarithms, the performed F-test is 10.62, which is slightly above the threshold of 10.[33]

[33] In econometrics, the rule of thumb to decide whether an instrumental variable is a strong instrument for the

When we introduce the interaction term as a second instrumental variable, as presented in Equation (11), we do not find any effect from the lottery earnings on total expenditures in levels (see column (5)). However, when we estimate the logarithmic specification (see column (6)), we do find that the lottery income shock increases total expenditures for those households located in winning postcodes, and the effect is statistically significant. More precisely, we find that total household expenditures increase by 5.89% in winning postcodes compared with non-winning ones due to the lottery income shock.[34] However, we need to test if this effect is statistically significant and, thus, check if the set of instrumental variables satisfy the relevance condition. Hence, the null hypothesis to test for the relevance condition is as follows:

Table 25: First stage estimation - Total household expenditures

	(1)	(2)	(3)	(4)	(5)	(6)
	Total expenditures	Total expenditures	Total expenditures	Total expenditures	Total expenditures	Total expenditures
win	-2.891*	1.788***	9537.8*	0.0847**	709.2	-0.124***
	(-2.22)	(164.08)	(2.41)	(3.26)	(0.17)	(-3.68)
prob_lottery	0.715	0.285			261527.1	0.708
	(0.98)	(1.68)			(1.63)	(0.97)
win× prob_lottery	19.67	0.564**			134722.0	4.899***
	(1.27)	(2.67)			(1.20)	(6.18)
Specification	Reduced-form £30.000 Lottery	Reduced-form Millions Lottery	Expenditures in levels	Expenditures in logarithms	Expenditures in levels	Expenditures in logarithms
F-test	1.39	134.60	5.81	10.62	1.54	19.42
p-value	0.2385	0.000	0.0159	0.0011	0.2137	0.0000
N	220817	222129	220817	220817	220817	220817
R-squared	0.0254	0.0002368	0.00901	0.0254	0.00902	0.0254

Here we report the estimates for the first stage regression. Variables *win*, *prob_lottery* and the interaction term between these two, $win \times prob\_lottery$ are already described in Table 21. In this case, we use *win* and $win \times prob\_lottery$ as instruments for total household expenditures. All specifications also include household, year and region fixed effects, as well as an interaction term between year and region fixed effects. Moreover, we also include in all regressions a set of household characteristics as, the age of the head of the household and its square, the race of the head and its marital status, the educational level and employment status of the head, the number of kids in the household and the household size. We also include the constant term, which in this case it does not report any inadvertent disclosure of personal information, given the regression specification previously described. We compute robust standard errors. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
The F-test performs the relevance condition test for the instrumental variables (*win* and $win \times prob\_lottery$), where the null hypothesis is that the set of instrumental variables for total expenditures are not relevant.

endogenous variable, is to check if the performed F-test is below or above 10. If it is below 10, then the instrumental variable is a weak one, but if it is above 10, then the instrument is strong.

[34]This effect is computed by taking the exponential of the estimated coefficients *win* and $win \times prob\_lottery$. Therefore, for a household that lives in a winning postcode, the estimated effect of the lottery earnings, on average, on total household expenditures is: $(e^{-0.124+4.899\times0.037} - 1) * 100 = 5.89\%$, where we use prob_lottery = 0.037, the average probability of playing the postcode lottery in our sample, as presented in Table 17.

Hypothesis 2:

$$H_0: \text{ The set of instrumental variables for total expenditures are not relevant.}$$

$$H_a: \text{ The set of instrumental variables for total expenditures are relevant.}$$

The F-test, in Table 25, shows that the set of instrumental variables are valid and strong under the logarithmic specifications only. Therefore, *Hypothesis 2* is rejected when we estimate the logarithm of total expenditures. However, we fail to reject it when we estimate total expenditures in levels. The exogeneity condition is satisfied under the identifying assumption that the winning postcode is independent of other household characteristics. Therefore, we adjust for log-expenditures in the second stage using win and the interaction term, $win \times prob\_lottery$, as instrumental variables. The reason to use both variables as instruments is because the resulting F-test is the highest one and thus, it reports the strongest set of possible instrumental variables for the logarithm of total expenditures.

Given the results from the first stage, the second stage focuses on the analysis of household log-consumption for the different types of goods. This procedure helps us to estimate the elasticity of household expenditures of specific goods on total household expenditures.

In Table 31 we report the estimated results for the second stage regression presented in Equation (12). Analyzing first the aggregates of durable and non-durable goods, we find that for the case of durable goods, the estimated effect is 1.38, which means that a 10% increase in household total expenditures leads to a 13.8% increase in household consumption for durable goods, and this effect is statistically significant. In this case the estimated effect is elastic, according to the theory, as it is grater that 1, in other words, durable goods are sensitive to a shock to total household expenditures. According to the theoretical explanations in Section 3, this finding is in line with theoretical results proposed by Grossman and Laroque (1990), where households who experience a large income shock use it to renew their durable goods and buy new ones, and Cerletti and Pijoan-Man (2014), where households use the income shock to increase durable goods consumption or payback their debts, in case households are liquidity constrained. However, according to Cerletti and Pijoan-Mas (2014), non-durable goods consumption should not react to income shocks and, thus, we should expect the estimated effect of total expenditures on non-durable goods to be inelastic - below 1. Nonetheless, according to our estimates, the estimated coefficient for non-durable goods is 1.04, implying a 10% increase in household total expenditures leads to an increase of 10.4% in household non-durable goods consumption.

However, we need to formally test whether the estimated elasticities for durable and non-durable goods are greater than one or not. The results of the elasticity test are shown in Table 31, where the null hypothesis is that household consumption expenditures of specific goods are unit elastic to total household expenditures. In this case we fail to reject the null hypothesis for durable and non-durable goods, which implies that durable and non-durable goods are unit elastic to total household expenditures. Given that, we need to test whether the estimated elasticities for durable and non-durable goods are different from each other or not. In this case, we perform a t-test under the null hypothesis that the estimated elasticities for durable and non-durable goods are the same. In this case we fail to reject the null hypothesis, as the estimated t-test is 1.09 with a p-value equals to 0.3048.

Analyzing the different types of goods separately, we observe that in general, there exist a positive effect of total expenditures on household consumption for most of the analyzed goods. Specifically, we find that an increase of 10% in total expenditures by household h leads to an increase of 12.94% in utility bills consumption, of 14.04% in mortgage payments, but also to a decrease in 54.54% in insurance payments, all these effects are statistically significant. What these three goods have in common is that all of them can be considered non-durable ones, as you pay them monthly or as you consume (like utility bills); therefore, one should expect these goods to not react sensitively to total expenditures and thus, have the total expenditures estimates below 1. However, this is not the case for British households, who report potential elastic estimates, implying that utility bills, mortgage and insurance payments are sensitive to a shock to total household expenditures. On the other hand, we find that a 10% increase in total expenditures leads to an increase of 2.39% in fuel consumption and of 5.67% in food consumption, but to a decrease of 4.99% in alcohol consumption and of 1.44% in monthly rent payments; where only the estimated effects for food and alcohol are statistically significant. In this case, we have that these four goods are inelastic to total expenditures, as one should expect.

When analyzing durable goods separately, we find that a 10% increase in total expenditures leads to an increase of 24.24% in the house value, of 66.16% in car values and of 53.8% in eating out, and all these effects are statistically significant. All these estimated effects are elastic, meaning that these three durable goods are sensitive to a shock to total household expenditures, satisfying our initial expectations regarding durable goods. Finally, we find that savings respond negatively to total expenditures: a 10% increase in total expenditures leads to a 17% decrease in household savings, this effect is statistically significant.

To confirm these findings with regards the elasticities, we need to look at the performed elasticity test in Table 31. In this case we reject the null hypothesis that the estimated elasticities for the different types of goods are unit-elastic to total household expenditures, except for utility bills.

Table 26: Second stage estimation - adjusting log-expenditures with two instrumental variables: win and $win \times prob\_lottery$

PANEL A	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Utility bills	Fuel	Food out of home	Food at home	Car value	Num. of cars	House value
log-expenditures	1.294***	0.239	5.380***	0.567***	6.616***	-0.956***	2.424***
	(5.66)	(1.85)	(34.50)	(5.53)	(22.73)	(-17.48)	(6.32)
Household controls	✓	✓	✓	✓	✓	✓	✓
Fixed effects	✓	✓	✓	✓	✓	✓	✓
Elasticity test	1.29	34.69	789.05	17.81	372.10	∅	13.76
p-value	0.1976	0.000	0.0000	0.0000	0.0000	∅	0.0002
N	220817	220817	220817	220817	220817	220817	220817
R-squared	0.0120	0.0118	0.0245	0.0732	0.450	0.132	0.0133
PANEL B	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Insurance	Mortgage	Alcohol	Rent	Savings	Durables	Non-durables
log-expenditures	-5.454***	1.404***	-0.499***	-0.144	-1.700***	1.380***	1.040***
	(-13.27)	(7.72)	(-3.56)	(-0.65)	(-12.94)	(4.69)	(10.26)
Household controls	✓	✓	✓	✓	✓	✓	✓
Fixed effects	✓	✓	✓	✓	✓	✓	✓
Elasticity test	246.49	4.93	114.28	31.47	422.30	1.66	0.15
p-value	0.0000	0.0264	0.0000	0.0000	0.0000	0.1976	0.6985
N	220817	220817	220817	220817	139127	220817	220817
R-squared	0.264	0.0176	0.0110	0.00600	0.0636	0.0224	0.0202

The $log - expenditures$ coefficient reports the estimates for total household expenditures, adjusted in the first stage using win and $win \times prob\_lottery$ as instrumental variables. This captures the elasticity effect of total household expenditures on household consumption expenditures for the different types of goods analyzed. All specifications also include household, year and region fixed effects, as well as an interaction term between year and region fixed effects. Moreover, we also include in all regressions a set of household characteristics as, the age of the head of the household and its square, the race of the head and its marital status, the educational level and employment status of the head, the number of kids in the household and the household size. We also include the constant term, which in this case it does not report any inadvertent disclosure of personal information, given the regression specification previously described. We compute robust standard errors. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of estimates available in Table 31 of the Appendix of this chapter. The reported Elasticity test examines the elasticity effect of expenditures towards household consumption, in words, whether the estimates are different from one.

Thus, the main finding of this chapter is that durable and non-durable goods are unit-elastic to total expenditures and thus, these react similarly to a positive shock caused to total household expenditures. This finding is surprising as it goes against the theoretical predictions previously described in Section 3, where we should expect only durable goods to exhibit a consumption increase but non-durable goods should smooth consumption.

Based on previous empirical studies, we find the case of the Dutch postcode lottery (see Kuhn et al., 2011). This paper focuses on a household consumption behavior analysis in levels where households that won the lottery significantly increase consumption expenditures only on durable goods (cars not included there) by 310€ and eating out by 18€ compared to non-winner households. Household expenditures for different types of non-durable goods do not react to the income shock. Moreover, comparing these results with the first chapter, the main difference is in the estimates for non-durable goods. That chapter also finds that the effect of total expenditures on household consumption for non-durable goods is 0.926, implying that for an increase of 10% in total expenditures, the increase in consumption for non-durable goods is 9.26%. Therefore, for the Spanish case non-durable goods are also unit-elastic to a shock to total household expenditures. But durable goods consumption does react sensitively to a shock to total household expenditures (with an estimate of 1.147). Thus, the the theoretical predictions from Cerletti and Pijoan-Mas (2014) hold only for the Dutch lottery, as durable goods consumption increases due to the lottery income shock, and non-durable goods consumption do not react to it.

Therefore, under the scenario presented in this chapter, where we analyze the effects of winning the Millions Postcode Lottery on household consumption expenditures for different types of goods, the behavior of British households for non-durable goods is different compared to what the theory predicts, or related empirical papers found.

7.2 Labor supply

Table 27 provides the estimated results for Equation 13. In this case, we observe that neither the Millions nor the £30.000 lottery income shocks significantly affect the allocation of hours worked by individuals that live in winning postcodes of the lotteries, relative to individuals that live in non-winning ones. Neither do we find a significant change, caused by the lottery income shocks, on the probability of being employed. This implies that the lottery income shocks do not have an impact on individuals' labor supply. This effect is confirmed when we perform the F-test under the following hypothesis:

Hypothesis 3:

$$H_0: \text{ The lottery income shock has no effect on labor supply.}$$

$$H_a: \text{ The lottery income shock has an effect on labor supply.}$$

Table 27: Reduced form estimation - Labor Supply

	(1)	(2)	(3)	(4)	(5)	(6)
	Hours Worked	Hours Worked	Extra Hours	Extra Hours	Employed	Employed
win	10.06	0.118	6.412	0.105	-0.0175	-1.029
	(1.41)	(0.05)	(1.41)	(0.03)	(-0.03)	(-0.92)
prob_lottery	-8.736***	-8.742***	-0.201	-0.203	-2.626***	-2.626***
	(-8.35)	(-8.35)	(-0.54)	(-0.55)	(-6.47)	(-6.47)
win× prob_lottery	-361.6	2.843	-188.9	-8.460	2.626	9.184
	(-1.38)	(0.13)	(-1.41)	(-0.28)	(0.12)	(1.02)
F-test	1.90	2.00	0.01	0.03	0.10	1.07
p-value	0.1676	0.1575	0.9043	0.8741	0.7538	0.2999
Lottery treatment	Millions	£30.000	Millions	£30.000	Millions	£30.000
N	409863	409863	409863	409863	409863	409863
R-Squared	0.000634	0.000620	0.0000243	0.00000244	0.000177	0.000177

Here we report the reduced-form estimates for the labor supply indicators. Variables *win*, *prob_lottery* and the interaction term between these two, $win \times prob\_lottery$ are already described in Table 21. All specifications also include household, year and region fixed effects, as well as an interaction term between year and region fixed effects. In these specifications, we only include the set of variables that belong to the lottery income shock, no other controls are included in this set. The constant term is not reported to avoid inadvertent disclosure of personal information. We compute robust standard errors, clustered at the household level. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
The F-test performs a joint significant test of the lottery income shock variables (*win* and $win \times prob\_lottery$) under the null hypothesis that the lottery income shock has no effect on individuals' labor supply.

When looking at the results from the F-tests, we observe that none of them present a statistically significant result, and thus, we fail to reject the null hypothesis proposed in *Hypothesis 3*. Therefore, we can conclude that individuals that live in winning postcodes of the Millions and the £30.000 Postcode Lotteries, do not alter their number of hours worked and neither do they change their employability status, compared to individuals that live in different postcodes.

8 Conclusions

This chapter uses the Millions Postcode Lottery organized by the People's Postcode Lottery in the United Kingdom as a monthly quasi-experiment to analyze household consumption behavior of potential winners of this lottery. Potential winner households are easy to locate, as they are clustered in one single postcode in each raffle; moreover, if more than one household in the winning postcode participates in the lottery, they share the prize, making the analysis heterogeneous, because more than one household can belong to the treatment group. Also, the prize from the Millions Postcode Lottery creates a large enough income shock to allow winner households to allocate more resources to consumption expenditures. Thus, it allows us to test whether households' increase their consumption due to the lottery winnings or they increase their savings.

Using Secure Data from the Understanding Society household survey provided by the UK Data Service, we run a fixed effects instrumental variable analysis, where we estimate the Engel curves using the lottery income shock as instrument for total household expenditures. We find evidence that durable goods are unit-elastic to total household expenditures and react positively to that. More precisely, we find that a 10% increase in total household expenditures leads to an increase of 13.8% in durable goods consumption expenditures. Unexpectedly, non-durable goods are also unit-elastic to total household expenditures. In this case, we find that an increase of 10% in total expenditures causes a 10.4% increase in household non-durable goods consumption. This effect implies that non-durable goods react on the same proportion to an increase on total household expenditures, as the estimated effect is not significantly different than 10%. In summary, the estimated results for durable goods consumption are in line with the theoretical predictions proposed by Cerletti and Pijoan-Mas (2014), however, the findings of this chapter for non-durable goods consumption, contradicts the results by this chapter, where these should not react to the income shock. The same occurs when we compare our estimates with related empirical studies for the Dutch postcode lottery (see Kuhn et al., 2011): durable goods consumption react to the income shock, but non-durable goods smooth consumption. Thus, this is the novelty of the chapter and what makes it unique: British households that live in winning postcodes of the lottery also exhibit changes on both: durable and non-durable goods consumption. Moreover, the elasticity findings are also in line with the results in the first chapter of this thesis.

Furthermore, we find evidence that the PIH is violated under the Millions Postcode Lottery scenario, as households that live in winning postcodes increase their consumption in durable and non-durable goods due to the lottery winnings, compared with households living in non-winning postcodes. However, the PIH holds under the proposed scenario of the £30,000 Postcode Lottery, as households in winning postcodes spend, on average, the same as those households living in non-winning ones.

Nonetheless, this chapter contains some limitations. The main drawback is that we do not know exactly which households won the Millions Postcode lottery. Therefore, we need to use the available information about where they live (which is accurate in the Secure Data Access) as a proxy to locate potential lottery winners of the lottery. However, when looking at the summary statistics by groups in Table 17, we notice that differences in household consumption expenditures are statistically significant, being those higher in winning postcodes. Hence, we have a high suspect that those households can be identified as real winners of the Millions Postcode Lottery. The other main limitation is that we do not know which individuals participate in the People's Postcode Lottery. This is also an important drawback of the chapter as it does not allow us to identify participants of the lottery and thus, potential households that can be part of the treatment group. However, we use the probability a household has of playing the Postcode Lottery each year in UK as a proxy to this drawback.

To conclude this chapter, as we aimed in chapter 2, we expect that the general increase in consumption on almost all goods, due to lottery winnings, leads to policy makers in the United Kingdom (or other developed countries) to design new fiscal policy measures, such tax rebates to stimulate consumption in all type of goods, or reductions in personal tax income, so that households can dispose of more resources to increase consumption.

Appendix Chapter 3

Proportion of households per region

Table 28: Proportion of Households included in the sample per region

Region	Percentage interviwed
North East	2.51%
North West	7.16%
Yorkshire and the Humber	5.39%
East Midlands	5.01%
West Midlands	4.70%
East of England	5.59%
London	4.36%
South East	7.97%
South West	5.55%
Wales	17.35%
Scotland	17.81%
Northern Ireland	16.33%

Source: Understanding Society, Public data base.

Tables of results

Table 29: Identification Strategy - Testing our identifying assumption

	(1)	(2)
	Win	Win
Age	0.000000696*	-1.86e-08
	(2.36)	(-0.07)
Married	-0.00000332	-0.00000288
	(-1.79)	(-1.26)
Job_status	0.000000340	-0.000000326
	(1.24)	(-1.42)
Has_sec_job	0.00000216**	0.00000360**
	(2.66)	(2.68)
numkids	-0.00000664**	0.0000288
	(-2.78)	(1.60)
Ethnic_group	0.000000283*	-0.000000807
	(2.57)	(-1.77)
Household size	-0.000000582	0.00000301
	(-1.01)	(0.55)
Lottery treatment	Millions	£30.000
N	220817	220817
R - Squared	0.0000320	0.0000651
F - test	0.800	0.800

We perform an exogeneity test in this table, in order to check if our identifying assumption presented in section 5 holds for both types of treatments we use in this chapter: the Millions and the £30.000 Postcode Lottery. We run a linear probability model of head of the household characteristics on the dummy variable that indicates whether a household lives in the winning postcode of the lottery or not. The constant term is not reported to avoid inadvertent disclosure of personal information. We compute robust standard errors, clustered at the household level. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
The F-test performs a joint significant test of all variables included in the regression under the null hypothesis that all estimated coefficients are equal to zero.

Table 30: First stage estimation: total log-expenditures

	(1) expenditures	(2) log-expenditures	(3) expenditures	(4) log-expenditures
win	9537.8*	0.0847**	709.2	-0.124***
	(2.41)	(3.26)	(0.17)	(-3.68)
win_lottery			134722.0	4.899***
			(1.20)	(6.18)
lottery			261527.1	0.708
			(1.63)	(0.97)
age	13408.5***	0.143***	13256.0***	0.143***
	(6.87)	(8.54)	(6.90)	(8.51)
age^2	-164.0***	-0.00162***	-164.0***	-0.00162***
	(-11.63)	(-25.13)	(-11.63)	(-25.14)
marital	-3496.0*	-0.0370***	-3371.1*	-0.0381***
	(-2.15)	(-5.46)	(-2.12)	(-5.61)
job_status	194.5	-0.00165**	194.5	-0.00165**
	(0.69)	(-2.61)	(0.69)	(-2.61)
has_sec_job	-369.0	0.0121*	-314.5	0.0117*
	(-0.18)	(2.40)	(-0.15)	(2.32)
num_kids	7117.1*	0.00562	7173.9*	0.00567
	(2.34)	(0.44)	(2.36)	(0.44)
race	139.0	-0.00742***	159.4	-0.00758***
	(1.10)	(-6.21)	(1.31)	(-6.33)
hhousehold size	11288.4***	0.146***	11267.5***	0.146***
	(4.50)	(13.17)	(4.49)	(13.17)
_cons	-27490.7	7.675***	-19510.7	7.692***
	(-0.37)	(9.49)	(-0.26)	(9.51)
F-test	5.81	10.62	1.54	19.42
p-value	0.0159	0.0011	0.2137	0.0000
N	220817	220817	220817	220817
R-Squared	0.00901	0.0254	0.00902	0.0254

Here we report the estimates for the first stage regression. Variables *win*, *prob_lottery* and the interaction term between these two, $win \times prob\_lottery$ are already described in Table 21. In this case, we use *win* and $win \times prob\_lottery$ as instruments for total household expenditures. All specifications also include household, year and region fixed effects, as well as an interaction term between year and region fixed effects. Moreover, we also include in all regressions a set of household characteristics as, the age of the head of the household and its square, the race of the head and its marital status, the educational level and employment status of the head, the number of kids in the household and the household size. We also include the constant term, which in this case it does not report any inadvertent disclosure of personal information, given the regression specification previously described. We compute robust standard errors. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The F-test performs the relevance condition test for the instrumental variables (*win* and $win \times prob\_lottery$), where the null hypothesis is that the set of instrumental variables for total expenditures are not relevant.

