# Winning the Tax Lottery: Evidence from a Superdraw on Christmas Eve

Panayiotis Nicolaides *

December 13, 2023

**Abstract**

This paper studies a tax lottery in Greece and documents an increase in VAT revenue. The lottery incentivises the use of electronic payments over cash to fight tax evasion by allocating €1 million in prizes every month. Tickets are awarded automatically when individuals complete retail transactions by electronic means. I exploit a superdraw at the start of the lottery in Christmas Eve 2017; participation was unanticipated and individuals could not influence their winning chances. I estimate that regional VAT revenue increased by 0.01% per additional winner (or by €2,700 compared to a €1,000 winning prize). This effect can be explained through winners, who increased their electronic consumption by 14%. Lasting for five months, this channel alone cannot explain the entire VAT effect. A second channel is documented through spillover effects from winners to non-winners. The lottery's positive outcome demonstrates the potential of incentives for electronic payments to raise tax revenue.

*JEL Classification*: H24, H26, H31, H83.

*Keywords*: Value-Added Tax; Tax Evasion; Tax Lottery; Electronic Payments.

*EU Tax Observatory, Paris School of Economics and Hertie School, Berlin (p.nicolaides@psemail.eu). I am indebted to the Greek Ministry of Finance, the Independent Authority for Public Revenue and the General Accounting Office of Greece, for making the data used in this study available in October 2018. I would like to thank Christian Traxler and Gabriel Zucman for help, comments and guidance and Manon Francois for a detail read of the manuscript. Seminar participants at the EU Tax Observatory, Paris School of Economics, IEB Barcelona and Independent Authority for Public Revenue in Greece provided useful comments and suggestions. Field studies in Greece were funded by the Hertie School through German Academic Exchange Service funds.

# 1 Introduction

During 2009-2018, Greece experienced a deep economic crisis; GDP collapsed and public debt peaked at 180% of GDP. In an attempt to raise revenue, Greece turned to digitalisation policies that sought to tackle a large VAT gap, estimated at €6.7 billion or 3.7% of GDP at the time (Poniatowski *et al.*, 2021). Switching from cash to electronic payments would induce economy-wide third-party information by generating electronic payment trails and, thus, improve tax compliance (Kleven *et al.*, 2011; Pomeranz, 2015).[1]

Initial support for the transition to electronic payments came from capital controls in 2015 (Danchev *et al.*, 2020).[2] In 2017, two additional measures were introduced: a tax incentive (studied in Nicolaides (2022)) and a tax lottery on electronic payments (or *lottery* hereafter). The lottery rewarded €1,000 to 1,000 individuals every month with tickets that corresponded to their aggregate volume of monthly electronic payments. This paper estimates the effect of the lottery on VAT revenue and identifies two mechanisms in which the effect took place; through winners and through spillovers to non-winners.

The identification strategy relies on a natural experiment: an unanticipated superdraw on Christmas Eve in 2017. The tax authority planned monthly draws to start in January 2017, but due to a technical delay, the lottery was announced in October 2017. Earmarked prizes of €9 million could only be allocated to winners until the end of the year. To utilise the available funds, 9 retroactive draws took place on the 24th of December 2017 with tickets corresponding to electronic spending completed in the months of January to September without prior anticipation by individuals. The timing prevented self-selection into the lottery: the retroactive setting of the superdraw meant that individuals could no longer influence their chances of winning. Conditional on the level of electronic consumption, the assignment of winning prizes was random.

I investigate the effect of the lottery on VAT revenue using three administrative datasets. Firstly, the universe of 9,000 lottery winners, including their monthly electronic consumption in 19 months, their annual income and their postcodes. Secondly, a sample of 50,000 non-winners randomly-drawn from the taxpayer population. Thirdly, aggregated monthly VAT revenue from 96 regional tax offices in Greece.[3] The postcodes of winners and non-winners allow matching of individuals to regional tax offices. Using the variation of winners, I estimate that one additional winner increased VAT revenue (at tax office level) by 0.01%. This is equivalent to about €2,700 per winner, eight months after the superdraw, or roughly triple the €1,000 winning prize.

[1] According to a study on payment attitudes by the European Central Bank, 80% of transactions in Greece were completed in cash. Yet, 86% of the Greek population had access to electronic payments and were issued with credit/debit cards (ECB, 2020).

