Getting Wealthier and Healthier? Evidence From the Spanish Christmas Lottery

Marcus Rundström

Lund University

June 1, 2023

Supervisor: Ana Rodríguez-González

Department of Economics

Contents

Getting Wealthier and Healthier? Evidence From the Spanish Christmas Lottery*

Marcus Rundström†

June 1, 2023

Abstract

This paper uses the Spanish Christmas Lottery to provide new evidence about the relationship between wealth and health for adults and infants. Using a difference-in-difference design across provinces and time, I explore how overall health status, hospitalization rates, neonatal health, and alcohol consumption are affected by lottery outcomes. For adults, my results suggest no immediate effect on the probability of reporting good or very good health status. Instead, there seems to be a long-lasting effect in provinces that have won the lottery at least once. Moreover, hospitalization rates within provinces also remain unchanged. I rule out that exogenous lottery wealth has an economically significant impact on infant birth weight. Finally, I find that alcohol consumption is affected on the intensive margin but not at the extensive margin.

Keywords: *Health, Birth weight, Wealth, Lottery, Difference-in-Difference, Causal*

---

* **Acknowledgement:** I am grateful for the comments and suggestions made by friends, and I am especially grateful to my friend Miguel for helping me with the Spanish. I thank Dr. Jan Bietenbeck for suggesting looking more into the Spanish Christmas Lottery. Thanks to the Data Science Lab at LUSEM for their assistance with the latest difference-in-difference literature. Finally, my deepest thanks go to my supervisor, Dr. Ana Rodríguez-González, for her patience and engagement when answering my many questions. Your supervision has been invaluable to me. I take responsibility for any remaining shortcomings.

† Lund University, Department of Economics - Master Essay I

1 Introduction

The positive correlation between health and wealth is well-known in economics but is subject to reverse causation and omitted variable bias (see: Bloom & Canning, 2000; Currie, 2009; Currie & Hyson, 1999; Deaton, 2002). This correlation is observed in every stage of life. For example, wealthier people often have better health and infants to wealthier parents are often born with higher birth weights. Although different approaches have been taken to provide evidence about the causal link between health and wealth, the direction of causality is ambiguous. Research has used macroeconomic fluctuations to study the relationship between wealth and health.[1] Later studies have used quasi-experimental approaches using random lottery windfall.[2] However, similar approaches have found mixed results on how wealth relates to health. Despite a large body of literature, the final answer about the causal link between wealth and health remains unsolved.

In a theoretical framework, Grossman (1972) suggests that individuals have a diminishing return to health and invest different amounts in their health, hence facing different returns in health outcomes due to different characteristics. The framework may serve as an explanation of why individuals are affected differently by wealth shocks. A common denominator for most studies has been the focus on individual outcomes and the neglect of the overall effect on society. One potential reason studies find contradicting results might be that individuals are affected by their surroundings. The social comparison, first mentioned by Festinger (1954), may lead to bias when self-estimating health and happiness.[3]

[1] A sample of these studies are Ruhm (2000) and Adda, Banks, and Von Gaudecker (2009) who found a positive relationship between mortality and income, but unhealthy habits such as smoking, obesity, and diet increased in booms and decreased in recessions. Both studies' results differ from most economists' view of health and wealth but align with other studies on unhealthy habits described by Cawley and Ruhm (2011). A sample of studies that found results align with most economists' view between health and income is Gerdtham and Johannesson (2004), Gerdtham and Johannesson (2005), and Schwandt (2018).

[2] A sample of these studies is Apouey and Clark (2015); Cesarini, Lindqvist, Notowidigdo, and Östling (2017); Cesarini, Lindqvist, Östling, and Wallace (2016); Kim and Koh (2021); Kuhn, Kooreman, Soetevent, and Kapteyn (2011); Lindahl (2005); Lindqvist, Östling, and Cesarini (2020); Östling, Cesarini, and Lindqvist (2020).

[3] Earlier studies suggest adults who move into a wealthier neighborhood from a high-poverty area will improve adult mental health, see Ludwig et al. (2012); Katz, Kling, and Liebman (2001).

In this paper, I use the Spanish Christmas Lottery to provide new evidence of the relationship between wealth and health. More specifically, I investigate the causal effect of living in a province awarded the top three prizes from the Spanish Christmas Lottery on adults' overall health status and infant birth weight. Provinces that win the top three prizes in the Christmas Lottery experience a positive wealth shock of approximately 3 percent, 1 percent, and 0.5 percent of their GDP, respectively. To the best of my knowledge, I am the first to use data from the Spanish Christmas Lottery to evaluate the causal link between wealth and health. One crucial difference in my study compared to previous lottery studies using a difference-in-difference design is that I address the latest concerns regarding the two-way fixed effects estimator, where I apply new robust estimators by De Chaisemartin and d'Haultfoeuille (2020) and De Chaisemartin and d'Haultfoeuille (2022).

This paper is part of a growing literature that exploits lotteries to study the effect of exogenous income shocks on health. For my analysis of adults, this paper is closely related to Lindahl (2005), Apouey and Clark (2015), Östling et al. (2020) and Kim and Koh (2021) that used self-estimated health as one of their primary outcome variables. Lindahl (2005) and Kim and Koh (2021) found a positive effect on overall health, whereas Apouey and Clark (2015) and Östling et al. (2020) found no effect. Thus, research has found different results over a range of environments.

For my analysis of infants' birth weight, my study relates most to Cesarini et al. (2016), who studied the effect of lottery windfall on infant outcomes, including birth weight but found no effect on birth weight or the fraction of infants born with low birth weight. A large body of research has investigated the linkage between parental socioeconomic status and infant birth weight. Currie (2009) reviews the impact of parental socioeconomic status on the number of children with low birth weight and concludes that infants from low-income families weigh less than infants from high-income families. Almond, Hoynes, and Schanzenbach (2011), Hoynes, Miller, and Simon (2015), and Amarante, Manacorda, Miguel, and Vigorito (2016) used payouts from welfare programs to explore the effect of income on birth weight. They concluded that payouts led to a decrease in the number of infants with low birth weight and a positive effect on birth weight.[4] However, as argued

[4] Bozzoli and Quintana-Domeque (2014) used macroeconomic fluctuation to investigate the effect on birth weight and found that only infants to low-educated mothers were affected by the macroeconomic conditions.

by Currie (2009), studies using welfare programs do not only capture an income effect due to changes in relative prices since the welfare programs affect other households. A second critique against studies using welfare programs to study the effect on infant birth weight is that those affected may be more sensitive to parental wealth shocks than others. A large number of studies have shown birth weight to be a determinant of individuals' long-run economic outcomes, which motivates research about how infant birth weight is affected.[5] In a setting similar to the Swedish lottery studies, no effect on self-estimated health status and hospitalization for adults is expected. Next, I expect no effect on infant birth weight in winning provinces since my setting allows for exogenous wealth shocks and not only for low-income families.

I use a difference-in-difference design to investigate the impact on several health outcomes of winning the top three prizes in the Spanish Christmas Lottery. For adults, self-estimated health was the primary variable of interest. Still, since this measure may be subjective, I also analyzed hospitalization rates as a robust measure of one's health within provinces. To investigate potential mechanisms, I first compare self-estimated health and hospitalization rates. As a supplementary analysis, I analyze how alcohol consumption affects winning provinces to explore potential mechanisms behind changes in adult health outcomes. For infants, I investigate changes in infant birth weight and the fractions of infants with low birth weight (birth weight $<$ 2500 grams) in winning provinces. Two key conditions regarding my study's validity were addressed. The first condition is regarding if lottery winnings are random. The second condition is the assumption of parallel trends. To assess these concerns, I provide evidence suggesting the assumptions may hold. To evaluate the concerns regarding time-varying heterogeneous effects, I apply new robust estimators that overcome this problem. The analysis uses data from the Spanish National Health Survey provided by Ministerio de Sanidad. I also use data from the Spanish Hospital Register and the Spanish Birth Register provided by Instituto Nacional de Estadística.

My results suggest that individuals living in a province that has won once are more likely to report good or very good health status by 2.5 percentage points. Next, I do not

[5] For research related to birth weight and long-lasting effects on various characteristics see Bharadwaj, Bietenbeck, Lundborg, and Rooth (2019); Bharadwaj, Lundborg, and Rooth (2017); Black, Devereux, and Salvanes (2007); Currie and Hyson (1999). All studies mentioned conclude that birth weight affects the future outcome of individuals.

find any effect on the probability of reporting good or very good health status the year after a lottery win, and it is valid for the extensive and intensive margin. By comparing changes in health status and hospitalizations, I present three reasons why I do not observe a short-run effect but an effect in the long run. First, a potential reason could be that an increased probability of reporting good or very good health status is a psychological response since no effect in hospitalization is observed, and it takes time to convince oneself about the living situation and feel worthy of the win. Second, due to large-scale lottery windfalls in a neighborhood, the social comparison may imply that individuals do not feel relatively richer. Third, health improvements may not be directly observed since one needs to invest in one's health stock, and in reality, such investments may take time. Next, I rule out the economically significant effect of lottery winnings on infant birth weight, and of the probability of being born with low birth weight in winning provinces. Finally, I provide some evidence that lottery winnings reduce alcohol consumption at the intensive margin but not at the extensive margin.

The remaining structure of the paper is as follows: Section 2 presents the institutional setup of the Spanish Christmas Lottery and the Spanish Health Care System. Section 3 describes the data, and Section 4 turns to the empirical strategy. Section 5 presents the results, including robustness checks, and Section 6 concludes. In the appendix, I provide additional figures and tables.

2 Background

2.1 Spanish Christmas Lottery

The Spanish Christmas Lottery (*Lotería de Navidad*) is the largest in the world and is held every year on December 22nd. The lottery has been organized since 1812 by the National Lottery Organization. Nowadays, the lottery draw attracts a wide television audience and is a popular Spanish tradition. The lottery tickets are expensive, so it is common to participate in the lottery together with others.[6]

[6]The expensive lottery tickets is a result of crimes associated with the lottery in the past. Thus, the government decided in 1862 to raise ticket prices to protect the society. Instead of deterring gambling, it had the opposite effect. Instead of fewer gamblers, the lottery became a widespread social tradition.

Bagues and Esteve-Volart (2016) provides a detailed explanation of the lottery setup. Lottery tickets have five-digit numbers. In 2004, 66,000 numbers played, and between 2005 and 2010, 85,000 numbers played. Each number is split into smaller units known as *series*. A *serie* consists of 10 fractions called *décimos*. Each fraction is divided into smaller shares called *participaciones*. The most common way to participate is to buy a fraction at €20 or a share between €2 and €5. Depending on how many fractions and shares each lottery number has between 1,500 to 15,000 ticket holders. Lottery tickets are sold by outlets around Spain, and lottery numbers are randomly allocated to outlets by a computer. Lottery tickets can also be sold legally or illegally by private persons. The first prize is called El Gordo, and a winning player who buys a fraction receives approximately €300,000, which is ten times higher than the average annual income in Spain. The top three prizes account for approximately half of the total winning amount. There are also thousands of smaller prizes. The top 3 prizes in the Christmas Lottery are sized at approximately 3 percent, 1 percent, and 0.5 percent of the provincial GDP, respectively.[7]

2.1.1 Player characteristics

Bagues and Esteve-Volart (2016) also explain the player characteristics. Spain's Centre for Sociological Research has surveyed 75 percent of the Spanish population aged 18 or older to participate in the lottery. 62 percent of the participants in the Christmas Lottery do not buy tickets otherwise. Only 10 percent of the players in the Christmas Lottery are frequent lottery players.[8] The average person in 2004 was planning to spend €40 to €60, and individuals tend to share tickets with friends, family, or co-workers. One explanation for the popularity of the Spanish Christmas Lottery is that it induces regret in non-participants.

[7] For a sample of studies describing the Spanish Christmas Lottery, see Bagues and Esteve-Volart (2016); Bermejo, Ferreira, Wolfenzon, and Zambrana (2020); Kent and Martínez-Marquina (2021).

[8] At Spain's Centre for Sociological Research web page, surveys regarding the Christmas Lottery are available. These surveys are the same as Bagues and Esteve-Volart (2016) studied (survey numbers 1779, 2274, 2316, 2406, 2587, and 2824).

2.2 The National Healthcare System of Spain

Borkan, Eaton, Novillo-Ortiz, Rivero Corte, and Jadad (2010) provide a detailed description of the Spanish National Healthcare System. Since 1978, Spain has expanded and improved its primary healthcare system. The Spanish constitution requires provinces to provide health protection to citizens and free access to the National Health System. The National Health System of Spain also provides foreign nationals in Spanish territory access to health care, including undocumented immigrants within Spanish territory, citizens of European Union states and Non-European states. Like Great Britain and the Scandinavian countries, Spain has a tax-based system that allows access for the entire population. Each Autonomous Community distributes healthcare services into Health Areas and Basic Health Zones to guarantee services approximately to users.[9] A main difference compared to, for example, Sweden is that public hospitals in Spain do not charge any fees for visiting or for overnight stays. According to OECD (2007) in 2005, Spain spent around 8.2 percent of its GDP on healthcare, which was below the average among the OCED countries (9 percent) and other Scandinavian countries such as Sweden (9.1), Denmark (9.1 percent), and Norway (9.1 percent). According to the Instituto Nacional de Estadística (2023), life expectancy across provinces differs, as shown in Figure A1. The Spanish life expectancy in 1995 was 77.98 years. Today life expectancy in Spain is the highest in Europe at 83.3 years, and the following country is Sweden at 83.1 years.

3 Data

3.1 Lottery data

I use the same lottery data as Bagues and Esteve-Volart (2016). Their data set contains provincial macroeconomic variables, expenditures related to the Spanish Christmas Lottery, and data on the top three prizes awarded to different provinces. The remaining prizes in the lottery are unobserved. Still, it is intuitive that the remaining prizes are proportionally distributed with the lottery expenditures by the nature of randomization

[9] Autonomous community is a larger area compared to provinces. In Spain, the autonomous community is the largest instance after the government. Provinces are the third largest instance, followed by municipalities.

by the lottery.[10] It was impossible to identify winning households. However, I use lottery winnings as exogenous wealth shocks to provinces, and all estimates are interpreted at the provincial level. The question about whether participants buy lottery tickets in their home province or not is examined by Bagues and Esteve-Volart (2016) who conclude that this is indeed the case.[11] In panel A at Table 1, I present descriptive statistics of the lottery data in winning and non-winning provinces.

3.2 Spanish National Health Survey

I obtained survey data from the Ministerio de Sanidad, more specifically from the Spanish National Health Survey which is a repeated cross-sectional survey containing a rich set of questions about one's health (Ministero de Sanidad, 2023). I use surveys for adults (individuals over 16 years old) from 1987, 1993, 1995, 1997, and 2001. I choose not to consider children because they answer the survey with a parent, which may influence their answers. Later survey waves did not allow observing the province of residence, only the autonomous community. Respondents were sampled by multi-stage sampling, using strata, and respondents were proportionally randomly selected within the strata.[12] Since randomization occurred within the strata, OLS will still produce unbiased estimates, and all estimators used are efficient. Respondents were interviewed over the whole year, and the date for the interview can be observed in the data. To ensure respondents did not answer the survey after the year's lottery draw, I used the interview date and dropped observations if they answered the survey on December 22 or later, the same year.

Self-estimated health (from now on referred to as health status) was measured on a scale of 1-5, where 1 is very good, and 5 is very bad.[13] A problem with this measure is

[10] Previous studies has also only observed the top three prices, for instance, see: Bagues and Esteve-Volart (2016); Bermejo, Ferreira, Wolfenzon, and Zambrana (2020); Kent and Martínez-Marquina (2021).

[11] This analysis is done since people may buy tickets together in social networks who live in other areas. If lottery tickets were not purchased in the province of residence, the analysis would fail to capture the causal effects of winning the lottery.

[12] The strata have been constructed by crossing all autonomous communities and dividing them into seven groups according to population size: less than or equal to 2000 inhabitants; from 2001 to 10 000; 10 001-50 000; 50 001-100 000; from 100 001 to 400 000; from 400 001-1 000 000; more than 1 000 000 inhabitants.

[13] A translated version of the question: In the last twelve months, that is, since the beginning of (month, year), would you say that your state of health has been very good, good, fair, bad, or very bad?

Table 1: Descriptive statistics

	Mean		Std. dev .		Min.		Max.		N	
	Win	No win	Win	No win	Win	No win	Win	No win	Win	No win
Panel A: Lottery data										
Expenditure (% GDP)	0.30	0.28	0.10	0.11	0.11	0.08	0.88	0.96	287	1,013
Top prizes (% GDP)	0.46	0.00	1.73	0.00	0.00	20.18	0.00	0.00	287	1,013
Expenditure per capita	0.05	0.04	0.02	0.02	0.01	0.01	0.19	0.20	287	1,013
Top prizes per capita	0.07	0.00	0.29	0.00	0.00	0.00	3.14	0.00	287	1,013
Panel B: Health data										
Good/very good health status	0.68	0.68	0.47	0.47	0	0	1	1	28511	46,719
Self-reported hospitalization	0.07	0.06	0.25	0.24	0	0	1	1	28511	46,719
Hospitalization rate	0.06	0.05	0.04	0.04	0.01	0.00	0.12	0.11	259	1,041
Birth weight (grams)	3,258.57	3276.88	506.16	504.76	300	300	6,570	6,550	1,220,876	2,123,063
Low birth weight (LBW)	0.06	0.05	0.23	0.22	0	0	1	1	1,220,876	2,123,063
Drinking alcohol last two weeks	0.52	0.5	0.5	0.5	0	0	1	1	28,511	46,719
Total units of alcohol each time	2.47	2.39	2.89	2.98	0	0	28	31	28,511	46,719

Notes: Table 1 shows descriptive statistics. Panel A shows descriptive statistics for the lottery data. Expenditure per capita and top prizes per capita are measured in thousands of euros. Panel B shows descriptive statistics for all outcome variables considered.

that health status is subjective. I try to overcome this problem by constructing a dummy variable denoting 1 if very good or good health is reported and 0 otherwise, similar to Kim and Koh (2021), and reduce the subjective interpretation of overall health status. To compare the probability of reporting good or very good health status with another health measure, I used respondents' answers to whether they had been hospitalized in the last twelve months.

To study alcohol consumption in provinces, I first constructed a dummy if the respondent had been drinking alcohol in the last two weeks. In the survey, respondents are also asked to estimate how much they drink when drinking different beverages. It was not possible to observe how much respondents spent on different beverages, but it is possible to observe the number of units of each sort when drinking. To measure alcohol consumption, I sum all units that each respondent reports. A unit of alcohol is defined as a glass of one particular alcoholic beverage, but the volume can vary. In panel B at Table 1, I present descriptive statistics for all outcome variables used in my regressions for winning and non-winning provinces.

A limitation of working with the National Health Survey is that the survey is not conducted yearly or continuously, which creates discontinuous gaps between the survey waves, and questions change from year to year. An ideal setting would have been to have survey data for all years with similar questions each year. I overcome the first issue by treating each survey wave as a period in my analysis.[14] Only questions that are the same or similar were considered. Later survey waves used in my study allowed for additional health measures, such as mental health, but the questions were inconsistent across the survey waves. Hence, the study has been limited to overall health status, self-reported hospitalizations, and alcohol consumption.