Table 31: Second stage estimation - adjusting log-expenditures with two instrumental variables: win and $win \times prob\_lottery$

PANEL A	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Utility bills	Fuel	Food out of home	Food at home	Car value	Num. of cars	House value
log-expenditures	1.294***	0.239	5.380***	0.567***	6.616***	-0.956***	2.424***
	(5.66)	(1.85)	(34.50)	(5.53)	(22.73)	(-17.48)	(6.32)
prob_lottery	-5.535***	-0.654	-6.001***	-3.434***	10.62***	5.793***	-0.883
	(-5.50)	(-1.14)	(-8.74)	(-7.02)	(8.35)	(19.59)	(-0.51)
age	-0.136***	-0.0149	-0.687***	-0.0457**	-0.880***	0.173***	-0.119*
	(-3.90)	(-0.79)	(-28.52)	(-2.82)	(-20.00)	(20.15)	(-1.99)
age^2	0.00147***	0.000125	0.00792***	0.000323*	0.00987***	-0.00193***	0.00122*
	(4.01)	(0.60)	(31.86)	(1.98)	(21.02)	(-21.75)	(1.98)
married	0.0560***	0.0119*	0.108***	-0.143***	0.322***	-0.0729***	0.0827***
	(4.84)	(1.97)	(13.04)	(-22.30)	(22.10)	(-25.85)	(4.36)
job_status	0.000706	0.000127	0.00689***	0.000238	0.00990***	-0.00253***	0.00249
	(0.77)	(0.28)	(10.87)	(0.53)	(8.43)	(-10.53)	(1.86)
has_sec_job	0.0125	-0.00387	-0.0555***	-0.000880	-0.0743***	0.00959***	-0.0230*
	(1.69)	(-1.19)	(-12.17)	(-0.26)	(-8.44)	(5.60)	(-2.06)
num_kids	0.0520**	0.0125	-0.120***	0.0232**	-0.146***	-0.227***	0.0261
	(3.04)	(1.20)	(-11.02)	(2.82)	(-6.59)	(-35.52)	(0.87)
race	0.00390	0.00233*	0.0220***	-0.0244***	0.0306***	-0.0124***	0.0154***
	(1.82)	(2.24)	(13.66)	(-15.87)	(11.12)	(-22.78)	(4.52)
hhsize	-0.136***	0.000870	-0.616***	0.0782***	-0.682***	0.446***	-0.0990
	(-3.72)	(0.04)	(-25.06)	(4.71)	(-14.76)	(45.47)	(-1.59)
_cons	-4.945*	-1.765	-40.00***	0.828	-52.51***	7.178***	-14.80***
	(-2.57)	(-1.71)	(-29.76)	(0.92)	(-21.65)	(15.03)	(-4.44)
N	220817	220817	220817	220817	220817	220817	220817
R-squared	0.0120	0.0118	0.0245	0.0732	0.450	0.132	0.0133
PANEL B	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Insurance	Mortgage	Alcohol	Rent	Savings	Durables	Non-durables
log-expenditures	-5.454***	1.404***	-0.499***	-0.144	-1.700***	1.380***	1.040***
	(-13.27)	(7.72)	(-3.56)	(-0.65)	(-12.94)	(4.69)	(10.26)
prob_lottery	-1.578	-1.374	0.341	0.675	4.847***	2.368	-3.568***
	(-0.75)	(-1.41)	(0.52)	(0.58)	(8.81)	(1.95)	(-7.58)
age	1.454***	-0.122***	0.101***	-0.0151	0.238***	-0.0277	-0.111***
	(18.56)	(-4.39)	(4.83)	(-0.43)	(12.09)	(-0.60)	(-6.92)
age^2	-0.0167***	0.00117***	-0.00116***	0.000278	-0.00271***	0.000259	0.00121***
	(-24.86)	(3.93)	(-5.14)	(0.77)	(-12.84)	(0.56)	(7.55)
married	-0.557***	0.0356***	-0.0630***	0.00190	-0.0746***	0.0264	0.000211
	(-27.03)	(4.07)	(-8.81)	(0.16)	(-9.70)	(1.60)	(0.04)
job_status	-0.00967***	-0.0000251	-0.00189***	-0.00290**	-0.00569***	-0.00124	0.000830
	(-6.92)	(-0.04)	(-3.48)	(-2.73)	(-8.18)	(-1.06)	(1.85)
has_sec_job	0.0972***	-0.0154**	0.00939*	0.0159*	0.00547	-0.0110	0.00153
	(9.20)	(-2.75)	(2.40)	(2.57)	(1.52)	(-1.30)	(0.42)
num_kids	0.00253	0.105***	-0.180***	-0.0734***	-0.310***	-0.0537**	0.0139
	(0.07)	(5.77)	(-16.79)	(-3.43)	(-33.23)	(-2.63)	(1.69)
race	-0.0412***	0.00736***	-0.0110***	-0.00230	-0.0131***	0.00289	-0.00268*
	(-11.11)	(4.43)	(-8.56)	(-1.06)	(-10.43)	(1.01)	(-2.26)
hhsize	0.994***	-0.101***	0.251***	-0.0683	0.665***	-0.00480	-0.0565***
	(15.25)	(-3.32)	(11.22)	(-1.83)	(31.73)	(-0.10)	(-3.41)
_cons	34.45***	-9.551***	5.689***	3.277	19.25***	-4.328	-1.743
	(7.95)	(-6.24)	(4.90)	(1.68)	(17.47)	(-1.66)	(-1.92)
N	220817	220817	220817	220817	139127	220817	220817
R-squared	0.264	0.0176	0.0110	0.00600	0.0636	0.0224	0.0202

The $log-expenditures$ coefficient reports the estimates for total household expenditures, adjusted in the first stage using win and $win \times prob\_lottery$ as instrumental variables. This captures the elasticity effect of total household expenditures on household consumption expenditures for the different types of goods analyzed. All specifications also include household, year and region fixed effects, as well as an interaction term between year and region fixed effects. Moreover, we also include in all regressions a set of household characteristics as, the age of the head of the household and its square, the race of the head and its marital status, the educational level and employment status of the head, the number of kids in the household and the household size. We also include the constant term, which in this case it does not report any inadvertent disclosure of personal information, given the regression specification previously described. We compute robust standard errors. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of estimates available upon request.

Testing the Parallel Trend Assumption

Table 32: Parallel Trend Test - using the Millions lottery

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	bills	fuel	food_out	food_in	car value	num_cars	hsval	insurance	mortgage	altob	rent	saved	durables	non-durables
F-test	2.68	0.19	2.80	13.25	2.15	3.27	1.22	0.06	0.46	2.85	1.47	3.63	0.11	0.83
p-value	(0.1014)	(0.6606)	(0.0942)	(0.0003)	(0.1430)	(0.0707)	(0.2690)	(0.8002)	(0.4987)	(0.0912)	(0.2246)	(0.0568)	(0.7457)	(0.3628)
N	220817	220817	220817	220817	220817	220817	220817	220817	220817	220817	220817	217943	220817	220817

p-values in parentheses.

In this table we present the results for the parallel trend assumption test, where we expect households that live in winning regions of the Millions Postcode Lottery to not anticipate the lottery income shock on their consumption of different goods. In this case, we report the F-test results and the p-values underneath in parenthesis. The main conclusions from this table are that households, who live in winning postcodes, do not alter their consumption in the pre-treatment period for almost all goods. We use the *time varying* treatment command, following the program designed in Cerulli and Ventura (2017).

Chapter 4

The Effect of Religious Constraints on Individual Labor Supply

Abstract

We study the effect of religious constraints on individuals' labor supply decisions in the context of Ramadan, one of the central pillars of Islam, consisting of an entire lunar month of fasting from sunrise to sunset. Using household panel data from Malawi, a country where 14% of its population are Muslims, for the years 2010, 2013 and 2016, we find that females reallocate their time from their jobs to household work. Moreover, for males we do not find such reallocation of hours, as they increase both; their hours worked at their jobs but also in the household. Further in this chapter, we use household data from Bangladesh to support the estimates on the extensive margins, where we do not find any effect on labor force or the likelihood of working during Ramadan. These findings show that we need to go beyond general belief that productivity is decreased in economic terms, as individuals keep being productive at home or working more hours when religious constraints appear.

Keywords: Labor Supply, Ramadan, Hours worked, Wooldridge Correction

JEL Classification: C01, C23, C26, C93, J22, J43 and O12.

1 Introduction

Religion plays an important role in human societies. All religions impose rules of behavior, or discipline, that constrain their followers, depending on the degree of strictness of each religion (see Campante- Yanagizawa-Drott, 2015). Hence, religious practices have an impact on individual behavior and thus, they can affect indirectly (or directly) economic outcomes in different ways: labor supply, productivity, growth, etc. Recent studies on this topic found a negative relationship between religious behavior and economic growth (see Barro and McCleary, 2003 and Barro and McCleary, 2006).

The main goal of this chapter is to analyze how culture and beliefs can affect human behaviors and economic outcomes, as labor supply, given that religions impose some rules of behaviors and practices that might constraint individuals in different ways. However, are religion culture impositions orthogonal to the economic outcomes or not? If the second option holds, then we need to understand why this is happening. In this case, using data from Malawi and Bangladesh, we analyze how religious constraints interfere with individuals' allocation of hours worked and the likelihood of being at work or in the labor force in these countries. Particularly, the focus is on how individuals reallocate their time under such constraints to see if there is any redistribution of time, from work to leisure or vice-versa, when there are constraints that might impose some restrictions on certain individuals/societies. Hence, the main research question of this chapter is *how do religious constraints affect the allocation of labor by individuals in developing countries?* In this case, the Muslim obligation to fast during the month of Ramadan is our religious constraint and the hypothesis we investigate is that Ramadan affects the allocation of hours worked during this period. In addition, we investigate how household composition influences work time allocation. Another aspect of this chapter is the analysis of the extensive margins to examine how Ramadan affects the likelihood of being employed and/or in the labor force.

According to van Ewijk (2011), Ramadan is one of the big Five Pillars of Islam and the *Holiest month of Islam*. It lasts for a period of 29-30 days, in which Muslims are not allowed to consume any type of food or drink, nor tobacco, nor have sexual activity from sunrise to sunset (see Çelen, 2015; Odabasi and Argan, 2009). Another important aspect of Ramadan is that it goes in line with the lunar calendar, in other words, every year it takes place in a different period, and it does not have a fixed date, as for example Christmas does for Christians on December 25. This is a key factor since it also allows us to test for seasonal fixed effects in our analysis. However, we also need to consider our identifying assumption, where individuals that are interviewed during the month of Ramadan do not systematically differ, on average, to individuals interviewed in a different period of the year. Thus, using a differences-in-differences estimation method, we can test whether the change in hours worked is because of a shock that has occurred in a given season (or individual and regional-year shocks) correlated with the Holy month, or because of Ramadan itself.

This chapter uses data from the Life Standard Measurement Survey (LSMS) from Malawi, from the Household Income and Expenditure Survey (HIES) and the Women's Life Choices and Attitudes Survey (WiLCAS) from Bangladesh. The reason why we use these countries is because of data availability, as not many surveys include samples taken during the Ramadan period. In this case, these datasets include surveys taken during the Holy month of Ramadan and thus, are valid ones for our study. Moreover, these two datasets also provide a good balance across treatment and control groups, satisfying our identifying assumption. Furthermore, we use the Malawian data to analyze the intensive and the extensive margins, whereas data from Bangladesh is used only for the estimation of the extensive margins, given that the survey only considers the number of hours worked per year and not per week.

To proceed with the data analysis, we estimate a Heckman model as it allows us to control for sample selection, as there might be individuals in the sample that decided not to work because of personal reasons or prefer to work in the household. The challenge in this chapter is the introduction of individual fixed effects, which is done by adding the Wooldridge correction to the model, allowing for robust-heteroskedastic standard errors clustered at the village level (see Wooldridge, 1995). The main finding in this chapter is that we do not find any evidence at the aggregate level that Ramadan has an impact on labor supply indicators; in other words, we do not observe that individuals that celebrate the Ramadan festivity increase/decrease their productivity during this period. However, when we analyze the effects at the individual level, Ramadan creates a reallocation of time for females. Specifically, we find that females reallocate their time from hours worked under a paid job to household work, increasing by 0.5 hours their housework and reduce by 2 hours their time under paid jobs. However, for Malawian males that take part in Ramadan, we find a general increase in 2.2 worked hours in total, compared to those that do not celebrate Ramadan. Such results are very important as it proves that individuals in Malawi do not reduce productivity as economic indicators may tell us, but they do more housework and dedicate more time to family and friends, thus, working for them.

We also estimate how family composition in Malawi affects labor supply, as we believe family composition also matters on the number of hours worked. This implies that more children at home might lead some members of the household to dedicate more time to them. This may also be the case when there are elderly people in the household that may need care. However, when we differentiate across genders, we find that for female and male subsamples, household composition matters and affects both subsamples similarly. One relevant fact is that having an additional adult or elderly female in the household leads to negative effects on individuals own labor supply allocation. Moreover, we proceed also with the analysis of the extensive margins in Malawi and Bangladesh. The idea is to see how the likelihood of working or being in the labor force is affected during the Holy month of Ramadan. However, we do not find any effects for either of the countries we analyze. These findings indicate that individuals do not enter or exit the labor force during Ramadan but reallocate their labor time across different types of activities.

Based on that, the literature finds that Ramadan has negative implications on economic indicators, as it negatively correlates with economic growth, productivity and labor markets (see

Campante and Yanagizawa-Drott, 2015 and Barro and McCleary, 2003). Specifically, Campante and Yanagizawa-Drott (2015) tests the impact of religious practices on macroeconomic indicators already mentioned and at the aggregate level for labor supply. This paper is the first one to analyze the economic effects of Ramadan on economic growth of all countries around the world, assuming Muslim countries to be the ones that celebrate this tradition. Campante and Yanagizawa-Drott (2015) found that Ramadan has negative effects on labor markets and economic growth, especially for those countries in which fasting takes longer. However, the authors also analyze the subjective well-being due to Ramadan in the Muslim population, where the effects are positive, especially for countries where fasting takes extra hours. They are not the only ones finding that Ramadan has a positive effect on happiness. Harvey et al. (2019) also find a positive effect of Ramadan in favor of pro-social behavior and social interactions through the 'Dictator game' in a lab experiment, where individuals that fast during the day are more generous than those that do not fast and thus, it leads to an increase in their own well-being. Hence, there exists a trade-off between economic growth and subjective well-being during the Ramadan period.

There are other studies analyzing the Ramadan effect on household consumption and labor behavior. Schofield (2014) analyses at the household level how both labor supply and calories consumed are affected during the Ramadan month in India. By using a combination of a natural experiment and a Randomized Control Trial (RCT), in which the treatment group receives food rich in calories, Schofield (2014) finds that religious constraints significantly reduce the labor productivity due to the low consumption of calories. However, households that belong to the treatment group and increase their calorie intake, do not reduce their productivity. Differently to our study, we do not use an RCT. Instead, we use a quasi-experiment where we test in detail the religious constraints effects, such as fasting during Ramadan, on labor supply at the individual level and how households adjust the intensive margins of labor when they have to adjust to cultural constraints, by performing an intrahousehold analysis on labor supply indicators. Thus, the uniqueness of this chapter, to our best knowledge, lies in the fact that we investigate the effects of cultural constraints on labor supply at the individual level, identifying the not only the number of hours worked under paid jobs but also hours worked in the household, to observe if there exists any reallocation of time devoted to labor during the Holy month of Ramadan.

The rest of the chapter is structured in six further sections. Section 2 gives a brief history of the Ramadan tradition in Malawi and Bangladesh. Section 3 gives the background of the labor market in Malawi and Bangladesh. In section 4 we describe the dataset we use. Sections 5 and 6 describe the econometric approach and the estimated results from the regressions, respectively. Finally, Section 7 concludes the chapter.

2 Background

Ramadan is the ninth month of the Islamic Calendar, consisting of mandatory fasting from sunrise to sunset. Fasting during the month of Ramadan is one of the big Five Pillars of Islam and Muslims are required to abstain from food, drinks, smoking and sexual activities during the sunlight hours for 29-30 days, depending on the length of the lunar month. As in every country where Ramadan takes place, food after fasting takes a principal role during the Holy month of Ramadan. There are two main meals that Muslims take during Ramadan, a big one during the sunset (*iftar*) and another one before the sunrise (*suhoor*).[35] However, there are some people exempted from fasting during this period. These groups of people are children under 12 years old, elderly people, those that are ill or women in the period of menstruation or who have recently given birth (see Almond and Mazumder, 2011).

Ramadan is also considered a period in which Muslims enhance their self-control, experience a personal growth and spirituality, developing empathy towards those in need and reinforce their connection with God, being mindful of their religious obligations by going to mosques and prayer-houses daily. Moreover, Ramadan is the time where Muslims choose to pay their annual charity tax (*zakat*), which is another pillar of Islam (see Demiroglu et al., 2017). However, the most important event of Ramadan comes at the end of the festivity. It is called *Eid al-Fitr* and consists of a three-day event where families meet together and break the fast. It is the first time after Ramadan when Muslims are allowed to not fast during the day (see Bone, 1982 and Malawi, CultureGrams 2018).

Ramadan fasting can lead to physical and/or psychological issues caused by the fact of fasting and not being allowed to ingest calories or liquids during the day. Medical research has shown that fasting can entail irritability caused by stress, sleep deprivation, headaches, dehydration, physical exhaustion, among other minor health problems but it rarely leads to major health problems (see Leipier and Molla, 2003). This can have effects on individual labor productivity. Some studies found that a significant number of people reported fatigue and aversion to work, in addition to a reduction in focus at work during the Holy month (see Afifi, 1997 and Karaagaoglu and Yucecan,

[35] One might find different names for the sunset and dawn meals in different papers, depending on the language they use.

2000). With that, many governments in Muslim countries do formally reduce the number of hours worked during the Holy month of Ramadan by one to three hours allowing workers to start and end their workday earlier (see Demiroglu et al., 2017).

Given that Ramadan follows the lunar calendar instead of the solar one, the period of Ramadan corresponds to a different set of dates on the solar calendar each year. Specifically, Ramadan starts about 11-12 days earlier each year according to the solar calendar (see Göçmen et al., 2004). In that aspect, some countries close to the Equator have the (dis)advantage of having the same number of hours of sunlight along the entire year, whereas countries in the Northern Hemisphere, such as Bangladesh, or in the Southern Hemisphere, such as Malawi, have different numbers of sunlight hours during the year. Thus, hours of fasting are not the same in all countries and can differ significantly depending on the location of the country and the season in which Ramadan takes place. Hence, fasting during Ramadan can have a different effect on individuals' labor supply behavior according to whether they live in the Northern or Southern Hemisphere.

Therefore, food and calorie intake also represents one of the most important aspects of Ramadan. Thus, Ramadan meals during sunset and sunrise are rich in calories and carbohydrates (see Shephard, 2012). This is important as individuals need to have energy during the day, as they want to keep their productivity in their jobs.

Hence, with all the restrictions that Ramadan presents, but also with all its opportunities for Muslims to increase their happiness and their social relations, Ramadan seems a good treatment to control for individual productivity and see its implications in the labor market at the individual level.

3 Labor Market Background

Another important aspect is the labor market situation in the countries we analyze in this chapter: Malawi and Bangladesh. This section gives some background about the labor market situation in these countries, the dominant labor sectors and the labor gender gap.

3.1 Malawi

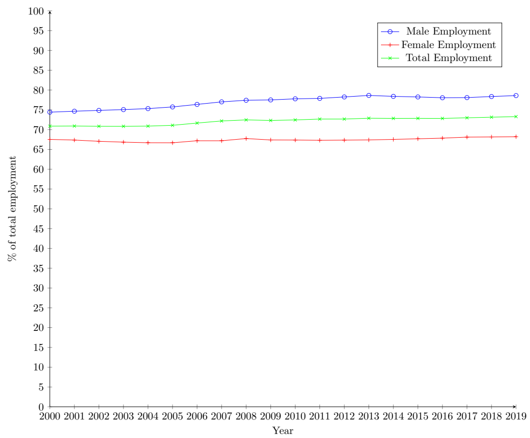
The agricultural sector is the most important one in the labor market of developing economies or sub-Saharan African countries. The same holds true for Malawi where around 71% of the population in 2019 worked in this sector (see indexmundi , checked on April 2020). Figure 1b shows the importance of the agricultural sector in the Malawian labor market, with a 70-80% share of total labor during the last 20 years.

Members of the household provide most of the agricultural labor in farming in Malawi. Around 60% of the farms in Malawi are small, with a land area less than 0.8 hectares. In contrast, just 0.4% of Malawian farmers operate in farms with an area greater than 4.5 hectares (see Julien et al., 2019). This implies that in large households, some household members will typically need to look for work outside the home because of an excess endowment of labor relative to land ownership. When analyzing intra-household allocation of labor, we need to consider that in developing countries, the head of the household make most of the decisions affecting household members; therefore, we can treat him as the social planner of the family (see Dercon and Krisnan, 2000). This means that if there are too many adult members in a household, the head of the household may employ some members to work on the household land (or farming activities) and send others to look for a job outside the home.

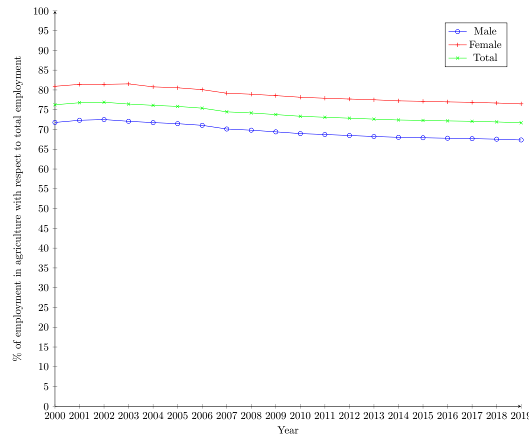
The main issue in sub-Saharan African countries is the investment limitations in irrigation infrastructures, where only 4% of the cropland is irrigated (see Sheahan and Barrett, 2017). As explained in Julien et al. (2019), some constraints are specific to some landowners, and some others are general to everyone. The first ones mainly affect small farms or farmers who produce to subsist. These constraints refer to new technology investments, the use of fertilizers or pesticides, but also to the access to credit markets, which is harder for owners of small lands/farms. Common constraints faced by all farms are climate change and environmental degradation. Each of these constraints affect production and whether directly or indirectly, the labor market.