[2] The capital controls in July 2015 followed a bank run incident. Banks remained closed and cash withdrawals were limited to €60 per individual per day. By contrast, credit/debit card payments remained unlimited.

[3] Tax offices in Greece (called DOYs) are regional administrative tax centers tasked with collecting taxes. Firms belong to one regional DOY, where VAT must be paid. Overall, 101 DOYs exist in Greece. For more information see Section 2.

What could be driving the VAT revenue increase in areas with more winners? One explanation could be an idiosyncratic winners' effect. Experiencing winning made the lottery salient for these individuals, resulting in an increase in their electronic payments. This generated additional third-party information, which increased VAT revenue. Yet another explanation could be spillover effects from winners to non-winners. The latter might have received information from winners in their area about their winning experience, thus making the lottery salient for them. Alternatively firms might have been adjusting to the increase in electronic payments by winners. I investigate changes to the electronic consumption of both winners and non-winners as possible channels of the increase in VAT revenue.

To examine the winners' response, I compare their electronic consumption in a difference-in-difference setting. A comparable group of non-winners with the same probability of winning, is constructed and used as a counterfactual. Winners initially increased their electronic consumption by 14% in the first month of receiving the prize. They gradually reverted back to pre-winning spending levels by the sixth month. The effect on payments behaviour was economically large for five months, albeit temporary.

Turning to non-winners, I investigate spillover effects from winners at the postcode level. I compare the electronic consumption of non-winners in postcodes which happened to experience many winners against postcodes with no or very few winners. Initially spillover effects are not statistically significant following the superdraw. Non-winners with many winners in their area increase their electronic consumption by up to 21% from the fifth month onwards. The data allow only for a short-term assessment.

Summing up, the evidence point to an increase in VAT revenue at the tax office level by 0.01% per additional winner, which can be decomposed in (a) an idiosyncratic effect from winners and (b) spillover effects to non-winners. Winners increase their electronic consumption for 5 months after winning. Non-winners residing in the winners' postcode increase their electronic consumption from the fifth month onwards.

Despite an increase in tax lotteries in later years, there is a surprisingly slim literature on the subject. In the EU alone, Bulgaria, Czechia, Greece, Italy, Latvia, Lithuania, Malta, Romania, Poland, Portugal, Slovenia and Slovakia introduced tax lotteries. Brazil, China, Georgia, South Africa and Taiwan also run their own versions. Due to different institutional and country characteristics there is a wide diversification in lotteries, as noted in Fooken *et al.* (2015). Varying institutional settings, information technology, prizes, tickets and participation criteria, can lead to successes or failures in practice. For instance, Romania and Georgia ended their schemes, while most of the countries proceeded to fine-tuning changes over the years. Little is known about what makes a lottery successful, which will require more evidence from existing schemes. Analysing the Greek scheme enhances our understanding of the institutional details and guides policy forward.

A notable contribution to the literature is Naritomi (2019), who analyses the Brazilian tax lottery.[4] The paper finds a 21% increase in reported sales by retail firms over 4 years after the lottery's introduction. Reported taxes increased at a lower level of 9.3%, due to firms adjusting their reported expenses. Whilst differences in the institutional structure and data availability do not allow for a direct comparison of the Greek and Brazilian lotteries, the increase in VAT revenue documented in this paper confirms the results in Naritomi (2019) of lotteries being fiscally-positive incentives mechanisms. This is an important finding, since the risks for a government in implementing one appear to be limited, with the potential revenue gains being economically significant.

A main contribution of this paper is the identification of two micro-mechanisms that lead to the increase in VAT revenue. Naritomi (2019) identifies whistle-blowing and collusion costs as potential mechanisms in driving the increase. A whistle-blowing option was a unique feature in the Brazilian lottery. In the Greek tax lottery the increase in VAT revenue appears to take place through changes in electronic consumption, which is the lottery's unique feature. Winners increase their electronic consumption temporarily as the lottery becomes more salient, but so do non-winners through spillover effects in regions with many winners. Evidence suggests that targeting electronic payments in the lottery is yet another channel through which third-party information can lead to a tax revenue increase.