3.3 Hospital data

I used hospital stays records between 1986-2011 in Spain from Instituto Nacional de Estadística (2023), which allows for repeated cross-sectional data and contains rich inpatient hospital stays information at the individual level. Individuals are only registered in the database if the patient had at least one overnight stay. The data set includes informa-

[14]For instance, survey wave 1987 is period 1, survey wave 1993 is period 2, survey wave 1995 is period 3, survey wave 1997 is period 4, and survey wave 2001 is period 5.

tion on the province of the hospital, residence of the patient, age, length of stay, and primary diagnosis, among others. The hospital data helps strengthen the results from the National Health Survey. A primary advantage of using the hospital register is that it gives information about all individuals hospitalized during a given period and allows for analysis at the province level. By counting the number of hospitalized in each province and year, I measured the hospitalization rate by the province of residence and year.[15] I dropped all observations hospitalized in December to ensure that last year's lottery prizes only affected observations in a given year.

3.4 Birth weight data

The birth data between 1986 and 2011 are drawn from Instituto Nacional de Estadística and allow for repeated cross-sectional data. The birth data set contains rich information about infants, such as birth weight, sex, and information about parents, among others. I constructed a dummy variable indicating low birth weight (birth weight $<$ 2500 grams). Only children that were born alive were considered. Due to a large amount of missing data, I limit the controls to sex, age of the mother, previous births, province, and year.[16] It would be plausible to control for parents' socioeconomic status as in previous studies on birth weight, but such data was not available.[17] The data set only allowed for observing specific labor sectors, making categorizing parents' socioeconomic status challenging. Finally, I drop all observations born in September, October, November, and December to only focus on ongoing pregnancies and avoid selection into pregnancy changes due to winning the lottery.

4 Empirical strategy

The objective was to study how lottery winnings affect health outcomes for adults and infants in winning provinces of the Spanish Christmas Lottery. I do so by studying health outcomes the year after lottery winnings are observed and by studying health outcomes

[15] Hospitalization rate $= \frac{\text{Number of hospitalized}}{\text{Province population}} \cdot 1000$.

[16] Two dummies for the mother's age were included. The first dummy indicated if the mother was younger than 20 years, and the second indicated if the mother was older than 35.

[17] Controls for mothers age and previous births are motivated by as controls for mother characteristics, similar to Cesarini et al. (2016).

for provinces that have won at least once. As noted in Section 1, the relationship between health and wealth is subject to reverse causality and omitted variable bias. The Spanish Christmas Lottery is ideal for a natural experiment by randomizing lottery winnings conditional on participation, allowing for studying a pure wealth effect on health. Due to the natural experimental setting with randomized lottery winnings, reverse causality could be ruled out and improve the validity of the results.

Due to provincial differences in life expectancy, assuming different health behaviors across provinces is reasonable. This issue was handled by controlling for the province of residence, which captures all unobserved time-invariant heterogeneity across provinces, and otherwise omitted variables such as local habits and cultural norms are now controlled for. In addition, year-fixed effects are included to control for time-varying trends. The macroeconomic control variables are included to isolate the treatment effect and serve as a robustness check for sensitivity.[18] Robust clustered standard errors by province were used to avoid underestimating those (Bertrand, Duflo, & Mullainathan, 2004). This step was crucial to ensure a valid estimation of standard errors.

The effect of lottery winnings is measured on different health outcomes, and the primary variable of interest was health status. Health status is an overall measure of health, which for different individuals may include several health parameters, such as mental and physical health. To compare the results, I used self-reported hospitalization by respondents in the National Health Survey to get an additional measure of one's health. In addition to self-reported hospitalization, I compare the hospitalization rate in the National Health Survey with the Spanish Hospital Register provided by Instituto Nacional de Estadística. The next health measure I consider is infant birth weight and if the infant is born with low birth weight as outcome variables. As a supplementary analysis, health behaviors related to alcohol were studied. Several factors may drive changes in health status, but only health behaviors related to alcohol could be obtained repeatedly over the survey waves.

I took three different approaches to studying the effect of lottery winnings. First, a binary setting is used where treatment is allowed to switch off. In the second setting, treatment was also allowed to switch off, but treatment was continuous. Finally, I per-

[18] As macroeconomic controls, I use provincial measures of GDP per capita, inflation, unemployment rate, and labor force participation following previous literature by Ruhm (2000), Gerdtham and Johannesson (2005), Bharadwaj et al. (2019) and Bozzoli and Quintana-Domeque (2014)

form a staggered difference-in-difference with binary treatment. My first two approaches capture the immediate effects of living in a winning province. The latter captures the effect of living in a province that once won and is necessary since units in a winning province might get long-term effects.[19]

A necessary condition for causal results is that lottery winnings are random. This is addressed as a robustness check, where I provide evidence that the assumption holds. I test it through the specification:

$$Win_{pt} = \mathbf{X}_{pt}\boldsymbol{\beta} + \boldsymbol{\lambda}_p + \boldsymbol{\mu}_t + \gamma_1 Exp_{GDP}pt + \varepsilon_{pt} \tag{1}$$

where I put the Win_{pt} as an outcome variable. Win_{pt} is a dummy indicating if the province p was awarded the top three prizes in year t; $\mathbf{X}_{pt}$ is a matrix of macroeconomic control variables; $\boldsymbol{\lambda}_p$ denotes province fixed effects p; $\boldsymbol{\mu}_t$ denotes time fixed effects; $Exp_{GDP}pt$ is lottery expenditures as percentage of GDP in province p at time t; and ε_{pt} is an error term in province p, at time t. I run these regressions at the province level since lottery winnings are observed at the province level. If one or several of these variables are significant, the assumption of randomized lottery winnings may be violated. As reported in Table A1, none of the coefficients are significant and appear robust and non-sensitive. Thus, the results support the assumption that lottery winnings are truly random and improve the validity of my study. A second necessary condition is the assumption of parallel trends. I provide evidence that this assumption may hold in Section 5.4.2 and discuss potential threats that may violate the assumption.

The most important threat in my difference-in-difference setting is heterogeneity between provinces and years. The literature about the two-way fixed effects estimator and the difference-in-difference methodology has been the subject of recent research.[20] The problem with the current difference-in-difference approach is the lack of a clear pre- and post-period as in the canonical 2x2 setting. Goodman-Bacon (2021), among others, shows that in a difference-in-difference setting, the two-way fixed estimator is the weighted average of all 2x2 differences-in-differences. However, at some point, the two-way fixed effects

[19] I use provincial averages for the regression using the Spanish Hospital Register and do not observe individual outcomes.

[20] A sample of these studies are Baker, Larcker, and Wang (2022), Callaway and Sant'Anna (2021), Goodman-Bacon (2021), Roth, Sant'Anna, Bilinski, and Poe (2023), De Chaisemartin and d'Haultfoeuille (2020), among others.

estimator uses Early-treated individuals as a control group for Late-treated individuals which changes counterfactual trends, which might invalidate the results.[21] The problem is often referred to as the problem of negative weights. To explore potential bias in the estimates, I applied the estimators by De Chaisemartin and d'Haultfoeuille (2020) and De Chaisemartin and d'Haultfoeuille (2022) as a robustness check where I compare the two-way fixed effects estimates with estimates using the robust estimators.

4.1 Binary treatment

First, I use a difference-in-difference setting to consider a binary treatment that could switch off to explore the effect of living in a province that won one of the top three prizes. The setting helps explore the immediate effects at the extensive margin of living in a province winning the top three prizes. The specification can be written as follows:

$$y_{ipt} = \mathbf{X}_{pt}\boldsymbol{\beta} + \mathbf{Z}_{ipt}\boldsymbol{\alpha} + \gamma_1 L_{pt-1} + \boldsymbol{\lambda}_p + \boldsymbol{\mu}_t + \gamma_2 Exp_{pt-1} + \varepsilon_{ipt} \tag{2}$$

where y_{ipt} is an outcome variable for individual i, in province p, at time t; $\mathbf{X}_{pt}$ is a matrix of control variables at the province; $\mathbf{Z}_{ipt}$ is a matrix of individual characteristics; L_{pt-1} is a dummy if the province has been awarded the top three prizes in year $t-1$; $\boldsymbol{\lambda}_p$ denotes province fixed effects; $\boldsymbol{\mu}_t$ denotes time fixed effects; Exp_{pt-1} is lottery expenditures per capita in province p at time $t-1$; and ε_{ipt} is an error term for individual i, in province p, at time t.

4.2 Continuous treatment

To explore the effect of the size of the Christmas lottery prize on the Christmas lottery I used a continuous difference-in-difference setting. Again, treatment turns off when a province is not awarded the top three prizes. This setting is similar to the analysis by Bagues and Esteve-Volart (2016). The setup allows for exploring the immediate effects at the intensive margin of winning the lottery and how the size of lottery windfall relates to several health outcomes. Mathematically the specification can be written as follows:

[21] New estimators have been developed such as Callaway and Sant'Anna (2021), De Chaisemartin and d'Haultfoeuille (2020), De Chaisemartin and d'Haultfoeuille (2022), among others who overcome this problem, both for staggered treatment and non-staggered treatment.

$$y_{ipt} = \mathbf{X}_{pt}\boldsymbol{\beta} + \mathbf{Z}_{ipt}\boldsymbol{\alpha} + \gamma_1 LotteryPrize_{pt-1} + \boldsymbol{\lambda}_p + \boldsymbol{\mu}_t + \gamma_2 Exp_{pt-1} + \varepsilon_{ipt} \tag{3}$$

where y_{ist} is an outcome variable for in province p, at time t; $\mathbf{X}_{pt}$ is a matrix of control variables at the province level; $\mathbf{Z}_{ipt}$ is a matrix of individual characteristics; $LotteryPrize_{pt-1}$ is the lottery windfall per capita in thousands of euros in year $t-1$ in province p; $\boldsymbol{\lambda}_p$ denotes province fixed effects; $\boldsymbol{\mu}_t$ denotes time fixed effects; Exp_{pt-1} is lottery expenditures per capita by thousands of euros in province p at time t; and ε_{ipt} is an error term for individual i, in province p, at time t.

4.3 Staggered binary treatment

In the third specification, a binary staggered difference-in-difference approach was applied across times and provinces in Spain to display the long-term effects of living in a province that has at least won once. Crucially, in this setting, treatment does not switch off. Mathematically, my method can be expressed in the following way:

$$y_{ipt} = \mathbf{X}_{pt}\boldsymbol{\beta} + \mathbf{Z}_{ipt}\boldsymbol{\alpha} + \gamma_1 L_{pt} + \boldsymbol{\lambda}_p + \boldsymbol{\mu}_t + \gamma_2 Exp_{pt-1} + \varepsilon_{ipt} \tag{4}$$

where y_{ipt} is an outcome variable for individual i, in province p, at time t; $\mathbf{X}_{pt}$ is a matrix of control variables at the province level; $\boldsymbol{Z}_{ipt}$ is a matrix of individual characteristics; L_{pt} is a dummy if the province has been awarded the top three prizes in any year before year t; $\boldsymbol{\lambda}_p$ denotes province fixed effects; $\boldsymbol{\mu}_t$ denotes time fixed effects; Exp_{pt-1} is lottery expenditures per capita in province p at time $t-1$; and ε_{ipt} is an error term for individual i, in province p, at time t.

5 Results

5.1 Adult health outcomes

Table 2 shows difference-in-difference estimates of adult health outcomes. The first two columns show estimates of good health status. Panel A and B report non-significant results close to zero, meaning that people are not more likely to report good or very good health status, and this is true for the binary and continuous settings, where treatment can switch off. In Panel C, where treatment does not switch off, I report evidence significant at

Table 2: Adult health outcomes from lottery windfall

	(1)	(2)	(3)	(4)	(5)	(6)
	Good health status		Self-reported hosp.		Hosp. rate	
	TWFE	TWFE	TWFE	TWFE	TWFE	TWFE
Panel A: Winning the top three prizes						
Win	0.00432	0.00614	-0.000109	-0.0000565	-0.000141	0.00000361
	(0.00951)	(0.00993)	(0.00291)	(0.00286)	(0.00131)	(0.00123)
Panel B: Size of winnings						
Lottery Prize	-0.0162	-0.0158	-0.0282***	-0.0274***	0.00118	0.00101
	(0.0169)	(0.0185)	(0.00600)	(0.00604)	(0.00311)	(0.00292)
Panel C: Living in a winning province						
Winning Once	0.0254*	0.0262**	0.000833	0.000801	-0.00325	-0.00339
	(0.0132)	(0.0129)	(0.00540)	(0.00549)	(0.00283)	(0.00278)
Observations	75,230	75,230	75,230	75,230	1,300	1,300
Fixed Effects	YES	YES	YES	YES	YES	YES
Controls	NO	YES	NO	YES	NO	YES
Non-win Mean	0.68		0.06		0.05	

Notes: Table 2 shows difference-in-difference estimates for adult health outcomes from lottery windfall. Panel A shows estimates using equation 2, how health outcomes are affected by living in a province winning the top three prizes in the previous year. Panel B shows estimates using equation 3 where the size of lottery winnings is considered. Panel C reports health outcomes if living in a province that has won at least once using equation 4. In the fixed effects, I include the province, year, and quarter of the interview, gender, and age. GDP per capita, inflation, unemployment rate, and labor force participation serves as additional controls. Standard errors within parentheses are clustered by province. I include GDP per capita, inflation, unemployment rate, and labor force participation as controls. *** p<0.01, ** p<0.05, * p<0.1.

the 10 percent level that people living in a province that has won the top three prizes once are more likely to report good or very good health by 2.5 percentage points. My findings in column 1 seem robust when including provincial macroeconomic control variables in column 2. In panel C, the significance becomes stronger in column 2, and the reason can be that additional controls isolate the treatment effect.

Columns 3 to 6 report results for hospitalization. Column 3 reports estimates for self-reported hospitalization. Panel A reports a non-significant estimate close to zero, whereas, in panel B, I find that an increase in lottery windfall per capita by thousand euros decreases the probability of reporting being hospitalized in the last twelve months by 2.8 percent. In panel C I report non-significant results close to zero. When including additional controls in column 4, the results seem robust due to small changes in the point estimates. All panels in columns 5 and 6 report coefficients close to zero, meaning that the results in columns 3 and 4 in Panel B contrast with the results for the provincial hospitalization rate in columns 5 and 6. In panel B the coefficients for hospitalization rates have the opposite sign but are non-significant. A substantial difference was not expected since the mean hospitalization rates reported in the National Health Survey and by the Hospital Register are similar in Table 1. One explanation for the results in panel B regarding hospitalization is that the two-way fixed effects estimator probably produced biased estimates, which I discuss in Section 5.4.1.

5.2 Infant birth weight

Since general adult health might be less sensitive to short-run fluctuations in wealth, I focus my attention on neonatal health in this Section. Table 3 shows the effect of lottery winnings on infant birth weight. In the first column of panel A, I report a non-significant effect on birth weight if living in a province exposed to the top three prizes. When including provincial macroeconomic variables in column 2, the coefficient appears significant at the 10 percent level. The estimate of the second column implies that infants in a province exposed to the top three prices weigh less by 1.9 grams the year after winning compared to non-winning provinces. In panel B, I report a non-significant increase in birth weight by 39 grams if the lottery windfall increase by thousands of euros. When including additional controls in column 2, my estimate changes down to 11 grams, which can be due to imprecise estimates. A similar story is true for estimates in panel C, the coefficients are non-significant, and there is a change from 12.9 grams to 3.5 grams in the point estimates when including additional controls in column 4.

Columns 3 and 4 report how lottery winnings affect the probability of infants being

Table 3: Impact from lottery windfall on infant birth weight

	(1)	(2)	(3)	(4)
	Birth weight		LBW	
	TWFE	TWFE	TWFE	TWFE
Panel A: Winning the top three prizes				
Win	-1.811	-1.918*	0.00178**	0.000333
	(3.420)	(1.085)	(0.000789)	(0.000541)
Panel B: Size of winnings				
Lottery Prize	38.51	11.03	-0.0110***	-0.00701**
	(24.78)	(13.60)	(0.00360)	(0.00306)
Panel C: Living in a winning province				
Winning Once	11.86	3.457	0.000424	-0.000314
	(7.509)	(5.187)	(0.00116)	(0.00141)
Observations	3,343,643	3,343,643	3,343,643	3,343,643
Fixed Effects	YES	YES	YES	YES
Controls	NO	YES	NO	YES
Non-win mean	3276.88		0.05	

Notes: Table 3 show difference-in-difference estimate regarding infant birth weight. Panel A shows the result for the binary difference-in-difference using equation 2. Panel B reports results from the continuous difference-in-difference where treatment can switch off using equation 3. Panel C show results using equation 2, where treatment does not switch off. Columns 1 and 2 show how birth weight in grams changes. Columns 3 and 4 report changes in the probability of being born with low birth weight. I include the province and year, and gender as fixed effects. GDP per capita, inflation, unemployment rate, and labor force participation serves as additional controls. Standard errors within parentheses are clustered by province. *** p<0.01, ** p<0.05, * p<0.1.

born with low birth weight. In panel A, I report a slight increase in the probability of being born with low birth weight by 0.2 percentage points which is significant at the 5 percent level. When including additional controls in column 4, I get less precise estimates resulting in non-significant results. In panel B I show that an increase of thousand euros slightly decreases the probability of low birth weight for infants by 1.1 percentage points. When including additional controls in column 4, the estimate seems robust. I report non-significant estimates close to zero in columns 3 and 4 of panel C. In Section 5.4.1, I address that the two-way fixed effects estimator probably produces biased estimates and reassesses my findings.

5.3 Supplementary analysis - Alcohol consumption

Table A2 presents difference-in-difference estimates of how alcohol consumption is affected by lottery winnings. Column 1 shows estimates of the probability of drinking alcohol in the last two weeks are affected by lottery winnings. In panel A, I report non-significant effects, but in panel B I present a significant decrease in the probability of drinking alcohol in the last two weeks by 4.8 percentage points if lottery windfall increases by thousand euros. Thus, it seems that the size of windfalls matters. In panel C, I present estimates of the staggered setting. The coefficients in column 1 are non-significant and seem robust when including additional controls in column 2. Hence, in the long run, lottery winnings seem not to affect the probability of drinking alcohol in the last two weeks.

Column 3 presents estimates of how lottery winnings affect total alcohol consumption. Panels A and C report non-significant results, while Panel B reports a significant decrease in total units consumed by 0.47 units if lottery prizes per capita increase by 1000 euros. The estimates in all panels do not seem sensitive when including additional controls in column 4. One explanation for the significant results could be that the two-way fixed effects estimator produces biased estimates, further discussed in Section 5.4.1.

5.4 Robustness checks

5.4.1 Heterogeneity analysis

To evaluate the estimates from the two-way fixed effects estimator in panel C of Tables 2, 3, and A2 and the bias from time-varying effects, I applied the Goodman-Bacon decomposition, which allows for observing which comparison has the most weight in the two-way fixed effects estimator.[22] Weights from the Bacon decomposition are reported in Table A3. In all panels, most weights come from the comparison of Always treated versus Timing groups. This may introduce bias in my estimates if there are dynamic treatment effects, and this can be since a large part of my sample was considered treated in 1987, my first year of observation. Another problematic comparison is the one between Late treatment versus Early treatment as a comparison, and together with the comparison with Always treated versus comparison groups, the weights of my estimates rely mainly on the problematic comparisons. These comparisons might suggest that the results from the two-way fixed effect estimator in specification 4 produce biased estimates. Next, I use robust estimators to evaluate potentially biased estimates of the staggered difference-in-difference setting.

I apply the robust estimators by De Chaisemartin and d'Haultfoeuille (2020) (DIDM) and De Chaisemartin and d'Haultfoeuille (2022) (DIDL) to estimate the average treatment effect on the treated (ATT).[23] Compared to the two-way fixed effects estimator, the DIDM and DIDL solve the problem of negative weights and produce unbiased and consistent estimates under heterogeneous effects. The difference between DIDM and DIDL is that the latter produces unbiased estimates under dynamic effects.[24] However, using a robust estimator forced a drop in all always-treated observations. In the Spanish Christmas Lottery setting and the data available, a large part of my sample is dropped because 17 provinces took part of the top three prizes in 1986 and were considered treated in 1987, which was the first year of observation.[25] Since the lottery has been ongoing since 1812, most provinces should always be considered treated, and with the data available, it required a strong assumption that all observations before 1987 had to be considered untreated. The ideal setting would have been a clear pre-period with no treated provinces before the first lottery draw observed.

If the two-way fixed effects estimator yielded biased results, the point estimates would

[22]The command `bacondecomp` is available in Stata using `ssc install bacondecomp` (Goodman-Bacon, Goldring, & Nichols, 2022).

[23]In Stata I used the command `did_multiplegt`, and can be downloaded with the command `ssc install did_multiplegt`.

[24]Dynamic effects imply that earlier treatment may affect current outcomes. To use the DIDL estimator, one should include the option `robust_dynamic`.

[25]Some of the provinces dropped was Madrid and Valencia who is among the largest provinces in Spain.

differ substantially, and if so, one should use the robust estimators instead. If the robust estimators present similar results as the two-way fixed effects estimator but provide less efficiency due to a large drop in the sample, one should use the two-way fixed effects estimator instead due to better precision. Difference-in-difference estimates using the robust estimators by De Chaisemartin and d'Haultfoeuille (2020) (DIDM) and De Chaisemartin and d'Haultfoeuille (2022) (DIDL) are presented in Table A4. All coefficients appear non-significant, but since the estimator is still unbiased and consistent, making point estimates are comparable to the two-way fixed effects estimator.

I start to consider heterogeneous time-varying effects for the analysis of adults. My estimates concerning good or very good health status in column 1 of Table 2 seem robust using DIDM and DIDL and can be concluded by comparing the point estimates in column 1 of Table 2 and column 1 of Table A4. The magnitude differs by about 2 percentage points when using the robust estimator DIDM, which can be considered small in my context. Hence, I argue that using the two-way fixed effects estimates regarding health status is valid. For self-reported hospitalization, I compare point estimates in the second column of Table 2 with estimates in the second column of Table A4. The significant results by the two-way fixed effects estimator in panel B of Table 2 differ in both size and sign compared to the estimates in Table A4, and hence considered spurious. The point estimates in panels A and C are close to zero in both Tables 2 and A4. For the binary and staggered settings, I can rule out that lottery winnings have an economically significant impact on the probability of being reported being hospitalized, which is also supported by the estimates regarding provincial hospitalization rates in column 5 of Table 2 and in column 3 of Table A4 which are close to zero. I do so based on the upper bound of the confidence interval.

For my analysis of infant birth weight, I compare point estimates in Table 3 with point estimates in Table A4. The coefficients in panel A of Table 3 appear non-robust since DIDL reports a coefficient with the opposite sign. Next, in panels B and C of Table 3, the estimates differ from the robust estimates, but the standard errors are large for both estimators. However, with the estimates given, it seems plausible to rule out that lottery wealth shocks have an economically significant impact on birth weight. Next, I consider infants with low birth weight. The results in panels A and B of Table 3 appear spurious when comparing the point estimates to column 5 in Table A4. The robust point estimates suggest a non-significant effect close to zero. Hence, I rule out that lottery winnings have an economically significant impact on the probability of being born with a low birth weight.

For my analysis of alcohol consumption, I compare point estimates in Table A2 and Table A4. In the first column of Table A2, my estimates differ from the point estimates in column 6 at Table A4, and the two-way fixed effects estimates are considered biased. Hence, I interpret the estimates in column 6 at Table A4 instead. The estimates in Table A4 are imprecise but not close to zero. Hence, I can only report non-significant results but not rule out the economic significance. Finally, for the total units of alcohol consumed, I compare point estimates in columns 3 and 4 at Table A2 and column 7 at Table A4. I conclude that the point estimates in panels A and C are biased in Table A2 since the two-way fixed effects estimator produces estimates close to zero while DIDM and DIDL report estimates further away from zero. The coefficient in panel B seems robust when comparing it to the point estimate by DIDL. Hence, exogenous lottery wealth shocks seem to affect alcohol consumption at the intensive margin.

5.4.2 Falsification tests

The results rely on two assumptions. First, lottery winnings are random, which I discuss in Section 4 and provide evidence that supports this assumption.

The second key identifying assumption is the parallel trends assumption, which implies that all provinces should follow the same trend in the absence of treatment. Compared to the textbook difference-in-difference methodology, my data does not have a pre- and post-period because the lottery has been running yearly since 1812. Data stretching back to 1812 was unavailable; thus, the Spanish Christmas Lottery setting threatens the parallel trends assumption. To check whether this assumption may hold, I follow suggestions by Roth et al. (2023) and perform placebo regressions to provide evidence of the *no anticipatory effects*, which implies that outcomes in period t are unaffected by future lottery winnings.[26] Hence, treatment is allowed to switch off. I perform two regressions:

$$y_{ipt} = \mathbf{X}_{pt}\boldsymbol{\beta} + \mathbf{Z}_{ipt}\boldsymbol{\alpha} + \gamma_1 PlaceboWin_{pt+j} + \boldsymbol{\lambda}_p + \boldsymbol{\mu}_t + \gamma_2 Exp_{pt+j} + \varepsilon_{ipt} \quad (5)$$

[26]These placebo regressions do not prove the assumption of parallel trends since one can not prove the assumption to be satisfied. The researcher can only provide evidence that the assumption might hold.

$$y_{ipt} = \mathbf{X}_{pt}\boldsymbol{\beta} + \mathbf{Z}_{ipt}\boldsymbol{\alpha} + \gamma_1 PlaceboLotteryPrizes_{pt+j} + \boldsymbol{\lambda}_p + \boldsymbol{\mu}_t + \gamma_2 Exp_{pt+j} + \varepsilon_{ipt} \quad (6)$$

where $j \in \{0, 1\}$; y_{ipt} is an outcome variable for individual i in province p at time t; $\mathbf{X}_{pt}$ is a matrix of provincial control variables; $\mathbf{Z}_{ipt}$ is a matrix of individual characteristics; $PlaceboWin_{pt+j}$ is a binary placebo treatment and $PlaceboLotteryPrizes_{pt+j}$ is a continuous placebo treatment in province p at time $t + j$.; $\boldsymbol{\lambda}_p$ denotes province fixed effects; $\boldsymbol{\mu}_t$ denotes time fixed effects; Exp is the lottery expenditures per capita in thousands of euros in province p at time $t + j$; ε_{ipt} is an error term for individual i in province p in year t.[27]

Previous concerns regarding biased estimates using the two-way fixed effects estimator, I perform the same regressions using DIDL. I present placebo regressions using the two-way fixed effects estimator with one lead in Table A5, where I report three significant placebo estimates. However, the health status and self-reported hospitalization estimates seem biased, and the robust estimator reports non-significant results close to zero in Table A6. The third significant placebo estimate is alcohol consumption last two weeks, which differs using the two-way fixed effects estimator and DIDL. In column 6 in Table A6, the point estimate is 0.0159, lower than 0.0385, produced by the two-way fixed effects estimator. The difference in magnitude in this setting is considered small.

Next, I present placebo estimates using the two-way fixed effects estimator for two leads in Table A7 and apply the robust estimator in Table A8. No coefficients appeared statistically significant, and almost all estimates speak for placebo effects close to zero in both tables. With only one robust significant coefficient in the t setting and coefficients close to zero for the remaining coefficients, I argue that the placebo regressions still speak for the assumption of parallel trends to hold, which supports the validity of my study.

As a final robustness check, I consider compositional changes in the long run. In this setting, the concern implies that individuals move to provinces that won. I choose not to address this concern for specifications 3 and 2 because individuals will not likely move to other provinces with such short notice. In the appendix, I provide a specification of the regressions. Table A9 does not show any sign that compositional changes are a problem in the long run.

[27]Note: when $j = 0$ all respondents have already answered the survey before the lottery draw, which should be considered as one lead and when $j = 1$ it is to be considered two leads.

5.5 Discussion

So far, I have analyzed several health outcomes and how these were affected by living in a winning province by first using the two-way fixed effects estimator and then comparing the point estimates to measures by robust estimators. Using the estimates from column 1 of Table 2, I do not find that lottery winnings affect the probability of reporting good or very good health status the year after a win, in line with my expectations. Next, I show that individuals living in an ever-winning province are 2.5 percent more likely to report good or very good health status, and the results appear robust. The result is significant at the 10 and 5 percent levels when additional controls are included in column 2 and is the opposite of what I expected. The effect is similar to that found by Lindahl (2005), who concluded an increase of 3 percentage points in improved health status in the long term. I see it as a strength that my estimates are similar to a study that observed individual winnings. My results contrast with the findings of Kim and Koh (2021), who found a positive effect on self-estimated health one year after an observed lottery win. Moreover, my results also contrast to Apouey and Clark (2015) and Östling et al. (2020), who found no long-term effects from lottery wealth on overall health status. After applying the robust estimators, I also rule out that exogenous lottery wealth economically impacts hospitalization in winning provinces. One explanation for the absence of changes is that overnight hospital stays in Spain are free. Hospitalization is also an extreme health measure and captures variations at other margins compared to self-reported health.

