Received 20 July 2023
Revised 1 December 2023
9 February 2024
Accepted 18 March 2024

Testing the permanent income hypothesis using the Spanish Christmas Lottery

Guillermo Cabanillas-Jiménez
Department of Economics, Facultad de Ciencias Económicas Empresariales y Turismo, Universidad de Alcalá, Alcalá de Henares (Madrid), Spain

Abstract

Purpose – This study aims to investigate the impact of local windfall gains from the Spanish Christmas lottery on household consumption behavior.

Design/methodology/approach – The study applies differences-in-differences to assess permanent income hypothesis (PIH) validity, examining pre- and postlottery consumption effects. Additionally, it also uses an instrumental variable regression, using the lottery shock as an instrument for total expenditures, to estimate the Engel curves.

Findings – The paper finds a PIH violation; households in winning region notably increase consumption on durable and nondurable goods compared to nonwinning ones. Moreover, durable goods consumption is responsive to lottery winnings, while nondurable goods consumption are unit-elastic to expenditure shocks.

Originality/value – To the best of the author's knowledge, this is the first paper analyzing the effects of winning regions of the Spanish Christmas lottery in all types of consumption goods, testing its consequences in the PIH and estimating its effects in the Engel curves.

Keywords Consumption, Durable goods, Nondurable goods, Winning region, Permanent income hypothesis (PIH)

Paper type Research paper

1. Introduction

Over recent years, research into consumer behavior has focused on how different income shocks affect the behavior of individual agents (see Berniell, 2018; 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

JEL classification – C01, C23, C26, C55, C93, D12, D14

This paper has benefited from all comments received in internal seminars at the University of Kent and UAH, and conferences presentations like the SAEe2022. The author wants to thank everyone who contributed with their ideas to improve the research quality of the paper. Also, a special thanks to Dr. Zaki Wahhaj for his support and advices along this project.

Applied Economic Analysis
Vol. 32 No. 96, 2024
pp. 232-259
Emerald Publishing Limited
2632-7627
DOI 10.1108/AEA-07-2023-0276

(see Adamopoulou and Zizza, 2017; Berger *et al.*, 2018). This prediction is known as the permanent income hypothesis (PIH), first proposed by Milton Friedman in 1957 (see Friedman, 1957). In 1978, Robert Hall (1978) 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 paper 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. However, there are some facts about the Spanish Christmas lottery to be taken into consideration in this study. First, 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. Second, the prize does not belong solely to one person – it is shared among all those individuals that bought the same lottery number [1]. Therefore, the shock affects more than one household, making the analysis more heterogeneous. Third, 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 winning regions react to a win: in other words, whether households in winning regions 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 questions we attempt to answer in this paper are: how the Spanish Christmas lottery income shock affects household consumption behavior in the Christmas lottery winning regions, in comparison with the nonwinning ones, and how sensitive household consumption is to these shocks. Our underlying hypothesis is that given the average size of the shock, the PIH holds and households in winning regions 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, rejecting the null hypothesis that living in a winning region of the lottery has no effect on household consumption. This has two direct implications: (i) winning regions of the Spanish Christmas lottery do increase

consumption after the shock and (ii) 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, to test the sensitivity of the lottery shock to household consumption.

When examining the direct effect of the lottery winnings on household consumption, we observe that households in winning regions increase their consumption in durable and nondurable goods due to the lottery shock. In this case, these estimates capture the average effect of regional lottery earnings 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 A2 that the average lottery earnings each year are €387, the year with the lowest average and €2,263, 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 households in winning regions of the lottery increase their consumption of both nondurable and durable goods. The effect on nondurable goods is unit-elastic, while for durable goods, it is elastic to total household expenditures. Specifically, a 10% increase in total expenditures leads to an 11.47% increase in household consumption of durable goods and a 9.26% increase in nondurable goods consumption. Such findings imply that the impact of the positive income shock on total expenditures (i.e. lottery income shock or living in a winning region) is similar for both durable and nondurable goods, contrary to theoretical predictions by Cerletti and Pijoan-Mas (2014) and Browning and Crossley (2009), which expected only durable goods to respond to lottery winnings.

On the second page, we are also interested in observing the effects that the lottery shock has on labor outcomes. We aim to test whether the income shock induces changes in the employability status and the number of hours worked by household heads. Consistent with established literature, prevalent expectations suggest minimal changes in the labor market dynamics. Winning regions often experience an increase in start-ups, with individuals dedicating more time to work than leisure and, thus, increase self-employment (see Bermejo *et al.*, 2019; Disney and Gathergood, 2009). Nevertheless, contrasting perspectives in the literature propose that winning the lottery may lead to a reduction in labor income and, consequently, a decline in employment levels (see Cesarini *et al.*, 2017; Imbens *et al.*, 2001).

This paper belongs to a growing literature exploring consumer behavior and household changes when experiencing an income shock. Previous studies have used the Spanish Christmas lottery as an exogenous shock affecting agents' decisions on their political vote (see Bagués and Esteve-Volart, 2016), entrepreneurship and self-employment (see Bermejo *et al.*, 2019) or the effects on CPI prices and unemployment rates (see Ghomi *et al.*, 2022), but also some of these papers find that winning regions experience a significant increase in cars registrations or purchases of these type of goods (see Bagués and Esteve-Volart, 2016; Ghomi *et al.*, 2022). However, this is the first paper analyzing the effects of winning regions of the Spanish Christmas lottery on all types of household consumption goods, testing its consequences in the PIH and estimating its effects in the Engel curves. The unique nature of this social game, involving a large portion of the Spanish population, allows most participants to be part of the treatment group and experience the lottery shock (see Garvía, 2007).

Theoretical findings in the literature suggest a violation of the PIH regarding agents' behavior when facing a positive income shock. This is because households spend income shock money on durable goods, showing sensitivity to income shocks and a significant

response to unexpected income increases (see Cerletti and Pijoan-Mas, 2014). In addition, receiving a one-time positive income shock leads agents to anticipate durable goods purchases, rather than waiting until the item becomes completely obsolete (see Grossman and Laroque, 1990). These results support the life-cycle hypothesis (LCH) [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 and Shapiro, 2015). 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.*, 2021). Despite these findings, changes in income lead to strong responses in consumption and play an important role in household decisions (see Krueger and Perri, 2011). Thus, when households experience positive income shocks, they have a larger propensity to consume; however, when the future is uncertain, individuals tend to save and postpone their consumption (see Kapplan 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.*, 2021), the UK (see Cabanillas-Jiménez, 2022) and Alaska (see Hsieh, 2003). As in this paper, previous ones use a difference-in-difference method to compare household decisions in winning places as compared to households located in nonwinning 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 paper 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 nonwinning 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 (i.e. households in winning regions). Furthermore, the obtained results in this paper are important, as contrary to what the literature finds, we encounter that households in winning regions react similarly to the lottery winnings toward consumption of durable and nondurable goods.