An additional strand of literature is that of third-party reporting through digital means. The effectiveness of third-party reporting in business-to-business transactions has been documented in a number of studies (Almunia and Rodriguez, 2014; Carrillo *et al.*, 2017; Pomeranz, 2015; Slemrod *et al.*, 2017; Waseem, 2020), including the use of information technology in Ethiopia (Ali *et al.*, 2021), in Hungary (Lovics *et al.*, 2019), in Peru (Bellon *et al.*, 2019), in Tajikistan (Okunogbe and Pouliquen, 2022) and, in Ghana (Dzansi *et al.*, 2022). Evidence of the effect of third-party reporting at the business-to-consumer (or retail) level has been a more recent subject of study. Das *et al.* (2022) examines a demonetization incident in India, Brockmeyer and Somarriba (2022) a VAT debit/credit card rebate programme in Uruguay, Adhikari *et al.* (2021) and Adhikari *et al.* (2022) study a requirement in some US cities to introduce credit card readers in small firms and in taxicabs. In line with Das *et al.* (2022), Adhikari *et al.* (2021) and Adhikari *et al.* (2022), this paper confirms the positive effect of electronic payments on tax revenue. Whilst a similar tax revenue effect is not present in the Uruguaian rebate programme in Brockmeyer and Somarriba (2022), the evidence corroborate with their findings on responses: individuals appear to be responsive to incentives that seek to increase electronic payments (either in the form of rebates or tax lotteries).

The remaining sections are structured as follows. Section 2 describes the lottery and Section 3 the data. Section 4 documents the effect of winning on VAT revenue. Section 5 and Section 6 investigate changes to the payment behaviour of winners and spillovers to non-winners, respectively. Finally, Section 7 concludes.

[4] An additional contribution, studied in Wan (2010), is the Chinese lottery.

## 2 Institutional Background

**Lottery** The Electronic Payments Tax Lottery is a scheme introduced in Greece in 2017, that provides incentives for individuals to use electronic payments instead of cash when completing retail transactions.[5] At the end of each month banks send to the tax authority the *aggregate volume* of electronic payments (but not single transactions) completed by each Greek tax resident.[6] All tax residents are included in the lottery by default, as long as they complete payments with electronic means.[7] The tax authority converts the amount of euros to tickets, using a diminishing euro-to-ticket scale as shown in Figure B.1.[8] Eligible payments that are converted into tickets are limited to everyday consumption expenses. Excluded are purchases of intangible or tangible assets, motor vehicles and payments of house rent, mortgages, taxes and fines. All other purchases award tickets if they are completed with credit cards, debit cards and e-payments.

**Prizes** Every month 1,000 winners win €1,000 each (€1 million in prizes per month). To ensure the fairness of the draws, the tax authority has implemented a double-blind draw system, where at first a research institute performs the draws and returns the winning numbers and then the tax authority applies a transformation to the numbers. In addition, individuals can only win once every month. For payments in a given month $m$, draws take place at the end of $m + 1$.[9] Winning tickets are announced to the public after the draw and winners are informed automatically via email and a text message to their mobile phones. They receive the prize in their bank accounts about a week after winning.[10] A dedicate website allows the public to view their tickets for all lottery months, as well as, any winning tickets.

[5] The lottery was included in Article 70 of L4446/2016 with the name Public Draws Programme. In the Greek public it became known as tax lottery ("Forolotaria").

[6] It is compulsory for all Greek tax residents above the age of 18 to acquire a tax ID, called AFM. This number acts as the main identifier of citizens by the state, much like an identity number. The matching of individuals between banks and the tax authority takes place through the tax ID. On one hand, when filing taxes individuals must declare their IBAN to complete the filing process. It is compulsory for all individuals above the age of 18 in Greece to file, even if they had no income during the financial year. To improve tax compliance during the economic crisis the filing process became completely electronic and automated with pre-filled information (paper declarations were eliminated). On the other hand, banks demand a tax ID when opening a bank account. This ensures matching when banks send the payment information to the tax authority.

[7] Individuals can opt out of the lottery by making a request to the tax authority. The request does not prevent banks from sending their payment information.