Figure 4: Employment Gender Gap in Malawi

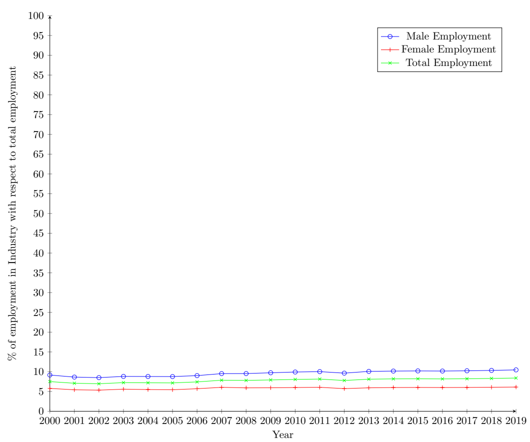
(a) Share of Total Employment



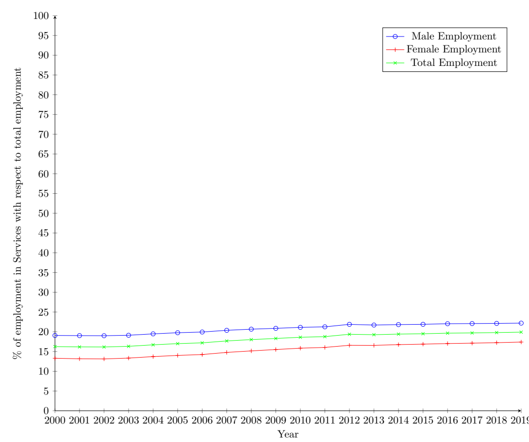
(b) Share of Total Employment in Agriculture



(c) Share of Total Employment in Insdustry



(d) Share of Total Employment in Services



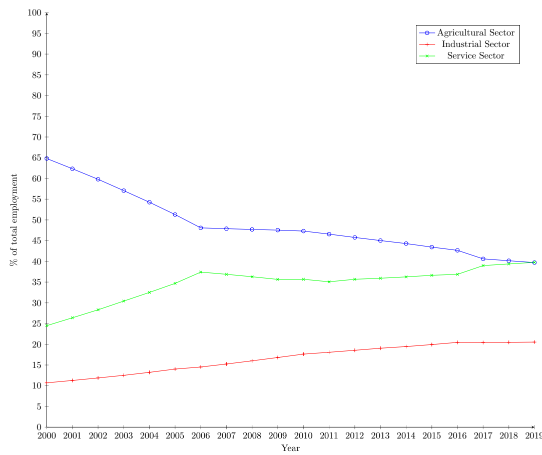
Source: index mundi, ILOSTAT.

Regarding the gender gap in the Malawian labor market, Figure 4 presents the share of males and females employed in the whole country and the share of employment in different sectors by gender. Subfigure 4a shows that the gender gap between males and females has increased over the years. In 2019, there was a 10% gap between males and females in the Malawian labor market. We observe that the majority of the population is employed in the agricultural sector, where the share of females employed in this sector is greater than the share of employed males (as we can also see in van Klaveren, 2009). This fact implies that most females are working on their own household land (see Julien et al., 2019). Nonetheless, when looking at the gender gaps per sector in subfigures 4b, 4c and 4d, we do not observe big differences across males and females.

3.2 Bangladesh

As in Malawi, the agricultural sector is the most important one in Bangladesh, as most of its population works in this sector - approximately 40% of the employed population (see Figure 5). However, the textile industry also plays an important role in the Bangladeshi economy, this industry being the largest one in the world. Indeed, after China, Bangladesh is the largest exporter in textiles, where approximately 80% of its total production is exported, divided between around 60% of total exports to Europe, and around 40% of total exports are to the American continent (see Owuor, 2019).

Figure 5: Share of Total Employment in Bangladesh by sectors



Source: index mundi, ILOSTAT

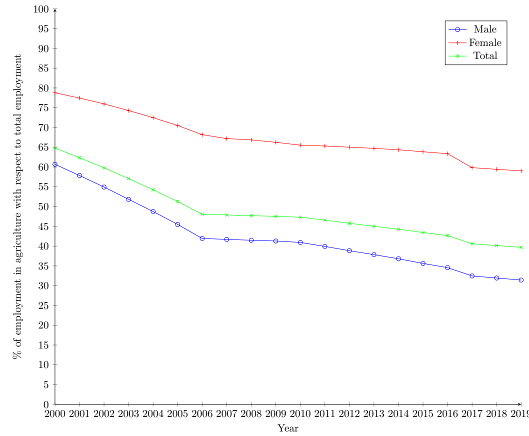
We need to be aware that there are two types of sectors in Bangladesh: the informal sector, which comprises most of the population (86.2%) and the formal sector composed of 13.8% of the population (see Rahman et al., 2018). According to *index mundi*, the informal sector includes *all jobs in unregistered and/or small-scale private unincorporated enterprises that produce goods or services meant for sale or barter.*

Figure 5 presents the evolution of the share of total employment by sectors in Bangladesh, where the industrial/manufacturing and the service sectors are gaining influence in the Bangladeshi labor market over the years and in contrast, the farming sector is losing its importance, even though it is still the one that employs most people in Bangladesh. However, we observe that the trends indicate that the services sector will become the largest employer in the near future.

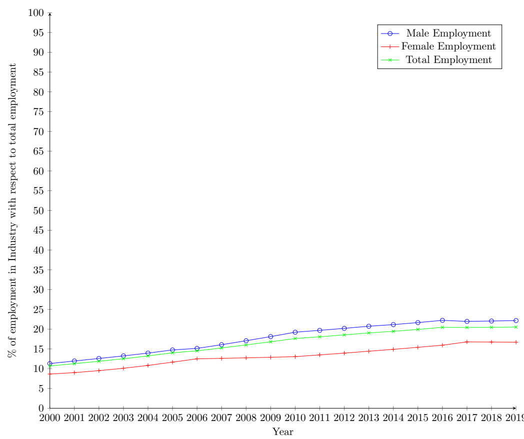
Figure 6: Employment Gender Gap in Bangladesh

(a) Share of Total Employment

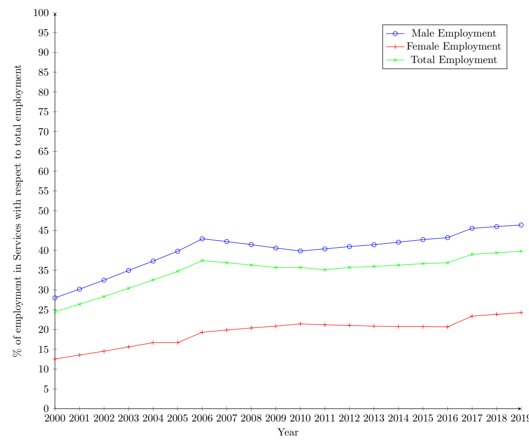
(b) Share of Total Employment in Agriculture



(c) Share of Total Employment in Industry



(d) Share of Total Employment in Services



Source: index mundi, ILOSTAT.

In terms of gender inequality, according to Asaduzzaman et al. (2015), the gender gap between male and female is substantial, where women are mainly engaged in household work and having no power (or very low power) in household decisions. Therefore, males are the ones who dominate household decisions. Despite that, approximately 88% of the females in Bangladesh contribute to increase household earnings (see Asaduzzaman et al., 2015). Going into more detail, Figure 6 shows the gender gap in the Bangladeshi labor market.

Contrary to what we observed for Malawi, Figure 6 shows a clear pattern of gender inequality in the labor market of Bangladesh. Specifically, this trend is clearly shown in subfigure 6a. There is a difference of more than 40% between male and female employment, as it is males who are mostly employed in Bangladesh. Analyzing sector by sector, we observe in subfigure 6b that it is females that are mostly employed in the agricultural sector, whereas only a small proportion of men are employed in this sector. This means that women are the ones who work on the land contrary to males, who nowadays work mainly in the service sector, as we can observe in subfigure 6d. However, subfigure 6c does not show a gender gap, meaning that the proportion of men and women employed in the industrial sector is similar.

4 Data

The first dataset we use in this chapter comes from the Malawian Integrated Household Panel Survey provided by the Malawi National Statistics Office (NSO). The second data source is the Household Income and Expenditure Survey (HIES) and the Women's Life Choices and Attitudes Survey (WiLCAS) from Bangladesh. The HIES survey is conducted by the Bangladesh Bureau of Statistics (BBS) in 2010 and the WiLCAS is conducted by the University of Kent and the University of Malaya in collaboration with Data Analysis and Technical Assistance (DATA) in 2014. We use Malawian data to analyze how Ramadan affects labor supply outcomes at the intensive and extensive margins and the Bangladeshi data is used only to study the extensive margins due to data availability.

The reasons for choosing Malawi as our country of interest are that it has a sizeable population of both Muslims (15%) and non-Muslims and because of the availability of a household panel data survey for which the period of interviews overlaps with the Ramadan period (28% of interviews conducted during the month of Ramadan - see Table 33). Hence, the survey includes a sizeable number of households that we classify as being treated: Muslims interviewed during the Ramadan period. A similar reasoning applies for the choice of Bangladesh, where we have panel data available for two years for households living in rural areas, in which a reasonable size of its total population is Muslim (87%) and a reasonable number of the interviews took place during Ramadan (19%). Thus, the dataset used for Bangladesh also includes a sizeable number of households that belong to the treatment group. In the coming subsections we describe the data structure and its collection in further detail, as well as the merging process between HIES and WiLCAS.

Both country datasets follow a panel data structure, which allows observation of individuals' behavior across different years. A common thing we need to do in both datasets is to establish those individuals that are part of our treatment group and those that belong to the control group. To do so, first we need to identify those individuals that are Muslims and those that are not. The other key feature that determines whether an individual is part of the treatment group is the date on which the interview took place. Muslims interviewed during the month of Ramadan belong to the treatment group; the rest of the sample is our control group.

4.1 Malawi

The data for Malawi constitutes a panel dataset in which individuals are observed in three different years: 2010, 2013 and 2016. The sample is composed of 3231 households randomly selected, of which 15% are Muslim. However, Malawi is a Christian country where, according to the last Population and Housing Census collected in 2018 by the NSO, 83% of the total population follows this religion. In contrast to that, 13.8% of the population is composed of Muslims, most of them living in the southern regions, and 2.1% of its population does not follow any religion.

The Malawian Integrated Household Panel Survey provides information about the number of hours the individuals have worked during the previous week to the survey, both in household work and paid jobs. We also have information about their consumption expenditures, gender, age, marital status, number of children they have, whether they live in a rural or urban area, and whether they have attended school, among other socio-demographic variables. Furthermore, we have access to the agricultural data, in which we have information from each household about the area of land they have in hectares (ha.), whether this is of its own property or not (and who has the land rights), their land production and its productivity livestock, among other variables.

This survey also asks individuals about their religion and beliefs, as well as keeping a record of when the interview took place. These last sources of information allow us to identify which individuals are part of the treatment and control groups. The identification of the treatment comes by interacting two facts: whether the individual follows Islam and whether the survey took place during the Ramadan period. If the value of the interaction term is equal to one, the individual is part of the treatment group; otherwise, he/she is part of the control group.

Another important addition to this data is the generation of the labor force participation group. This will help us later to control for any sample selection when analyzing the intensive margins in the selection equation - discussed later in section 5. In this case, we say that *an individual belongs to the labor force if he/she is actively working or belongs to the working age range.* The working age range is defined as between 15 and 65 years old. Table 33 shows the summary statistics of the main variables this chapter focuses on:

Table 33: Summary statistics

Variable	Mean	Std. Dev.
Housework Hours	5.375	10.933
Paid Job Hours	10.106	17.588
Total Hours	15.48	24.157
Proportion Working	0.671	0.47
Labor Force Participation	0.885	0.32
Islam	0.154	0.36
Ramadan	0.284	0.451
Islam $\times$ Ramadan	0.059	0.237
Age	22.587	18.072
Male	0.489	0.5
School	0.858	0.349
Urban	0.259	0.438
Labor force	0.51	0.5
Household size	5.975	2.385
Land area (ha.)	2.357	9.183
Num. adults	2.355	1.248
Proportion of male children	0.186	0.177
Proportion of female children	0.187	0.177
Proportion of male adults	0.215	0.194
Proportion of female adults	0.227	0.178
Proportion of old males	0.043	0.103
Proportion of old females	0.050	0.103

From Table 33, we observe that on average, people work around 5.4 hours per week in the household and 10 hours in their jobs (those who work outside home). However, the standard deviation is almost twice the mean in both cases, meaning that there is lot of variability across individuals on the allocation of working hours, where some individuals might not work, and some others might double the hours worked from the average. Moreover, when observing at the proportion of people working and that belongs to the labor force in our sample, we detect that 67% are working and 88% of the individuals that took part in the interview belong to the labor force.[36]

Another fact from Table 33 is that only 6% of the sample belongs to the treatment group ($Islam \times Ramadan$). The sample is composed of young people, 22 years old on average, and almost equally divided between males and females: in total there are 49% males and 51% females. We also observe that most of the people have attended school (86%) and live in rural areas (74%). Observing the labor force participation, 51% of the sample is part of it, meaning that half the sample decided not to work for several reasons (they have young children, they decided to stay at home and work for the household, they have many children to take care of, among other reasons). Another important fact from the summary statistics table is that on average there are two adults per household and the average number of people in each household is six. This means that on average there are around four children per household. Finally, the average land area in the household is 2.4 hectares; however, notice that the standard deviation is almost four times higher than the average. According to Julien et al. (2019), Malawi is characterized by not having big areas of land; more precisely, 60% of the farms are small with an area of less than 0.8 hectares, and only 0.4% of the farms are greater than 4.5 hectares.

Finally, we also show the proportion of children, adults and elder people in the household, by gender. On average, we observe that 36% of the household is composed by children and 9% by elderly people, equally divided by males and females. On the other hand, 45% of the household is composed by adult people, where 22% are males and 23% are females.[37] This data will be useful for later, when we estimate the intrahousehold allocation affecting the labor supply during the Ramadan.

4.1.1 Balance Tables - Malawi

The idea of showing balance tables and normalized differences, jointly with the orthogonality test instead of the classic t-test for differences across groups is to check how correlated the main

[36]When we analyze the proportion of males and females that are working and in the labor force in our sample, we find that 68% of the females are working and 88% of them belong to the labor force; whereas 67% of males are working and 69% of them are part of the labor force. Such results are in line with the data presented in Figure 4a.

[37]In this case, we assume that children are composed by individuals under 18 years old. Adults are those individuals who are between 18 and 65 years old, and finally the elderly population is that one above 65 years old.

control variables of our study are within those individuals interviewed during the Ramadan month and those that were not and check how balanced the sample is across the groups of interest (see McKenzie, 2017). In words, we analyze whether households that have been interviewed during the month of Ramadan differ systematically from those that were interviewed in other months. The approach used in this mechanism is the one proposed by Imbens and Rubin (2015) in which a *normalized* difference of 0.25 or less would imply a good signal of balanced data. In this case, a *normalized* difference is defined as *the difference between treatment and control groups means, over the square root of half of the sum of the treatment and control group variances* (see McKenzie, 2017). This helps to see how well distributed the sample across groups is and if there are many differences across them.

We observe in Table 34 that the sample is well-balanced. The proportion of Muslims is greater under the treatment group (21% of the individuals who are interviewed under Ramadan are Muslims, whereas only 13% of the ones interviewed under the control group are Muslims) and the size of land is also significantly different across groups. Therefore, the question should be whether this difference in land size is because of the fact that there is a greater proportion of Muslims in the treatment group or because there are richer individuals in the control group. In order to test that, Table 35 presents the balance summary statistics only for the subsample of Muslims in the dataset.

In this case, when controlling for Muslims only, the panel is well-balanced, and the normalized differences of the land size are not significant anymore. The remaining variables are also balanced except for the area where they live, urban or rural towns. Thus, we can say that our identifying assumption that *individuals interviewed during the month of Ramadan are, on average, identical to those interviewed outside the Ramadan period, in a given year*, is satisfied. Thus, if there are changes in individual labor supply, these should be due to the Ramadan effect rather than to individual effects or region-year shocks correlated with the selection of fasting in a given year.

Table 34: Balance Summary Statistics - Malawi

	(1)		(2)		t-test	Normalized
	Non-Ramadan Month		Ramadan Month		Difference	difference
Variable	N	Mean/SE	N	Mean/SE	(1)-(2)	(1)-(2)
Islam	12018	0.131	4756	0.210	-0.079***	-0.218
		(0.003)		(0.006)		
Age	12018	22.610	4756	22.528	0.082	0.005
		(0.164)		(0.265)		
Male	12018	0.488	4756	0.492	-0.005	-0.009
		(0.005)		(0.007)		
School	10038	0.864	3946	0.845	0.019***	0.054
		(0.003)		(0.006)		
Urban	12018	0.256	4756	0.266	-0.010	-0.022
		(0.004)		(0.006)		
Num. children	12018	2.283	4756	2.315	-0.032	-0.019
		(0.015)		(0.024)		
Household size	12018	5.961	4756	6.010	-0.049	-0.021
		(0.021)		(0.036)		
Land area (ha.)	10519	2.480	4004	2.033	0.447***	0.049
		(0.102)		(0.067)		
Marital status	12018	4.196	4756	4.195	0.001	0.001
		(0.021)		(0.034)		

Notes: The value displayed for t-tests are the differences in the means across the groups. Standard errors are robust. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level. The F-test for the joint orthogonality is 16.44 with a corresponding p-value equal to 0.000.

Table 35: Balance Summary Statistics for Muslims - Malawi

	(1) Non-Ramadan Month		(2) Ramadan Month		T-test Difference	Normalized difference
Variable	N	Mean/SE	N	Mean/SE	(1)-(2)	(1)-(2)
Age	1577	22.478	998	21.893	0.585	0.031
		(0.473)		(0.594)		
Male	1577	0.502	998	0.500	0.002	0.004
		(0.013)		(0.016)		
School	1310	0.736	801	0.718	0.018	0.041
		(0.012)		(0.016)		
Urban	1577	0.219	998	0.147	0.071***	0.182
		(0.010)		(0.011)		
Num. children	1577	2.370	998	2.466	-0.096	-0.062
		(0.039)		(0.048)		
Household size	1577	5.955	998	6.056	-0.101	-0.043
		(0.058)		(0.076)		
Land area (ha.)	1422	1.872	905	1.936	-0.064	-0.018
		(0.088)		(0.125)		
Marital status	1577	4.200	998	4.159	0.040	0.018
		(0.058)		(0.073)		

Notes: The value displayed for t-tests are the differences in the means across the groups. Standard errors are robust. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level. The F-test for the joint orthogonality is 3.90 with a corresponding p-value equal to 0.1%.

4.2 Bangladesh

The HIES and the WiLCAS projects collect household data, where individuals are randomly selected across Bangladesh. The first project collected data in 2010, while the second one takes track only of those households living in rural areas, who were interviewed previously in the HIES. The WiLCAS is a nationally representative survey developed in Bangladesh that took place in 2014 focusing on women between the ages of 20 and 39 to analyze their access to school, education, and childbirth facilities among other variables.[38] To merge both datasets, we need first to identify those individuals in the HIES that will be eligible in the WiLCAS survey (i.e., individuals living in rural areas). After that, we use household and individual identifiers, gender and age to proceed with the merge. The reason for using gender and age (adjusted to the corresponding time-line) is because it helps to match individuals in each household more accurately. After merging both datasets, the sample is composed of 5781 households in both years, leading to a well-balanced panel data structure for the extensive margins' analysis.

According to the Population and Housing Census in Bangladesh, collected in 2011 by the BBS, approximately 90% of the total population in Bangladesh is Muslim, and the country has the fourth-largest Muslim population in the world, after India, Indonesia and Pakistan. The second most followed religion in Bangladesh is Hinduism, where 8.5% of its population follows it, and only 0.6% are Buddhists and 0.4% are Christians.

We need to apply the same modifications to the dataset, as we did previously for Malawi, in other words, generate the treatment and control groups following the same procedure as in Malawi, as well as identifying those individuals who are part of the labor force. For the Bangladeshi case, given data disposal, we say that *an individual takes part in the labor force in Bangladesh if he or she is actively working or looking for a job.*

The reason why we focus only on the extensive margins is because the number of hours worked for each individual are given in a yearly form; therefore, we cannot distinguish between hours worked during and outside of Ramadan. However, we know, on a weekly basis, whether an individual is working or not and whether he/she belongs to the labor force, which makes it feasible to observe the Ramadan effect in labor market outcomes. In words, we know whether the individual works or not and if they are actively looking for a job. This dataset also provides information on such individual characteristics as gender, age, and household size, among other socio-demographic variables.

[38] For more information one can look here: WiLCAS website.

Table 36: Summary statistics

Variable	**Mean**	**Std. Dev.**	**N**
Work	0.36	0.48	52272
Labor Force	0.364	0.4814	52272
Islam	0.879	0.326	63689
Ramadan	0.189	0.392	63689
Islam $\times$ Ramadan	0.167	0.373	63689
Age	26.807	19.213	63614
Male	0.477	0.499	63614
Num. adults	4.207	2.355	63689
Household size	6.299	2.673	63689

Table 36 presents the summary statistics for Bangladesh. From the extensive margins' variables, we observe that 36% of the sample works and 36.4% of the total sample is part of the labor force. Looking at other variables, we observe that most of the sample is composed of Muslims (88%), and 16% of the total sample belongs to the treatment group. As in Malawi, we have a young sample where the average age is 26 years old, and the sample is well divided between males (47%) and females (53%). Finally, we observe that on average, the household size is composed of six people per household, and out of these six individuals, four are adults.