The paper is structured as follows: Sections 2 and 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; Section 7 is based on the robustness checks of the results; and, finally, Section 8 concludes the paper.

2. Background: the Spanish Christmas lottery

The Spanish Christmas lottery, known as *Lotería de Navidad*, is a national lottery game organized by the national lottery organization. It has been played since 1812 and is considered the largest lottery event worldwide. Approximately 75% of the Spanish population participates in this game, making it more of a social network event than a traditional gambling activity. It is common for friends, family, or colleagues to share the purchase of lottery tickets. Unlike other lotteries in Spain, most Christmas lottery players only participate in this particular game.

Each lottery ticket costs €20, and the whole series (ten tickets) costs €200. There are also shares and participations available at lower prices, typically ranging from €2 to €5, with €1

usually going to charity. On average, the Spanish population spends around €64 per person on the Spanish Christmas lottery, based on recent years' data [4]. A survey conducted in 2004 by the center for sociological research showed that individuals planned to spend between €40 and €60, with only 8% of the respondents planning to spend over €150.

The lottery tickets for the Christmas lottery consist of five-digit numbers. Since 2011, a total of 100,000 numbers have been played in the draw [5], ranging from 100,000 to 99,999. Each ticket number is divided into 160 series, and each series is further divided into ten fractions called *décimos*. These fractions can be divided into smaller units known as *participaciones*. Out of the 100,000 numbers played, 1,807 will receive a prize. However, the probability of winning the lottery is extremely low, with a 0.001% chance of winning the first prize, called *El Gordo*. The probability of this scenario is relatively low, as in any lottery game. Hence, the likelihood that households in the treatment group are actual lottery winners is low. Further elaboration on this point is provided in Section 3, where we extensively discuss the average lottery earnings per region each year and the count of potential winners in each region.

The distribution of prizes and the associated amount of money for each ticket bought are shown in Table 1. In addition, if someone's ticket contains the last numbers of *El Gordo*, they also receive a proportional amount of money based on their investment. Overall, 70% of the revenue is allocated to prizes, while the remaining 30% represents the commission paid to the ticket outlets. Prizes above €2,500 are subject to a 20% tax, which goes to the Treasury. Despite the tax, winners after 2013 still earn more than before, with €320,000 after tax per ticket, which is approximately 12 times higher than the average household income in Spain (€26,092) [6].

The Christmas lottery has the characteristic of being clustered, with only a few towns winning the lottery each year and sometimes only one city is awarded as the winner. This is because each lottery outlet is randomly assigned the numbers it sells. This clustering allows for easier monitoring of the regions that win and observing any changes in their consumption expenditures. The lottery's syndicate nature, where people in the same network play the same number, contributes to this clustering phenomenon (see Bagués and Esteve-Volart, 2016).

Two exceptional outlets where the Christmas lottery is sold are worth mentioning. The first is *doña Manolita*, a renowned outlet in Madrid. The second outlet is located in Sort, a town in the province of Lleida, known as *La Bruixa d'Or* (the gold witch). These outlets have gained fame for selling winning numbers over the years, attracting people from all over the country due to superstitions and beliefs associated with their success.

3. Data

The main data source we use in this paper comes from the *Encuesta de Presupuestos Familiares* (EPF) [7], provided by the *Instituto Nacional de Estadística* (INE) [8]. This is the Spanish Family Income Survey, in which households are randomly selected to take part.

Prize	No. 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	€300 per euro
Pedrea	1,974	€100	€5 per euro

Source: ABC – lottery prizes

Table 1. Distribution of the lottery prizes

The sample is composed of 22,346 households in total, covering the years from 1998 to 2016 [9], where surveys in each wave can take place in any given period of the year. The 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 presents 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 [10] 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 Christmas 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. 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 nonwinning regions in respect of the Spanish Christmas lottery and, thus, have a general idea of the average consumption pattern between regions. In Table A1 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 populations 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 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 that 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. Moreover, to refine the treatment group, the EPF includes data on whether a household is located in the capital of province. Consequently, if the winning outlet is in the capital of province, we can restrict the treatment group exclusively on households residing in both a winning region and a provincial capital. Conversely, if the prize is not awarded at an outlet in a provincial capital, we can limit the treatment group to households in winning regions outside provincial capitals. Thus, it allows us to reduce the potential noise/spillovers of the treatment of covering the entire region. 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 provides a descriptive summary statistic of the sample, where 15% of the interviewed households live in a region that won *El Gordo* 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, with a standard deviation of 28 years. In addition, most of our sample is either single or married – 43% are single, and 47% are married; the remaining are

Variable	Mean	SD	*n*
Lottery × win_region	0.072	0.258	305,550
Win_region	0.152	0.36	305,550
Lottery	0.433	0.496	305,550
Lottery expenditure	91.869	36.16	273,098
Age	31.165	28.394	239,434
Single	0.426	0.495	305,550
Married	0.468	0.499	305,550
Education	3.514	1.629	278,490
Employed	0.424	0.494	278,490
Retired	0.337	0.473	305,550

Notes: Education reports the average level of completed studies by the head of the household; if we disentangle the different categories in this variable, indeed, we find that the majority have achieved Levels 3 and 4, which are secondary and A-levels (31.1% and 23.1% of the heads, respectively). On top of that, 15% of the heads completed only primary school and 13.5% completed a bachelor's degree, jointly with 7% who did a master's degree and only 0.8% completed a PhD. The remaining 8.9% have no studies
Source: Instituto Nacional de Estadística (INE)

Table 2.
Summary statistics

widowed (6%) or separated/divorced (4%). Only 42% of the sample is used 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 (31% of the heads completed secondary school) and four (23% of the heads completed A-levels), 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 presents a summary statistics of household consumption expenditures by winning and nonwinning 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 nonwinning 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 – Table A2 in the append provides the average earnings per year. 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.

We conducted an exogeneity test on the treatment group using a linear probability model. The model uses a binary variable as the dependent variable, indicating whether households lived in winning or nonwinning regions, on various household characteristics and control variables. The resulting *F*-test is 0.66, indicating that we fail to reject the null hypothesis that all estimated regressors are zero. This suggests that our identifying assumption, described in Section 4, holds (see Table A3).

Analyzing Table 3, it is evident that winning regions of the Christmas lottery tend to have lower expenditures compared to nonwinning regions. Specifically, households in winning regions spent €531 less on durable goods, €794 less on nondurable goods and €1,325 less on total household expenditures. These differences are statistically significant at the 1% level, indicating variations in household consumption behavior across regions. Moreover, households in winning regions saved more than those in nonwinning regions, with an additional monthly savings of €30. Despite being a small difference, this was

Variable	Winning regions Mean	SD	Non-winning regions Mean	SD	Testing differences *t*-test difference	*p*-value
Total expenditures	6,142.035	9,116.764	7,466.984	8,062.632	−20.93	0.000
Durables	2,600.31	3,999.109	3,130.787	3,574.774	−19.09	0.000
Non_durables	3,541.725	5,186.04	4,336.198	4,608.617	−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 are presented in Euros

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

statistically significant at the 1% level. Overall, households in winning regions did not exhibit higher consumption levels, except in a few categories such as car value, education and holidays. Although the differences were minor, the *t*-tests showed their statistical significance.