[8] Note that the scale is public knowledge. At the introduction of the lottery it was rewarding 1 ticket per €1 for the first €100 spent; 1 ticket per €2 for the additional €400 (i.e. from €100 to €500); 1 ticket per €3 for the additional €500 (i.e. from €500 to €1,000); and 1 ticket per €4 for any payments above €1,000. For example, suppose that in a given month an individual spends €200 in electronic payments. The individual would receive 150 tickets (100 for the first €100 and 50 for the rest). Lastly, note that the diminishing scale was introduced as a safeguard to limit individuals with high electronic consumption from winning more often. In practice it has proven insufficient and the tax authority revised the euro-to-ticket scale twice since the lottery's introduction. The tax authority has also placed a ceiling at the eligible payments that can be converted to tickets at €5,000 per month. However, for this paper only the initial scale applies.

[9] For example, for all payments completed in October, banks collect payment information from October 1st to 31st, aggregate them and send them to the tax authority early November. Payments are converted to eligible tickets and the draw take place at the end of November. Winning numbers are announced immediately after the draw. The same procedure applies for the rest of the months.

[10] Winning prizes are not subject to tax and are protected against confiscation.

**Superdraw** At Christmas Eve in 2017 a unique and unexpected superdraw took place with 9,000 winners and €9 million in prizes. Since the lottery was initially planned to begin in January 2017, the tax authority budgeted €12 million in prizes for the entire year, €1 million for each month. However, a 9-month technical delay in implementing the draws followed, resulting in a public announcement on the 9th of October 2017.[11] The first lottery took place at the 30th of November with payments completed in October and a second lottery was planned for 30th of December for payments completed in November. A €9 million earmarked amount corresponding to lotteries in the previous months remained unused and could only be allocated before the end of the financial year.[12] On the 24th of December 2017, the tax authority decided to run 9 consecutive draws, each corresponding to monthly payments completed from January 2017 to September 2017.[13]

An example of the history of draws, tickets and winning prizes for a typical individual is shown in Figure B.3 in Appendix B.1. The column "E-Consumption Period" and "E-Consumption Amount" correspond to the period and amount the spending has taken place. A "Number of Tickets" column shows the converted amount of tickets according to the euro-to-ticket scale. Importantly, the "Draw Date" column includes the superdraw lotteries with the same date (24th of December 2017) and each draw corresponds to payments made in months from January to September 2017. Winning tickets for each lottery are indicated by the red numbers (in this example, the individual did not win).

The superdraw resembles closely a natural experiment and can be exploited as an identification event. Firstly, draws were not announced in advance making the policy unexpected for individuals. Electronic payments completed in those months corresponded to their payment behaviour absent of the tax lottery's expectation. This ensures that individuals did not self-select into the policy (i.e., spending more and increasing their winning chances), which would have been the case had the lottery been announced in advance. As a result, one is not faced with individual unobservable attitudes towards the lottery. Secondly, the draws took place retroactively based on past payments; individuals could not alter their winning chances after the superdraw's announcement.

## 3 Data

**Winners** Three administrative datasets are used in this study. Firstly, the universe of 9,000 winners in the superdraw (corresponding to lotteries in January to September 2017) and complemented by an additional set of 10,000 winners in 10 regular monthly draws (from October 2017 to July 2018). The data include 19 consecutive months of aggregated monthly electronic payments as transferred from the banks to the tax authority (12 months before the superdraw and 7 months

[11]The announcement took place with a Ministerial Decision 1161, published in the official gazette at 3657/2017. A copy (in Greek) can be found in the following link https://www.aade.gr/sites/default/files/2017-11/pol1161.pdf.

[12]This was because of budgetary reasons. Accrual amounts to individual winners could only be made until 31st of December, even in payments took place a few days into the new financial year. As with any public organisation, the budget is annual and earmarked amounts cannot be transferred to the following year.

[13]A visual illustration of the lottery's timeline in 2017 is shown in Figure B.2 in Appendix B.1

after).[14] For each individual, one can determine the tickets they received in each draw. I complement these data with tax returns information from 2017, which includes the individual's income, postcode and employment category (wage-earner, pensioner, business-owner, agricultural worker or zero-income).[15]

**Non-winners** Secondly, a random sample of 50,000 individuals from the taxpayer population who did not win the tax lottery (or *non-winners* hereafter). The non-winners' information are identical to those of winners. It includes their aggregated monthly electronic payments as transferred from the banks to the tax authority in 19 months (January 2017 to July 2018). Through their monthly spending one can determine the amount of tickets they received. Similar to winners, I complement their spending with tax return information from 2017.