A potential explanation for this finding could be a psychological rather than a physical improvement. According to Deaton (2002), unexpected wealth shocks can reduce stress and improve health. The interpretation is in line with previous studies related to mental health that found positive effects on mental health (see: Apouey & Clark, 2015; Lindahl, 2005; Lindqvist et al., 2020). This explanation is also plausible when considering the health care system in Spain, which is similar to the ones in Great Britain and Sweden. In this context, similar to other countries in Western Europe, there is a stronger safety net compared to Singapore, which was studied by Kim and Koh (2021), who found positive effects on health outcomes up to one year after lottery winnings. Hence, social differences seem vital for one's health status, which was also stressed by Kim and Koh (2021). The explanation of a psychological response is consistent with Kuhn et al. (2011) that argues that after winning, the winners felt that they did not deserve the win, but years afterward, they convinced themselves that they did deserve the win. Thus, the explanation by Kuhn et al. (2011) also explains why I do not observe an effect in the short run.

Another possible explanation for an increased probability of reporting good or very good health status is the social comparison theory by Festinger (1954). Due to large-scale lottery windfall, the social comparison might affect health status in winning provinces differently than if one were the only winner in a town. Thus, one may not feel richer in relative terms. A final explanation can be used through the Grossman model, where individuals invest in their health stock (Grossman, 1972). In reality, returns from such investments may take time to observe.

For the analysis of birth weight, I rule out that exogenous lottery wealth has an economically significant impact on birth weight and the fraction of low birth weight, in line with my expectations. The findings are similar to Cesarini et al. (2016) that lottery winnings do not affect infant birth weight which stands in contrast to findings by Almond et al. (2011), Hoynes et al. (2015) and Amarante et al. (2016). This finding speaks to the fact that lottery wealth differs from other wealth types and does not directly affect socioeconomic status since it contradicts findings related to parents' socioeconomic status and the birth weight of their children. A key difference in my study compared to Almond et al. (2011), Hoynes et al. (2015) and Amarante et al. (2016) is that I am first able to capture a pure wealth effect and second, I analyze the outcome for all infants regardless of family background. Their studies capture changes in relative prices due to financing behind the programs, and welfare programs capture individuals who are already poor. Moreover, my analysis would have benefited if I controlled for parental occupation since the previous literature addressed it as a potential determinant for infant birth weight.

For alcohol consumption in the last two weeks, I report non-significant coefficients but cannot rule out economically significant effects from exogenous lottery wealth on the extensive margin. My regressions regarding total alcohol consumption produce biased estimates except for the continuous analysis that speaks for an effect at the intensive margin. I conclude that if lottery prizes increase by thousands of euros per capita, individuals in the winning provinces consume 0.47 units less than individuals in non-winning provinces. Less drinking may be a driving factor that people living in a province that has won at least once are likelier to report good or very good health. However, the estimate is small, and one should be careful since much of my evidence suggests that weekly or total alcohol consumption is not affected by lottery winnings.

6 Conclusion

The relationship between wealth and health is associated with every stage in life. However, despite a large body of research, the causal nature of this relationship remains unsolved. Earlier studies using exogenous wealth shocks to explore the causal relationship between wealth and health for adults and children have found mixed results. In this study, I use the setting of the Spanish Christmas Lottery to investigate the health outcomes of living in a winning province for adults and infants. I find no effect on the probability of reporting good or very good health status the year after a win for both the extensive och intensive margin. In the long run, winning provinces have an increased probability of reporting good or very good health status by 2.5 percentage points. I also conclude that lottery winnings do not affect neonatal health, such as birth weight and the probability of being born with a low birth weight. Finally, for alcohol consumption, I did not find any effect on the extensive margin, but I provide some evidence that total alcohol consumption decreases at the intensive margin.

To address the study's internal validity concerns, I performed robustness tests, such as placebo regressions, and provided evidence that lottery winning is random. One of the essential features of this paper is the use of robust estimators to investigate potentially biased estimates produced by the two-way fixed effects estimator. As with many natural experiments, my study might suffers from low external validity. A common concern for lottery studies that may threaten external validity is that lottery players may have specific characteristics. Still, Spaniards' social tradition of participating in the Christmas Lottery may improve external validity. Another common concern that threatens the external validity is that lottery wealth differs from other wealth shocks. Shocks like entrepreneurial income or inheritance are often not entirely exogenous, and lottery wealth may be spent frivolously. This concern is assessed by Cesarini et al. (2016) and Cesarini et al. (2017), who conclude that this is not the case. Instead, lottery winners invest money in safe assets. The estimated effect on Spaniards may differ from other countries, especially since Spain has one of the highest life expectancies worldwide, and one may find different effects depending on the country. Finally, a positive wealth shock does not necessarily imply the same magnitudes as a negative one. Negative wealth shocks such as unemployment and stock market crashes imply other societal effects. Hence it may be difficult to compare results from studies investigating negative wealth shocks, such as Schwandt (2018), with studies like mine investigating positive wealth shocks.

This study has several limitations. First, I could only observe lottery windfall at the province level in my empirical strategy. Investigating individual outcomes like in previous lottery studies would have been preferable. Instead, interpretation was made at the province level. Second, with more variables, for example, individual socioeconomic status, mental health, and more health behavior measures, I may have been able to outline potential mechanisms. Instead, I could only suggest potential mechanisms that may drive the results. In addition, the scarcity of control variables can be a potential source of omitted variable bias. Third, due to the data available, many provinces were considered treated in the first year of observation, making a large part of my sample always treated, causing a large drop in observations when applying the robust estimators. Finally, due to the data limitations, I was not able to provide more evidence that the parallel trend assumption is satisfied. I do provide evidence that speaks for the assumption, but more evidence, such as an event study with a clear pre-period, would have further strengthened the internal validity.

In the future, research could adopt robust estimators to ensure robust results when using a difference-in-difference design. Next, research could use large-scale lotteries similar to the Spanish Christmas Lottery, which benefits external validity. The results might differ if many in one's surroundings receive lottery windfall compared to lotteries with only single winners. The concern about how the effects will stand when implementing a policy on a larger scale has been discussed by economists for a long time. If research used large-scale lottery windfalls, the research could get closer to a scenario where a large part of the population is affected by the treatment. Finally, the link between wealth and health remains unsolved, and more evidence is needed from different settings. Future research can build on this paper to further prove how wealth and health are related.

References

Adda, J., Banks, J., & Von Gaudecker, H.-M. (2009). The impact of income shocks on health: evidence from cohort data. *Journal of the European Economic Association*, *7*(6), 1361–1399.

Almond, D., Hoynes, H. W., & Schanzenbach, D. W. (2011). Inside the war on poverty: The impact of food stamps on birth outcomes. *Review of Economics and Statistics*, *93*(2), 387–403.

Amarante, V., Manacorda, M., Miguel, E., & Vigorito, A. (2016). Do cash transfers improve birth outcomes? evidence from matched vital statistics, and program and social security data. *American Economic Journal: Economic Policy*, *8*(2), 1–43.

Apouey, B., & Clark, A. E. (2015). Winning big but feeling no better? the effect of lottery prizes on physical and mental health. *Health Economics*, *24*(5), 516–538.

Bagues, M., & Esteve-Volart, B. (2016). Politicians' luck of the draw: Evidence from the spanish christmas lottery. *Journal of Political Economy*, *124*(5), 1269–1294.

Baker, A. C., Larcker, D. F., & Wang, C. C. (2022). How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics*, *144*(2), 370–395.

Bermejo, V. J., Ferreira, M. A., Wolfenzon, D., & Zambrana, R. (2020). Entrepreneurship and regional windfall gains: Evidence from the spanish christmas lottery. *CEPR Discussion Paper No. DP14638*.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, *119*(1), 249–275.

Bharadwaj, P., Bietenbeck, J., Lundborg, P., & Rooth, D.-O. (2019). Birth weight and vulnerability to a macroeconomic crisis. *Journal of Health Economics*, *66*, 136–144.

Bharadwaj, P., Lundborg, P., & Rooth, D.-O. (2017). Birth weight in the long run. *Journal of Human Resources*, *53*(1), 189–231.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2007). From the cradle to the labor market? the effect of birth weight on adult outcomes. *The Quarterly Journal of Economics*, *122*(1), 409–439.

Bloom, D. E., & Canning, D. (2000). The health and wealth of nations. *Science*, *287*(5456), 1207–1209.

Borkan, J., Eaton, C. B., Novillo-Ortiz, D., Rivero Corte, P., & Jadad, A. R. (2010). Renewing primary care: lessons learned from the spanish health care system. *Health Affairs*, *29*(8), 1432–1441.

Bozzoli, C., & Quintana-Domeque, C. (2014). The weight of the crisis: evidence from newborns in argentina. *Review of Economics and Statistics*, *96*(3), 550–562.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230.

Cawley, J., & Ruhm, C. J. (2011). The economics of risky health behaviors. In M. V. Pauly, T. G. Mcguire, & P. P. Barros (Eds.), *Handbook of health economics* (Vol. 2, pp. 95–199). Elsevier.

Cesarini, D., Lindqvist, E., Notowidigdo, M. J., & Östling, R. (2017). The effect of wealth on individual and household labor supply: evidence from swedish lotteries. *American Economic Review*, *107*(12), 3917–3946.

Cesarini, D., Lindqvist, E., Östling, R., & Wallace, B. (2016). Wealth, health, and child development: Evidence from administrative data on swedish lottery players. *The Quarterly Journal of Economics*, *131*(2), 687–738.

Currie, J. (2009). Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development. *Journal of Economic Literature*, *47*(1), 87–122.

Currie, J., & Hyson, R. (1999). Is the impact of health shocks cushioned by socioeconomic status? the case of low birthweight. *American Economic Review*, *89*(2), 245–250.

Deaton, A. (2002). Policy implications of the gradient of health and wealth. *Health Affairs*, *21*(2), 13–30.

De Chaisemartin, C., & d'Haultfoeuille, X. (2022). Difference-in-differences estimators of intertemporal treatment effects. *National Bureau of Economic Research*.

De Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, *110*(9), 2964–2996.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*(2), 117–140.

Gerdtham, U.-G., & Johannesson, M. (2004). Absolute income, relative income, income inequality, and mortality. *Journal of Human Resources*, *39*(1), 228–247.

Gerdtham, U.-G., & Johannesson, M. (2005). Business cycles and mortality: results from swedish microdata. *Social science & medicine*, *60*(1), 205–218.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277.

Goodman-Bacon, A., Goldring, T., & Nichols, A. (2022). Bacondecomp: Stata module to perform a bacon decomposition of difference-in-differences estimation. *Statistical Software Components*.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy*, *80*(2), 223–255.

Hoynes, H., Miller, D., & Simon, D. (2015). Income, the earned income tax credit, and infant health. *American Economic Journal: Economic Policy*, *7*(1), 172–211.

Instituto Nacional de Estadística. (2023). *Database for demography and population.* `https://www.ine.es/en/index.htm`. (Accessed: 2023-03-30)

Katz, L. F., Kling, J. R., & Liebman, J. B. (2001). Moving to opportunity in boston: Early results of a randomized mobility experiment. *The Quarterly Journal of Economics*, *116*(2), 607–654.

Kent, C., & Martínez-Marquina, A. (2021). When a town wins the lottery: Evidence from spain. *Available at SSRN 3960784*.

Kim, S., & Koh, K. (2021). The effects of income on health: Evidence from lottery wins in singapore. *Journal of Health Economics*, *76*, 102414.

Kuhn, P., Kooreman, P., Soetevent, A., & Kapteyn, A. (2011). The effects of lottery prizes on winners and their neighbors: Evidence from the dutch postcode lottery. *American Economic Review*, *101*(5), 2226–2247.

Lindahl, M. (2005). Estimating the effect of income on health and mortality using lottery prizes as an exogenous source of variation in income. *Journal of Human Resources*, *40*(1), 144–168.

Lindqvist, E., Östling, R., & Cesarini, D. (2020). Long-run effects of lottery wealth on psychological well-being. *The Review of Economic Studies*, *87*(6), 2703–2726.

Ludwig, J., Duncan, G. J., Gennetian, L. A., Katz, L. F., Kessler, R. C., Kling, J. R., & Sanbonmatsu, L. (2012). Neighborhood effects on the long-term well-being of low-income adults. *Science*, *337*(6101), 1505–1510.

Ministero de Sanidad. (2023). *Data base of the spanish national health survey 1987, 1993, 1995, 1997, and 2001.* `https://www.sanidad.gob.es/home.htm`. (Accessed:

2023-03-30)

OECD. (2007). *Health at a glance 2007*. OECD Publishing Paris.

Roth, J., Sant'Anna, P. H., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *Journal of Econometrics*.

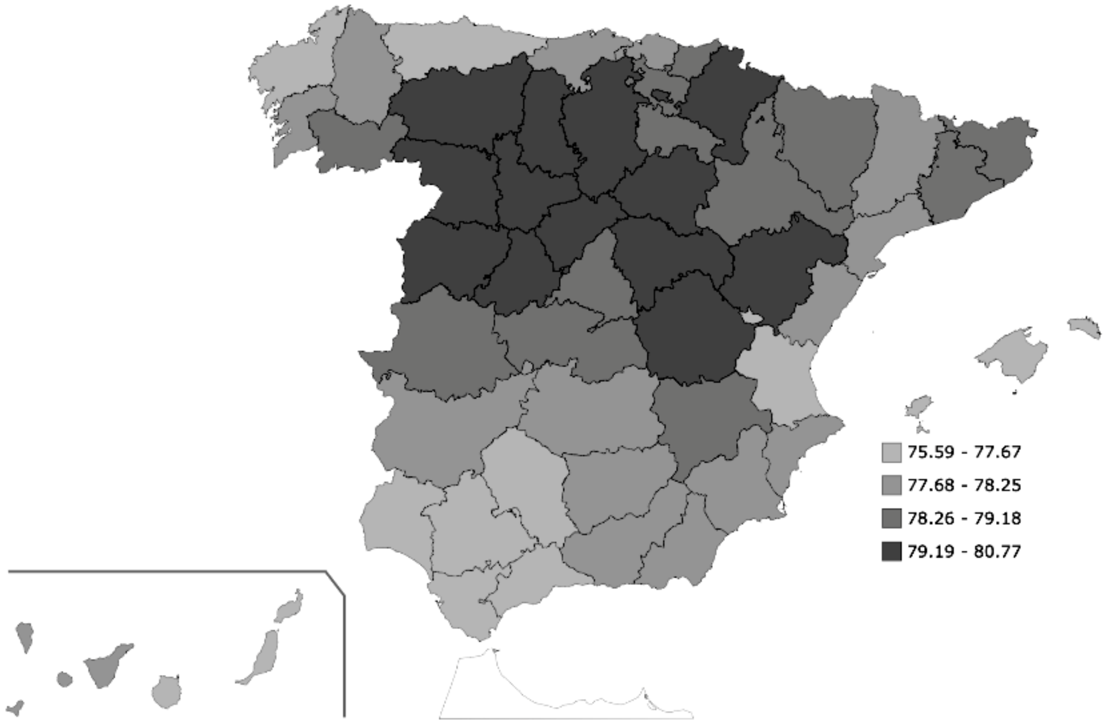
Ruhm, C. J. (2000). Are recessions good for your health? *The Quarterly Journal of Economics*, *115*(2), 617–650.

Schwandt, H. (2018). Wealth shocks and health outcomes: Evidence from stock market fluctuations. *American Economic Journal: Applied Economics*, *10*(4), 349–77.

Östling, R., Cesarini, D., & Lindqvist, E. (2020). Association between lottery prize size and self-reported health habits in swedish lottery players. *JAMA Network Open*, *3*(3), e1919713–e1919713.

Appendix

Figure A1: Life expectancy in Spain 1995



Note: Life expectancy in Spain is plotted by provinces using data from 1995 obtained from Instituto Nacional de Estadística. All numbers are given in years.

Table A1: Validity check - Are lotteries truly random?

	Exposed to the top three prizes		
	(1)	(2)	(3)
Lottery Expenditure (% of GDP)	-0.0583	-0.0000264	-0.171
	(0.183)	(0.192)	(0.329)
GDP		0.00000170	0.00000208
		(0.00000129)	(0.00000132)
Inflation		0.0106	0.00115
		(0.0183)	(0.0258)
Unemployment Rate		-0.000796	0.00242
		(0.00393)	(0.00544)
Labor Force Participation		0.000771	-0.00150
		(0.00377)	(0.00518)
Observations	1,300	1,300	1,300
Fixed Effects	YES	YES	YES
Controls	NO	YES	YES
Weighted	NO	NO	YES

Notes: Table A1 reports results if winning the top three prizes is random using equation 1. A dummy was created to test if prizes are truly random, indicating if a province was exposed to the top three prizes. The first column shows the amount of total expenditure do impact the probability of winning; the second column includes GDP per capita, inflation, unemployment rate, and labor force participation as additional controls; the third column includes population weights to test the sensitivity of the results in columns 1 and 2. Clustered standard errors within parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A2: The effect of lottery windfall on alcohol consumption

	(1)	(2)	(3)	(4)
	Alcohol last two weeks		Units of alcohol	
	TWFE	TWFE	TWFE	TWFE
Panel A: Winning the top three prizes				
Win	-0.0126	-0.0138	0.0174	-0.00449
	(0.0121)	(0.0124)	(0.103)	(0.102)
Panel B: Size of winnings				
Lottery Prize	-0.0481**	-0.0423**	-0.465***	-0.418**
	(0.0191)	(0.0189)	(0.167)	(0.172)
Panel C: Living in a winning province				
Winning Once	-0.0211	-0.0199	-0.0455	-0.0405
	(0.0211)	(0.0218)	(0.135)	(0.129)
Observations	75,230	75,230	75,230	75,230
Fixed Effects	YES	YES	YES	YES
Controls	NO	YES	NO	YES
Non-win mean	0.50		2.39	

Notes: Table A2 show the difference-in-difference results regarding alcohol consumption. Panel A shows the result for the binary difference-in-difference using equation 2. Panel B reports results from the continuous difference-in-difference using equation 3 where treatment can switch off. Panel C show results using equation 4, where treatment does not switch off. Columns 1 and 2 show self-reported consumption of alcohol in the last two weeks. Columns 3 and 4 show results about the total units of alcohol consumed each time drinking. Gender, age, and quarter are included in all regressions. GDP per capita, inflation, unemployment rate, and labor force participation serves as additional controls. Standard errors within parentheses are clustered by province. *** p<0.01, ** p<0.05, * p<0.1.

Table A3: Bacon decomposition

Outcome variable	Comparison	Weight	Avg. DiD estimate
Panel A: Good Health Status	Early T vs Late T	0.155	0.001
	Late T vs Early T	0.161	0.005
	Always T vs Timing	0.505	0.020
	Never T vs Timing	0.178	0.074
Panel B: Self-reported hospitalization	Early T vs Late T	0.155	0.001
	Late T vs Early T	0.161	0.000
	Always T vs Timing	0.505	0.006
	Never T vs Timing	0.178	-0.009
Panel C: Hospitalization rate	Early T vs Late T	0.145	0.000
	Late T vs Early T	0.746	0.014
	Always T vs Timing	-	-
	Never T vs Timing	0.108	0.001
Panel D: Birth weight	Early T vs Late T	0.141	0.713
	Late T vs Early T	0.256	4.012
	Always T vs Timing	0.518	13.323
	Never T vs Timing	0.087	9.419
Panel E: LBW	Early T vs Late T	0.143	-0.000
	Late T vs Early T	0.253	-0.001
	Always T vs Timing	0.517	-0.002
	Never T vs Timing	0.087	-0.002
Panel F: Alcohol last two weeks	Early T vs Late T	0.155	0.008
	Late T vs Early T	0.161	0.037
	Always T vs Timing	0.506	0.005
	Never T vs Timing	0.178	-0.038
Panel G: Units of alcohol	Early T vs Late T	0.176	0.129
	Late T vs Early T	0.1450	-0.281
	Always T vs Timing	0.505	0.188
	Never T vs Timing	0.178	-0.566

Notes: Table A3 provides weights for the different types of comparisons and the weighted coefficient for each group that my estimates in equation 4 rely on, using the `bacondecomp` command.

Table A4: Difference-in-difference estimates using robust estimators

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Good health status	Self-reported hosp.	Hosp. rate	Birth weight	LBW	Alcohol	Units of alcohol
Panel A: Binary							
Win	-0.0180	-0.00194	0.00261	3.737	-0.000734	-0.0533	-0.411
	(0.0338)	(0.103)	(0.00147)	(4.025)	(0.00164)	(0.0640)	(0.270)
Observations	29959	29959	303	681180	681180	27142	27142
Panel B: Continuous							
Lottery Prize	-0.0307	0.0137	0.00261	3.737	-0.000734	0.00727	-0.505
	(0.0288)	(0.0115)	(0.00232)	(2.882)	(0.00235)	(0.0596)	(0.305)
Observations	17792	17792	303	535055	535055	17792	17792
Panel C: Living in a winning province							
Winning Once	0.0451	-0.0155	0.00261	3.737	-0.000734	-0.00969	-0.282
	(0.0270)	(0.00130)	(0.00273)	(4.757)	(0.00114)	(0.0435)	(0.299)
Observations	14,144	14,144	303	535,055	535,055	14,144	14,144

Notes: Table A4 shows difference-in-difference estimates using the estimator by De Chaisemartin and d'Haultfoeuille (2020) (DIDM) and De Chaisemartin and d'Haultfoeuille (2022) (DIDL). Panels A and B show results using the DIDL estimator where I consider equation 2 and 3. Panel C show results using the DIDM estimator where I consider equation 4. All regressions control for lottery expenditure per capita. Columns 1, 2, 6, and 7 also control for gender and age. Columns 4 and 5 control for gender, age of the mother, and if the mother has given previous births. Both DIDM and DIDL, excluding already treated units. Standard errors within parentheses are clustered by province. *** p<0.01, ** p<0.05, * p<0.1.

Table A5: Placebo estimates one lead - TWFE

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Good health status	Self-reported hosp.	Hosp. rate	Birth weight	LBW	Alcohol	Units of alcohol
Panel A: Binary							
Placebo Win	-0.00403	0.00568	-0.00136	-1.861	0.00107	0.0385***	0.0604
	(0.00895)	(0.00489)	(0.00113)	(1.943)	(0.000754)	(0.0115)	(0.0814)
Panel B: Continuous							
Placebo Lottery Prize	-0.0598*	0.0518***	0.000384	6.687	0.00258	0.0701	0.318
	(0.0332)	(0.0150)	(0.00246)	(8.579)	(0.00213)	(0.0868)	(0.464)
Observations	75,230	75,230	1,300	3,343,643	3,343,643	75,230	75,230
Controls	YES	YES	YES	YES	YES	YES	YES
Fixed Effects	YES	YES	YES	YES	YES	YES	YES

Notes: Table A5 shows results from the placebo estimates with one lead. Panel A exploits the binary approach using equation 5, and panel B exploits the continuous approach using equation 6. All regressions for control lottery expenditure per capita. Columns 1, 2, 6, and 7 also control for gender and age. Columns 4 and 5 control for gender, age of the mother, and if the mother has given previous births. GDP per capita, inflation, unemployment rate, and labor force participation serves as additional controls. Standard errors within parentheses are clustered by province. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A6: Placebo estimates one lead- Robust estimators

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Good health status	Self-reported hosp.	Hosp. rate	Birth weight	LBW	Alcohol	Units of alcohol
Panel A: Binary							
Placebo Win	0.0166	0.00991	-0.00115	-2.367	0.000700	0.00820	0.0135
	(0.0233)	(0.00706)	(0.00143)	(3.396)	(0.00129)	(0.0241)	(0.189)
Observations	29,959	29,959	270	856,871	856,871	29,959	29,959
Panel B: Continuous							
Placebo Lottery Prizes	0.00526	0.00734	-0.00158	-2.252	0.00102	-0.0125	0.0398
	(0.0182)	(0.00935)	(0.00154)	(4.498)	(0.00153)	(0.0172)	(0.177)
Observations	28,159	28,159	253	815,998	815,998	28,159	28,159
Controls	YES	YES	YES	YES	YES	YES	YES
Fixed Effects	YES	YES	YES	YES	YES	YES	YES

Notes: Table A6 show placebo estimates with one lead using the estimator. Panel A and B uses the estimator by De Chaisemartin and d'Haultfoeuille (2022) (DIDM). Panel C uses the estimator by De Chaisemartin and d'Haultfoeuille (2020) (DIDL). All regressions for control lottery expenditure per capita. Columns 1, 2, 6, and 7 also control for gender and age. Columns 4 and 5 control for gender, age of the mother, and if the mother has given previous births. Both DIDM and DIDL, excluding already treated units. GDP per capita, inflation, unemployment rate, and labor force participation serves as controls. Standard errors within parentheses are clustered by province. *** p<0.01, ** p<0.05, * p<0.1.

Table A7: Placebo estimates two leads - TWFE

	(1) Good health status	(2) Self-reported hosp.	(3) Hosp. rate	(4) Birth weight	(5) LBW	(6) Alcohol	(7) Units of alcohol
Panel A: Binary							
Placebo Win	0.00113	0.00131	0.000113	0.246	-0.0000133	-0.00146	-0.0204
	(0.00708)	(0.00388)	(0.00105)	(1.457)	(0.000416)	(0.0168)	(0.106)
Panel B: Continuous							
Placebo Lottery Prize	0.000575	-0.00561	0.000524	10.39	-0.000567	-0.104	-1.108
	(0.0323)	(0.0203)	(0.00153)	(7.279)	(0.00224)	(0.146)	(0.706)
Observations	75,230	75,230	1,300	3,343,643	3,343,643	75,230	75,230
Controls	YES	YES	YES	YES	YES	YES	YES
Fixed Effects	YES	YES	YES	YES	YES	YES	YES

Notes: Table A7 shows results from the placebo estimates using the two-way fixed effects estimator with two leads. Panel A exploits the binary approach using equation 5, and panel B exploits the size of winnings using equation 6. All regressions control for lottery expenditure per capita. Columns 1, 2, 6, and 7 also control for gender and age. Columns 4 and 5 control for gender, age of the mother, and if the mother has given previous births. GDP per capita, inflation, unemployment rate, and labor force participation serves as controls. Standard errors within parentheses are clustered by province. *** p<0.01, ** p<0.05, * p<0.1.

Table A8: Placebo estimates two leads- Robust estimators

	(1) Good health status	(2) Self-reported hosp.	(3) Hosp. rate	(4) Birth weight	(5) LBW	(6) Alcohol	(7) Units of alcohol
Panel A: Binary							
Placebo Win	0.00466	-0.00921	-0.00303	-1.915	0.000913	0.0240	-0.00268
	(0.0192)	(0.00742)	(0.00200)	(5.385)	(0.00144)	(0.0405)	(0.237)
Observations	25,867	25,867	315	769,076	769,076	25,867	25,867
Panel B: Continuous							
Placebo Lottery Prizes	0.00235	-0.00567	-0.00315	2.683	0.000122	-0.0260	-0.124
	(0.0196)	(0.0128)	(0.00174)	(4.563)	(0.00188)	(0.0378)	(0.377)
Observations	20,219	20,219	312	670,220	670,220	20,219	20,219
Controls	YES	YES	YES	YES	YES	YES	YES
Fixed Effects	YES	YES	YES	YES	YES	YES	YES

Notes: Table A8 show placebo estimates with two leads using the estimator by De Chaisemartin and d'Haultfoeuille (2022) in panel A and B. Panel C uses the estimator by De Chaisemartin and d'Haultfoeuille (2020). All regressions control for lottery expenditure per capita. Columns 1, 2, 6, and 7 also control for gender and age. Columns 4 and 5 control for gender, age of the mother, and if the mother has given previous births. Both DIDM and DIDL, excluding already treated units. GDP per capita, inflation, unemployment rate, and labor force participation serves as controls. Standard errors within parentheses are clustered by province. *** $p<0.01$, ** $p<0.05$, *

Compositional changes

I run regressions using pre-determined variables as the outcome variable. Next, the population is investigated to see if lottery windfall affects the population count. The specification can be expressed as follows:

$$Characteristics = \gamma_1 L_{pt} + \gamma_2 Exp_{pt-1} + \boldsymbol{\lambda}_p + \boldsymbol{\mu}_t + \varepsilon \tag{7}$$

where $Characteristics$ is an individual or provincial characteristic; L_{pt} is a dummy indicating if a province has won at least once before period t; Exp_{pt-1} is the lottery expenditure per capita in province p at period $t-1$; $\boldsymbol{\lambda}_p$ denotes province fixed effects; $\boldsymbol{\mu}_t$ denotes province fixed effects; and ε is the individual or provincial error term, depending of characteristics.

Table A9: Robustness check - Compositional changes

	(1)	(2)	(3)	(4)	(5)	(6)
	National Health Survey		Birth data			Lottery data
	Gender	Age	Gender	Young	Old	Population
Winning Once	-0.0465	-0.0727	0.00128	-0.0108	0.00513	-17.05
	(0.116)	(0.312)	(0.00649)	(0.00725)	(0.00487)	(47.18)
Observations	75,230	75,230	3,343,643	3,343,930	3,343,930	3,343,930
Fixed Effects	YES	YES	YES	YES	YES	YES
Controls	NO	NO	NO	NO	NO	NO

Notes: Table A9 shows results using equation 7 to investigate compositional changes with the two-way fixed effects estimator. Standard errors within parentheses are clustered by province. *** p<0.01, ** p<0.05, * p<0.1.