Therefore, we observe that households that live in winning regions behave differently than those that live in nonwinning regions. A reason that can explain these negative differences IS that winning regions are likely to be poorer than nonwinning regions. However, when we look at the standard deviation, we observe that for the aggregates (durable and nondurable 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.

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:

(1) only households that participate in the Christmas lottery can experience such a shock; and

(2) 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 nonzero 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 and, whether this household lives in a capital of province or not – in case the lottery is awarded in a capital of province 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 paper 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 relationship between living in a winning region of the Christmas lottery and household consumption expenditures for specific goods. First, we run a regression to analyze whether residing in a winning region significantly impacts households' consumption behavior, incorporating various head of the household characteristics as control variables, jointly with region and year fixed-effects. Next, we estimate the Engel curves for different categories of goods using an instrumental variable regression approach to address potential endogeneity arising from total expenditures in the Engel curves estimation. The endogeneity arises because households can infer from resource allocation and decide how much to spend each month, leading to unobserved factors influencing expenditure decisions that may 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) [11]. We also extend our analysis by analyzing the effects of the lottery income shock on the labor market.

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. Equation (1) is used to estimate the direct impact of the lottery income shock on household consumption, with the incorporation of household and region controls to assess the robustness of the estimates in the presence of household characteristics and fixed effects. In addition, we introduce two demographic variables: GDP per capita in each region and the average lottery expenditures per region. This ensures that the income shock is entirely random, as regions with higher lottery expenditures may have greater chances of winning the lottery. Given that regions with larger populations are more likely to win the lottery, as discussed in Section 3, this approach enables us to account for all potential disparities presented in Tables 3, A1 and A2. Hence, the regression is as follows:

$$\ln\left(c_{h,t}^{g}\right) = \beta_0 + \beta_1 win\_region_{h,t-1} + \beta_2 lottery_{h,t-1} + \beta_3 win_{h,t-1} \times lottery_{h,t-1} + X_{h,t}^{'}\boldsymbol{\beta}_4 + \left(gdp_{r,t}, \log\left(lot\_exp_{r,t-1}\right)\right)^{'}\boldsymbol{\beta}_5 + \eta_{h,t} + \tau_t + u_{h,t} \quad (1)$$

where $c_{h,t}^{g}$ 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. $X_{h,t}^{'}$ 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, $lot\_exp_{r,t-1}$. 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. The inclusion of $\eta_{h,t}$ and τ_t are explained by the fact that regions are systematically different to each other and average earnings of the lottery differ per region and year and by changes in probabilities of being a winner in each region (given its characteristics) every year. The term uh,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 nonwinning 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.

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 good categories. 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:

$$\ln(exp_{h,t}) = \beta_0 + \left(win\_region_{h,t-1}, lottery_{h,t-1}, win_{h,t-1} \times lottery_{h,t-1}\right)' \boldsymbol{\beta}_1 + X_{i,t}\boldsymbol{\beta}_2 + \left(gdp_{r,t}, lot\_exp_{r,t-1}\right)' \boldsymbol{\beta}_3 + \eta_{h,t} + \tau_t + \nu_{h,t} \quad (2)$$

where the included coefficients are the same as those included in Equation (1). Once equation (2) is estimated, we check that the relevance condition for the instruments holds. This can easily be tested by computing the F-test for instrumental variables.

Nonetheless, equations (1) and (2) 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. 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) [12]. 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; 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) [13]. This is something to take into account 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).

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 paper. Hence, the second stage regression is as follows:

$$\ln\left(c^g_{h,t}\right) = \gamma_0 + \gamma_1 \ln(exp_{h,t}) + \left(X_{h,t}\right)' \boldsymbol{\gamma}_2 + \left(gdp_{r,t}, lot\_exp_{r,t-1}\right)' \boldsymbol{\gamma}_3 + \eta_{h,t} + \tau_t + u_{i,t} \quad (3)$$

where the $\ln(exp_{h,t})$ is the logarithm of total expenditures, estimated previously in equation (2). Our coefficient of interest is γ_1, because it captures the elasticity effect of total household expenditures on household consumption behavior in winning regions of the lottery. 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:

$$\begin{bmatrix} employed_{h,t} \\ num\_hours_{h,t} \end{bmatrix} = \beta_0 + \beta_1 win\_region_{h,t-1} + \beta_2 lottery_{h,t-1} + \beta_3 win_{h,t-1} * lottery_{h,t-1} + u_{h,t} \quad (4)$$

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 (4) 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:

$$\begin{bmatrix} employed_{h,t} \\ num\_hours_{h,t} \end{bmatrix} = \beta_0 + \beta_1 win\_region_{h,t-1} + \beta_2 lottery_{h,t-1} + \beta_3 win_{h,t-1} * lottery_{h,t-1} + X'_{h,t}\boldsymbol{\beta}_4 + \left(gdp_{r,t}, lot\_exp_{r,t-1}\right)' \boldsymbol{\beta}_5 + \eta_{h,t} + \tau_t + u_{h,t} \quad (5)$$

In this case, the vector of individual controls, $X_{h,t}$, does not include the employment status in equation (5), as this is one of our dependent variables.

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.

6.1 Consumption expenditures estimation

Table 4 presents the results for the household consumption presented in equation (1). Observing, first, the interaction term estimates, *lottery* × *win_region*, we find a positive effect on the average consumption of durable and nondurable 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 4.9% in the consumption of durable goods and an increase of 4.5% in the consumption of nondurable goods, in comparison with those households that did not win the lottery.

Looking at specific goods in more detail, we find a statistically significant effect in all goods except for *house durables* and *communication* (or technology) goods, where the effect of the interaction term is either close to zero or nonstatistically significant. On the other hand, we find a negative and significant effect for *alcohol*, *clothes*, *car value*, *transport*, *leisure*, *education*, *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 increase household consumption, in general, rather than to increase savings. This might indicate that the PIH is not satisfied, as individuals do not keep their consumption allocation prior to the income shock.

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* × *win_region*), on

Table 4.
Household consumption expenditures

Panel A	(1) Food at home	(2) Alcohol	(3) Clothes	(4) House rent	(5) House investments	(6) Health	(7) Car value	(8) Transport
Win_region	−0.148*** (−14.92)	0.0398 (1.84)	−0.0367 (−1.88)	−0.168*** (−15.60)	−0.166*** (−11.99)	0.0448* (2.05)	0.253*** (13.66)	−0.0316 (−1.81)
Lottery	0.165*** (27.75)	1.052*** (80.55)	0.842*** (71.66)	0.0610*** (9.40)	0.303*** (36.32)	1.194*** (90.63)	1.449*** (129.78)	0.702*** (66.75)
Lottery × win_region	0.0900*** (6.88)	−0.341*** (−11.90)	−0.206*** (−7.98)	0.0333* (2.34)	0.000498 (0.03)	−0.332*** (−11.48)	−0.460*** (−18.74)	−0.159*** (−6.90)
_cons	8.476*** (16.06)	10.03*** (8.67)	4.043*** (3.88)	10.80*** (18.78)	11.90*** (16.12)	1.982(1.70)	1.711 (1.73)	6.927*** (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	211,096	211,096	211,096	211,096	211,096	211,096	211,096	211,096
R-squared	0.831	0.518	0.612	0.874	0.802	0.507	0.212	0.693
Panel B	(9) Communication	(10) Leisure	(11) Education	(12) Food out home	(13) Holidays	(14) Savings	(15) Durables	(16) Nondurables
Win_region	−0.172*** (−12.51)	−0.116*** (−6.16)	0.0903*** (3.45)	−0.0652*** (−3.61)	0.102*** (4.47)	0.174*** (7.68)	−0.201*** (−15.62)	−0.164*** (−13.95)
Lottery	−1.006*** (23.22)	0.164*** (83.49)	0.742*** (109.84)	0.291*** (74.72)	0.906*** (98.07)	0.856*** (73.22)	−0.486*** (49.70)	−0.275*** (54.87)
Lottery × win_region	−0.0114 (−0.63)	−0.129*** (−5.18)	−0.438*** (−12.64)	−0.209*** (−8.74)	−0.355*** (−11.77)	−0.359*** (−11.99)	0.0487** (2.87)	0.0445** (2.87)
_cons	6.625*** (9.03)	3.009** (3.00)	2.440*** (6.15)	−1.738(−1.80)	−10.16*** (−8.35)	32.24*** (26.73)	12.54*** (18.31)	10.81*** (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	211,096	211,096	211,096	211,096	211,096	211,096	211,096	211,096
R-squared	0.814	0.633	0.308	0.675	0.347	0.391	0.867	0.855

Notes: 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

Sources: *Instituto Nacional de Estadística (INE)*. 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 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, clustered at the region level. *t*-statistics in parentheses: *$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$. A full set of estimates is available upon request

durable and nondurable goods, households in winning regions spend 15% less on durable goods and 12% less on nondurable goods compared to those in nonwinning regions. Hence, we observe that households that live in winning regions spend less on household consumption of the goods analyzed in comparison with those that live in nonwinning regions. 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 4, under the null that *living in winning lottery regions has no impact on household consumption*. The results from the *F*-test rejects the null hypothesis in all goods, indicating that the estimated effect of the lottery income shock remains statistically significant even when controlling for household and demographic characteristics. Thus, the study finds a violation of the PIH as households in winning regions do not smooth consumption but instead, they increase consumption expenditures for both durable and nondurable goods, 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.

These results are not in line with the theoretical findings in Cerletti and Pijoan-Mas (2014), where a positive income shock primarily should lead to increased consumption of durable goods or debt repayment, while nondurable goods consumption should not show a reaction to the income shock. However, these estimates do not capture the elasticity effect of the income shock on consumption behavior. To estimate this, we need to estimate the Engel curves, as proposed in equation (3).

We also test the parallel trend assumption. Figure C1 shows no evidence of households in winning regions anticipating the lottery income shock and increasing consumption before the draw celebration. Therefore, the parallel trend assumption is satisfied. Such results can also be confirmed in Table A7 in the Appendix.

Table 5 presents results for the first-stage regression in equation (2), showing that households living in winning regions participating in the Christmas lottery experience a 5% increase in total household expenditure, indicating potential winners might increase their consumption. However, the *win_region* coefficient appears negative, possibly influenced by a higher participation in lower-income regions.

The results from the *F*-test for the relevance of the 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 considered 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 6 shows the results for the second stage regression in equation (3). 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 nondurable goods is below one. This implies, from a theoretical point of view, that nondurable goods are not sensitive to a shock in total expenditures and, thus, the estimated effect is inelastic. Specifically,

	(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	211,096
R-squared	0.858

Notes: 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*

Sources: *Instituto Nacional de Estadística (INE)*. 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 4. 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: $*p < 0.05$, $**p < 0.01$ and $***p < 0.001$. Full set of estimates available in Table A4 of the Appendix of this paper

Table 5.
First stage regression – total household expenditures

increasing total household expenditures by 10% implies an increase of 9.26% in household consumption of nondurable 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 nondurable goods, and to check if the estimates are significantly different from each other.

To test whether the estimated effects for durable and nondurable 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 6 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 nondurable 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 nondurable goods are unit elastic to total household expenditures.

For the second test, we perform a *t*-test under the null hypothesis that the estimated elasticities for durable and nondurable 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 equal to zero. However, looking at the estimated elasticities, we observe that the difference between coefficients is of 0.22, which is small. Thus, the estimates of total household expenditures to durable and nondurable goods 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

Table 6. Second stage estimation – household consumption behavior

Panel A	(1) Food at home	(2) Alcohol	(3) Clothes	(4) House rent	(5) House investments	(6) Health	(7) Car value	(8) Transport
Log expenditures	0.725*** (14.75)	0.505*** (4.68)	0.684*** (7.06)	0.979*** (18.27)	1.037*** (15.09)	0.453*** (4.17)	−0.571*** (−6.19)	0.549*** (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	211,096	211,096	211,096	211,096	211,096	211,096	211,096	211,096
R-squared	0.831	0.517	0.612	0.874	0.802	0.506	0.211	0.693
Panel B	(9) Communication	(10) Leisure	(11) Education	(12) Food out home	(13) Holidays	(14) Savings	(15) Durables	(16) Nondurables
Log expenditures	1.101*** (16.13)	1.009*** (10.80)	0.406*** (3.85)	0.870*** (9.69)	0.146(1.29)	−0.294** (−2.62)	1.147*** (17.99)	0.926*** (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	211,096	211,096	211,096	211,096	211,096	211,096	211,096	211,096
R-squared	0.814	0.633	0.331	0.675	0.346	0.390	0.867	0.855

Sources: Instituto Nacional de Estadística (INE). 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: *$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$. A full set of estimates is available in Table A5 of the Appendix of this paper. The reported elasticity test examines the elasticity effect of expenditures toward household consumption, in words, whether the estimates are different from one

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 these can be considered as nondurable 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 6, we observe that monthly house rent, house investments, communication goods, leisure, eating out and nondurable 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 research that uses lottery earnings as income shocks, we observe similar results regarding the violation of the PIH. The Postcode Millions lottery in the UK shows an increase in household consumption of both durable and nondurable goods due to the lottery income shock in winning postcodes. Moreover, the elasticity estimates in the UK 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 nondurable 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 (see Cabanillas-Jiménez, 2022). Similarly, the Dutch Postcode lottery led to a €310 increase in household durable goods consumption for lottery winners, but did not show statistically significant effects on nondurable goods consumption (see Kuhn *et al.*, 2011). Moreover, Ghomi *et al.* (2022) find that at the aggregate macro level, consumption in winning regions of the Spanish Christmas lottery also increases during the six months after having won the lottery for both, durable and nondurable goods.

Hence, given the results presented in this paper, we conclude two important things:

(1) that the PIH is violated under the scenario that the Spanish Christmas lottery presents; and

(2) economically speaking, the estimated elasticities of total household expenditures for durable and nondurable 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 nondurable goods.

6.2 Labor supply estimation

Table A6 in the Appendix presents the results for the estimated regression presented in equations (4) and (5). 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 nonwinning 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 h more than the head of a household in a nonwinning region [14]. Looking at the *F*-test, we find this effect to be statistically significant.

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 [15]. This is confirmed by 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*.

As mentioned in the introduction, these results are in line with the previous work done by Bermejo *et al.* (2019) and Disney and Gathergood (2009), where these papers find that lottery winnings do not lead to a decrease in employment or labor market outcomes. Indeed, Bermejo *et al.* (2019) finds that winning lottery regions in Spain do increase the self-employment by increasing the number of start-ups in these regions.

7. Robustness exercises

To validate the findings presented in this paper, we conducted various exercises to assess the robustness of our results, all documented in the *Online Supplementary material file* of the paper. Initially, we reexamined the analysis by excluding the regions of Madrid and Cataluña, as explained in Section 2, where two outlets in these regions have traditionally sold the winning number multiple times, attracting lottery ticket purchases from the entire Spanish population. The exclusion aimed to mitigate potential spillovers from these regions. The results, as displayed in the tables in Section C1 of the *Online Supplementary material file*, persist without significant alterations, even with the omission of these two regions from the analysis. While the PIH remains violated in this scenario, we observe a loss of significance in the coefficient of the interaction term between lottery and region when these two regions are excluded. This trend persists in the estimation of household consumption in real terms. Such results suggest that the household consumption effects of potential winners may be primarily influenced by these two regions, both of which have been recurrent recipients of the first prize in the lottery over the years.

The second robustness check involves narrowing the analysis to single-province regions. This choice arises from the fact that the lottery shock operates at the municipal level, a unique aspect of the game. Consequently, if a town in Valencia wins *El Gordo* in a given year, households in Alicante and Castellon would also be considered treated, introducing potential biases and noise to the estimated coefficients. To enhance result reliability, we restrict the sample to autonomous communities that function as provinces, following the methodology used by Bagués and Esteve-Volart (2016) and Bermejo *et al.* (2019), who worked with province-level data. Once again, the results from Section C2 of the *Online Supplementary material file* do not deviate from the main findings presented in the previous section.

The third exercise performed pertains to testing the PIH using household consumption in real terms. This methodology allows us to calculate real household consumption for each good, yielding a reliable measure that mitigates the potential impact of inflation or price

increases in winning regions, as elaborated in Ghomi *et al.* (2022). Once again, the results obtained from measuring household consumption in real terms, presented in Section C3 of the *Online Supplementary material file*, confirm the persistent violation of the PIH and nondurable goods remain unit-elastic. However, durable goods this time are elastic to a shock to total household expenditures when we estimate consumption measures in real terms. This highlights the necessity of considering nominal effects and the monetary expenses of households due to these distinctions. Furthermore, the estimated coefficients for the PIH demonstrate some variation compared to nominal estimates, particularly when analyzing the entire sample, a pattern not observed when excluding Cataluña and Madrid from the analysis. As previously noted, this suggests that the household consumption effects of potential winners are largely influenced by these two regions.

Consequently, in none of the robustness checks applied to the results of this paper do we observe significant alterations to the main results. This further reinforces the estimates obtained in this paper and strengths the reliability of the effects of the lottery income shock on household consumption. However, we observe that the increase in household consumption is primarily due to the effects caused in Cataluña and Madrid.

8. Conclusion

The Spanish Christmas lottery provides a unique opportunity to study the effects of lottery prizes on household consumption behavior in winning regions. Its advantages lie in the significant economic impact it generates for the winning region, with a 3.5% increase in the local GDP. Moreover, the population of Spain spends around 0.3% of the national GDP on the Christmas lottery. In addition, the lottery tickets can be shared among multiple individuals, making the analysis more heterogeneous, as different individuals can be part of the treatment group and over 75% of the population participates in the game, which makes it a syndicate game and not a lottery for gamblers (see Garvía, 2007). Finally, winners are clustered and easy to locate because each winning number is typically sold by one outlet.

The research uses a fixed effect instrumental variable analysis to causally examine the impact of lottery winnings, a completely randomized and exogenous income shock, on household consumption behavior. The results reveal, in first instance, that the PIH is violated under the scenario presented by the Spanish Christmas lottery. These findings challenge existing theoretical expectations and diverge from the outcomes of a study on the Dutch Postcode lottery, where lottery winners primarily increased their consumption of durable goods. Second, we find that the consumption of household durable goods is sensitive to changes in total household expenditures, as an increase of 11.47% in the consumption of these goods. In contrast, the consumption of nondurable goods shows a unit-elastic response to total expenditures, because an increase of 10% in total household expenditures led to an increase of 9.26% in the consumption of household nondurable goods.

Statistically, these estimates are significantly different from each other, but economically, the elasticities are similar to one another. This suggests that households respond similarly to positive expenditure shocks in both durable and nondurable goods. This is the novelty of our paper, leading to a contradiction of the theoretical results in Cerletti and Pijoan-Mas (2014), where households that experience a positive income shock increase their consumption of durable goods and smooth their consumption of nondurable goods.

However, when we analyze the implications of the lottery income shock on labor supply, we do not find any evidence that lottery winnings induce heads of households to change their employment status or the number of hours worked. However, the paper acknowledges a limitation concerning the identification of individual winners. Instead, we rely on the region of residence to approximate the winning households. Due to variations in population

density and geographic size, this approach may limit the accuracy of the analysis, such as Madrid or Andalucía. To address this limitation, the paper proposes conducting a more focused regional analysis to better understand consumption behavior in winning regions.

The findings of increased consumption expenditures in various goods following the lottery income shock could be valuable for policymakers. The paper suggests that policy measures such as tax rebates or reductions in personal income taxes (IRPF) may encourage household consumption, especially in durable goods.

Notes

1. This is a syndicate game, where most of individuals share tickets with friends, colleagues or relatives.
2. The average lottery earnings per household is equal to 160 series of each number times ten 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 $= \frac{160 \times 10 \times 320{,}000}{\text{Total households per region}}$.
3. The 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 paper focuses on the PIH only.
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. *Source:* INE, 2014.
7. Family income survey.
8. Spanish National Statistics Institute.
9. Along the years the amount of households might vary slightly, some years the survey presents around 21,000, others more, others less. But in total, the sample covers 22,346.
10. 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.
11. Some authors include past-time consumption in the analysis or other nonlinearities, 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.
12. 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 pretreatment period, where none of the observations in the sample are treated and a posttreatment period, where some individuals in the sample will be treated after that given year.
13. According to Goodman-Bacon (2021), the TWFE method assigns a weighted average treatment effect to the 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.
14. This effect is obtained by adding the estimated coefficients for *win_region* and *lottery* × *win_region*.
15. The aggregate effect of *win_region* and *lottery* × *win_region* is close to zero.

References

Adamopoulou, E. and Zizza, R. (2017), "Regular versus Lump-Sum payments in union contracts and household consumption", ECB Working Paper No. 2013.

Arellano, M., Blundell, R. and Bonhomme, S. (2017), "Earnings and consumption dynamics: a nonlinear panel data framework", *Econometrica*, Vol. 85 No. 3, pp. 693-734.

Bagués, M. and Esteve-Volart, B. (2016), "Politicians' luck of the draw: evidence from the spanish christmas lottery", *Journal of Political Economy*, Vol. 124 No. 5, pp. 1269-1294.

Baker, A.-C., Larcker, D.-F. and Wang, C.-Y. (2021), "How much should we trust staggered difference-in-differences estimates?", *Journal of Financial Economics*, Vol. 144 No. 2, pp. 370-395.

Banks, J., Blundell, R. and Lewbel, A. (1997), "Quadratic Engel curves and consumer demand", *Review of Economics and Statistics*, Vol. 79 No. 4, pp. 527-539.

Berger, D., Guerrieri, V., Lorenzoni, G. and Vavra, J. (2018), "House prices and consumer spending", *The Review of Economic Studies*, Vol. 85 No. 3, pp. 1502-1542.

Bermejo, V.J., Ferreira, M.A., Wolfenzon, D. and Zambrana, R. (2019), "Entrepreneurship and regional windfall gains: evidence from the Spanish Christmas lottery", available at SSRN 3297158.

Berniell, I. (2018), "Pay cycles: individual and aggregate effects of paycheck frequency (No. 221)", *Documento de Trabajo*.

Blundell, R., Pashardes, P. and Weber, G. (1993), "What do We learn about consumer demand patterns from micro data?", *The American Economic Review*, Vol. 83 No. 3, pp. 570-597.

Browning, M. and Collado, D. (2007), "Habits and heterogenity in demands: a panel data analysis", *Journal of Applied Econometrics*, Vol. 22 No. 3, pp. 625-640.

Browning, M. and Crossley, T.-F. (2009), "Shocks, stocks and socks: smoothing consumption over a temporary income loss", *Journal of the European Economic Association*, Vol. 7 No. 6, pp. 1169-1192.

Cabanillas-Jiménez, G. (2022), "Consumption responses to income shocks: evidence from the british postcode lottery", available at SSRN 4144768.

Callaway, B. and Sant'Anna, P.-H. (2021), "Difference-in-differences with multiple time periods", *Journal of Econometrics*, Vol. 225 No. 2, pp. 200-230.

Cerletti, E.A. and Pijoan-Mas, J. (2014), "Durable goods, borrowing constraints and consumption insurance", CEPR Discussion Paper DP9035.

Cerulli, G. and Ventura, M. (2019), "Estimation of pre-and posttreatment average treatment effects with binary time-varying treatment using Stata", *The Stata Journal: Promoting Communications on Statistics and Stata*, Vol. 19 No. 3, pp. 551-565.

Cesarini, D., Lindqvist, E., Notowidigdo, M.J. and Östling, R. (2017), "The effect of wealth on individual and household labor supply: evidence from Swedish lotteries", *American Economic Review*, Vol. 107 No. 12, pp. 3917-3946.

Chaisemartin, C. and D'Haultfoeuille, X. (2020), "Two-way fixed effects estimators with heterogeneous treatment effects", *American Economic Review*, Vol. 110 No. 9, pp. 2964-2996.

Deaton, A. (2005), "Franco Modigliani and the life cycle theory of consumption", Available at SSRN 686475.

Disney, R. and Gathergood, J. (2009), "Housing wealth, liquidity constraints and selfemployment", *Labour Economics*, Vol. 16 No. 1, pp. 79-88.

Fagereng, A., Holm, M.B. and Natvik, G.J. (2021), "MPC heterogeneity and household balance sheets", *American Economic Journal: Macroeconomics*, Vol. 13 No. 4, pp. 1-54.

Friedman, M. (1957), *The Permanent Income Hypothesis. In a Theory of the Consumption Function*, Princeton University Press, Princeton, pp. 20-37.

Garvía, R. (2007), "Syndication, institutionalization, and lottery play", *American Journal of Sociology*, Vol. 113 No. 3, pp. 603-652.

Ghomi, M., Micó-Millán, I. and Pappa, E. (2022), "The sentimental propagation of lottery winnings: evidence from the Spanish Christmas lottery", CEPR Discussion Paper No. DP17173.

Goodman-Bacon, A. (2021), "Difference-in-differences with variation in treatment timing", *Journal of Econometrics*, Vol. 225 No. 2, pp. 254-277.

Grossman, S.J. and Laroque, G. (1990), "Asset pricing and optimal portfolio choice in the presence of illiquid durable consumption goods", *Econometrica*, Vol. 58 No. 1, pp. 25-51.

Hall, R.E. (1978), "Stochastic implications of the life cycle-permanent income hypothesis: Theory and evidence", *Journal of Political Economy*, Vol. 86 No. 6, pp. 971-987.

Haushofer, J. and Shapiro, J. (2015), "Household response to income changes: evidence from an unconditional cash transfer program in Kenya", *Massachusetts Institute of Technology*, Vol. 24 No. 5, pp. 1-57.

Hsieh, C.T. (2003), "Do consumers react to anticipated income changes? Evidence from the Alaska permanent fund", *American Economic Review*, Vol. 93 No. 1, pp. 397-405.

Imbens, G.W., Rubin, D.B. and Sacerdote, B.I. (2001), "Estimating the effect of unearned income on labor earnings, savings, and consumption: evidence from a survey of lottery players", *American Economic Review*, Vol. 91 No. 4, pp. 778-794.

Kapplan, G. and Violante, G.-L. (2014), "A model of the consumption response to fiscal stimulus payments", *Econometrica*, Vol. 82 No. 4, pp. 1199-1239.

Krueger, D. and Perri, F. (2011), "How do households respond to income shocks?", Consortium on Financial Systems and Poverty.

Kuhn, P., Kooreman, P., Soetevent, A. and Kapteyn, A. (2011), "The effects of lottery prizes on winners and their neighbours: evidence from the Dutch postcode lottery", *American Economic Review*, Vol. 101 No. 5, pp. 2226-2247.