To allow for a meaningful comparison of winners and non-winners, one has to account for the different sampling of the two datasets. The non-winners sample was drawn randomly from the population of taxpayers conditional on not having won. The winners were explicitly drawn. To arrive at a sample that represents the baseline population of Greek taxpayers, I expand (or re-weight) the non-winners such that they match (a) the overall number of taxpayers in the population and (b) the overall number of lottery tickets (i.e., the aggregated amount of electronic payments). The details of this approach is described in Appendix B.2.

**Tax Offices** Thirdly, a dataset of monthly VAT revenue recorder in Greek tax offices. The main VAT rate in Greece was 24% in 2017 and collection was tasked to 101 regional tax offices, administering an area and acting as points of contact between taxpayers and the tax authority. Each company belongs to a single tax office, according to its established location and must declare VAT revenue in that particular office. Listed companies that may operate nationally belong to 3 dedicated national tax offices. There are also 2 local tax offices in the sample with incomplete information. For this analysis I use information from the 96 local tax offices which deal with small and medium enterprises (excluding listed companies).[16]

For each tax office I observe the aggregated amount of VAT revenue recorded from August 2017 to August 2018 (5 months before the superdraw and 8 months after). Declarations by firms are either completed every month or three months, based on their legal form and size.[17] Figure A.1 in

[14]This information was rounded to the nearest €10 by the tax authority.

[15]Note, firstly, that the tax returns data were rounded to the nearest €5. Secondly, that the tax returns information were submitted by taxpayers in the first half of 2018 for the economic year 2017. They were obtained from the tax authority in October 2018, after undergoing assessment. Thirdly, note that the zero-income category includes individuals who declared zero income in their tax returns. This group includes students, unemployed individuals or non-economically active who earned no income in 2017, yet they had to declare since this is compulsory in Greece for everyone above the age of 18.

[16]The tax offices classify VAT revenue and report the aggregated amounts to the Ministry of Finance. In particular, this dataset records Income Code 1119, which is defined as value-added taxes on any activities except: (a) those of electronic services collected by other EU member states (b) new buildings and houses (c) collected through customs (d) fuel products (e) tobacco products. Focusing on income code 1119 aides the analysis since some of these categories, such as the VAT on new buildings were excluded from receiving tickets in the lottery as discussed in Section 2.

[17]Smaller firms that declare VAT every three months are limited partnerships, general partnerships, sole proprietorships and who have annual turnover below €1.5 million. Larger firms not belonging in these categories declare monthly.

Appendix A.1 shows the mean tax office VAT revenue recorded every month. Large firms declare monthly, while smaller firms declare quarterly, leading to spikes in recorded VAT revenue in March, June, September and December. Since I observe one VAT revenue value per tax office per month, the declarations of monthly and quarterly firms cannot be distinguished. Mean monthly VAT revenue range from €1 to 3 million monthly and from €6 to 9.5 million quarterly. The overall revenue from the 96 tax office in the sample in 13 months is €5 billion, equivalent to 2.78% of the GDP of Greece in 2017.[18]

The three datasets can be combined by matching winners and non-winners to tax offices using their postcodes. The combination produces a single dataset where at the tax office level, one can determine the absolute number of winners or the winners as a percentage of the representative population. From the winners and non-winner's sample I exclude individuals with zero consumption, since by not spending they did not participate in the lottery, as well as, business owners who might have used their personal bank accounts for business transactions. The overall sample includes 7,748 winners, 44,383 non-winners and 96 tax offices. Summary statistics are presented in Table A.1.

# 4 Effect of Winning on VAT Revenue

## 4.1 Identification Strategy

The objective of this analysis is to quantify the effect of winning the lottery on VAT revenue. Recall that the superdraw was unanticipated and that, conditional on electronic consumption, 9,000 prizes in the superdraw were allocated randomly. Since each individual resides in a geographical area supervised by a tax office, I exploit the variation of winners across tax offices to estimate the effect on VAT revenue.

The variation of winners as a percentage of the representative tax office population is shown in Figure A.2 in Appendix A.1. The mean and median number of winners per tax office (as a percentage of tax office population) were 0.233% and 0.226% respectively, equivalent to 1 in every 400 individuals per tax office experiencing winning. Winners per tax office ranged from 0.153%, or 1 in every 650 individuals, to 0.37%, or 1 in every 270 individuals.