4.2.1 Balance Tables - Bangladesh

Table 37 presents the balance summary statistics and the normalized differences across groups for Bangladesh, following the same procedure as we did for Malawi. Looking at the normalized differences, we observe that all of them are below the threshold of 0.25 and thus, we can conclude that the sample for Bangladesh is well-balanced between those individuals interviewed during Ramadan and those that were not - except for the household size that has a normalized difference greater than 0.25 in absolute values. Despite that, all control variables present a well-balanced normalized difference and thus, we can conclude that our identifying assumption holds.[39] Given

[39]Our identifying assumption is the same one as we stated for Malawi: *individuals interviewed during the month of Ramadan are, on average, identical to those interviewed outside the Ramadan period, in a given year.*

that we perform the analysis only for the Muslim sample for Bangladesh, we directly present the balance table restricted for this subsample.

Table 37: Balance Summary Statistics - Bangladesh

	(1)		(2)		T-test	Normalized
	Non-Ramadan Month		Ramadan Month		Difference	difference
Variable	N	Mean/SE	N	Mean/SE	(1)-(2)	(1)-(2)
Age	45360	26.466	10575	27.116	-0.649***	-0.034
		(0.090)		(0.184)		
School	45324	3.405	10575	3.519	-0.113***	-0.029
		(0.018)		(0.037)		
Num. children	45382	2.111	10628	2.184	-0.073***	-0.050
		(0.007)		(0.015)		
Household size	45382	6.145	10628	7.048	-0.904***	-0.337
		(0.012)		(0.028)		
Marital status	45360	1.618	10575	1.628	-0.011	-0.017
		(0.003)		(0.006)		

Notes: The value displayed for t-tests are the differences in the means across the groups. Standard errors are robust. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level. The F-test for the joint orthogonality is 106.35 with a corresponding p-value equal to 0.000.

5 Empirical Strategy

We start by examining the direct effect of time allocation of working hours decided by individuals under the month of Ramadan. We assume Ramadan to be a cultural constraint, which occurs every year but on a non-fixed date, as explained in earlier sections. This has a crucial implication in household/individual decisions: Ramadan is seasonal, implying that it does not occur on a fixed date every year; thus, in the summer months, fasting lasts for longer hours, contrary to when it takes place in the winter months. Therefore, longer fasting hours means the ingestion of fewer calories during the day and it might lead to a decrease in an individual's productivity at work. Whereas shorter periods of fasting should not have such an effect on productivity and thus, allocations of working hours by individuals should not be significantly altered. Thus, it allows us to include year and seasonal fixed effects in our regression to control for some possible unobserved effects outside Ramadan to the allocation of hours worked.

The approach we use in this chapter is a differences-in-differences estimation with individual, seasonal and year fixed effects testing whether or not there exists a change in the hours worked between those individuals who celebrate Ramadan relative to those who do not. Because Ramadan is understood as a religious constraint where individuals who take part of it cannot decide on the date when it occurs, we assume this cultural constraint to be randomly assigned. Hence, there is no obvious reason why it would be correlated with other region-year shocks.

However, before implementing this identification strategy, it is important to know the immediate impact of Ramadan on the labor supply and check whether there are substantial differences in a first test. Nevertheless, this specification has a lack of fixed effects, censoring (due to the number of zeros observed in hours worked) and sample selection, as there might be certain aspects that will determine the decision of working. Along this section, we disentangle these issues and give a final argument as to why the Heckman model is the preferred one for the intensive margins' analysis.

Equation (14) shows the first set of labor supply estimations of this project:

$$\vec{H_{i,t}} = \alpha + (islam, ramadan, islam \times ramadan)'_{i,t}\beta + u_{i,t} \tag{14}$$

where $\vec{H_{i,t}}$ considers the three different dependent variables we are interested to estimate in this chapter, which are: $Hours_{i,t}$ as the number of total hours that individual i has worked during the previous week in year t, $Housework_{i,t}$ represents the number of hours that individual i worked in the previous week in his/her household in year t and $Prod.Hours_{i,t}$ is the number of hours that individual i worked in his/her job along the past week in year t. Regarding the controls added in Equation (14),$islam_{i,t}$ represents a dummy variable indicating whether individual i is Muslim or not; $ramadan_{i,t}$ is another dummy variable with value one if the individual was interviewed during the Ramadan period and zero otherwise; $islam \times ramadan_{i,t}$ is the interaction term between the previous two variables, where if the resulting outcome is one, it means that the individual belongs to the treatment group and if zero, then he/she is part of the control group. Finally, $u_{i,t}$ is the error term of the regression.

The idea of adding the interaction term in this study is because the estimated coefficient reports the *real effect* of Ramadan into labor supply. In words, how the fact of being Muslim and having been interviewed during Ramadan affects the number of hours worked for this given individual.

After estimating Equation (14), we proceed to apply the identification strategy proposed at the beginning of this section, by adding individual, seasonal and time-fixed effects. Therefore, we implement it in the following regression equation:

$$\vec{H_{i,t}} = \alpha + (islam, ramadan, islam \times ramadan)'_{i,t}\beta + \vec{X_{i,t}}'\theta + \eta_i + \tau_t + u_{i,t} \tag{15}$$

where $\vec{X_{i,t}}$ is a vector of individual characteristics, which includes age and its square, marital status, educational level, whether the individual lives in an urban or rural area, the number of children living at home and the household size. η_i represents the individual fixed effects and finally, τ_t represents the time and seasonal fixed effects.

However, Equation (15) suffers from censoring and sample selection bias. The fact of analyzing the number of hours worked might report zeros for certain individuals. This might be because of censoring and thus, we should re-estimate our linear fixed effect model by using a Tobit model with individual, seasonal and time-fixed effects. This solves the issue, allowing us to control for all the number of zeros that we might have in our data. However, it does not solve the whole problem; sample selection bias. In this case, there may be some factors that will have some implications in the decision of being part of the labor force (or labor market), such as the household size, the number of children one is responsible for within the household, the amount of land someone owns or the number of adults living in the same household, among other things. Hence, to control for this issue we need to estimate a Heckman model with fixed effects.

To introduce fixed effects into non-linear models, like Tobit and Heckman, we need to follow the approach proposed in Wooldridge (1995).

5.1 Tobit Model with Fixed Effects

The Tobit model is used in regression analysis to deal with a significant number of zeros that we might observe in our dependent variable(s), and thus, correct for downward bias due to censoring we may face in the observed number of hours worked in our sample.

In this case, following Wooldridge (1995) and Dustmann and Rochina-Barrachina (2007), we need to apply the *Wooldridge correction.* It consists of the following; instead of estimating the individual fixed effects coefficient, η, we need to include a vector of individual means over time of each variable that we include in the regression analysis. In this case, we express the vector of means as $\vec{\bar{X}}_i$, without subindex t, since we are extracting the average of the variables for each individual across years. This vector of means also includes the averages for $islam_i$, $ramadan_i$ and $islam \times ramadan_i$, jointly with the seasonal and time-fixed effects average, τ. Therefore, the regression we want to analyze is as follows:

$$\vec{H_{i,t}} = \alpha + (islam, ramadan, islam \times ramadan)_i'\beta + \vec{X_{i,t}}'\theta + \vec{\bar{X}}_i'\eta + \tau_t + u_{i,t} \qquad (16)$$

Hence, the set of estimators we obtain using the Tobit estimation method are ($\alpha,\ \beta,\ \theta,\ \eta$). With that, we are controlling for censoring, and we are solving the issue of controlling for fixed effects in a non-linear model, using the matrix of averages $\vec{\bar{X}}_i$.

5.2 Heckman Model with Fixed Effects

The use of the Heckman model in this set-up helps to solve the problem of sample selection. In this case, we first need to design the selection equation and decide which factors make an individual be part of the labor force or not, in other words, analyze the factors that might have influence on the decision of working (or being potentially looking for a job) for a given individual.

Given the country we are analyzing in this chapter, we need to highlight that Malawi belongs to a developing economy. Thus, personal decisions, like labor, are made at the household level, meaning that it is the head of the household who makes the decisions about which members of the household work at home and which ones need to look for a job outside the household (see Nagler and Naudé, 2014).

As mentioned in Section 3, the agricultural sector is a key one in the Malawian labor market. This implies that land ownership should have an impact on the decision of being part of the labor force. In fact, there are studies that show that the area of land each household has is a key determinant in whether to look for a job outside the home or to stay and work for the household (see Julien et al., 2019). Another important aspect that determines criteria for being part of the labor force or not is the number of adults living in the household, in other words, if many adult members live in the same household the chance that one works on the household's own land is lower (as enough people are working there) and thus, this person has more opportunities to look for a job outside the home and be more likely to be part of the labor market (see Nagler and Naudé, 2014).

Knowing this information, we use the ratio of area of land in a household divided by the number of adult members living in the same household as one of the instrumental variables for our selection equation. Moreover, we expect the exclusion restriction to be fulfilled since the ratio of land over adults should not affect the number of hours worked since the time one dedicates to labor is established by a contract or it is simply a social norm: one works the same hours as his/her neighbors do.

We are also aware that gender, age and marital status can be crucial determinants for being part of the labor force in developing countries. Thus, we interact these variables with the ratio of land area over the number of adults living in the same household and use them as instrumental variables for the selection equation. Hence, the selection equation for the Heckman model is as follows:

$$\begin{aligned}\mathbb{P}(LFP_{i,t} = 1 \mid X) &= \Phi((islam, ramadan, islam \times ramadan)'_{i,t}\gamma_1 + \gamma_2 land\_person_{i,t} + \vec{Z_{i,t}}\gamma_{\mathbf{3}} + \\ &\quad + \vec{X_{i,t}}'\gamma_{\mathbf{4}} + \bar{\vec{Z}}_i'\eta_{\mathbf{1}} + \bar{\vec{X}}_i'\eta_{\mathbf{2}} + \tau_t + \varepsilon_{i,t})\end{aligned} \tag{17}$$

where $\Phi(.)$ defines the cumulative density function of the Probit model. $LFP_{i,t}$ is a dummy variable taking value one if the individual is part of the labor force and zero otherwise, $land\_person_{i,t}$ is the ratio between the amount of land owned by the household over the number of adults in the household, $\vec{Z_{i,t}}$ is the set of individual characteristics (age, gender and marital status dummies) interacted with the ratio of land area in the household over adults in the same household, used as instrumental variable. $\vec{X_{i,t}}$ is the set of control variables, which are the same ones as in Equation (16), and $\bar{\vec{X}}_i$ is the vector of means of the variables included in the regression. Finally,$\varepsilon_{i,t}$ is the error term of the selection equation.

Following the theory, we estimate Equation (17) through a Probit model (see Cameron and Trivedi, 2005; Heckman, 1974). In this case, the set of instruments, $land\_person$ and $\vec{Z_{i,t}}$, satisfy the relevance property - as we see in Table 41, some of the instruments are statistically significant while some others are not, but the joint significant test shows that the set of instrumental variables are statistically significant and provides evidence that the set of instruments used are strong and valid.[40] As mentioned above, the exclusion restriction is also satisfied, as the owned land should not affect the number of hours worked. Thus, the selection equation is estimated using the Wooldridge (1995) technique, previously explained in subsection 5.1 - this also applies to the second step estimation of the Heckman model in Equation (18).

Once the selection equation is estimated, we calculate the inverse of the Mills' Ratio as a function

[40] Other studies that analyze labor supply effects, like Mroz (1987), are using household size and number of children as the set of instrumental variables for the selection equation. In our case, these two variables seem to be endogenous violating the exogeneity condition. However, in developed countries like the US, these two factors seem to be key determinants for the labor force participation.

of the controls and the obtained coefficients in Equation (17).

$$\lambda(x_{i,t}\hat{\gamma}) = \frac{\phi(x_{i,t}\hat{\gamma})}{\Phi(x_{i,t}\hat{\gamma})}$$

where x_i represents all variables included in the selection equation and γ englobes all the estimated coefficients in Equation (17).

Next, we compute the second-stage equation using a Pooled Ordinary Least Squares (OLS) method. In this case, we need to add the Inverse of Mills' Ratio to the set of controls. Therefore, the regression to estimate is as follows:

$$\vec{H_{i,t}} = \alpha + (islam, ramadan, islam \times ramadan)_i'\beta + \vec{X_{i,t}}'\theta + \vec{\vec{X}}_i'\eta + \psi\lambda_i + \tau_t + u_{i,t} \tag{18}$$

where we introduce the Inverse of Mills' Ratio (λ) as a control. The next step is to test for sample selection, where we test the null hypothesis that $\psi = 0$. If the null is rejected, then we have sample selection and Heckman is the appropriate model to use.

The way to obtain the coefficients and to test for sample selection seems easy. However, this is not the case for the standard errors, as we need to compute for heteroskedasticity-robust ones. To get them, we need to compute the Asymptotic Variance or robust variance matrix for the estimated coefficients: $Avar(\alpha, \beta, \theta, \eta, \psi)$. The way to estimate the Asymptotic Variance, using a method of moments, can be followed in the Appendix of Wooldridge (1995). Once this is done, the next step is to compute the root square of the diagonal of the Asymptotic Variance matrix to obtain the standard errors. Moreover, we also cluster the standard errors at the village level. The main reason for clustering at the village level and not at the household one is because shocks to labor and other unobserved determinants of labor outcomes may be correlated across individuals that live in the same village - notice that the Muslim population in Malawi is concentrated in the southern regions.

Nonetheless, the proposed regressions in this chapter consider individuals treated in different years, which implies that the difference-in-difference estimation method is with staggered treatment, which means that the number of treated observations in each period of the survey is varying across years; in other words, the size of the treatment changes along the periods and individuals that are treated in 2010 might not be treated in 2013 but can be treated again in 2016. This is because individuals can be randomly interviewed in any moment of the year. Thus, the treatment effect can be heterogeneous, given the size change in each year, leading to a violation of the constant treatment effect assumption (see de Chaisemartin and D'Haultfoeuille, 2020). [41] Therefore, when

[41] The constant treatment effect assumption imposes that the treatment effect should be constant across groups

this happens, the TWFE estimates might be biased and/or inconsistent, as there might be some individuals that are part of the treatment and control group along our analysis, which implies that the treatment group will be heterogeneous (see de Chaisemartin and D'Haultfoeuille, 2020; Callaway and Sant'Anna, 2021; and Goodman-Bacon, 2021). If this is the case in our study, we should expect to have a downward bias from the TWFE estimates, given the negative weights that the TWFE method assigns to periods with larger amounts of treated individuals- this is because of the difference in treatment sizes across periods (see de Chaisemartin and D'Haultfoeuille, 2020). [42] This is something to consider later in our results, as our treatment is staggered and it can lead to Type-I and Type-II errors (see Baker et al., 2021).

5.3 Intra-Household Allocation

We introduce in this section the intra-household effects on labor supply in our analysis. The idea is to include a set of triple interactions to Equation (18) that allows us to investigate whether the effect of Ramadan on labor supply varies by household composition. Thus, the set of triple interactions is based on the number of children, adults and older people living in the household respectively, differentiating by males and females, interacted with the variables *islam*, *ramadan* and $islam \times ramadan$. Therefore, the resulting regression we want to estimate under the Heckman method with fixed effects is:

$$\begin{aligned} \vec{H_{i,t}} \quad = \quad & \alpha + (islam, ramadan, islam \times ramadan)'_{i,t}\beta_{\mathbf{1}} + islam \times ramadan \vec{\times} hh\_composition_{i,t}{}'\beta_{\mathbf{2}} \\ & + islam \times hh\_\vec{composition}_{i,t}{}'\beta_{\mathbf{3}} + ramadan \times hh\vec{\_composition}_{i,t}{}'\beta_{\mathbf{4}} + \vec{X_{i,t}}{}'\theta + \vec{\bar{X}}'_i\eta + \psi\lambda_i + \tau_t + u_{i,t} \end{aligned} \tag{19}$$

where the $hh\_composition_{i,t}$ considers all the categories described above about the different types of groups that can be part of the household: male and female children; male and female adults, and male and female aged members. In this case, we interact it with the treatment interaction term: $islam \times ramadan$, and with *islam* and *ramadan* itself. To facilitate the interpretation of the interaction terms, we demean the variables included in the $hh\_composition_{i,t}$, which is simply

and over the years. This implies that there is a pre-treatment period, where none of the observations in the sample are observed during the Ramadan month, and a post-treatment period, where some individuals in the sample will be celebrating the Ramadan festivity after a given year.

[42] According to Goodman-Bacon (2021), the TWFE method assigns a weighted average treatment effect to the TWFE difference-in-difference (TWFEDD) estimators that compares timing groups to each other. If the constant treatment effect assumption holds, the TWFEDD estimates should not be biased. But if it is violated, the variance of the TWFEDD estimator might be incorrectly estimated and lead to inconsistent estimates, or the estimated treatment effect will be downward biased.

the subtraction of the mean from each of these variables prior to the estimation. Hence, the estimated coefficients of the interaction terms are interpreted as the additional effects of Ramadan when deviating from the "average" household (in terms of demographic composition). Moreover, $\vec{X_{i,t}}$ also takes into account each of these categories that are part of the household composition. Before estimating Equation (19), we introduce the same set of interactions to the selection equation described in Equation (17).

5.4 Extensive Margins

We are also interested in estimating the extensive margins to observe the effect of Ramadan on the probability that an individual has of working and in being part of the labor force, respectively. For this last case, the approach is similar to the one used in the selection equation presented in Equation (17). To estimate both likelihoods under the Ramadan period in Malawi and Bangladesh, we use a conditional Logit model with fixed effects with clustered standard errors at the village level (see Chamberlain, 2010). Therefore, the set of regressions we estimate are:

$$work_{i,t} = \Lambda(\alpha + (islam, ramadan, islam \times ramadan)'_i\beta + X'_{i,t}\theta + \eta_i + \tau_t + u_{i,t}) \tag{20}$$

$$lfp_{i,t} = \Lambda(\alpha + (islam, ramadan, islam \times ramadan)'_i\beta + X'_{i,t}\theta + \eta_i + \tau_t + u_{i,t}) \tag{21}$$

where $\Lambda(.)$ represents the logistic cumulative distribution function. In Equations (20) and (21), we are analyzing just the pure effect of Ramadan; however, we also run another set of regressions in which we introduce the set of triple interactions described in the previous subsection to test the effect of household composition during the month of Ramadan into the extensive margins.

6 Results

This section presents the estimated results for the proposed regressions in the previous section. We show results for the reduced form, for the fixed effects and Tobit models, as well as for the Heckman model with fixed effects. Complete tables of results are in Appendix 7. The intensive margins estimations for the Heckman model are also done for the female and male subsamples.

6.1 The Effect of Ramadan on Labor Supply on the Intensive Margin

Table 38 shows the results for the Pooled OLS estimation presented in Equation (14). There, we find that none of the coefficients of interest ($Ramadan$ and $Islam \times Ramadan$) are statistically significant in any of the cases. In particular, if we observe one of the coefficients of interest, $Islam \times Ramadan$ (the interaction term, which estimates the effect for a Muslim interviewed during the Ramadan period) it has no effect on labor supply, not in reduced form estimation nor in the extended one. The same occurs with the estimates of $Ramadan$. From these results we observe that Ramadan, understood as a religious constraint that imposes a cultural/religious barrier to work by restricting nutrition at certain times of the day, does not have any effect on the number of hours worked. In addition, we perform an F-test to test the pure effect of Ramadan into labor supply by adding $Ramadan$ and $Islam \times Ramadan$ coefficients. Hence, the null hypothesis is as follows:

Hypothesis 1:

$$H_0: \text{ The pure Ramadan effect has no effect on labor supply.}$$

$$H_a: \text{ The pure Ramadan effect has an effect on labor supply.}$$

The results of the joint-test for all the regressions run under the Pooled OLS method fail to reject the null hypothesis and, thus, these indicate that labor supply is not affected under the Ramadan month. However, this method is biased since we are not considering censoring or sample selection.

When controlling for unobserved heterogeneity in Table 39, we do not observe many changes from the Pooled OLS estimation. The estimated coefficients do not change the direction of the effect towards labor supply, in other words, all of them are positive but not statistically significant. Therefore, from Table 39, we conclude that by introducing individual fixed effects into the regression analysis, someone who follows Ramadan does not see his/her number of hours worked affected. Furthermore, looking at the pure effect of Ramadan, under Hypothesis 1, we fail again to reject the null hypothesis. However, again we are not controlling for censoring or sample selection, due to non-participation in the labor force.

Table 38: Labor Supply Allocation in Malawi - Pooled OLS

	(1)	(2)	(3)	(4)	(5)	(6)
	Hours	Hours	Housework	Housework	Prod. Hours	Prod. Hours
Islam	-1.718	-0.754	0.239	-0.0192	-1.964	-0.731
	(-1.27)	(-0.81)	(0.45)	(-0.05)	(-1.86)	(-0.90)
Ramadan	1.051	0.451	0.507	0.135	0.543	0.365
	(1.02)	(0.50)	(1.49)	(0.45)	(0.66)	(0.49)
Islam × Ramadan	1.071	2.011	0.416	0.671	0.664	1.290
	(0.48)	(1.12)	(0.47)	(1.01)	(0.41)	(0.95)
_cons	16.39***	15.22***	5.362***	5.790***	11.03***	9.412***
	(23.17)	(11.78)	(24.60)	(10.12)	(19.29)	(9.76)
F-test	1.09	3.48	1.31	1.77	0.67	1.70
p-value	0.2993	0.0624	0.2554	0.1858	0.4141	0.1956
Individual Controls	✗	✓	✗	✓	✗	✓
N	10686	8186	10694	8192	10705	8200
R-Squared	0.000816	0.0204	0.000782	0.00696	0.00128	0.0225

The *Islam* coefficient reports the effect that following Islam as a religion or not, for a given individual, has on labor supply. *Ramadan* estimates how the fact of participating (or not) in the Ramadan festivity affects the number of weekly hours worked during the Holy month. Finally, *Islam* × *Ramadan*, is the interaction term between the previous two variables. This coefficient captures the effect of a Muslim individual who celebrates Ramadan on individual labor supply, compared to individuals who do not follow Islam, have not been interviewed during Ramadan or neither one. In the specifications presented in columns (1), (3) and (5), we only include the set of variables that belong to the Ramadan cultural barrier; no other controls are included in this set. On specifications included in columns (2), (4) and (6), we also control for individual characteristics: age and its square, whether or not the individual has gone to school, whether he/she lives in an urban or rural area, the number of children in the household and the household size. Moreover, we also include the set of marital status dummies. We compute robust standard errors clustered at the village area level. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The full set of estimated coefficients are reported in Table 47 of the Appendix to this chapter.

The F-test performs a joint significant test of the Ramadan cultural barrier variables (*Ramadan* and *Islam* × *Ramadan*), on individuals' weekly hours worked. The null hypothesis of the F-test is that *individuals who are face the cultural constraint of the Ramadan festivity do not alter their number of weekly hours worked during the Holy month of Ramadan.*

Table 39: Labor Supply Allocation in Malawi - Panel Data FE

	(1) Hours	(2) Hours	(3) Housework	(4) Housework	(5) Prod. Hours	(6) Prod. Hours
Islam	2.307	2.375	1.659	1.049	0.669	1.390
	(0.68)	(0.63)	(0.86)	(0.58)	(0.36)	(0.63)
Ramadan	0.600	0.685	0.203	0.0420	0.388	0.725
	(0.60)	(0.61)	(0.62)	(0.13)	(0.49)	(0.77)
Islam× Ramadan	1.482	0.687	0.840	0.525	0.642	0.0467
	(0.66)	(0.33)	(0.93)	(0.63)	(0.42)	(0.03)
_cons	14.88***	18.64**	5.023***	5.928*	9.861***	13.29**
	(25.76)	(3.30)	(16.98)	(2.48)	(28.61)	(2.85)
F-test	1.04	0.50	1.57	0.49	0.57	0.36
p-value	0.3110	0.4790	0.2126	0.4853	0.4537	0.5502
Individual Controls	✗	✓	✗	✓	✗	✓
N	13902	10686	13911	10694	13936	10705
R-Squared	0.000636	0.00790	0.000913	0.00465	0.000284	0.00760

In this table we present the fixed effects estimates. The $Islam$, $Ramadan$ and $Islam \times Ramadan$ coefficients are as explained before in Table 38. In the specifications presented in columns (1), (3) and (5), we include the set of variables that belong to the Ramadan cultural barrier, but we also control for individual, year and seasonal fixed effects. On specifications included in columns (2), (4) and (6), we also control for individual characteristics: age and its square, whether the individual has gone to school or not, whether he/she lives in an urban or rural area, the number of children in the household and the household size. We also include the set of marital status dummies. Moreover, we also include individual, year and seasonal fixed effects. We compute robust standard errors clustered at the village area level. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The full set of estimated coefficients are reported in Table 48 of the Appendix to this chapter.

The F-test performs a joint significant test of the Ramadan cultural barrier variables ($Ramadan$ and $Islam \times Ramadan$), on individuals weekly hours worked. The null hypothesis of the F-test is that individuals who are facing the cultural constraint of the Ramadan festivity do not alter their number of weekly hours worked during the Holy month of Ramadan.

Table 40 presents the results for the Tobit model, presented in Equation (16). In this case, we introduce the Wooldridge correction into the regression to control for individual fixed effects in non-linear models. Again, when observing the estimates of our coefficients of interest ($Ramadan$ and $Islam \times Ramadan$), we observe that both coefficients have a positive effect on the hours worked for those individuals who celebrate Ramadan, relative to those who do not. This result is confirmed when we run the F-test under Hypothesis 1, where we fail to reject the null hypothesis that the celebration of Ramadan has no impact on the number of hours worked, neither in the household nor at a paid job, across groups. However, under the Tobit model, we are only controlling for censoring, but we are not considering sample selection; therefore, to correct for that we need to estimate a Heckman model.

Table 41 shows the results for the selection equation presented in Equation (17), estimated under the Probit model. We perform three different types of analysis to get a better understanding of the labor market situation in Malawi when the Ramadan month takes place: we estimate the effect for the entire sample, but we also do the analysis for females and males subsamples separately. Analyzing the variables itself, we observe that when estimating the selection equation for the entire sample, the amount of land over the number of adults in the household itself positively affects the chances of being part of the labor force, and it is statistically significant. We also find a positive effect on the interaction of the marital status categories of polygamy and divorce, but in this case, the estimated effects are not statistically significant. On the other hand, the remaining set of instrumental variables negatively affect the likelihood of being in the labor force, but we only find a statistically significant effect on the interaction with gender, as well as with separated and widowed individuals. However, when we split the sample between males and females, labor force participation for males is not significantly affected by any of the proposed instrumental variables. The reason why the proposed instruments are not significant is because, the proportion of males working in Malawi is high, as we discussed in Section 4, where 69% belong to the labor force. Hence, we can adopt a simple fixed effects model for the male subsample in the second stage.

On the other hand, when we analyze the labor force participation for females, we find that the size of land divided by total adults in the household negatively affects the likelihood of being in the labor force when the woman is separated or widowed, and this effect is statistically significant. We also find a negative effect on the interaction with married females, but in this case, the effect is not statistically significant.

Nevertheless, to check whether the set of proposed instrumental variables influence the probability of being (or not) in the labor force, we need to test for the relevance condition of the instrumental variables, as well as the exogeneity one. The null hypothesis for the relevance condition of instrumental variables is as follows:

Table 40: Labor Supply Allocation in Malawi - Tobit

	(1)	(2)	(3)	(4)	(5)	(6)
	Hours	Hours	Housework	Housework	Prod. Hours	Prod. Hours
Islam	-1.625	1.193	1.161	2.366	-3.054	-0.579
	(-0.43)	(0.28)	(0.43)	(0.78)	(-0.86)	(-0.14)
Ramadan	1.289	0.461	0.743	0.198	0.632	0.222
	(1.01)	(0.31)	(1.18)	(0.30)	(0.40)	(0.12)
Islam × Ramadan	0.759	0.222	0.182	0.277	1.064	0.0261
	(0.21)	(0.07)	(0.11)	(0.16)	(0.29)	(0.01)
[1em] _cons	6.831***	8.325***	-4.222***	-1.350	-4.688***	-2.388
	(9.31)	(4.02)	(-8.27)	(-0.93)	(-4.91)	(-1.10)
F-test	0.37	0.05	0.34	0.09	0.24	0.01
p-value	0.5409	0.8309	0.5590	0.7639	0.5590	0.9383
Individual Controls	✗	✓	✗	✓	✗	✓
N	13902	10686	13911	10694	13936	10705

In this Table we present the Tobit model estimates with fixed effects, using the *Wooldridge correction* proposed in Wooldridge (1995). The $Islam$, *Ramadan* and $Islam \times$ *Ramadan* coefficients are as explained before in Table 38. In the specifications presented in columns (1), (3) and (5), we include the set of variables that belong to the Ramadan cultural barrier, but we also control for individual, year and seasonal fixed effects. On specifications included in columns (2), (4) and (6), we also control for individual characteristics: age and its square, whether the individual has gone to school or not, whether he/she lives in an urban or rural area, the number of children in the household and the household size. We also include the set of marital status dummies. Moreover, we also include individual, year and seasonal fixed effects. We compute robust standard errors clustered at the village area level. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The full set of estimated coefficients are reported in Table 49 of the Appendix to this chapter.

The F-test performs a joint significant test of the Ramadan cultural barrier variables (*Ramadan* and $Islam \times$ *Ramadan*), on individuals weekly hours worked. The null hypothesis of the F-test is that *individuals who are facing the cultural constraint of the Ramadan festivity do not alter their number of weekly hours worked during the Holy month of Ramadan.*

Table 41: Selection Equation - Labor Force Participation

	(1)	(2)	(3)	(4)	(5)	(6)
	Labor Force	Labor Force	Labor Force	Labor Force	Labor Force	Labor Force
land_person × age	0.00196	0.00156	0.00213	0.00224* *	0.000837	0.00121
	(1.587)	(1.498)	(1.869)	(1.974)	(0.645)	(0.693)
land_person × gender	-0.0698* *	-0.0431				
	(-2.395)	(-1.684)				
land_person*married	-0.0928	-0.0514	-0.0832	-0.0767	-0.00879	-0.00960
	(-1.776)	(-1.153)	(-1.648)	(-1.275)	(-0.251)	(-0.362)
land_person × polygamy	0.116	0.165	3.133	2.627	0.140	0.135
	(0.705)	(1.012)	(1.952)	(1.721)	(0.930)	(0.788)
land_person × separated	-0.274** **	-0.214* *	-0.233* *	-0.238	-0.0459	-0.0406
	(-2.738)	(-2.235)	(-2.151)	(-1.348)	(-0.443)	(-0.194)
land_person × divorced	0.0864	0.272	0.146	0.301	0.0529	-0.0486
	(0.512)	(1.180)	(0.806)	(1.384)	(0.468)	(-0.832)
land_person × widowed	-0.205* *	-0.155	-0.187* *	-0.177* *	0.292	0.181
	(-2.207)	(-1.829)	(-2.117)	(-2.085)	(1.189)	(0.696)
land_person	0.0480** **	0.0240	0.00941	-0.00385	-0.0285	-0.0313
	(2.681)	(1.358)	(0.379)	(-0.177)	(-1.028)	(-1.054)
_cons	-2.239*** ***	-2.925*** ***	-1.623*** ***	-2.342*** ***	-1.444*** ***	-1.207*** ***
	(-12.17)	(-15.09)	(-11.21)	(-9.663)	(-7.920)	(-3.937)
F-test	18.92	14.52	14.75	14.56	4.70	4.09
p-value	0.0153	0.0692	0.0394	0.0421	0.6969	0.7692
Sample	Whole	Whole	Female	Female	Male	Male
Triple interaction	✗	✓	✗	✓	✗	✓
N	11768	11768	6059	6059	5709	5709

In this Table we present the results for the selection equation, estimated under a Probit model specification with fixed effects. We present the results for the entire sample, but we also split the analysis between males and females. The variables presented here are our set of instrumental variables for labor force participation, which is our selecting variable. The *land_person* coefficient is understood as the area of land an individual has in his/her household divided by the number of adults within the household. Thus, this estimate tells us how the size of land in a household over the total number of adults there affects the probability of being in the labor force. Moreover, we interact this coefficient with age, gender and marital dummies, to check if any of these individual characteristics, jointly with the area of land over the number of adults in the household can interfere with the likelihood of an individual's participation in the labor force. In all specifications, we control for individual characteristics: age and its square, whether the individual has gone to school or not, whether he/she lives in an urban or rural area, the number of children in the household and the household size. We also include the set of marital status dummies. Moreover, we also include individual, year and seasonal fixed effects. In addition to that, in specifications (2), (4) and (6) we also include the set of triple interactions. We compute robust standard errors clustered at the village area level. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The full set of estimated coefficients are reported in Table 50of the Appendix to this chapter.
The F-test checks the relevance condition for instrumental variables. The null hypothesis of the F-test is that *the land size over total adults in the household and its interaction with different individual characteristics have no effect on labor force participation.*

Hypothesis 2:

H_0: The set of instrumental variables has no effect on self-selection to participate in the labor force.
H_a: The set of instrumental variables has an effect on self-selection to participate in the labor force.

After performing the F-test for the three samples, we observe that the relevance condition is satisfied for the whole sample and the female subsample estimates, rejecting the null hypothesis proposed in Hypothesis 2; but for males we fail to reject the null hypothesis that the set of instruments are not relevant. Therefore, we conclude that for the entire sample and females, the choice set of instrumental variables is valid and strong, as the resulting F-test are 18.92 and 14.75, respectively.

We also analyze the selection equation under the set of triple interactions proposed in subsection 5.3. In this case, we observe that we still do not find any significant effects of the instrumental variables on labor force for males, and neither is the null hypothesis in Hypothesis 2 rejected. However, when we analyze the whole sample and the female subsample, we observe that the significance of some coefficients is lost in both cases. For the entire sample, we only find that the size of land negatively affects the likelihood of being part of the labor force if the individual is separated. This is the only estimation that is statistically significant; the rest of the coefficients do not reveal any relevant impact on labor force. When analyzing the female subsample, we find that the only coefficient that keeps significantly (and negatively) affecting labor force is the area of land in the household over adults for widowed females. The rest of the coefficients do not have a significant effect on the chances of being part of the labor force for females. However, observing at the performed F-test under Hypothesis 2, again, the relevance condition is satisfied for the whole sample and the female subsample only. In both cases, the set of instrumental variables are valid and strong, as the obtained F-tests are greater than 10; more precisely, 14.52 for the whole sample and 14.56 for females. However, for males, the set of instruments do not affect labor force participation, as we fail to reject the null hypothesis.

After estimating the selection equation, we compute the Inverse of Mills' Ratio and include it into the second-stage regression to control for sample selection. Results are in Table 42 for the entire sample, Table 43 for the female subsample and Table 44 for the male subsample.

Table 42: Heckman second-stage Analysis - Entire Sample Estimation

	(1) Hours	(2) Hours	(3) Housework	(4) Housework	(5) Prod. Hours	(6) Prod. Hours
Islam	1.7109 ***	3.5071 ***	2.4722 ***	3.4821 ***	-0.7468	0.0401
	(9.40)	(32.50)	(6.18)	(41.93)	(-1.65)	(1.23)
Ramadan	0.7800 ***	0.0071	-0.1279	-0.1846 ***	0.9331 ***	0.1761 ***
	(14.59)	(0.12)	(-1.39)	(-8.14)	(8.36)	(5.12)
Islam*Ramadan	-0.2251	-0.4207	0.2110	-0.2045 ***	-0.4156 **	-0.1619 *
	(-1.66)	(-1.69)	(0.82)	(-6.51)	(-3.18)	(-2.48)
Islam*Ramadan*Kids_male		13.3843 ***		1.6198 ***		11.6650 ***
		(26.81)		(11.66)		(34.51)
Islam*Ramadan*Adult_male		-9.5640 ***		-1.0293 ***		-8.8144 ***
		(-19.94)		(-11.75)		(-11.09)
Islam*Ramadan*Old_male		67.6752 ***		34.6047 ***		32.8909 ***
		(66.35)		(181.90)		(66.38)
Islam*Ramadan*Adult_female		-28.6123 ***		-7.1923 ***		-22.1345 ***
		(-19.84)		(-106.69)		(-55.93)
Islam*Ramadan*Old_female		-58.3492 ***		-35.2805 ***		-23.3228 ***
		(-35.15)		(-190.84)		(-19.34)
lambda	-0.3528 **	-0.1575 **	0.8201 ***	0.0581 **	-1.1850 ***	-0.2066 ***
	(-3.10)	(-2.64)	(20.26)	(16.16)	(-10.05)	(-7.14)
F-test	19.68	416.12	0.04	304.14	9.03	11.30
p-value	0.000	0.000	0.9801	0.000	0.0109	0.000
Triple interaction	✗	✓	✗	✓	✗	✓
N	5732	5732	5732	5732	5732	5732
R-Squared	0.0334	0.0465	0.0162	0.0264	0.0372	0.0469

In this table we present the results of the second step for the Heckman model with fixed effects, using the *Wooldridge correction* proposed in Wooldridge (1995). The $Islam$, $Ramadan$ and $Islam \times Ramadan$ coefficients are as explained before in Table 38. We control for individual characteristics: age and its square, whether the individual has gone to school or not, whether he/she lives in an urban or rural area, the number of children in the household, the household size and the set of marital status dummies. Besides, we control as well for sample selection, including the Inverse of Mills' Ratio (the lambda coefficient). Moreover, we also include individual, year and seasonal fixed effects. In addition to that, in specifications (2), (4) and (6) we also include the set of triple interactions. We compute heteroskedastic-robust standard errors clustered at the village area level. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The full set of estimated coefficients are reported in Table 51 of the Appendix to this chapter.

Focusing first on the estimates for the entire sample of Equation (18), we find a significant change with respect to previous estimations. Starting with the estimations for the whole sample, we observe that being part of the treatment group significantly affects labor supply behavior, meaning that those individuals that take part in Ramadan do change their allocation of hours during the Holy month of Ramadan. Specifically, we observe that individuals interviewed during Ramadan increase, on average, their number of hours worked during Ramadan by 0.55 hours, compared with individuals interviewed in a different period of the year. This effect is obtained by aggregating the estimated coefficients for *Ramadan* and $Islam \times Ramadan$, as it provides us with the total effect of Ramadan on labor supply. Moreover, after performing the F-test, we find that the aggregate effect of Ramadan on total hours worked is statistically significant, in other words, we reject the null hypothesis proposed in Hypothesis 1. For housework, we do not find significant effects of Ramadan on hours worked there. Finally, under hours of paid work, we find that individuals interviewed during Ramadan increase their number of hours worked by 0.52 hours, relative to individuals interviewed outside the Ramadan period. Again, this effect is obtained by adding the *Ramadan* and $Islam \times Ramadan$ estimates. We also find this effect to be statistically significant and thus, we do reject the null hypothesis proposed under Hypothesis 1.

When we look at the household composition effects proposed in Equation (19), we observe that for the entire sample, Ramadan itself has a positive and statistically significant effect on total hours worked (see column (2)), as well as the interaction term. However, these estimates per se do not tell anything, as these assume that the household is composed by 1 individual only, which is unlikely (given the summary statistics presented in Table 33). Therefore, we need to also look at the estimates of the triple interactions to know the average effect of Ramadan on hours worked. In this case, we consider the set of interactions that include the *Ramadan* coefficient with the demeaned household composition variables and the *Islam* variable, and the *Ramadan* variable itself. When we add up the estimates, we find that an individual that does not deviate from the average household composition works 15 hours less during Ramadan, compared to those individuals that do not celebrate Ramadan,[43] and this effect is statistically significant when looking at the F-test in column (2). When looking at the set of triple interactions in detail, we find that one extra male child in an average household has a positive impact on the total number of hours worked, as well as having an extra old male individual at home. The rest of interactions have a statistically significant and negative impact on total hours worked, meaning that an extra adult (male or female) or an extra old female in an average household leads to a decrease of hours worked

[43] We compute the average effect of Ramadan in total hours worked as follows:

$$\begin{aligned} \mathbb{E}(hours \mid Ramadan = 1 \& Islam = 1) \quad = \quad & 0,0071 - 0,4207 + 13,3843 - 9,5640 + 67,6752 - 28,6123 \\ & -58,3492 - 1,2639 - 1,7570 + 11,7336 - 2,5601 - 5,7728 \approx -15 \end{aligned}$$

for a given individual.

Table 43: Heckman second-stage Analysis - Female subsample Estimation

	(1)	(2)	(3)	(4)	(5)	(6)
	Hours	Hours	Housework	Housework	Prod. Hours	Prod. Hours
Islam	1.7109 ***	3.5071 ***	2.4722 ***	3.4821 ***	-0.7468	0.0401
	(9.40)	(32.50)	(6.18)	(41.93)	(-1.65)	(1.23)
Ramadan	0.7800 ***	0.0071	-0.1279	-0.1846 ***	0.9331 ***	0.1761 ***
	(14.59)	(0.12)	(-1.39)	(-8.14)	(8.36)	(5.12)
Islam*Ramadan	-0.2251	-0.4207	0.2110	-0.2045 ***	-0.4156 **	-0.1619 *
	(-1.66)	(-1.69)	(0.82)	(-6.51)	(-3.18)	(-2.48)
Islam*Ramadan*Kids_male		13.3843 ***		1.6198 ***		11.6650 ***
		(26.81)		(11.66)		(34.51)
Islam*Ramadan*Adult_male		-9.5640 ***		-1.0293 ***		-8.8144 ***
		(-19.94)		(-11.75)		(-11.09)
Islam*Ramadan*Old_male		67.6752 ***		34.6047 ***		32.8909 ***
		(66.35)		(181.90)		(66.38)
Islam*Ramadan*Adult_female		-28.6123 ***		-7.1923 ***		-22.1345 ***
		(-19.84)		(-106.69)		(-55.93)
Islam*Ramadan*Old_female		-58.3492 ***		-35.2805 ***		-23.3228 ***
		(-35.15)		(-190.84)		(-19.34)
lambda	-0.3528 **	-0.1575 **	0.8201 ***	0.0581 **	-1.1850 ***	-0.2066 ***
	(-3.10)	(-2.64)	(20.26)	(16.16)	(-10.05)	(-7.14)
F-test	5.54	4082.79	14.75	7.51	4.70	10384.58
p-value	0.0626	0.000	0.000	0.000	0.0953	0.1293
Triple interaction	✗	✓	✗	✓	✗	✓
N	2919	2919	2919	2919	2919	2919
R-squared	0.0302	0.0527	0.0231	0.0402	0.0292	0.0431

In this table we present the results of the second step for the Heckman model with fixed effects for the female subsample, using the *Wooldridge correction* proposed in Wooldridge (1995). The $Islam$, *Ramadan* and $IslamRamadan$ coefficients are as explained before in Table 38. We control for female individual characteristics: age and its square, whether she has gone to school or not, whether she lives in an urban or rural area, the number of children in the household, the household size and the set of marital status dummies. Besides, we control as well for sample selection, including the Inverse of Mills' Ratio (the lambda coefficient). Moreover, we also include individual, year and seasonal fixed effects. In addition to that, in specifications (2), (4) and (6) we also include the set of triple interactions. We compute heteroskedastic-robust standard errors clustered at the village area level. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The full set of estimated coefficients are reported in Table 52 of the Appendix to this chapter.

For the case of hours worked at home and under paid jobs, we observe a similar behavior regarding the Ramadan effects and the set of triple interactions. The average individual interviewed during the Holy month significantly decrease their hours worked at home by 12.35 hours per week and by 3.73 hours at their jobs.[44] When looking at the effects of the triple interactions in more detail, we find that in both cases an extra adult or elder female in an average household drastically decreases the hours worked under a paid job or at home for individual i. We find these effects to be statistically significant. On the other hand, having an extra child or elder male significantly increases the allocation of hours worked in both places.

When we estimate the effect of Ramadan on hours worked for females (see Table 43), we do find that females that are interviewed during the Holy month of Ramadan work 1.08 hours less in total and 1.64 hours less in paid work, compared with females interviewed outside the Ramadan period.[45] When we perform the joint hypothesis test of the aggregate effect of Ramadan on labor supply, we find that at the 5% significance level it is not significant and, thus, we do not reject the null hypothesis in Hypothesis 1. However, at the 10% level of significance, we do find both estimated effects statistically significant and, thus, females interviewed during the Holy month of Ramadan do significantly reduce their total number of hours worked and hours in a paid job. However, when we analyze the effect on housework, we find that females interviewed during Ramadan increase their time allocation by 0.67 hours.[46] This effect is statistically significant at the 5% significance level, when we perform the F-test. Based on these results, there is evidence that females that celebrate the Holy month of Ramadan do reallocate their hours by working more at home and less at their jobs, compared with females that do not celebrate Ramadan or are interviewed outside this period. This is one of the main findings in this chapter; individual productivity does not decrease during Ramadan, but individuals do reallocate their labor hours from their paid job to housework.

When estimating the intra-household effects of Ramadan for females on labor supply, we find that females that live in an average household and celebrates Ramadan increase her housework by 16.86 hours, relative to other females that do not celebrate it. On the other hand, females that celebrate Ramadan do decrease on average their hours worked by 13.42 hours in relation to those that do not celebrate it.[47] Hence, in total we find that the average Ramadan effect leads to an increase of

[44] The way we compute such effects is the same one as in footnote 43, but using the estimates for housework and productive hours respectively.

[45] These effects are the total effect of Ramadan on labor supply, which are obtained by aggregating the estimated coefficients for $Ramadan$ and $Islam \times Ramadan$

[46] Again, this is the total effect that the Holy Month of Ramadan has on labor supply. This is computed by adding the estimates of $Ramadan$ and $Islam \times Ramadan$.

[47] The way we compute the effects for housework and hours worked is the same one as in footnote 43, but using the estimates for females housework and productive hours respectively, shown in Table 43.

4.61 total hours worked. We find these effects to be statistically significant.

Table 44: Fixed effects analysis - Male subsample Estimation

	(1)	(2)	(3)	(4)	(5)	(6)
	Hours	Hours	Housework	Housework	Prod. Hours	Prod. Hours
Islam	-1.023	4.666	0.747	2.071	-1.726	2.716
	(-0.28)	(1.08)	(0.41)	(0.97)	(-0.57)	(0.73)
Ramadan	0.651	-0.322	-0.521	1.009	1.309	-1.099
	(0.42)	(-0.13)	(-1.02)	(1.08)	(1.06)	(-0.57)
Islam× Ramadan	-0.0300	-6.824	-0.169	-2.897	-0.0242	-4.158
	(-0.01)	(-0.89)	(-0.21)	(-1.17)	(-0.01)	(-0.73)
Islam× Ramadan× Kids_male 27.46		4.214		22.69		
		(1.50)		(0.56)		(1.89)
Islam× Ramadan× Adult_male		9.309		5.184		4.219
		(0.67)		(1.17)		(0.40)
Islam× Ramadan× Old_male		44.88		15.54		29.96
		(1.95)		(1.87)		(1.56)
Islam× Ramadan× Adult_female		-7.623		2.894		-10.04
		(-0.51)		(0.51)		(-0.86)
Islam× Ramadan× Old_female		-59.78*		-21.22*		-38.14
		(-2.16)		(-2.26)		(-1.73)
_cons	13.53	3.276	4.192	3.119	10.33	1.227
	(1.64)	(0.39)	(1.23)	(0.83)	(1.61)	(0.20)
F-test	1.13	0.05	0.67	1.02	0.11	0.52
p-value	0.2898	0.8287	0.4143	0.3157	0.7388	0.4721
Triple interaction	✗	✓	✗	✓	✗	✓
N	2721	2721	2721	2721	2721	2721
R-squared	0.0486	0.0689	0.0195	0.0311	0.0567	0.0769

In this table we present the results for the male subsample for their hours worked, using a fixed effects model estimation. In this case, as most of males are working, we do not need to use a sample selection model - as we also observe from the selection equation estimation. The $Islam$, $Ramadan$ and $Islam \times Ramadan$ coefficients are as explained before in Table 38. We control for male individual characteristics: age and its square, whether he has gone to school or not, whether he lives in an urban or rural area, the number of children in the household, the household size and the set of marital status dummies. Besides, we control as well for sample selection, including the Inverse of Mills' Ratio (the lambda coefficient). Moreover, we also include individual, year and seasonal fixed effects. In addition to that, in specifications (2), (4) and (6) we also include the set of triple interactions. We compute heteroskedastic-robust standard errors clustered at the village area level. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The full set of estimated coefficients are reported in Table 53 of the Appendix to this chapter.

When looking at the set of triple interactions individually, we observe that an extra female in an average household (either adult or elder) significantly decreases the number of hours worked in the household or under a paid job for females that celebrate Ramadan, compared to those that do not. On the other hand, we find a positive and statistically significant effect on living with one extra old male in an average household.

The effect of living with an extra adult male in an average household leads to a positive and significant effect on housework hours, but to a negative one in hours worked. The effect of having an extra child in an average household leads to a significant increase in housework hours and a decrease of hours allocated to their jobs, meaning that when there is an extra child in the average household, women spend more time taking care of children and working more hours in the household. This reinforces the evidence found before, where women reallocate their labor hours to working more at home for their families, instead of working at their jobs.

When analyzing the results for males (see Table 44), using a simple fixed effects models, as explained above; we find that males increase their total hours worked in their jobs by 1.29 hours per week, but they work 0.69 hours less at home. These estimates lead to a total increase in 0.62 hours per week for males that celebrate the Ramadan festivity, [48] compared with individuals that do not celebrate the Ramadan. However, we do not find any statistically significant effects of Ramadan into hours worked - either at the household or under paid jobs. Such effect is confirmed when we perform the joint hypothesis test under Hypothesis 1, where we fail to reject the null hypothesis. Therefore, for males, we do not find that Ramadan leads to time reallocation of hours worked, as we did find for females.

If we analyze the family composition effect during Ramadan, we find a decrease of 1.13 housework hours and 0.34 hours in their jobs and, thus, a total decrease of 1.46 hours due to the Ramadan celebration. [49] However, none of these effects are statistically significant.

When looking at the set of triple interactions, we find that for males that do celebrate Ramadan, an extra male in the household leads to an increase in hours working at home at their jobs. However, the fact of having an extra female in households that take part in Ramadan leads to a decrease in all individuals' labor supply estimates, except for housework, where an extra adult female affects

[48] These effects are computed by adding the estimated coefficients of $Ramadan$ and $Islam \times Ramadan$, in order to get the total effect that Ramadan has on housework hours.

[49] The way we compute such effects is the same one as in Footnote 43, but using the estimates for males total hours worked, housework and productive hours respectively, shown in Table 44.

positively to the hours worked there. Nevertheless, these effects are not statistically significant, for any of the labor supply estimators.

Finally, lambda, also known as Inverse of Mills' Ratio, controls for sample selection. More precisely, with the introduction of the lambda factor, we are testing for sample selection under the following hypothesis, previously stated in section 5.2:

Hypothesis 3:

$$H_0 : \ \psi = 0 \text{ No sample selection.}$$

$$H_a : \ \psi \neq 0 \text{ There is sample selection.}$$

We find that the lambda coefficient is statistically significant in all cases. This means that sample selection exists when we estimate the labor supply outcomes for the whole sample and the female subsample, as we do reject the null hypothesis stated under Hypothesis 3 in almost all cases. Therefore, the Heckman model is the correct estimation method for labor supply in Malawi.

Moreover, we also test for the parallel trend assumption, to check that individuals do not anticipate the potential treatment effects into labor supply. In this case, we observe that the parallel trend assumption holds at the 5% significance level (see Table 56 in the appendix) for the male subsample and the entire sample and, thus, individuals do not anticipate the change in hours worked in the pre-treatment period. However, we observe that the parallel trends assumption does not hold for the female subsample, when we analyze the hours worked in their paid jobs and the total hours worked. Hence, it seems that females anticipate the Ramadan and plan their number of hours worked at their jobs in advance. This result should not be surprising, as Ramadan is not a shock - they know when the Ramadan time comes and may plan to do extra work done before it arrives to have more time to allocate to the housework. This might lead to a downward bias in the estimates of Ramadan for productive hours for females, as they do plan ahead the festivity and tend to work for more hours prior to the arrival of Ramadan.

6.2 The Effect of Ramadan on Labor Supply on the Extensive Margin

In this subsection we analyze the obtained results for the extensive margins presented in Equations (20) and (21). We use not only data from Malawi, but also data from Bangladesh to examine whether the observed behavior in the Malawian labor market is similar to the Bangladeshi one when the Holy month of Ramadan takes place. For Bangladesh we restricted the analysis for the

Muslim subsample only, as they cover most of the sample; therefore, to avoid multicollinearity problems we include only sets of double interactions between Ramadan and the ratio of children, adults and old males and females within the household over the total household size.

Table 45: Extensive Margins - Malawi

	(1) work	(2) work	(3) work	(4) Labor Force	(5) Labor Force	(6) Labor Force
Islam	-0.233	-0.243	-0.247	0.619	-0.0935	-0.877
	(-0.95)	(-0.75)	(-0.61)	(1.68)	(-0.13)	(-0.66)
Ramadan	0.114	0.0447	0.161	0.102	0.151	0.525
	(1.60)	(0.49)	(0.80)	(0.85)	(0.81)	(0.94)
Islam × Ramadan	-0.0445	-0.101	0.0307	-0.384	-0.300	-0.249
	(-0.28)	(-0.55)	(0.07)	(-1.45)	(-0.62)	(-0.16)
Individual Controls	✗	✓	✓	✗	✓	✓
Triple interaction	✗	✗	✓	✗	✗	✓
Aggregate Ramadan effect			0.37			-0.19
F-test	0.24	0.12	0.06	1.47	0.11	0.40
p-value	0.6247	0.7271	0.7990	0.2251	0.7394	0.5246
N	6220	4273	5114	2268	1990	1990

In this table we present the results for the extensive margins in Malawi, using a conditional Logit model, as proposed in Chamberlain (2010). The $Islam$, *Ramadan* and $Islam \times$ *Ramadan* coefficients are as explained before in Table 38. We control for individual characteristics: age and its square, whether the individual has gone to school or not, whether he/she lives in an urban or rural area, the number of children in the household, the household size and the set of marital status dummies. Moreover, we also include individual, year and seasonal fixed effects. In addition to that, in specifications (3) and (6) we also include the set of triple interactions. We compute robust standard errors clustered at the village area level. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The full set of estimated coefficients are reported in Table 54 of the Appendix to this chapter.

The F-test performs a joint significant test of the Ramadan cultural barrier variables (*Ramadan* and $Islam \times$ *Ramadan*), on the extensive margins. The null hypothesis of the F-test is that *individuals who participate in the Ramadan festivity do not alter their probability of being employed or in the labor force during the Holy month.*

Table 45 offers the estimated results for Malawi. We do not find any Ramadan effect on the likelihood of working or being in the labor force. Such effect is confirmed when we compute the joint hypothesis test for the aggregate effect of Ramadan on the likelihood of working or being in the labor force, where we aggregate all the marginal effects of the estimated coefficients of interest shown in Table 54 (see columns (3) and (6)). We also perform the joint hypothesis test under the null hypothesis that the Holy month of Ramadan has no effect on the extensive margins.

Table 46: Extensive Margins - Bangladesh

	(1)	(2)	(3)	(4)	(5)	(6)
	work	work	work	Labor Force	Labor Force	Labor Force
Ramadan	1.354***	-0.260*	-0.118	0.542***	-0.250***	-0.0524
	(8.01)	(-2.27)	(-0.42)	(6.64)	(-3.50)	(-0.30)
Individual Controls	✗	✓	✓	✗	✓	✓
Triple interaction	✗	✗	✓	✗	✗	✓
Aggregate Ramadan effect			-0.24			-0.08
F-test			0.42			1.51
p-value			0.5192			0.2188
N	7844	7844	7844	9792	9776	9776

In this table we present the results for the extensive margins in Bangladesh, using a conditional Logit model, as proposed in Chamberlain (2010). The $Islam$, *Ramadan* and $Islam \times$ *Ramadan* coefficients are as explained before in Table 38. We control for individual characteristics: age and its square, whether the individual has gone to school or not, whether he/she lives in an urban or rural area, the number of children in the household, the household size and the set of marital status dummies. Moreover, we also include individual, year and seasonal fixed effects. In addition to that, in specifications (3) and (6) we also include the set of triple interactions. We compute robust standard errors clustered at the village area level. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The full set of estimated coefficients are reported in Table 54 of the Appendix to this chapter.
The F-test performs a joint significant test of the Ramadan cultural barrier variables (*Ramadan* and the set of double interactions presented in Table 55), on the extensive margins. The null hypothesis of the F-test is that *individuals who participate in the Ramadan festivity do not alter their probability of being employed or in the labor force during the Holy month.*

However, when we observe the estimated effects for Bangladesh in Table 46, we observe that individuals who are interviewed during the Holy month of Ramadan, experience a negative effect on the likelihood of working or being part of the labor force participation (see columns (2) and (4)), and it is statically significant. Nevertheless, when we control for household composition effects (see

Table Table 55), we observe that the *Ramadan* coefficient loses all its statistical significance, but it still affects both probabilities negatively. However, to estimate the overall effect of Ramadan on the extensive margins, we need to aggregate all the estimates of double interactions in Table 55 and perform a joint significant test, to see if there is an effect in the extensive margins or not.

Observing at the set of double interactions in Table 55, none of them presents statistically significant results for the probability of working. Nevertheless, if we look at the labor force, we find that one extra adult female in the household reduces the likelihood of being in the labor force in Bangladesh during the Holy month of Ramadan.

When we aggregate the estimates of triple interaction estimates from Malawi, and the estimates of double ones for Bangladesh reported in Tables 54 and 55 respectively, and we observe the total effect of Ramadan on the likelihood of working or being in the labor force, we find that, on average, an individual has 37% more chance of working and 19% fewer options of being in the labor force during Ramadan in Malawi. Whereas an individual has 24% less chance of being employed and 8% less chance of being in the labor force during the Ramadan period in Bangladesh. However, the performed F-tests show that none of the effects are statistically significant at the 5% significance level. With that, we can affirm that at the extensive margins there are no relevant effects of Ramadan, neither in Malawi nor in Bangladesh.

7 Conclusion

This chapter analyzes the effect that religious practices have on individual labor supply decisions. Using household data from Malawi, available in the LSMS, we find evidence that individuals do reallocate their labor time resources, from paid jobs to housework, during the Holy month of Ramadan; a period where Muslims experience fasting during the sunlight hours - from sunrise to sunset. Therefore, individuals tend to spend more time at home, most likely preparing family/friends meetings to have meals all together at the allowed times. This effect is accentuated when we do the analysis for the female subsample. Moreover, we find that males do also increase their time allocation to hours worked in both housework and paid jobs. But only the increase in paid jobs is statistically significant.

As mentioned in the introduction, to our best knowledge, this is the first chapter to analyze such effects at the individual level and to analyze the reallocation of hours between paid jobs and household work, also at the intra-household level. We performed a Heckman model using the

Wooldridge (1995) correction to control for potential sample selection and allowing for individual fixed effects. In summary, we can assert that females work less at their jobs, but this effect is compensated for by household work. Therefore, it can be concluded that females that celebrate Ramadan do not work less, but they spend more time at home, most probably preparing activities for their family and friends during the Holy month of Ramadan.

The last thing done in this chapter was the analysis of the extensive margins for Malawi, and we also introduced a new country to support the analysis. In this case we used survey data from Bangladesh, and we can observe that Ramadan has no effect on the extensive margins for any of the countries analyzed. These results support the evidence found in the intensive margins analysis in Malawi: labor is not reduced during Ramadan, there is only a reallocation of time between paid jobs and housework.

Finally, we would like to emphasize the fact that we could not apply the intensive margins analysis for Bangladesh due to data restrictions, as all the variables needed for the analysis are given on a yearly basis, and other countries that were analyzed do not provide surveys taking place during Ramadan or the data analysis does not provide evidence of any Ramadan effect on labor supply.

Our findings contribute to other studies in which religious events have a relevant impact in economic activity and economic decisions (see Barro and McCleary, 2003; Campante and Yanagizawa-Drott, 2016; Demiroglu et al., 2017). In our case, we provide evidence showing that Ramadan has a direct impact in labor decisions for those individuals that are under its practice. Therefore, this might lead to some policy implications in terms of readjustments of labor market structures in countries where the practice of Ramadan is popular, as it has a significant effect on an important proportion of the world population.

Appendix Chapter 3

Table 47: Labor Supply allocation in Malawi - Pooled OLS

	(1)	(2)	(3)	(4)	(5)	(6)
	Hours	Hours	Housework	Housework	Prod. Hours	Prod. Hours
Islam	-1.718	-0.754	0.239	-0.0192	-1.964	-0.731
	(-1.27)	(-0.81)	(0.45)	(-0.05)	(-1.86)	(-0.90)
Ramadan	1.051	0.451	0.507	0.135	0.543	0.365
	(1.02)	(0.50)	(1.49)	(0.45)	(0.66)	(0.49)
Islam× ramadan	1.071	2.011	0.416	0.671	0.664	1.290
	(0.48)	(1.12)	(0.47)	(1.01)	(0.41)	(0.95)
Age		0.0270		-0.0309		0.0590
		(0.42)		(-1.05)		(1.29)
Age^2		-0.00115		0.0000439		-0.00120*
		(-1.68)		(0.15)		(-2.44)
School		1.350*		0.0971		1.276**
		(2.14)		(0.36)		(2.69)
Urban		4.595***		1.245*		3.385***
		(3.66)		(2.56)		(3.73)
Num. children		-0.736***		-0.127		-0.591***
		(-3.44)		(-1.22)		(-4.03)
Household size		-0.395*		-0.116		-0.288*
		(-2.41)		(-1.82)		(-2.34)
_cons	16.39***	15.22***	5.362***	5.790***	11.03***	9.412***
	(23.17)	(11.78)	(24.60)	(10.12)	(19.29)	(9.76)
N	8186	10686	8192	10694	8200	10705
R-Squared	0.000816	0.0204	0.000782	0.00696	0.00128	0.0225

This Table presents the extended results of Table 38, where we described in detail the specifications and the estimation process. We are also controlling for marital status dummies. We compute robust standard errors clustered at the village area level. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 48: Labor Supply allocation in Malawi - Panel Data FE

	(1)	(2)	(3)	(4)	(5)	(6)
	Hours	Hours	Housework	Housework	Prod. Hours	Prod. Hours
Islam	2.307	2.375	1.659	1.049	0.669	1.390
	(0.68)	(0.63)	(0.86)	(0.58)	(0.36)	(0.63)
Ramadan	0.600	0.685	0.203	0.0420	0.388	0.725
	(0.60)	(0.61)	(0.62)	(0.13)	(0.49)	(0.77)
Islam× ramadan	1.482	0.687	0.840	0.525	0.642	0.0467
	(0.66)	(0.33)	(0.93)	(0.63)	(0.42)	(0.03)
Land area		-0.00836		0.00859		-0.0167
		(-0.35)		(1.82)		(-0.74)
Age		0.0889		0.0472		0.0185
		(0.33)		(0.48)		(0.08)
Age^2		-0.00704*		-0.00175		-0.00513*
		(-2.59)		(-1.42)		(-2.52)
School		1.177		0.00786		1.179
		(1.24)		(0.02)		(1.68)
Urban		-0.673		-0.770		0.208
		(-0.26)		(-0.64)		(0.12)
Num. children		-1.010*		-0.421*		-0.609
		(-2.28)		(-2.09)		(-1.79)
Household size		-0.224		-0.0241		-0.219
		(-0.82)		(-0.22)		(-0.98)
_cons	14.88***	18.64**	5.023***	5.928*	9.861***	13.29**
	(25.76)	(3.30)	(16.98)	(2.48)	(28.61)	(2.85)
N	13902	10686	13911	10694	13936	10705
R-Squared	0.000636	0.00790	0.000913	0.00465	0.000284	0.00760

This Table presents the extended results of Table 39, where we described in detail the specifications and the estimation process. We compute robust standard errors clustered at the village area level. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 49: Labor Supply allocation in Malawi - Tobit

	(1)	(2)	(3)	(4)	(5)	(6)
	Hours	Hours	Housework	Housework	Prod. Hours	Prod. Hours
Islam	-1.625	1.193	1.161	2.366	-3.054	-0.579
	(-0.43)	(0.28)	(0.43)	(0.78)	(-0.86)	(-0.14)
Ramadan	1.289	0.461	0.743	0.198	0.632	0.222
	(1.01)	(0.31)	(1.18)	(0.30)	(0.40)	(0.12)
Islam$\times$ ramadan	0.759	0.222	0.182	0.277	1.064	0.0261
	(0.21)	(0.07)	(0.11)	(0.16)	(0.29)	(0.01)
Land area		-0.0319		0.00913		-0.0578
		(-1.73)		(0.58)		(-1.27)
Age		0.743		0.0960		0.965
		(1.50)		(0.32)		(1.73)
Age^2		-0.0116**		-0.00493		-0.0120**
		(-2.79)		(-1.91)		(-2.88)
School		2.921*		0.711		2.509
		(2.25)		(0.89)		(1.92)
Urban		-0.421		-2.275		-0.450
		(-0.14)		(-1.25)		(-0.16)
Num. children		-1.450*		-0.597		-1.837**
		(-2.29)		(-1.44)		(-2.98)
Household size		-0.390		-0.0626		-0.504
		(-1.07)		(-0.33)		(-1.25)
_cons	6.831***	8.325***	-4.222***	-1.350	-4.688***	-2.388
	(9.31)	(4.02)	(-8.27)	(-0.93)	(-4.91)	(-1.10)
N	13902	10686	13911	10694	13936	10705

This Table presents the extended results of Table 40, where we described in detail the specifications and the estimation process. We compute robust standard errors clustered at the village area level. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 50: Selection Equation

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	Labor Force	Labor Force	Labor Force	Labor Force	Labor Force	Labor Force
area/adults*age	0.00196	0.00156	0.00213	0.00224*	0.000837	0.00121
	(1.587)	(1.498)	(1.869)	(1.974)	(0.645)	(0.693)
area/adults*male	-0.0698*	-0.0431				
	(-2.395)	(-1.684)				
area/adults*married	-0.0928	-0.0514	-0.0832	-0.0767	-0.00879	-0.00960
	(-1.776)	(-1.153)	(-1.648)	(-1.275)	(-0.251)	(-0.362)
area/adults*polygamia	0.116	0.165	3.133	2.627	0.140	0.135
	(0.705)	(1.012)	(1.952)	(1.721)	(0.930)	(0.788)
area/adults*separated	-0.274**	-0.214*	-0.233*	-0.238	-0.0459	-0.0406
	(-2.738)	(-2.235)	(-2.151)	(-1.348)	(-0.443)	(-0.194)
area/adults*divorced	0.0864	0.272	0.146	0.301	0.0529	-0.0486
	(0.512)	(1.180)	(0.806)	(1.384)	(0.468)	(-0.832)
area/adults*widowed	-0.205*	-0.155	-0.187*	-0.177*	0.292	0.181
	(-2.207)	(-1.829)	(-2.117)	(-2.085)	(1.189)	(0.696)
area/adults	0.0480**	0.0240	0.00941	-0.00385	-0.0285	-0.0313
	(2.681)	(1.358)	(0.379)	(-0.177)	(-1.028)	(-1.054)
islam	0.163	0.163	0.328	0.174	-0.174	-0.412
	(1.304)	(0.979)	(1.898)	(0.711)	(-0.979)	(-1.514)
ramadan	0.0335	-0.179	0.0164	-0.309*	0.0425	0.288
	(0.916)	(-1.780)	(0.320)	(-2.235)	(0.872)	(1.780)
Islam*Ramadan	0.00950	0.279	0.0677	0.255	-0.0128	-0.345
	(0.143)	(1.647)	(0.726)	(0.593)	(-0.113)	(-1.040)
Islam*Ramadan*Kids_male		0.249		0.650		1.231
		(0.504)		(0.788)		(1.421)
Islam*Ramadan*Adult_male		-0.422		-0.149		0.308
		(-1.036)		(-0.156)		(0.489)
Islam*Ramadan*Old_male		-0.0419		0.0452		1.812
		(-0.0490)		(0.0242)		(1.618)
Islam*Ramadan*Adult_female		-0.898		-1.336		-0.155
		(-1.700)		(-0.989)		(-0.174)
Islam*Ramadan*Old_female		-0.875		-0.431		-1.896
		(-1.052)		(-0.367)		(-1.551)

Selection Equation - continued

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	Labor Force	Labor Force	Labor Force	Labor Force	Labor Force	Labor Force
Ramadan*Kids_male		0.0286		0.157		-0.777*
		(0.134)		(0.422)		(-2.345)
Ramadan*Adult_male		0.330		0.530		-0.114
		(1.665)		(1.235)		(-0.358)
Ramadan*Old_male		0.0650		-0.134		-0.975
		(0.154)		(-0.129)		(-1.457)
Ramadan*Adult_female		0.607*		0.949*		-0.254
		(2.121)		(1.978)		(-0.771)
Ramadan*Old_female		-0.291		-0.168		0.100
		(-0.863)		(-0.343)		(0.187)
age	-0.0319**	-0.0154	0.0137	0.0413**	-0.00570	0.00806
	(-3.086)	(-1.328)	(0.977)	(2.728)	(-0.528)	(0.679)
married	0.578***	0.660***	1.005***	1.372***	-0.102	-0.141
	(5.202)	(5.895)	(6.258)	(6.664)	(-0.848)	(-0.929)
polygamia	0.248	0.293	-0.647	-0.312	-0.303	-0.297
	(0.893)	(1.066)	(-0.897)	(-0.446)	(-1.122)	(-0.990)
separated	0.588**	0.537**	0.881***	1.150***	-0.350*	-0.652***
	(3.102)	(2.754)	(3.999)	(4.357)	(-2.552)	(-3.887)
divorced	0.286	0.131	0.466	0.590*	0.145	0.0244
	(1.240)	(0.474)	(1.684)	(2.136)	(0.635)	(0.0744)
widowed	0.391*	0.425*	0.806***	1.066***	-0.821*	-0.712
	(2.145)	(2.396)	(3.951)	(4.334)	(-2.166)	(-1.541)
school	-0.250***	-0.288***	-0.323***	-0.376***	-0.191**	-0.199**
	(-5.683)	(-6.058)	(-5.684)	(-6.025)	(-3.097)	(-2.794)
urban	0.167	0.192* *	0.170	0.168	0.235	0.241
	(1.839)	(2.052)	(1.249)	(1.192)	(1.687)	(1.604)
num_children	-0.251***	-0.209***	-0.251***	-0.184***	-0.252***	-0.242***
	(-12.96)	(-9.727)	(-7.620)	(-3.456)	(-9.521)	(-5.663)
hhsize	0.0480***	0.0661***	0.0527***	0.0746**	0.0392**	0.0194
	(4.441)	(5.420)	(3.487)	(2.987)	(2.987)	(0.946)
Constant	-2.239***	-2.925***	-1.623***	-2.342***	-1.444***	-1.207***
	(-12.17)	(-15.09)	(-11.21)	(-9.663)	(-7.920)	(-3.937)
Sample	Whole	Whole	Female	Female	Male	Male
N	11768	11768	6059	6059	5709	5709

This Table presents the extended results of Table 41, where we described in detail the specifications and the estimation process. We compute robust standard errors clustered at the village area level. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 51: Heckman second-stage analysis - Entire Sample Estimation

	(1) Hours	(2) Hours	(3) Housework	(4) Housework	(5) Prod. Hours	(6) Prod. Hours
Islam	1.7109 ***	3.5071 ***	2.4722 ***	3.4821 ***	-0.7468	0.0401
	(9.40)	(32.50)	(6.18)	(41.93)	(-1.65)	(1.23)
Ramadan	0.7800 ***	0.0071	-0.1279	-0.1846 ***	0.9331 ***	0.1761 ***
	(14.59)	(0.12)	(-1.39)	(-8.14)	(8.36)	(5.12)
Islam*Ramadan	-0.2251	-0.4207	0.2110	-0.2045 ***	-0.4156 **	-0.1619 *
	(-1.66)	(-1.69)	(0.82)	(-6.51)	(-3.18)	(-2.48)
Islam*Ramadan*Kids_male		13.3843 ***		1.6198 ***		11.6650 ***
		(26.81)		(11.66)		(34.51)
Islam*Ramadan*Adult_male		-9.5640 ***		-1.0293 ***		-8.8144 ***
		(-19.94)		(-11.75)		(-11.09)
Islam*Ramadan*Old_male		67.6752 ***		34.6047 ***		32.8909 ***
		(66.35)		(181.90)		(66.38)
Islam*Ramadan*Adult_female		-28.6123 ***		-7.1923 ***		-22.1345 ***
		(-19.84)		(-106.69)		(-55.93)
Islam*Ramadan*Old_female		-58.3492 ***		-35.2805 ***		-23.3228 ***
		(-35.15)		(-190.84)		(-19.34)
Ramadan*Kids_male		-1.2639 **		-1.0894 ***		-0.2362 *
		(-2.87)		(-8.33)		(-2.09)
Ramadan*Adult_male		-1.7570 ***		-2.8349 ***		1.3261 ***
		(-4.38)		(-82.02)		(8.32)
Ramadan*Old_male		11.7336 ***		2.5335 ***		9.0673 ***
		(22.96)		(31.44)		(41.05)
Ramadan*Adult_female		-2.5601 ***		-3.0816 ***		1.1631 ***
		(-6.70)		(-104.07)		(4.62)
Ramadan*Old_female		-5.7728 ***		-0.2172 *		-5.3452 ***
		(-16.83)		(-2.24)		(-24.99)
Kids_male		1.5192 ***		-1.8634 ***		3.2283 ***
		(5.35)		(-60.13)		(20.92)
Adult_male		2.2404 ***		0.3864 **		1.8735 ***
		(4.15)		(15.44)		(14.35)
Old_male		-2.3804 ***		-1.4919 **		-0.9717 ***
		(-5.05)		(-12.27)		(-4.45)
Adult_female		6.1419 ***		3.4651 ***		2.8132 ***
		(8.53)		(152.86)		(9.32)
Old_female		17.8363 ***		2.0603 ***		15.7562 ***
		(23.48)		(12.34)		(46.62)

Heckman second-stage analysis - Entire Sample Estimation (continued)

	(1)	(2)	(3)	(4)	(5)	(6)
	Hours	Hours	Housework	Housework	Prod. Hours	Prod. Hours
age	-0.2288 **	-0.4902 ***	-0.0077	0.1486 ***	-0.2106 **	-0.6246 ***
	(-3.04)	(-5.12)	(-0.31)	(8.03)	(-3.29)	(-22.21)
age2	0.0010 ***	0.0011 ***	0.0021 ***	0.0023 ***	-0.0012 ***	-0.0012 ***
	(4.79)	(7.51)	(24.16)	(53.72)	(-4.48)	(-10.86)
married	5.4403 ***	0.4280 ***	1.8526 ***	0.2219 ***	3.4533 ***	0.1816 ***
	(75.43)	(11.09)	(33.17)	(11.90)	(16.74)	(5.54)
polygamia	12.6087 ***	5.9709 ***	4.2791 ***	1.9231 ***	8.2055 ***	3.9043 ***
	(185.95)	(13.73)	(45.10)	(97.00)	(41.94)	(36.37)
separated	4.4094 ***	13.8103 ***	2.2479 ***	4.3524 ***	2.0643 ***	9.3329 ***
	(51.89)	(33.79)	(17.51)	(208.95)	(12.74)	(82.47)
divorced	5.6005 ***	3.6323 ***	2.2768 ***	1.9054 ***	3.2261 ***	1.6274 ***
	(40.50)	(11.76)	(36.70)	(100.23)	(15.21)	(19.34)
widowed	4.9291 ***	4.7453 ***	1.6898 ***	2.2186 ***	3.1254 ***	2.3994 ***
	(45.51)	(12.81)	(20.25)	(115.62)	(13.90)	(30.85)
school	3.0264 ***	5.9696 ***	-0.0056	1.6936 ***	3.0083 ***	4.1592 ***
	(65.46)	(19.26)	(-0.07)	(56.59)	(58.05)	(43.21)
urban	1.1680 ***	2.9361 ***	-0.5085 **	-0.0488 ***	1.7242 ***	2.9548 ***
	(8.47)	(58.26)	(-2.81)	(-3.43)	(8.72)	(84.47)
num_children	-0.4984 ***	0.6330	-0.3779 ***	-0.6059 ***	-0.1066 ***	1.3079 **
	(-17.58)	(1.43)	(-22.02)	(-75.29)	(-3.66)	(88.24)
hhsize	-0.7228 ***	-0.4299 ***	0.0271	-0.1989 ***	-0.7549 ***	-0.1927 ***
	(-41.93)	(-4.60)	(1.95)	(-27.68)	(-54.06)	(-4.56)
lambda	-0.3528 **	-0.1575 **	0.8201 ***	0.0581 **	-1.1850 ***	-0.2066 ***
	(-3.10)	(-2.64)	(20.26)	(16.16)	(-10.05)	(-7.14)
cons	42.6497 ***	-0.7191 ***	9.4973 ***	0.1299 ***	33.1639 ***	-0.8604 ***
	(125.71)	(-6.17)	(33.53)	(4.29)	(72.68)	(-12.72)
N	5732	5732	5732	5732	5732	5732
R-Squared	0.0334	0.0465	0.0162	0.0264	0.0372	0.0469

This Table presents the extended results of Table 42, where we described in detail the specifications and the estimation process. We compute robust standard errors clustered at the village area level. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 52: Heckman second-stage analysis - Female subsample Estimation

	(1)	(2)	(3)	(4)	(5)	(6)
	Hours	Hours	Housework	Housework	Prod. Hours	Prod. Hours
Islam	1.7109 ***	3.5071 ***	2.4722 ***	3.4821 ***	-0.7468	0.0401
	(9.40)	(32.50)	(6.18)	(41.93)	(-1.65)	(1.23)
Ramadan	0.7800 ***	0.0071	-0.1279	-0.1846 ***	0.9331 ***	0.1761 ***
	(14.59)	(0.12)	(-1.39)	(-8.14)	(8.36)	(5.12)
Islam*Ramadan	-0.2251	-0.4207	0.2110	-0.2045 ***	-0.4156 **	-0.1619 *
	(-1.66)	(-1.69)	(0.82)	(-6.51)	(-3.18)	(-2.48)
Islam*Ramadan*Kids_male		13.3843 ***		1.6198 ***		11.6650 ***
		(26.81)		(11.66)		(34.51)
Islam*Ramadan*Adult_male		-9.5640 ***		-1.0293 ***		-8.8144 ***
		(-19.94)		(-11.75)		(-11.09)
Islam*Ramadan*Old_male		67.6752 ***		34.6047 ***		32.8909 ***
		(66.35)		(181.90)		(66.38)
Islam*Ramadan*Adult_female		-28.6123 ***		-7.1923 ***		-22.1345 ***
		(-19.84)		(-106.69)		(-55.93)
Islam*Ramadan*Old_female		-58.3492 ***		-35.2805 ***		-23.3228 ***
		(-35.15)		(-190.84)		(-19.34)
Ramadan*Kids_male		-1.2639 **		-1.0894 ***		-0.2362 *
		(-2.87)		(-8.33)		(-2.09)
Ramadan*Adult_male		-1.7570 ***		-2.8349 ***		1.3261 ***
		(-4.38)		(-82.02)		(8.32)
Ramadan*Old_male		11.7336 ***		2.5335 ***		9.0673 ***
		(22.96)		(31.44)		(41.05)
Ramadan*Adult_female		-2.5601 ***		-3.0816 ***		1.1631 ***
		(-6.70)		(-104.07)		(4.62)
Ramadan*Old_female		-5.7728 ***		-0.2172 *		-5.3452 ***
		(-16.83)		(-2.24)		(-24.99)
Kids_male		1.5192 ***		-1.8634 ***		3.2283 ***
		(5.35)		(-60.13)		(20.92)
Adult_male		2.2404 ***		0.3864 **		1.8735 ***
		(4.15)		(15.44)		(14.35)
Old_male		-2.3804 ***		-1.4919 **		-0.9717 ***
		(-5.05)		(-12.27)		(-4.45)
Adult_female		6.1419 ***		3.4651 ***		2.8132 ***
		(8.53)		(152.86)		(9.32)
Old_female		17.8363 ***		2.0603 ***		15.7562 ***
		(23.48)		(12.34)		(46.62)

Heckman second-stage analysis - Female subsample Estimation (continued)

	(1) Hours	(2) Hours	(3) Housework	(4) Housework	(5) Prod. Hours	(6) Prod. Hours
age	-0.2288 **	-0.4902 ***	-0.0077	0.1486 ***	-0.2106 **	-0.6246 ***
	(-3.04)	(-5.12)	(-0.31)	(8.03)	(-3.29)	(-22.21)
age2	0.0010 ***	0.0011 ***	0.0021 ***	0.0023 ***	-0.0012 ***	-0.0012 ***
	(4.79)	(7.51)	(24.16)	(53.72)	(-4.48)	(-10.86)
married	5.4403 ***	0.4280 ***	1.8526 ***	0.2219 ***	3.4533 ***	0.1816 ***
	(75.43)	(11.09)	(33.17)	(11.90)	(16.74)	(5.54)
polygamia	12.6087 ***	5.9709 ***	4.2791 ***	1.9231 ***	8.2055 ***	3.9043 ***
	(185.95)	(13.73)	(45.10)	(97.00)	(41.94)	(36.37)
separated	4.4094 ***	13.8103 ***	2.2479 ***	4.3524 ***	2.0643 ***	9.3329 ***
	(51.89)	(33.79)	(17.51)	(208.95)	(12.74)	(82.47)
divorced	5.6005 ***	3.6323 ***	2.2768 ***	1.9054 ***	3.2261 ***	1.6274 ***
	(40.50)	(11.76)	(36.70)	(100.23)	(15.21)	(19.34)
widowed	4.9291 ***	4.7453 ***	1.6898 ***	2.2186 ***	3.1254 ***	2.3994 ***
	(45.51)	(12.81)	(20.25)	(115.62)	(13.90)	(30.85)
school	3.0264 ***	5.9696 ***	-0.0056	1.6936 ***	3.0083 ***	4.1592 ***
	(65.46)	(19.26)	(-0.07)	(56.59)	(58.05)	(43.21)
urban	1.1680 ***	2.9361 ***	-0.5085 **	-0.0488 ***	1.7242 ***	2.9548 ***
	(8.47)	(58.26)	(-2.81)	(-3.43)	(8.72)	(84.47)
num_children	-0.4984 ***	0.6330	-0.3779 ***	-0.6059 ***	-0.1066 ***	1.3079 **
	(-17.58)	(1.43)	(-22.02)	(-75.29)	(-3.66)	(88.24)
hhsize	-0.7228 ***	-0.4299 ***	0.0271	-0.1989 ***	-0.7549 ***	-0.1927 ***
	(-41.93)	(-4.60)	(1.95)	(-27.68)	(-54.06)	(-4.56)
lambda	-0.3528 **	-0.1575 **	0.8201 ***	0.0581 **	-1.1850 ***	-0.2066 ***
	(-3.10)	(-2.64)	(20.26)	(16.16)	(-10.05)	(-7.14)
cons	42.6497 ***	-0.7191 ***	9.4973 ***	0.1299 ***	33.1639 ***	-0.8604 ***
	(125.71)	(-6.17)	(33.53)	(4.29)	(72.68)	(-12.72)
N	2919	2919	2919	2919	2919	2919
R-Squared	0.0302	0.0527	0.0231	0.0402	0.0292	0.0431

This Table presents the extended results of Table 43, where we described in detail the specifications and the estimation process. We compute robust standard errors clustered at the village area level. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 53: Heckman second-stage analysis - Male subsample Estimation

	(1)	(2)	(3)	(4)	(5)	(6)
	Hours	Hours	Housework	Housework	Prod. Hours	Prod. Hours
Islam	-1.023	4.666	0.747	2.071	-1.726	2.716
	(-0.28)	(1.08)	(0.41)	(0.97)	(-0.57)	(0.73)
Ramadan	0.651	-0.322	-0.521	1.009	1.309	-1.099
	(0.42)	(-0.13)	(-1.02)	(1.08)	(1.06)	(-0.57)
Islam*Ramadan	-0.0300	-6.824	-0.169	-2.897	-0.0242	-4.158
	(-0.01)	(-0.89)	(-0.21)	(-1.17)	(-0.01)	(-0.73)
Islam*Ramadan*Kids_male 27.46		4.214		22.69		
		(1.50)		(0.56)		(1.89)
Islam*Ramadan*Adult_male		9.309		5.184		4.219
		(0.67)		(1.17)		(0.40)
Islam*Ramadan*Old_male		44.88		15.54		29.96
		(1.95)		(1.87)		(1.56)
Islam*Ramadan*Adult_female		-7.623		2.894		-10.04
		(-0.51)		(0.51)		(-0.86)
Islam*Ramadan*Old_female		-59.78*		-21.22*		-38.14
		(-2.16)		(-2.26)		(-1.73)
Ramadan*Kids_male		-8.132		-3.715		-4.001
		(-1.36)		(-1.34)		(-0.90)
Ramadan*Adult_male		0.450		-1.033		1.278
		(0.08)		(-0.41)		(0.30)
Ramadan*Old_male		5.539		0.0936		5.088
		(0.57)		(0.02)		(0.60)
Ramadan*Adult_female		8.602		-2.834		11.09
		(0.97)		(-0.70)		(1.74)
Ramadan*Old_female		8.676		0.388		7.786
		(0.84)		(0.09)		(1.01)
Kids_male		4.794		-2.392		7.343*
		(1.03)		(-1.16)		(2.24)
Adult_male		0.934		-0.944		1.818
		(0.30)		(-0.62)		(0.83)
Old_male		-5.359		-2.284		-3.059
		(-0.68)		(-0.89)		(-0.46)
Adult_female		8.169		4.512		3.662
		(1.56)		(1.69)		(1.20)
Old_female		3.924		1.927		2.062
		(0.64)		(0.58)		(0.41)

Heckman second-stage analysis - Male subsample Estimation (continued)

	(1)	(2)	(3)	(4)	(5)	(6)
	Hours	Hours	Housework	Housework	Prod. Hours	Prod. Hours
Age	0.499	0.642	0.0831	0.0946	0.368	0.497
	(1.17)	(1.57)	(0.53)	(0.57)	(1.12)	(1.62)
Age2	-0.00899	-0.00764	-0.000661	-0.0000982	-0.00794*	-0.00714*
	(-1.96)	(-1.63)	(-0.33)	(-0.05)	(-2.42)	(-2.19)
Married	-2.315	-3.335	-0.716	-0.922	-1.628	-2.397
	(-0.71)	(-1.10)	(-0.45)	(-0.60)	(-0.68)	(-1.06)
Polygamia	3.010	2.719	0.711	0.587	2.292	2.162
	(0.55)	(0.47)	(0.29)	(0.23)	(0.61)	(0.55)
Separated	3.497	4.186	1.793	1.532	1.695	2.641
	(0.73)	(0.87)	(0.92)	(0.76)	(0.45)	(0.75)
Divorced	-5.938	-6.612	-1.707	-1.217	-4.199	-5.268
	(-0.82)	(-1.06)	(-0.63)	(-0.44)	(-0.76)	(-1.20)
Widowed	6.115	5.598	3.197	3.281	2.876	2.291
	(1.19)	(1.07)	(1.03)	(1.01)	(0.73)	(0.58)
School	2.349	2.251	0.592	0.668	1.780	1.606
	(1.69)	(1.57)	(1.05)	(1.13)	(1.68)	(1.50)
Urban	0.445	0.519	0.251	0.248	0.183	0.272
	(0.13)	(0.15)	(0.22)	(0.21)	(0.07)	(0.10)
Num. children -0.279	-0.0171	-0.126	0.128	-0.200	-0.217	
	(-0.45)	(-0.02)	(-0.48)	(0.39)	(-0.42)	(-0.44)
Household size	-0.679	-0.237	-0.130	-0.198	-0.569	-0.0614
	(-1.60)	(-0.50)	(-0.81)	(-1.14)	(-1.64)	(-0.16)
_cons	13.53	3.276	4.192	3.119	10.33	1.227
	(1.64)	(0.39)	(1.23)	(0.83)	(1.61)	(0.20)
N	2721	2721	2721	2721	2721	2721
R-Squared	0.00797	0.0212	0.00436	0.0129	0.00985	0.0251

This Table presents the extended results of Table 44, where we described in detail the specifications and the estimation process. We compute robust standard errors clustered at the village area level. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 54: Extensive margins - Malawi

	(1)	(2)	(3)	(4)	(5)	(6)
	work	work	work	Labor Force	Labor Force	Labor Force
Islam	-0.233	-0.243	-0.247	0.619	-0.0935	-0.877
	(-0.95)	(-0.75)	(-0.61)	(1.68)	(-0.13)	(-0.66)
Ramadan	0.114	0.0447	0.161	0.102	0.151	0.525
	(1.60)	(0.49)	(0.80)	(0.85)	(0.81)	(0.94)
Islam*Ramadan	-0.0445	-0.101	0.0307	-0.384	-0.300	-0.249
	(-0.28)	(-0.55)	(0.07)	(-1.45)	(-0.62)	(-0.16)
Age		0.0331	-2.083***		-0.287***	-0.309
		(1.29)	(-14.96)		(-4.76)	(-1.72)
School		0.226*	0.244**		-1.142***	-1.230***
		(2.38)	(2.60)		(-3.97)	(-4.21)
Urban		-0.380	-0.510*		0.0901	0.0590
		(-1.52)	(-2.34)		(0.31)	(0.19)
Num. children		-0.105*	-0.113*		-1.008***	-1.016***
		(-2.32)	(-2.34)		(-5.31)	(-4.61)
Household size		-0.0128	-0.0116		0.199**	0.249**
		(-0.49)	(-0.40)		(3.12)	(3.05)
Islam*Ramadan*Children_male			-1.522			0.412
			(-1.32)			(0.19)
Islam*Ramadan*Adult_male			-0.657			2.284
			(-0.69)			(0.70)
Islam*Ramadan*Old_male			0.753			-4.006
			(0.43)			(-0.72)
Islam*Ramadan*Adult_female			1.265			-3.529
			(1.32)			(-1.29)
Islam*Ramadan*Old_female			-0.707			1.354
			(-0.41)			(0.26)

Extensive margins - Malawi (continued)

	(1)	(2)	(3)	(4)	(5)	(6)
	work	work	work	Labor Force	Labor Force	Labor Force
[1em] Ramadan*Children_male			0.349			-0.348
			(0.72)			(-0.32)
Ramadan*Adult_male			0.163			-1.019
			(0.39)			(-0.72)
Ramadan*Old_male			0.261			1.162
			(0.34)			(0.70)
Ramadan*Adult_female			-0.640			0.310
			(-1.33)			(0.33)
Ramadan*Old_female			-0.0625			-1.477
			(-0.09)			(-0.98)
N	6220	4273	5114	2268	1990	1990

This Table presents the extended results of Table 45, where we described in detail the specifications and the estimation process. We compute robust standard errors clustered at the village area level. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 55: Extensive margins - Bangladesh

	(1)	(2)	(3)	(4)	(5)	(6)
	work	work	work	Labor Force	Labor Force	Labor Force
ramadan	1.354***	-0.260*	-0.118	0.542***	-0.250***	-0.0524
	(8.01)	(-2.27)	(-0.42)	(6.64)	(-3.50)	(-0.30)
hhsize		0.00116	0.00140		0.000416	0.000708
		(0.73)	(0.89)		(0.40)	(0.68)
num_childre		0.00187	0.000838		0.00152	0.000367
		(0.57)	(0.26)		(0.76)	(0.18)
age		0.114	0.104		0.0500	0.0243
		(0.20)	(0.11)		(0.15)	(0.03)
age2		-0.00885***	-0.00945***		-0.00335***	-0.00233***
		(-18.64)	(-14.99)		(-9.98)	(-6.23)
education		-0.0247	-0.0255		0.0480**	0.0496**
		(-0.79)	(-0.83)		(2.69)	(2.81)
Ramadan*Child_male			-0.238			0.0682
			(-0.28)			(0.15)
Ramadan*Adult_male			0.644			0.459
			(0.99)			(1.09)
Ramadan*Old_male			1.235			-0.702
			(0.93)			(-0.79)
Ramadan*Adult_female			-0.834			-0.957**
			(-1.33)			(-2.65)
Ramadan*Old_female			-1.684			-0.801
			(-1.16)			(-0.92)
N	7844	7844	7844	9792	9776	9776
ll	-2489.6	-1090.2	-1076.3	-3337.7	-2717.5	-2669.2

This Table presents the extended results of Table 46, where we described in detail the specifications and the estimation process. We compute robust standard errors clustered at the village area level. t-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 56: Testing the Parallel Trend Assumption - Malawi

	Housework	Prod. Hours	Hours
Total sample	1.65	1.80	2.59
	(0.1997)	(0.1801)	(0.1073)
Females	0.81	5.93	4.67
	(0.3683)	(0.0150)	(0.0307)
Males	0.61	0.07	0.03
	(0.4360)	(0.7893)	(0.8648)

In this table we present the results for the parallel trend assumption test, where we expect that individuals do not anticipate the Ramadan effects on their working hours. In this case, we report the test effects and the p-values in parenthesis. The main conclusions from this table are that, at the 1% significance level, individuals do not alter their hours worked, prior to the Ramadan and, thus, the parallel trend assumption is satisfied. We use the *time varying* treatment command, following the analysis in Cerulli and Ventura, 2017.

Chapter 5

Conclusion of the Thesis

Along this thesis we have disentangled three different scenarios that the economy can present in different countries and analyzed how individuals and households behave under such situations. The first proposed scenario analyzed how household consumption expenditures behave across winning and non-winning regions of the Spanish Christmas lottery. Similarly, we proposed a second scenario that tests how local windfalls gains from the British Postcode Lottery affects consumption behavior for those households that live in winning postcodes, compared with those households that live in non-winning postcodes. Our third proposed scenario is based on how religious constraints affect individuals' behavior in the labor market, in other words, how Ramadan interferes in the number of hours worked on individuals that take part of this festivity.

Using different household data surveys, we find that using the Spanish Household Survey, households that live in winning regions of the Spanish Christmas Lottery significantly increase their expenditures in durable and non-durable goods. Moreover, we find that the estimated elasticities of these two goods to total household expenditures, adjusted in the first stage with the lottery income shock, are similar. Using secure access data from the Understanding Society, that British households that live in winning postcodes of the Millions Postcode lottery increase consumption in durable and non-durable goods, compared with households that live in non-winning postcodes. In this case we find that Spanish and British households behave similarly under a scenario where they live in an area that won the lottery. Thus, the PIH is violated in both cases. Moreover, when we estimate the elasticities of durable and non-durable goods to total household expenditures, again we find that these are similar. Finally, when we analyze data coming from the Malawian Integrated Household Panel Survey, we find that individuals that celebrate Ramadan do not decrease their total number of hours worked during this period. Women redistribute their hours worked, from their jobs to housework, whereas men do increase their hours worked at their jobs. In this chapter, we also use the HiES and the WiLCAS datasets from Bangladesh to examine the extensive margins effects of Ramadan. In this case, we do not observe any changes in the likelihood of being employed or belonging to the labor force due to the celebration of Ramadan.

However, research done along this thesis present some limitations we need to address. Regarding the Spanish and British datasets, these do not include information about which households have won the lottery and which ones have not. However, these datasets contain information about the region where households live for the Spanish case (also whether they live in a capital of province or not), and about the postcodes where British households live (using secure data access). Therefore, we can use this information as a proxy to know the potential winners of the Spanish Christmas Lottery and the People's Postcode Lottery. Regarding the third chapter, we encountered data for many countries that did not collect surveys during Ramadan or that, like the case of the HiES and the WiLCAS, reported the yearly number of hours worked, but not the weekly one as we need for the analysis of the intensive margins in that chapter.

Finally, we aim that the findings of this thesis can have some policy implications. Regarding the first two chapters, we aim that our results can contribute to tax rebates to stimulate consumption, given the general increase in consumption caused by the lottery income shocks to households that live in winning areas of the lottery. And we also aim that Governments in countries that celebrate Ramadan can readjust labor market structures, in terms of hours worked during the Holy month, as there is a significant percentage of the world population that takes part of this festivity.

References

[1] Adamopoulou, E. & Zizza, R., *Regular versus Lump-Sum Payments in Union Contracts and Household Consumption*, ECB Working Paper No. 2013, 2017, 58 pages.

[2] Afifi, Z. E. M., *Daily practices, study performance and health during the Ramadan fast*, Journal of the Royal Society of Health, Vol. 117, num. 4, 1997, pp. 231-235.

[3] Almond, D. & Mazumder, B., *Health capital and the prenatal environment: the effect of Ramadan observance during pregnancy*, American Economic Journal, Vol. 3, num. 4, 2011, pp. 56-85.

[4] Arellano, M., Blundell, R. & Bonhomme, S., *Earnings and Consumption Dynamics: A Nonlinear Panel Data Framework*, Econometrica, Vol. 85, num. 3, 2017, pp. 693-734.

[5] Asaduzzaman, M., Kabir, R. A. & Radović-Marković, M., *Gender inequality in Bangladesh*, Journal of Women's Entrepreneurship and Education, Vol. 3, num. 4, 2015, pp. 54-64.

[6] Attanasio, O. & Lechene, V., *Tests of Income Pooling in Household Decisions*, Review of Economic Dynamics, Vol. 5, 2002, pp. 720-748.

[7] Bagués, M. & Esteve-Volart, B., *Politicians' Luck of the Draw: Evidence from the Spanish Christmas Lottery*, Journal of Political Economy, Vol. 124, num. 5, 2016, pp. 1269-1294.

[8] Baker, A.-C., Larcker, D.-F. & Wang, C.-Y., *How much should we trust staggered difference-in-differences estimates?.*, Journal of Financial Economics, Vol. 144, num. 2, 2021, pp. 370-395.

[9] Barro, R. J. & McCleary, R., *Religion and economic growth across countries* , American Sociological Review, Vol. 68, 2003, pp. 760-781.

[10] Barro, R. J. & McCleary, R., *Religion and economy*, Journal of Economic Perspectives, Vol. 20, 2006, pp. 49-72.

[11] Berger, D., Guerrieri, V., Lorenzoni, G. & Vavra, J., *House Prices and Consumer Spending*, NBER Working Paper No. 21667, 2015.

[12] Bermejo, V.J., Ferreira, M.A., Wolfenzon, D. & Zambrana, R., *Entrepreneurship and Regional Windfall Gains: Evidence from the Spanish Christmas Lottery*, Working Paper Available at SSRN 3297158, 2019.

[13] Berniell, I., *Individual and Aggregate Effects of Paycheck Frequency*, Job Market Paper, 2016.

[14] Blundell, R., Pistaferri, L. & Saporta-Eksten, I., *Consumption Inequality and Family Labor Supply*, American Economic Review, Vol. 106, num. 2, 2016, pp. 387-435.

[15] Blundell, R., Pashardes, P. & Weber, G., *What Do We Learn About Consumer Demand Patterns from Micro Data?*, American Economic Review, Vol. 83, num. 3, 1993, pp. 570-597.

[16] Bone, D.S., *Islam in Malawi1*, Journal of Religion in Africa, Vol. 13, num. 2, 1982, pp. 126-138.

[17] Browning, M. & Collado, D., *Habits and Heterogenity in Demands: A Panel Data Analysis*, Journal of Applied Econometrics, Vol. 22, num. 3, 2007, pp. 625-640.

[18] Browning, M. & Crossley, T.-F., *Shocks, stocks and socks: smoothing consumption over a temporary income loss*, Journal of the European Economic Association, Vol. 7, num. 6, 2009, pp. 1169-1192.

[19] Callaway, B. & Sant'Anna, P.-H., *Difference-in-differences with multiple time periods* , Journal of Econometrics, Vol. 225, num. 2, 2021, pp. 200-230.

[20] Campante, F. & Yanagizawa-Drott, D., *Does Religion Affect Economic Growth and Happiness? Evidence from Ramadan*, The Quarterly Journal of Economics, Vol. 130, num. 2, 2015, pp. 615-658.

[21] Campbell, J.Y. & Cocco, J.F., *How do prices affect consumption? Evidence from micro data*, Journal of Monetary Economics, Vol. 54, num. 3, 2007, pp. 591-621.

[22] Cameron, Colin A. & Trivedi, Pravin K., *Microeconometrics Methods and Applications*, first edition, Ed. Cambridge, USA, 2005.

[23] Çelen, A., *Influence of Holy Month Ramadan on Alcohol Consumption in Turkey*, Journal of religion and health, Vol. 54, num. 6, 2015, pp. 2122-2133.

[24] Cerletti, E.A. & Pijoan-Mas, J., *Durable Goods, Borrowing Constraints and Consumption Insurance*, 2014, CEPR Discussion Paper DP9035.

[25] Cerulli, G., & Ventura, M., *Estimation of pre-and posttreatment average treatment effects with binary time-varying treatment using Stata*, The Stata Journal, Vol. 19, num. 3, 2019, pp. 551-565.

[26] Chaisemartin, C., & d'Haultfoeuille, X., *Two-way fixed effects estimators with heterogeneous treatment effects*, American Economic Review , Vol. 110, num. 9, 2020, pp. 2964-96.

[27] Chamberlain, G., *Binary response models for panel data: Identification and information*, Econometrica, Vol. 78, num. 1, 2010, pp. 159-168.

[28] *Culture Grams, World Edition, "Malawi"*, ProQuest. 2018. Web. 9 May. 2020: www.culturegrams.com.

[29] Danzer, A.M., *Labor supply and consumption smoothing when income shocks are non-insurable*, IZA discussion paper, n.5499, 2011.

[30] Deaton, A., *Understanding Consumption*, Oxford University Press, USA, 1992.

[31] Deaton, A., *Franco Modigliani and the life cycle theory of consumption* , Working Paper - Available at SSRN 686475.

[32] Demiroglu, C., Ozbas, O., Silva, R. C. & Ulu, M. F., *Does religion affect economic decisions? Evidence from Ramadan Loans*, Working Paper, 2017.

[33] Dercon, S. & Krishnan, P., *In sickness and in health: Risk sharing within households in rural Ethiopia*, Journal of Political Economy, Vol. 108, num. 8, 2000, pp. 668-727.

[34] Disney, R., Gathergood, J. & Henley, A., *House price shocks, negative equity, and household consumption in the United Kingdom*, Journal of the European Economic Association, Vol. 8, num. 6, 2010, pp. 1179-1207.

[35] Dustmann, C. & Rochina-Barrachina, M. E., *Selection correction in panel data models: An application to the estimation of females' wage equations*, The Econometrics Journal, Vol. 10, num. 2, 2007, pp. 263-293.

[36] Fagereng, A., Holm, M.B. & Natvik, G.J., *MPC heterogeneity and household balance sheets*, available at SSRN: https://ssrn.com/abstract=2861053, 2016, pp. 1-41.

[37] Friedman, M., *The permanent income hypothesis. In A theory of the consumption function*, Princeton University Press, USA, 1957, pp. 20-37.

[38] Fisher, M. & Kandiwa, V., *Can agricultural input subsidies reduce the gender gap in modern maize adoption? Evidence from Malawi*, Food Policy, Vol. 45, 2014, pp. 101-111.

[39] Garvía, R., *Syndication, Institutionalization, and Lottery Play*, American Journal of Sociology, Vol. 113, num. 3, 2007, pp. 603-652.

[40] Göçmen, E., Koç, M., Tez, M., Yoldas, Ö., Bilgin, A. & Keskek, M., *Effect of Ramadan on surgical emergencies*, Annals of emergency medicine, Vol. 44, num. 3, 2004, pp. 283-285.

[41] Goodman-Bacon, A., *Difference-in-differences with variation in treatment timing*, Journal of Econometrics, Vol. 225, num. 2, 2021, pp. 254-277.

[42] Grossman, S.J. & Laroque, G., *Asset Pricing and Optimal Portfolio Choice in the Presence of Illiquid Durable Consumption Goods*, Econometrica, Vol. 58, num. 1, 1990, pp. 25-51.

[43] Hall, R.E., *Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence*, Journal of Political Economy, Vol. 86, num. 6, 1978, pp. 971-987.

[44] Harvey, E.E., Ioannou, C.A. & Golshirazi, F., *The Religious Observance of Ramadan and Prosocial Behavior*, Economic Inquiry, Vol. 56, num. 1, 2018, pp. 226-237.

[45] Heathcote, J., Storesletten, K., & Violante, G. L., *Consumption and labor supply with partial insurance: An analytical framework*, American Economic Review, Vol. 104, num. 7, 2014, pp. 2075-2126.

[46] Heckman, J.J., *Life cycle consumption and labor supply: An explanation of the relationship between income and consumption over the life cycle*, The American Economic Review, Vol. 24, num. 1, 1974, pp. 188-194.

[47] Hsieh, C.T., *Do Consumers React to Anticipated Income Changes? Evidence from the Alaska Permanent Fund*, American Economic Review, Vol. 93, num. 1, 2003, pp. 397-405.

[48] Imbens, G. W. & Rubin, D. B., *Causal inference in statistics, social, and biomedical sciences*, Cambridge University Press, 2015.

[49] Julien, J., Bravo-Ureta, B.E. & Rada, N.E., *Assessing farm performance by size in Malawi, Tanzania, and Uganda*, Food Policy, Vol. 84, 2019, pp. 153-164.

[50] Kapplan, G., & Violante, G.-L., *A model of the consumption response to fiscal stimulus payments*, Econometrica, Vol. 82, num. 4, 2014, pp. 1199-1239.

[51] Krueger, D. & Perri, F., *How do households respond to income shocks?*, Consortium on Financial Systems and Poverty, 2011.

[52] Kuhn, P., Kooreman, P., Soetevent, A. & Kapteyn, A., *The effects of lottery prizes on winners and their neighbours: evidence from the Dutch Postcode Lottery.*, American Economic Review, Vol. 101, num. 5, 2011, pp. 2226-47.

[53] Leiper, J. B. & Molla, A. M, *Effects on health of fluid restriction during fasting in Ramadan* , European journal of clinical Nutrition, Vol. 57, num. 2, 2003, pp. S30-S38.

[54] Luengo-Prado, M.J., *Durables, nondurables, down payments and consumption excess.*, Journal of Monetary Economics, Vol. 53, num. 7, 2006, pp. 1509-1539.

[55] Martin, R.F., *Consumption, Durable Goods, and Transaction Costs*, 2003, FRB International Finance Discussion Paper No. 756. Available at SSRN: https://ssrn.com/abstract=382061 or http://dx.doi.org/10.2139/ssrn.382061

[56] McKeinze, D., *Should we require balance t-tests of baseline observables in randomized experiments?*, World Bank Blogs, 2017.

[57] McKeinze, D., *When should you cluster standard errors? New wisdom from the econometrics oracle*, World Bank Blogs, 2017.

[58] Michael, R.T. & Becker, G.S., *On the new theory of consumer behavior.*, The Swedish Journal of Economics, 1973, pp. 378-396.

[59] Mroz, T. A., *The sensitivity of an empirical model of married women's hours of work to economic and statistical assumptions*, Econometrica: Journal of the econometric society, 1987, pp. 765-799.

[60] Musaiger, A. O., *Socio-Cultural and Economic Factors Affecting Food Consumption Patterns in the Arab Countries*, Journal of the Royal Society of Health, Vol. 113, num. 2, 1993, pp. 68.

[61] Nagler, P & Naudé, W., *Labor productivity in rural African enterprises: empirical evidence from the LSMS-ISA*, IZA Discussion Paper No. 8524. Available at SSRN: https://ssrn.com/abstract=2508579, 2014.

[62] Odabasi, Y. & Argan, M., *Aspects of Underlying Ramadan Consumption Patterns in Turkey*, Journal of International Consumer Marketing, Vol. 21, num. 3, 2009, pp. 203-218.

[63] Pariona, A., *The economy of Bangladesh*, World Atlas, 2017.

[64] Pew Research Center, The, *How religious commitment varies by country among people of all ages*, PEW Forum, 2018.

[65] Rahman, M., Bhattacharya, D. & Al Hasan, M., *The role of the informal sector in inclusive growth: A state of knowledge study from policy perspectives*, 2018.

[66] Schofield, H., *The economic costs of low caloric intake: Evidence from India*, Unpublished Manuscript, Harvard University, 2014.

[67] Sheahan, M. & Barrett, C. B., *Ten striking facts about agricultural input use in Sub-Saharan Africa*, Food Policy, Vol. 67, 2017, pp. 12-25.

[68] Stock, J.H. & M.W., *Introduction to econometrics.*, third edition, Ed. Pearson, USA, 2007.

[69] University of Essex, Institute for Social and Economic Research, *Understanding Society: Waves 1-10, 2009-2019 and Harmonised BHPS: Waves 1-18, 1991-2009: Secure Access*, 11th Edition, UK Data Service. SN: 6676, http://doi.org/10.5255/UKDA-SN-6676-11, 2020.

[70] van Ewijk, R., *Long-term health effects on the next generation of Ramadan fasting during pregnancy*, Journal of Health Economics, Vol. 30, num. 6, 2011, pp. 1246-1260.

[71] van Klaveren, M., *An Overview of Women's Work and Employment in Malawi: Decisions for Life MDG3 Project*, Country Report No. 6, AIAS, 2009.

[72] Wetmore, D., *Why Bangladesh is poor*, The Borgen Project, 2017.

[73] Wooldridge, J.M., *Selection corrections for panel data models unders conditional mean independence assumptions*, Journal of Econometrics, Vol. 68, 1995, pp. 115-132.