Further reading

Blanchard, O., Amighini, A. and Giavazzi, F. (2010), *Macroeconomics. A European Perspective*, Pearson Education, London.

Blundell, R., Pistaferri, L. and Saporta-Eksten, I. (2016), "Consumption inequality and family labor supply", *American Economic Review*, Vol. 106 No. 2, pp. 387-435.

Supplementary material

The supplementary material for this article can be found online.

Corresponding author

Guillermo Cabanillas-Jiménez can be contacted at: guillermo.cabanillas@uah.es

Appendix 1. Average lottery earnings

Region	Proportion of interviewed people (%)	Estimation of potential winners (%)	Region (%)	Proportion of interviewed 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

Source: Table created by author, based on the data provided from the Instituto Nacional de Estadística (INE). This table presents the proportion of households that take part of our sample by region, as well as an expectation of the proportion of individuals in each region that can be potential winners of the lottery. In other words, of those individuals that were interviewed 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

Table A1.
Proportion of people included in the sample and potential winners per region

Year	Average	Year	Average
2016	387.72	2006	731.38
2015	224.12	2005	240.2
2014	1,507.5	2004	428.34
2013	753.81	2003	914.65
2012	860.88	2002	535.82
2011	2,263.4	2001	847.62
2010	1,641.59	2000	1,157.63
2009	387.72	1999	348.95
2008	679.22	1998	667.95
2007	918.29		

Source: Table created by author, based on the data provided from the Instituto Nacional de Estadística (INE). This table presents the average lottery earnings per year. The way we compute this mean is by multiplying the number of series (160) and tickets in each series (ten) times the awarded prize (320,000), over the number of households in the winning region of the Christmas lottery. The variability on earnings differs along the years, as the winning region changes each year and population size is different across regions

Table A2.
Average lottery earnings per region

Appendix 2. Estimation results

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	

Source: Instituto Nacional de Estadística (INE). Here we present an exogeneity test of the treatment group. The model used to test this regression is a linear probability one, using the treatment variable as our dependent variable on all household characteristics and control variables. We use the *F*-test to check for exogeneity, under the null that all estimated regressors are equal to zero. Given that the *F*-test is 0.66, we fail to reject the null hypothesis and the identifying assumption of the paper holds. We compute robust standard errors. *t*-statistics in parentheses: *$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$

Table A3. Identification strategy

	(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	211,096
R-squared	0.858

Source: Instituto Nacional de Estadística (INE). This table presents the extended results of Table 5, where we described in detail the specifications and the estimation process. We compute robust standard errors. *t*-statistics in parentheses: *$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$

Table A4. First stage regression – total household expenditures

Table A5.
Second stage estimation – household consumption behavior

Panel A	(1) Food at home	(2) Alcohol	(3) Clothes	(4) House rent	(5) House durable	(6) Health	(7) Car value	(8) Transport
Log expenditures	0.725*** (14.75)	0.505*** (4.68)	0.684*** (7.06)	0.979*** (18.27)	1.037*** (15.09)	0.453*** (4.17)	−0.571*** (−6.19)	0.549*** (6.34)
Lottery	−0.108*** (14.75)	0.780*** (4.68)	0.525*** (7.06)	−0.326*** (18.27)	−0.114*** (15.09)	0.945*** (4.17)	1.586*** (−6.19)	0.449*** (6.34)
Log lottery expenditures	0.0968 (1.33)	−0.781*** (−4.90)	−0.428** (−2.98)	−0.120 (−1.52)	−0.0846(−0.83)	−0.687*** (−4.27)	−1.199*** (−8.80)	−0.539*** (−4.21)
Log GDP	0.104 (0.51)	−0.672 (−1.51)	1.557*** (3.89)	−0.0302 (−0.14)	−0.362(−1.28)	1.713*** (3.82)	−0.889* (−2.34)	0.134 (0.37)
Age	−0.0132*** (−6.04)	0.0966*** (20.21)	0.0621*** (14.44)	−0.0148*** (−6.25)	0.0112*** (3.66)	0.0527*** (10.93)	0.000892 (0.22)	0.138*** (35.81)
Age^2	0.000140*** (6.94)	−0.000890*** (−20.18)	−0.000835*** (−21.04)	0.000162*** (7.37)	−0.000106*** (−3.78)	−0.000383*** (−8.61)	−0.0000932* (−2.47)	−0.00162*** (−45.73)
Marital	0.0138*** (3.56)	−0.0613*** (−7.21)	−0.0600*** (−7.84)	−0.00317 (−0.75)	−0.0172** (−3.18)	−0.0724*** (−8.44)	−0.0315*** (−4.33)	−0.0920*** (−13.43)
Education	−0.00160 (−0.78)	0.00410 (0.91)	0.0296*** (7.29)	−0.00856*** (−3.81)	0.000268 (0.09)	0.0461*** (10.11)	0.0113** (2.92)	0.0523*** (14.37)
Employed	−0.0142* (−2.30)	−0.0339* (−2.50)	0.0675*** (5.53)	−0.00483 (−0.72)	−0.00389 (−0.45)	0.0410** (2.99)	0.0127 (1.10)	0.100*** (9.16)
Retired	−0.00144 (−0.08)	−0.439*** (−11.12)	0.122*** (3.44)	−0.00799 (−0.41)	0.0339 (1.35)	−0.260*** (−6.51)	−0.483*** (−14.30)	0.151*** (4.76)
_cons	−0.302 (−0.40)	3.573* (2.17)	−4.492** (−3.03)	−1.126 (−1.37)	−0.768 (−0.73)	−3.836* (−2.31)	8.294*** (5.88)	0.0774 (0.06)
N	211,096	211,096	211,096	211,096	211,096	211,096	211,096	211,096
R-squared	0.831	0.517	0.612	0.874	0.802	0.506	0.211	0.693