As a first step, I reconstruct aggregate monthly electronic consumption at the tax office level. This is necessary because the tickets (and winning chances) are proportional to the electronic consumption as explained in Section 2. Let electronic consumption in tax office $i$ and time $t$, be represented by $C_{i,t}$. Recall that the representative population can be obtain from the winners and non-winners sample by multiplying (or expanding) the non-winners population by a factor $\omega = 129$, as explained Appendix B.2. To arrive at the monthly electronic consumption per tax office, the non-winners electronic consumption is expanded by 129 and added to the winners electronic consumption. For a

[18] Note that the overall VAT amount was €12 billion. Listed companies belonged to 3 dedicated tax offices as mentioned in Section 2, recorded the remaining €7 billion.

tax office with winners $W$ and non-winners $NW$ the monthly aggregate consumption per tax office becomes:

$$C_{i,t} = C^{W}_{i,t} + \omega C^{NW}_{i,t}$$

At a second step, I regress the number of winners on VAT revenue, controlling for electronic consumption, tax office and time fixed effects. For tax office, $i$, let $R_{i,t}$ be the VAT revenue recorded at time $t$ and $W_i$ be the number of superdraw winners. Spending that generated tickets took place in time-lags $\ell \in \mathcal{L}$ from the superdraw.[19] Lagged electronic consumption is represented by $\sum_{\ell}^{\mathcal{L}} C_{i,t-\ell}$. The regression equation takes the following form:

$$\underbrace{R_{i,t}}_{VAT\ Revenue} = \alpha + \overbrace{\beta W_i \times Post_t}^{Winners\ variation} + \underbrace{\sum_{\ell}^{\mathcal{L}} \gamma_{t-\ell} C_{i,t-\ell}}_{Consumption\ Period} + \delta_i + \lambda_t + \epsilon_{i,t} \tag{1}$$

The first term on the right-hand side is the main parameter of interest and captures the VAT revenue effect after the superdraw using the winners' variation in tax offices. The second term controls for spending that took place during the period where tickets were generated (January 2017 to September 2017). Tax office and time-invariant factors are controlled for by $\delta_i$ and $\lambda_t$, respectively.

## 4.2 Results

Regression estimates of the effect of winners on VAT revenue are presented in Table 4.1.[20] The post-superdraw period is 8 months (from January 2018 to August 2018), therefore, the effect can only be assessed in the short-term. The results show that one additional winner increased VAT revenue at the tax office level by 0.01%. In fiscal terms this is equivalent to €2,700 of VAT revenue per superdraw winner, which is almost triple the winning prize of €1,000.[21]

The results remain robust to a number of specifications. All regressions include tax office and time-fixed effects. Robust standard errors are clustered at the tax office level due to the possibility of information sharing between firms or individuals in close proximity. Columns (1) to (3) are estimated using monthly observations, whilst for columns (4) to (6) quarterly observations are used.[22] Columns (1) and (4) exclude the lagged consumption period as controls, whilst (2) and (5) include only

[19] For example, electronic spending in January 2017 took place with a 12-month lag and for September 2017 with a 4-month lag.

[20] Detailed estimates for the controls are shown in Table A.2.

[21] A back-of-the-envelop calculation is as follows. Total VAT proceeds following the lottery, from January 2018 to August 2018, were €2.5 billion. This is equivalent to €27 million per tax office on average. An increase of 0.01% on €27 million is equivalent to €2,655.

[22] Quarterly observations avoid large spikes every quarter originating from the reporting requirements of smaller firms, as was discussed in Section 2.

observations after the superdraw (i.e., 8 months in 2018). The effect remains economically positive, statistically significant at the 99% level and largely unchanged in all specifications. Columns (3) and (6) present the main estimates of Regression 1, including all time periods (August 2017 to August 2018), all tax offices (96 in total) and with lagged electronic consumption corresponding to lottery months as controls.

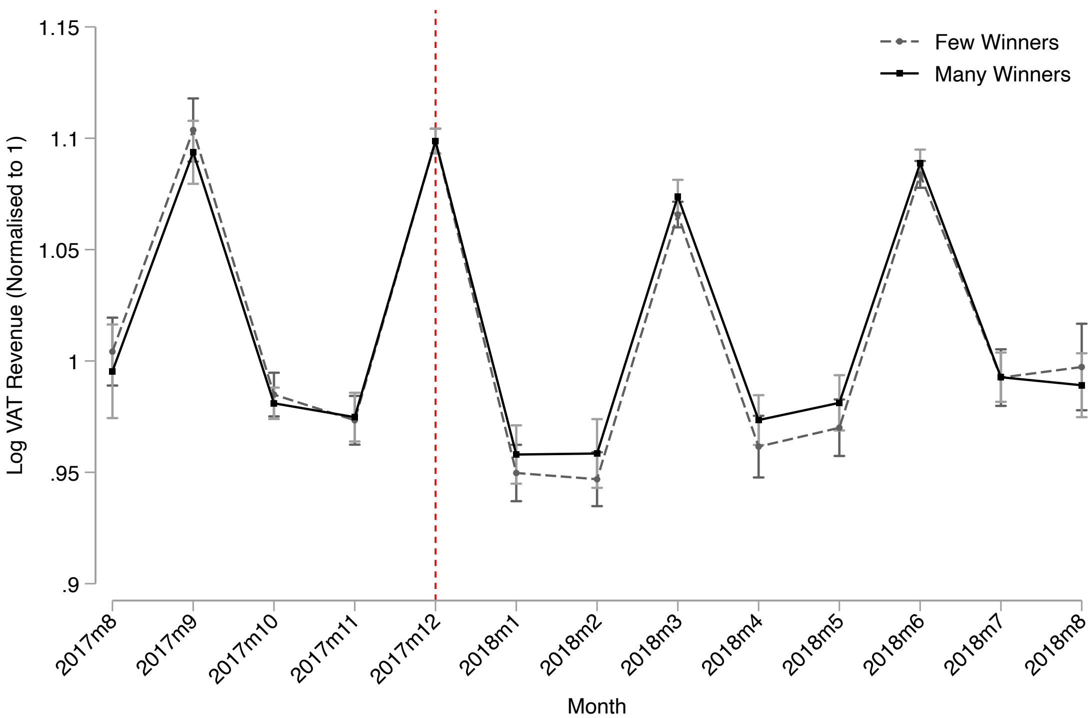
**Table 4.1** Effect of Winning on VAT Revenue

| | Monthly | | | Quarterly | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | Log Revenue | Log Revenue | Log Revenue | Log Revenue | Log Revenue | Log Revenue |
| Winner's Effect | 0.0017*** | 0.0012*** | 0.0010** | 0.0023*** | 0.0012*** | 0.0012*** |
| | (0.0004) | (0.0004) | (0.0004) | (0.0005) | (0.0004) | (0.0004) |
| Controls | No | No | Yes | No | No | Yes |
| Tax Office FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Time FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Mean | 14.50 | 14.32 | 14.32 | 15.90 | 15.76 | 15.76 |
| Observations | 1248 | 768 | 768 | 384 | 192 | 192 |
| Tax Offices | 96 | 96 | 96 | 96 | 96 | 96 |
| Adj. R-squared | 0.90 | 0.90 | 0.90 | 0.89 | 0.96 | 0.96 |

*Notes:* The table presents estimates from Regression 1. "Winner's Effect" corresponds to the variation of winners in tax offices following the superdraw and captures the effect of one additional winner on VAT revenue at the tax office level. For all regressions tax office fixed effects, time fixed and robust standard errors clustered at the tax office level are used. Columns (1), (2) and (3) use 13 months of VAT observations in 96 tax offices. Columns (4), (5) and (6) use 4 quarterly observations in 96 tax offices. All regressions present estimates of the association of winners and VAT revenue in logarithmic form. Column (1) and (4) include regressions without lagged electronic consumption values (no controls). Columns (2) and (5) include time observations *only* after the superdraw (i.e., in the months or quarters in 2018). Columns (3) and (6) correspond to the full specification of Regression 1 at the monthly and quarterly level respectively. These include lagged e-consumption values, resulting in the same observations as Columns (2) and (5) in the monthly and quarterly regression respectively.

To illustrate this result graphically, I create deciles of tax offices ranked from lowest to the highest based on the percentage of winners in their population and, I compare VAT revenue between the lowest and highest 20%. