

Online Supplementary material file

Testing the Permanent Income Hypothesis using the Spanish Christmas Lottery.

Guillermo Cabanillas-Jiménez¹

¹Universidad de Alcalá (UAH), guillermo.cabanillas@uah.es

March 28, 2024

This document forms part of an online appendix to the paper. It provides a detailed description of the variables used throughout the paper, offers evidence for the data obtained from the Sociedad Estatal de Loterías y Apuestas del Estado (SELAE), and conducts several robustness checks to address various issues: (i) excluding the regions of Cataluña and Madrid in the analysis due to the presence of outlets in these areas that serve the entire country's population with lottery tickets, (ii) considering only regions with a single province in the analysis, (iii) adjusting household consumption in real terms, and finally, (iv) exploring whether the lottery income shock might also influence household composition by impacting decisions regarding the number of children.

A Map of the variables used

In this section we aim to provide a list of all the variables used in the analysis of the paper, jointly with the description provided in the questionnaire of the EPF.

- **Win_region:** This a dummy variable that says that if the household is located in a winning region of the Christmas Lottery and, the lottery was awarded in a capital of province and the household is located there too, this variable takes value one. In case the lottery is awarded outside a capital of province and the household is located outside it (but in the winning region), this variable takes also value one. Otherwise, the variable takes value zero.
- **Lottery:** The EPF asks about consumption in the Christmas Lottery and other lottery games. This is a dummy variable that takes value one if consumption is greater than 0.
- **Lottery expenditure:** Collected from the *Sociedad Estatal de Loterías y Apuestas del Estado* (SELAE). Provides the average Christmas Lottery expenditures per year in each region.

- **Age:** Age of the head of the household.
- **Marital_status:** Categorical variable indicating the marital status of the head of the household. Value 1 is single, value 2 is married, value 3 is widowed, value 4 is separated and value 5 means that the head is divorced.
- **Education:** Categorical variable indicating the educational level of the head of the household.
- **Employed:** Dummy variable that takes value one if the head of the household is employed.
- **Retired:** Dummy variable that takes value one if the head of the household is retired.
- **Consumption expenditures, considering the different good categories:** Total amount of the annual expenditure of the household (monetary and non-monetary, adjusted for temporary and population factors) (for in-kind salary, both the amount paid and the bonus received are accounted for).
- **Total Expenditures:** Aggregate of household consumption expenditures of all different goods.
- **Durables:** Aggregate of household consumption expenditures of all durable goods.
- **Non_durables:** Aggregate of household consumption expenditures of all non-durable goods.

B SELAE Data

Apart from the EPF, we also use regional and national data on Christmas Lottery expenditures, available in the SELAE. In addition, we use the national GDP and GDP per capita, both sources available in the INE, to measure the average lottery expenditure per year, relative to the Spanish GDP. Table S1 presents a summary statistics for all these variables of interest covering from 1998 to 2016.

From Table S1, we observe that the Christmas Lottery expenditure relative to GDP is equal to 0.3%. This fact has been stable not only throughout the years under analysis, but also in the previous two decades (see Bagués and Esteve-Volart, 2016). However, when we look at lottery expenditures in levels, this increased through the years until 2006, when it became stagnant and fell during the following years - coinciding with the economic recession in 2008. After these years lottery consumption was, on average, around 50€, but started to increase again. In 2018 consumption rose to 67.58€, according to the SELAE.

Table S1: Lottery Expenditures

Year	GDP	GDP pc	Lottery exp. per capita	% of lottery exp. to GDP pc
1998	118386400€	140236.8€	39.87€	0.29%
1999	121493500€	142041.3€	41.83€	0.30%
2000	125689700€	145465.5€	42.42€	0.29%
2001	130972800€	149243.1€	44.19€	0.30%
2002	136616500€	153629.1€	47.76€	0.31%
2003	142270900€	157438.8€	50.09€	0.32%
2004	147994900€	161861.8€	50.88€	0.32%
2005	154340900€	165057.9€	53.13€	0.33%
2006	160380700€	169275.2€	53.56€	0.32%
2007	165626800€	173197.8€	54.55€	0.32%
2008	171188800€	175453.3€	51.89€	0.30%
2009	162610600€	164495.8€	50.37€	0.31%
2010	162272700€	164185.9€	49.82€	0.31%
2011	162326500€	164724.6€	49.27€	0.30%
2012	160492369€	159618.1€	52.26€	0.33%
2013	158887445€	158660.4€	63.20€	0.40%
2014	160794095€	161040.3€	61.60€	0.38%
2015	167869035€	168287.2€	62.70€	0.37%
2016	173576582€	173840.6€	63.80€	0.37%

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

C Robustness checks

C.1 Excluding Cataluña and Madrid

To ensure the robustness of our findings, we conducted an additional analysis excluding the regions of Cataluña and Madrid. This decision was motivated by the presence of two major outlets in both regions, which attract the entire population due to superstitious beliefs in buying lottery tickets. Consequently, if these outlets happen to be the winners in a given year, it could potentially skew the results and not accurately reflect the impact of the income shock, as explained in Section 2 of the paper. When examining the income lottery shock as outlined in Equation (1), we observed minimal deviations from the results presented in Table 3 in the paper.

Table S2: Household consumption expenditures - Excluding Cataluña and Madrid

PANEL A	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Food at home	Alcohol	Clothes	House Rent	House Investments	Health	Car Value	Transport
win_region	-0.122*** (-11.79)	0.0265 (1.13)	-0.0414 (-1.91)	-0.147*** (-13.91)	-0.140*** (-9.44)	0.0130 (0.55)	0.187*** (12.42)	-0.0131 (-0.69)
Lottery	0.170*** (26.66)	1.028*** (77.26)	0.832*** (69.91)	0.0716*** (10.00)	0.313*** (35.74)	1.178*** (87.60)	1.470*** (118.24)	0.698*** (65.17)
Lottery × win_region	0.0496*** (3.62)	-0.325*** (-11.47)	-0.197*** (-7.74)	0.00992 (0.68)	-0.0328 (-1.80)	-0.282*** (-9.87)	-0.364*** (-14.12)	-0.163*** (-7.21)
.cons	22.32*** (10.26)	45.26*** (10.62)	18.95*** (4.85)	36.40*** (15.12)	39.36*** (13.53)	20.74*** (4.89)	3.181 (0.90)	30.08*** (8.55)
F-test	46.06	208.64	176.31	136.46	162.59	171.84	57.39	119.49
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	176505	176505	176505	176505	176505	176505	176505	176505
R-squared	0.799	0.478	0.570	0.845	0.766	0.464	0.210	0.653
PANEL B	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Communication	Leisure	Education	Food out Home	Holidays	Savings	Durables	Non-durables
win_region	-0.173*** (-11.37)	-0.140*** (-6.48)	0.0284 (1.13)	-0.108*** (-5.31)	0.0571* (2.53)	0.0924*** (3.99)	-0.173*** (-13.62)	-0.144*** (-12.18)
Lottery	0.209*** (23.86)	0.947*** (83.11)	1.709*** (104.02)	0.816*** (74.64)	1.344*** (93.38)	0.977*** (68.63)	0.395*** (46.48)	0.393*** (51.12)
Lottery × win_region	-0.00797 (-0.42)	-0.107*** (-4.25)	-0.359*** (-10.20)	-0.181*** (-7.61)	-0.262*** (-8.64)	-0.317*** (-10.67)	0.00766 (0.44)	0.0129 (0.79)
.cons	26.83*** (9.26)	21.41*** (5.79)	2.795*** (6.78)	2.410 (0.67)	-25.64*** (-5.94)	110.9*** (25.41)	38.75*** (13.60)	31.15*** (11.99)
F-test	178.30	200.98	130.17	314.15	71.15	95.79	138.82	100.12
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	176505	176505	176505	176505	176505	176505	176505	176505
R-squared	0.776	0.590	0.290	0.635	0.316	0.375	0.838	0.826

Data Source: 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. *t*-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of estimates available upon request.

The *F*-test performs a joint significant test of the lottery income shock variables (*win_region* and *Lottery × win_region*), 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.

Table S2 reveals that excluding Cataluña and Madrid from the analysis continues to show a decrease in household consumption across various goods, both durable and non-durable, when comparing winning regions to non-winning ones. However, when focusing on potential lottery winners -those residing in winning regions who purchased lottery tickets -we maintain the positive effects observed in Table 5 of the paper. Specifically, households in winning regions that bought lottery tickets exhibit an increase of 0.008 percentage points in durable goods consumption and 0.013 percentage points in non-durable goods consumption compared to the control group. This confirms that indeed, households that can be potential winners of the lottery (i.e., buyers of lottery tickets and living in a winning region) increase their consumption,

leading to a potential violation of the PIH. This is further corroborated by an F -test, rejecting the null hypothesis that living in winning lottery regions has no impact on household consumption. This violation of the PIH persists across all goods even when excluding the regions of Cataluña and Madrid from the analysis.

Table S3: First stage regression - Total household expenditures, excluding Cataluña and Madrid

	(1) Total Expenditures
win_region	-0.154*** (-10.28)
lottery	0.396*** (46.77)
lottery \times win_region	0.0195 (0.97)
_cons	35.76*** (13.20)
Specification	Expenditures in logarithms
F -test for the IV	76.77
p -value	0.0000
N	176505
R-squared	0.843

Data Source: 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 \times win_region*, as instruments for total household expenditures. The coefficients *win_region*, *lottery* and *lottery \times win_region*, are as previously described in Table S2. We present the effect of the lottery income shock on total household expenditures in logarithms; we also did the estimations for total expenditures in levels, but the results were showing a negative impact of the lottery income shock on total expenditures, thus we avoid using expenditures in levels, as the estimates go against our expectations. Both specifications include as control variables the age of the head of the household and its square, the marital status of the head of the household and his/her educational level, employment status and whether he/she is retired or not. Moreover, we also include the logarithm of lottery expenditures per region and the regional log-GDP as demographic controls, and both regional and year fixed-effects. We compute robust standard errors. t -statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of estimates available upon request.

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

In addition, we conduct a robustness exercise focusing on the Engel curves. In the first stage analysis, as reflected in Table 6 of the paper, we find that the lottery income shock significantly impacts total household expenditures. Specifically, when excluding Cataluña and Madrid, households in winning regions participating in the Christmas Lottery experience a 0.02 percentage point increase in total household expenditure, suggesting that potential winners may contribute to an increase in consumption. However, the *win_region* coefficient displays a negative trend, potentially influenced by higher participation rates in lower-income regions. The F -test, assessing the relevance of the instruments, returns a value of 76.77. Like the analysis of the entire sample, this test confirms that the relevance condition is fulfilled. Therefore, the

relevance condition is satisfied in jointly with orthogonality, automatically maintained due to the completely random assignment of winning regions, as previously explained.

Table S4: Second stage estimation - Household consumption behavior, excluding Cataluña and Madrid

PANEL A	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Food at home	Alcohol	Clothes	House Rent	House Investments	Health	Car Value	Transport
log expenditures	0.678*** (10.53)	0.814*** (5.93)	0.845*** (6.80)	0.964*** (13.55)	1.021*** (11.36)	0.695*** (5.01)	-0.236* (-2.02)	0.513*** (4.59)
Household controls	✓	✓	✓	✓	✓	✓	✓	✓
Fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Elasticity test	25.00	1.83	1.55	0.26	0.05	4.83	112.6	18.89
<i>p</i> -value	0.0000	0.1765	0.2128	0.6078	0.8151	0.0280	0.0000	0.0000
N	176505	176505	176505	176505	176505	176505	176505	176505
R-squared	0.816	0.502	0.595	0.858	0.785	0.489	0.218	0.675
PANEL B	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Communication	Leisure	Education	Food out Home	Holidays	Savings	Durables	Non-durables
log expenditures	1.222*** (13.57)	1.221*** (10.16)	0.756*** (5.70)	1.203*** (10.40)	0.414** (2.91)	0.315* (2.22)	1.149*** (13.60)	0.930*** (12.12)
Household controls	✓	✓	✓	✓	✓	✓	✓	✓
Fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Elasticity test	6.06	3.38	3.40	3.09	16.90	23.29	3.11	0.82
<i>p</i> -value	0.0138	0.0659	0.0652	0.0789	0.0000	0.0000	0.0778	0.3646
N	176505	176505	176505	176505	176505	176505	176505	176505
R-squared	0.794	0.614	0.322	0.657	0.333	0.394	0.858	0.848

Data Source: 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: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of estimates available upon request. The reported Elasticity test examines the elasticity effect of expenditures towards household consumption, in words, whether the estimates are different from one.

The estimates from the second stage analysis align closely with those in Table 7 in the paper. Upon examining the results in Table S4, even with the exclusion of Cataluña and Madrid, durable and non-durable goods continue to exhibit unit-elastic responses to a shock in total household expenditures, as confirmed by the *F*-test, under the null hypothesis that the estimated elasticity coefficient differs from one. More precisely, a 10% increase in household total expenditures leading to an 11.49% rise in household expenditures for durable goods and a 9.30% increase in household consumption of non-durable goods, both effects are statistically significant. Therefore, in line with the findings presented in the paper, the results still demonstrate unit elasticity to a shock in total household expenditures, and the estimated elasticities for durable and non-durable goods not significantly different from each other.

Therefore, results are still consistent with the obtained ones in the paper when excluding Cataluña and Madrid from the analysis, for both: the PIH test and the Engel curves estimates.

C.2 Using single-provined regions

The second robustness check undertaken in this study involves replicating the analysis, but this time focusing on regions that consist of a single province. This decision arises from the fact that the lottery shock operates at the municipal level, a feature of the game. Consequently, if a town in Valencia wins *El Gordo* in a given year, households in Alicante and Castellon would also be treated, posing potential questions and introducing bias or noise to the estimated coefficients. To enhance the reliability of the results, we restrict the sample to autonomous communities that also function as provinces. This approach follows the methodology employed by Bagués and Esteve-Bolart (2016) and Bermejo et al. (2020), who used data available at the province level. Once more, when examining the income lottery shock as specified in Equation (1), we once again identify minimal deviations from the results presented in Table 5 of the paper.

Table S5: Household consumption expenditures - Including only single-provined regions

PANEL A	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Food at home	Alcohol	Clothes	House Rent	House Investments	Health	Car Value	Transport
win_region	-0.0951*** (-4.24)	0.0350 (0.68)	-0.0474 (-0.99)	-0.0967*** (-4.32)	-0.0856** (-2.67)	0.0527 (1.00)	0.195*** (5.63)	-0.0540 (-1.30)
Lottery	0.201*** (18.25)	1.113*** (48.61)	0.865*** (42.24)	0.0895*** (7.40)	0.374*** (24.85)	1.198*** (51.99)	1.529*** (69.93)	0.689*** (37.97)
Lottery × win_region	0.0859*** (3.50)	-0.156** (-2.93)	-0.0450 (-0.92)	0.0118 (0.49)	0.0181 (0.55)	-0.153** (-2.84)	-0.363*** (-7.70)	-0.0115 (-0.27)
.cons	-29.56*** (-8.16)	-60.85*** (-6.98)	-74.50*** (-9.54)	-50.94*** (-13.56)	-52.63*** (-9.88)	-88.72*** (-10.18)	-10.73 (-1.41)	-67.53*** (-10.04)
<i>F</i> -test	0.26	13.45	9.84	20.47	9.63	9.38	17.01	6.42
<i>p</i> -value	0.6072	0.0002	0.0017	0.0000	0.0019	0.0022	0.0000	0.0113
N	56188	56188	56188	56188	56188	56188	56188	56188
R-squared	0.812	0.520	0.605	0.862	0.784	0.511	0.215	0.689
PANEL B	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Communication	Leisure	Education	Food out Home	Holidays	Savings	Durables	Non-durables
win_region	0.0421 (-1.44)	-0.00978 (-0.22)	0.0174 (0.32)	0.0264 (0.60)	-0.00749 (-0.14)	0.0961 (1.87)	-0.123*** (-4.55)	-0.115*** (-4.52)
Lottery	0.205*** (14.09)	0.929*** (48.43)	1.838*** (64.52)	0.804*** (43.34)	1.332*** (53.12)	1.077*** (43.65)	0.449*** (31.21)	0.416*** (31.80)
Lottery × win_region	-0.0481 (-1.60)	-0.0825 (-1.82)	-0.318*** (-4.84)	-0.118** (-2.69)	-0.0732 (-1.26)	-0.241*** (-4.16)	0.0583* (1.99)	0.0752** (2.68)
.cons	-64.39*** (-13.18)	-69.74*** (-9.58)	4.348*** (5.57)	-92.84*** (-13.24)	-107.6*** (-11.69)	84.22*** (9.10)	-53.31*** (-11.71)	-43.85*** (-10.57)
<i>F</i> -test	18.75	11.17	35.07	12.45	3.88	11.91	8.16	3.41
<i>p</i> -value	0.0000	0.0008	0.0000	0.0004	0.0488	0.0006	0.0043	0.0647
N	56188	56188	56188	56188	56188	56188	56188	56188
R-squared	0.813	0.638	0.318	0.677	0.357	0.374	0.854	0.842

Data Source: Instituto Nacional de Estadística (INE). Our dependent variables are the logarithm of consumption expenditures of each good category. The coefficients presented in this table, *win_region*, *Lottery* and *lottery × win_region*, are as described in Table S2. In this specification, we also include as control variables the age of the head of the household and its square, the marital status of the head of the household and his/her educational level, employment status and whether he/she is retired or not. Moreover, we also include the logarithm of lottery expenditures per region and the regional log-GDP as demographic controls, and both regional and year fixed-effects. We compute robust standard errors. *t*-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of estimates available upon request. The *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.

In Table S5, the repeated analysis for the subsample of single-province regions indicates that the estimates consistently reveal a decrease in household consumption across various goods, both durable and non-durable, when comparing winning regions to non-winning ones. However, when focusing on potential lottery winners -those residing in winning regions who purchased lottery tickets- we observe the sustained positive and significant effects outlined in Table 5 of the paper. Specifically, households in winning regions that bought lottery tickets show an increase of 0.058 percentage points in durable goods consumption and 0.075 percentage points in non-durable goods consumption compared to the control group. This verifies that households with the potential to win the lottery (i.e., buyers of lottery tickets living in a winning region) increase their consumption, potentially violating the Perfectly Insured Household (PIH). This is further affirmed by an *F*-test that rejects the null hypothesis that living in winning lottery regions has no impact on household consumption.

This violation of the PIH persists across all goods, even when considering the subsample of single-province regions exclusively.

Table S6: First stage regression - Total household expenditures, including only single-provined regions

(1)	
Total Expenditures	
win_region	-0.122*** (-3.29)
lottery	0.434*** (26.06)
lottery × win_region	0.0969* (2.24)
_cons	-42.57*** (-6.90)
Specification	Expenditures in logarithms
<i>F</i> -test for the IV	5.42
<i>p</i> -value	0.0044
N	56188
R-squared	0.859

Data Source: 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 S2. We present the effect of the lottery income shock on total household expenditures in logarithms; we also did the estimations for total expenditures in levels, but the results were showing a negative impact of the lottery income shock on total expenditures, thus we avoid using expenditures in levels, as the estimates go against our expectations. Both specifications include as control variables the age of the head of the household and its square, the marital status of the head of the household and his/her educational level, employment status and whether he/she is retired or not. Moreover, we also include the logarithm of lottery expenditures per region and the regional log-GDP as demographic controls, and both regional and year fixed-effects. We compute robust standard errors. *t*-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of estimates available upon request.

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.*

Additionally, we perform the robustness exercise with a focus on the Engel curves. In the first stage analysis, as showed in Table 6 of the paper, we find that the lottery income shock significantly influences total household expenditures. Specifically, when analyzing the subsample of single-province regions in Table S6, households in winning regions participating in the Christmas Lottery experience a 0.097 percentage point increase in total household expenditure, indicating a potential contribution of potential winners to a consumption increase. However, the *win_region* coefficient exhibits a negative trend, potentially influenced by higher participation rates in lower-income regions. The *F*-test, evaluating the relevance of the instruments, is of 5.42. In line with the analysis of the entire sample, this test confirms the fulfillment of the relevance condition. Therefore, the relevance condition is satisfied concurrently with orthogonality, automatically maintained due

to the completely random assignment of winning regions, as previously detailed.

Table S7: Second stage estimation - Household consumption behavior, including only single-provined regions

PANEL A	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Food at home	Alcohol	Clothes	House Rent	House Investments	Health	Car Value	Transport
log expenditures	0.803** (3.27)	0.0369 (0.07)	0.216 (0.46)	0.782** (2.95)	0.805* (2.36)	-0.445 (-0.85)	-1.536*** (-3.30)	0.452 (1.08)
Household controls	✓	✓	✓	✓	✓	✓	✓	✓
Fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Elasticity test	0.65	3.44	2.79	0.68	0.32	7.64	29.66	1.72
<i>p</i> -value	0.4217	0.0637	0.0948	0.4110	0.5689	0.0057	0.0000	0.1900
N	56188	56188	56188	56188	56188	56188	56188	56188
R-squared	0.829	0.545	0.629	0.875	0.802	0.539	0.222	0.710
PANEL B	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Communication	Leisure	Education	Food out Home	Holidays	Savings	Durables	Non-durables
log expenditures	0.219 (0.67)	-0.244 (-0.55)	-0.363 (-0.70)	-0.0754 (-0.18)	-0.0910 (-0.16)	-0.747 (-1.36)	0.959** (3.03)	0.932** (3.22)
Household controls	✓	✓	✓	✓	✓	✓	✓	✓
Fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Elasticity test	5.77	7.83	6.91	6.27	3.88	10.10	0.02	0.06
<i>p</i> -value	0.0163	0.0051	0.0086	0.0122	0.0489	0.0015	0.8957	0.8129
N	56188	56188	56188	56188	56188	56188	56188	56188
R-squared	0.830	0.661	0.347	0.699	0.376	0.394	0.909	0.901

Data Source: 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: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of estimates available upon request. The reported Elasticity test examines the elasticity effect of expenditures towards household consumption, in words, whether the estimates are different from one.

The results derived from the second stage analysis closely align with those illustrated in Table 7 in the paper. After reviewing the findings in Table S7, even when considering the analysis of the subsample of single-province regions, durable and non-durable goods maintain their unit-elastic responses to a shock in total household expenditures, as validated by the *F*-test under the null hypothesis that the estimated elasticity coefficient deviates from one. More specifically, a 10% increase in household total expenditures results in a 9.59% increase in household expenditures for durable goods and a 9.32% rise in household consumption of non-durable goods - both effects are statistically significant. Hence, consistent with the findings presented in the paper, the results persist in demonstrating unit elasticity to a shock in total household expenditures, and the estimated elasticities for durable and non-durable goods do not show statistically significant differences.

Therefore, again, results are still consistent with the obtained ones in the paper when considering only the subsample of single-province regions in the analysis, for both: the PIH test and the Engel curves estimates.

C.3 Testing the PIH, adjusting consumption in real terms

Table S8: Household consumption expenditures in real terms

PANEL A	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Food at home	Alcohol	Clothes	House Rent	House Investments	Health	Car Value	Transport
win_region	-0.166*** (-16.28)	-0.224*** (-14.23)	-0.240*** (-16.33)	-0.216*** (-17.32)	-0.266*** (-19.92)	-0.235*** (-14.46)	-0.248*** (-5.54)	-0.271*** (-18.15)
Lottery	0.180*** (28.93)	0.203*** (20.48)	0.242*** (26.03)	0.136*** (16.70)	0.274*** (31.60)	0.208*** (20.56)	0.0919*** (3.72)	0.267*** (28.95)
Lottery × win_region	0.107*** (8.08)	0.0919*** (4.63)	0.128*** (6.81)	0.0498** (2.97)	0.133*** (7.49)	0.105*** (5.17)	0.101* (2.10)	0.177*** (9.36)
.cons	20.72*** (10.15)	30.35*** (9.56)	29.68*** (10.12)	29.35*** (11.16)	25.08*** (9.03)	28.57*** (8.90)	26.77*** (4.68)	42.31*** (14.30)
<i>F</i> -test	33.71	80.51	62.50	155.96	92.18	76.29	45.40	43.21
<i>p</i> -value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	211096	211096	211096	211096	211096	211096	211096	211096
R-squared	0.793	0.841	0.829	0.821	0.823	0.843	0.843	0.825
PANEL B	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Communication	Leisure	Education	Food out Home	Holidays	Savings	Durables	Non-durables
win_region	-0.221*** (-16.54)	-0.264*** (-17.75)	-0.218*** (-9.16)	-0.232*** (-15.94)	-0.163*** (-8.14)	-0.127*** (-6.96)	-0.247*** (-18.91)	-0.176*** (-15.56)
Lottery	0.122*** (13.82)	0.411*** (44.94)	0.193*** (15.20)	0.221*** (23.80)	0.108*** (9.26)	0.00659 (0.55)	0.498*** (57.52)	0.420*** (57.30)
Lottery × win_region	0.0691*** (3.92)	0.197*** (10.39)	0.138*** (4.93)	0.108*** (5.80)	0.0345 (1.43)	-0.0357 (-1.57)	0.101*** (5.80)	0.0526*** (3.51)
.cons	30.44*** (10.81)	28.75*** (9.94)	9.467*** (26.38)	35.82*** (12.18)	7.316* (2.04)	28.42*** (7.50)	31.67*** (11.47)	24.25*** (10.27)
<i>F</i> -test	121.19	21.86	19.30	77.47	61.10	97.40	115.53	110.86
<i>p</i> -value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	211096	211096	211096	211096	211096	211096	211096	211096
R-squared	0.835	0.822	0.884	0.834	0.863	0.862	0.829	0.820

Data Source: Instituto Nacional de Estadística (INE). Our dependent variables are the logarithm of consumption expenditures of each good category. The coefficients presented in this table, *win_region*, *Lottery* and *lottery × win_region*, are as described in Table S2. In this specification, we also include as control variables the age of the head of the household and its square, the marital status of the head of the household and his/her educational level, employment status and whether he/she is retired or not. Moreover, we also include the logarithm of lottery expenditures per region and the regional log-GDP as demographic controls, and both regional and year fixed-effects. We compute robust standard errors. *t*-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of estimates available upon request. The *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.

The final robustness exercise performed in this paper assesses the PIH test using household consumption in real terms. To perform this test, we initially calculate the real interest rate using the formula outlined in the manual by Blanchard et al. (2010), which involves the difference between the nominal interest rate and the inflation rate. After computing this for each type of good, we adjust household consumption for different goods by dividing it by the real interest rate. This allows us to obtain the real household consumption for each good and, consequently, obtain a reliable measure that mitigates the potential impact of inflation or price increases in winning regions, as demonstrated in Ghomi et al. (2023).

Table S9: First stage regression - Total household expenditures in real terms

	(1) Total Expenditures
win_region	-0.191*** (-16.41)
lottery	0.419*** (55.12)
lottery \times win_region	0.061*** (3.98)
_cons	27.18*** (11.11)
Specification	Expenditures in logarithms
<i>F</i> -test for the IV	169.02
<i>p</i> -value	0.0000
N	211096
R-squared	0.825

Data Source: 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* \times *win_region*, as instruments for total household expenditures. The coefficients *win_region*, *lottery* and *lottery* \times *win_region*, are as previously described in Table S2. We present the effect of the lottery income shock on total household expenditures in logarithms; we also did the estimations for total expenditures in levels, but the results were showing a negative impact of the lottery income shock on total expenditures, thus we avoid using expenditures in levels, as the estimates go against our expectations. Both specifications include as control variables the age of the head of the household and its square, the marital status of the head of the household and his/her educational level, employment status and whether he/she is retired or not. Moreover, we also include the logarithm of lottery expenditures per region and the regional log-GDP as demographic controls, and both regional and year fixed-effects. We compute robust standard errors. *t*-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of estimates available upon request.

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

Table S8 displays the outcomes of the PIH test using real household consumption. The results consistently point to a decline in household consumption across various goods, in real terms, including durable and non-durable goods, when comparing winning regions to non-winning ones. However, a closer examination of potential lottery winners -those residing in winning regions who purchased lottery tickets- reveals sustained positive and statistically significant effects, mirroring the findings in Table 5 of the paper. Specifically, households in winning regions that bought lottery tickets experience an increase of 0.101 percentage points in durable goods real consumption and 0.061 percentage points in non-durable goods real consumption compared to the control group. This highlights that households with the potential to win the lottery increase their real consumption, potentially violating the PIH. Further support is derived from an *F*-test rejecting the null hypothesis that residing in winning lottery regions has no impact on household consumption. The PIH violation persists across all goods; thus, when adjusting for inflation, we still get the same results as in the paper.

In table S9 we observe the results for the first-stage regression, indicating a significant impact of the lottery income shock on total household expenditures. With an F -test statistic of 169.02, it is evident that the instrumental variables meet the relevance condition, maintaining their efficacy even when analyzing expenditures in real terms. Hence, it is still a good set of instruments even if we estimate expenditures in real terms.

Table S10: Second stage estimation - Household consumption behavior in real terms

PANEL A	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Food at home	Alcohol	Clothes	House Rent	House Investments	Health	Car Value	Transport
log expenditures	0.697*** (16.81)	0.444*** (4.95)	0.628*** (7.64)	0.937*** (22.03)	0.997*** (17.57)	0.421*** (4.63)	-0.565*** (-8.22)	0.526*** (7.27)
Household controls	✓	✓	✓	✓	✓	✓	✓	✓
Fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Elasticity test	25.99	0.51	0.90	4.68	4.00	0.45	0.01	0.66
p -value	0.0000	0.4863	0.3572	0.0459	0.0628	0.5125	0.9209	0.4271
N	211096	211096	211096	211096	211096	211096	211096	211096
R-squared	0.793	0.841	0.829	0.821	0.823	0.843	0.843	0.825
PANEL B	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Communication	Leisure	Education	Food out Home	Holidays	Savings	Durables	Non-durables
log expenditures	1.037*** (18.18)	0.961*** (11.91)	0.407*** (4.86)	0.807*** (10.55)	0.129 (1.41)	-0.324*** (-3.56)	1.098*** (21.53)	0.888*** (18.55)
Household controls	✓	✓	✓	✓	✓	✓	✓	✓
Fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Elasticity test	6.79	3.07	0.72	4.34	1.06	2.60	17.14	1.76
p -value	0.0092	0.0798	0.3968	0.0372	0.3043	0.1067	0.0000	0.1851
N	211096	211096	211096	211096	211096	211096	211096	211096
R-squared	0.835	0.822	0.862	0.834	0.863	0.862	0.829	0.820

Data Source: 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: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of estimates available upon request. The reported Elasticity test examines the elasticity effect of expenditures towards household consumption, in words, whether the estimates are different from one.

Table S10 shows the results for the two-stage regression analysis of household consumption behavior in real terms. The log expenditures coefficient is positively significant for all but two categories of goods: holidays and savings. This implies that an increase in total household expenditures in real terms corresponds to increased consumption across most goods. Additionally, the t -test is significant for all categories except holidays. Specifically, a 10% increase in total real household expenditures leads to a 10.98% increase in real durable expenditures and an 8.88% increase in real non-durable expenditures. While elasticity tests reject the null hypothesis for durable goods, they do not for non-durable goods. Therefore, contrary to nominal measures, durable goods are elastic to changes in total household expenditures in real terms.

C.3 Analysing consumption in real terms, excluding Cataluña and Madrid

Table S11: Household consumption expenditures in real terms - Excluding Cataluña and Madrid

PANEL A	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Food at home	Alcohol	Clothes	House Rent	House Investments	Health	Car Value	Transport
win_region	-0.122*** (-11.79)	0.0265 (1.13)	-0.0414 (-1.91)	-0.147*** (-13.91)	-0.140*** (-9.44)	0.0130 (0.55)	0.187*** (12.42)	-0.0131 (-0.69)
Lottery	0.170*** (26.66)	1.028*** (77.26)	0.832*** (69.91)	0.0716*** (10.00)	0.313*** (35.74)	1.178*** (87.60)	1.470*** (118.24)	0.698*** (65.17)
Lottery × win_region	0.0496*** (3.62)	-0.325*** (-11.47)	-0.197*** (-7.74)	0.00992 (0.68)	-0.0328 (-1.80)	-0.282*** (-9.87)	-0.364*** (-14.12)	-0.163*** (-7.21)
.cons	28.10*** (12.91)	51.04*** (11.97)	24.73*** (6.33)	42.18*** (17.52)	45.14*** (15.51)	26.52*** (6.25)	8.958* (2.54)	35.86*** (10.19)
F-test	34.65	61.36	52.73	106.68	83.19	53.84	31.93	42.45
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	176505	176505	176505	176505	176505	176505	176505	176505
R-squared	0.777	0.825	0.812	0.804	0.806	0.825	0.828	0.809
PANEL B	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Communication	Leisure	Education	Food out Home	Holidays	Savings	Durables	Non-durables
win_region	-0.173*** (-11.37)	-0.140*** (-6.48)	0.0284 (1.13)	-0.108*** (-5.31)	0.0571* (2.53)	0.0924*** (3.99)	-0.173*** (-13.62)	-0.144*** (-12.18)
Lottery	0.209*** (23.86)	0.947*** (83.11)	1.709*** (104.02)	0.816*** (74.64)	1.344*** (93.38)	0.977*** (68.63)	0.395*** (46.48)	0.393*** (51.12)
Lottery × win_region	-0.00797 (-0.42)	-0.107*** (-4.25)	-0.359*** (-10.20)	-0.181*** (-7.61)	-0.262*** (-8.64)	-0.317*** (-10.67)	0.00766 (0.44)	0.0129 (0.79)
.cons	32.61*** (11.26)	27.19*** (7.35)	8.573*** (20.81)	8.188* (2.27)	-19.87*** (-4.61)	116.7*** (26.74)	44.53*** (15.63)	36.93*** (14.21)
F-test	91.16	21.92	17.14	60.51	38.31	58.01	99.14	86.19
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	176505	176505	176505	176505	176505	176505	176505	176505
R-squared	0.818	0.807	0.872	0.817	0.846	0.848	0.813	0.803

Data Source: Instituto Nacional de Estadística (INE). Our dependent variables are the logarithm of consumption expenditures in real terms 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, at the region level. *t*-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of estimates available upon request.

The *F*-test performs a joint significant test of the lottery income shock variables (*win_region* and *Lottery × win_region*), 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.

In this subsection we repeat the exercise of analyzing consumption in real terms, but excluding the regions of Cataluña and Madrid. In this case, table S11 shows that winning the lottery has a positive and statistically significant impact on the household consumption in real terms of various goods and services. Specifically, households residing in winning regions and participating in the lottery game exhibit, on average, a 0.7% increase in durable goods consumption and a 1.29% increase in non-durable goods consumption in real terms. However, these effects are not statistically significant.

However, when performing the PIH test, we find that the null hypothesis of the test is rejected for all goods, meaning that still when excluding Cataluña and Madrid and measuring consumption in real terms, households do alter consumption not according to Friedman's theory. These results do not differ from the ones presented in nominal terms in Table S2.

Table S12: First stage regression - Total household expenditures in real terms, excluding Cataluña and Madrid

	(1) Total Expenditures
win_region	-0.163*** (-12.17)
lottery	0.425*** (49.69)
lottery × win_region	0.023 (1.26)
_cons	40.87*** (14.09)
Specification	Expenditures in logarithms
<i>F</i> -test for the IV	104.50
<i>p</i> -value	0.0000
N	176505
R-squared	0.808

Data Source: 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 S2. We present the effect of the lottery income shock on total household expenditures in logarithms; we also did the estimations for total expenditures in levels, but the results were showing a negative impact of the lottery income shock on total expenditures, thus we avoid using expenditures in levels, as the estimates go against our expectations. Both specifications include as control variables the age of the head of the household and its square, the marital status of the head of the household and his/her educational level, employment status and whether he/she is retired or not. Moreover, we also include the logarithm of lottery expenditures per region and the regional log-GDP as demographic controls, and both regional and year fixed-effects. We compute robust standard errors. *t*-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of estimates available upon request.

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

Table S12 presents the estimated results for the first-stage regression, with the exclusion of Cataluña and Madrid from the analysis. Here, we observe a positive and statistically significant impact of the lottery income shock on total household expenditures, with an *F*-test score of 104.05. This reaffirms that the relevance condition of the instrumental variables passed, indicating their effectiveness even when expenditures are estimated in real terms and these two regions are excluded from the analysis. Comparing these results with those presented in Table S3 in nominal terms reveals no significant variation in coefficients.

Table S13: Second stage estimation - Household consumption in real terms, excluding Cataluña and Madrid

PANEL A	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Food at home	Alcohol	Clothes	House Rent	House Investments	Health	Car Value	Transport
log expenditures	0.662*** (12.64)	0.695*** (6.14)	0.793*** (7.68)	0.929*** (16.94)	0.995*** (13.76)	0.667*** (5.82)	-0.243** (-2.65)	0.519*** (5.69)
Household controls	✓	✓	✓	✓	✓	✓	✓	✓
Fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Elasticity test	22.50	0.80	0.68	2.92	7.55	0.30	0.86	1.13
p-value	0.0000	0.3697	0.4085	0.0876	0.0060	0.5809	0.3535	0.2885
N	176505	176505	176505	176505	176505	176505	176505	176505
R-squared	0.777	0.825	0.812	0.804	0.806	0.825	0.828	0.809
PANEL B	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Communication	Leisure	Education	Food out Home	Holidays	Savings	Durables	Non-durables
log expenditures	1.146*** (15.66)	1.191*** (11.74)	0.737*** (6.85)	1.186*** (12.35)	0.328** (2.82)	0.246* (2.12)	1.104*** (16.76)	0.900*** (14.71)
Household controls	✓	✓	✓	✓	✓	✓	✓	✓
Fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Elasticity test	5.68	0.61	0.04	1.45	1.27	1.93	8.09	0.97
p-value	0.0171	0.4358	0.8428	0.2280	0.2599	0.1643	0.0044	0.3251
N	176505	176505	176505	176505	176505	176505	176505	176505
R-squared	0.818	0.807	0.847	0.817	0.846	0.848	0.813	0.803

Data Source: 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: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of estimates available upon request. The reported Elasticity test examines the elasticity effect of expenditures towards household consumption, in words, whether the estimates are different from one.

Table S13 shows the second stage estimation of household consumption in real terms, excluding Cataluña and Madrid. The results indicate statistically significant elasticity of total household expenditures across all categories of consumption. This implies that a shock to total household expenditures in real terms leads to an increase in household consumption of the respective goods or services, also in real terms. Specifically focusing on durable and non-durable goods, a 10% increase in total real household expenditures results in an 11% increase in real durable expenditures and a 9% increase in real non-durable expenditures. Similar to the entire sample, the elasticity tests in these cases reject the null hypothesis for durable goods but not for non-durable goods. Thus, once again, contrary to measures in nominal terms, durable goods are elastic to a shock to total household expenditures in real terms. However, the estimated coefficients for these two goods remain consistent with estimates in nominal terms.

C.4 Intergenerational analysis

When households experience a windfall, such as winning the lottery, their thinking about family composition tends to be affected and they increase the number of children, because they have more money or potential savings. In this subsection we analyze the effect of the income shock caused by the Christmas Lottery on family composition; in other words, we want to test whether living in a lottery winning region increases the number of children in the household or not. To test this idea, we estimate Equation (1) and Equation (2) using the forecast of two periods ahead for dependent children at

home as our dependent variable:

$$child_{h,t+2} = \beta_0 + \beta_1 win_region_{h,t-1} + \beta_2 lottery_{h,t-1} + \beta_3 win_{h,t-1} * lottery_{h,t-1} + u_{h,t} \quad (1)$$

where $child_{h,t+2}$ represents the number of children in the household, two periods after the lottery shock. Equation (1) presents the reduced-form estimation, whereas Equation (2) presents the extended regression, where the robustness check is performed by adding household characteristics, plus region and time fixed-effects as control variables to the regression, previously specified in subsection 5.1 of the paper:

$$child_{i,t+2} = \beta_0 + \beta_1 win_region_{h,t-1} + \beta_2 lottery_{h,t-1} + \beta_3 win_{h,t-1} * lottery_{h,t-1} + X'_{h,t} \beta_4 + (gdp_{r,t}, lot_exp_{r,t-1})' \beta_5 + \eta_{h,t} + \tau_t + u_{h,t} \quad (2)$$

Table S14 in the appendix reports the estimated results for Equation (1) and Equation (2). We do not find any changes in family composition resulting from the lottery income shock. Therefore, living in the winning region does not affect the number of children, between 0 and 18 years old, in the household. When performing the robustness check by controlling for individual characteristics, demographic controls and fixed effects, the effect of the income shock is still not relevant for household composition in the periods after the income shock happened.

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

Table S14: Intergenerational analysis

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

Data Source: Instituto Nacional de Estadística (INE). The coefficients presented in this Table, *win_region*, *Lottery* and *lottery × win_region*, are as described in Table S2. In this case, we are interested in how the lottery income shock affects family composition: whether it leads to households having more children, or not. The specification presented in column (1) represents the reduced-form estimation, which only includes the variables presented in this table, whereas the specification in column (2) controls as well for the age of the head of the household and its square, the marital status of the head of the household and his/her educational level, employment status and whether he/she is retired or not. Moreover, in these two specifications, we also include the logarithm of lottery expenditures per region and the regional log-GDP as demographic controls, and regional and year fixed-effects. Thus, given the completeness of this estimation, we are interested in the results of this last one, instead of the reduced-form one. We compute robust standard errors. *t*-statistics in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of estimates available upon request.

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