Panel B	(9) Communication	(10) Leisure	(11) Education	(12) Food out home	(13) Holidays	(14) Savings	(15) Durables	(16) Nondurables
Log expenditures	1.101*** (16.13)	1.009*** (10.80)	0.406*** (3.85)	0.870*** (9.69)	0.146 (1.29)	−0.294** (−2.62)	1.147*** (17.99)	0.926*** (15.89)
Lottery	−0.253*** (16.13)	0.514*** (10.80)	1.233*** (3.85)	0.421*** (9.69)	1.215*** (1.29)	1.042*** (−2.62)	−0.0672* (17.99)	0.0244 (15.89)
Log lottery expenditures	0.0939 (0.93)	0.215 (1.56)	−0.505** (−3.24)	0.0705 (0.53)	−1.449*** (−8.66)	−0.954*** (−5.75)	−0.0403 (−0.43)	−0.00506 (−0.06)
Log GDP	1.371*** (4.87)	2.024*** (5.25)	3.753*** (8.62)	3.418*** (9.23)	5.733*** (12.28)	−10.75*** (−23.20)	−0.0751 (−0.29)	0.171 (0.71)
Age	0.00656* (2.16)	0.0910*** (21.95)	0.0983*** (20.99)	0.0922*** (23.15)	0.0856*** (17.02)	0.118*** (23.61)	0.00337 (1.19)	0.00139 (0.54)
Age^2	−0.0000240 (−0.86)	−0.00108*** (−28.32)	−0.00116*** (−26.88)	−0.00113*** (−30.84)	−0.00107*** (−23.16)	−0.000521*** (−11.33)	−0.00000599 (−0.23)	−0.0000315 (−1.32)
Marital	−0.0274*** (−5.08)	−0.105*** (−14.29)	−0.173*** (−20.80)	−0.100*** (−14.12)	−0.0609*** (−6.81)	0.0334*** (3.76)	−0.0106* (−2.11)	−0.000621 (−0.13)
Education	0.0172*** (6.00)	0.0656*** (16.76)	0.0484*** (10.95)	0.0610*** (16.20)	0.0982*** (20.69)	−0.0923*** (−19.60)	−0.00467 (−1.75)	0.00351 (1.44)
Employed	0.0511*** (5.94)	0.0767*** (6.52)	−0.0197(−1.48)	0.169*** (14.95)	0.154*** (10.83)	−0.136*** (−9.59)	−0.00105 (−0.13)	0.00517 (0.70)
Retired	−0.00470 (−0.19)	0.0798* (2.33)	−0.821*** (−21.21)	0.0819* (2.49)	0.316*** (7.62)	−0.384*** (−9.34)	−0.0129 (−0.55)	0.0163 (0.76)
_cons	−6.844*** (−6.56)	−9.427*** (−6.60)	−10.62*** (−6.58)	−12.54*** (−9.14)	−12.24*** (−7.07)	35.53*** (20.69)	−1.433 (−1.47)	−0.460 (−0.52)
N	211,096	211,096	211,096	211,096	211,096	211,096	211,096	211,096
R-squared	0.814	0.633	0.331	0.675	0.346	0.390	0.867	0.855

Source: Instituto Nacional de Estadística (INE). This table presents the extended results of Table 6, where we described in detail the specifications and the estimation process. We compute robust standard errors. *t*-statistics in parentheses: *$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$

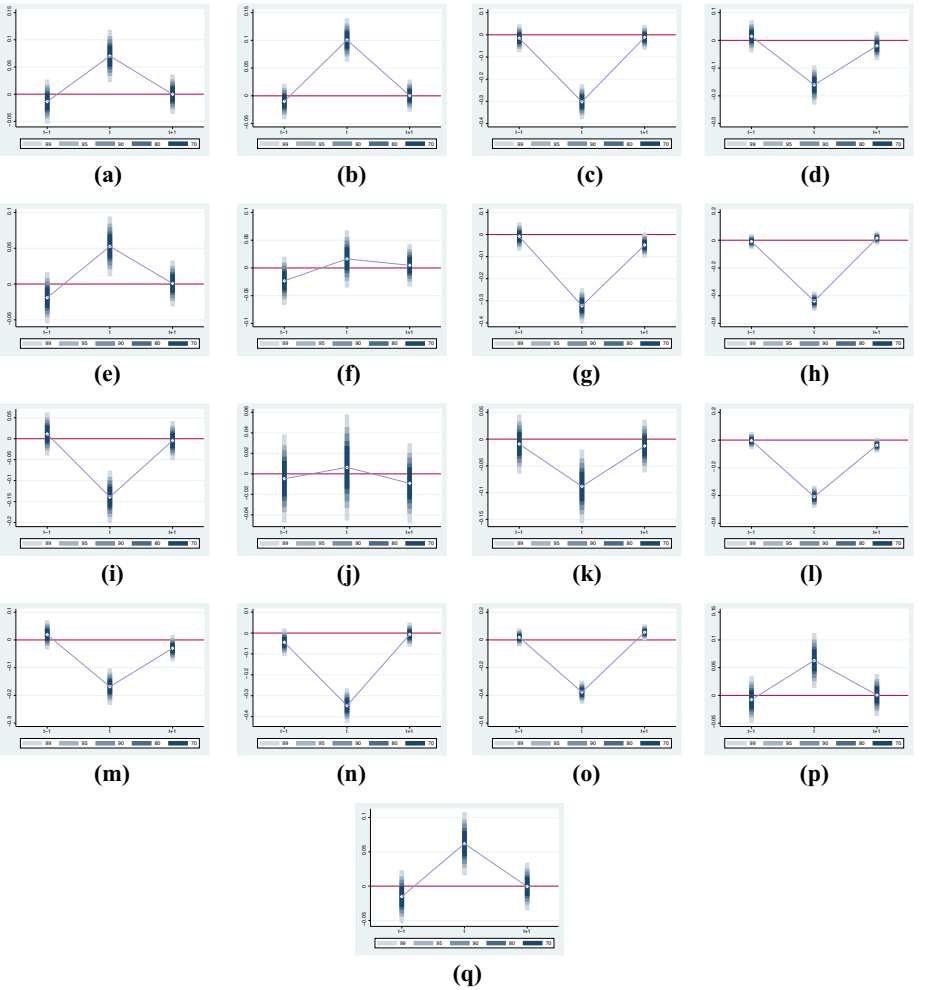
	(1) Employed	(2) Employed	(3) Hours worked	(4) Hours worked
Win_region	−0.00684* (−2.05)	−0.00639 (−1.73)	−0.123*** (−3.61)	−0.0605 (−1.43)
Lottery	0.0124*** (6.43)	0.00678** (3.05)	−0.329*** (−18.53)	0.000878 (0.03)
Lottery × win_region	0.00219 (0.46)	−0.00281 (−0.58)	0.271*** (5.50)	0.0564 (1.00)
_cons	0.419*** (321.75)	−1.117*** (−5.68)	4.206*** (348.30)	0.800 (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	278,490	211,096	174,706	95,840
R-squared	0.000225	0.136	0.00197	0.000276

Notes: 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*

Source: *Instituto Nacional de Estadística (INE)*. The coefficients presented in this table, *win_region*, *Lottery* and *lottery* × *win_region*, are as described in table. 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 Columns (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: *$p < 0.05$; **$p < 0.01$ and ***$p < 0.001$. Full set of estimates available upon request

Table A6. Labor supply analysis

Appendix 3. Testing the parallel trend assumption



Notes: (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) nondurables
Source: Instituto Nacional de Estadística (INE)

Figure A1.
Parallel trends - treatment effect

Table A7. Testing the parallel trend assumption

Panel A	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
F-test	0.61	0.35	0.44	1.81	1.87	0.12	0.20	0.32
	Food at home	Alcohol	Clothes	House rent	House investments	Health	Car value	Transport
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	Nondurables
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

Source: Instituto Nacional de Estadística (INE). 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 (2019)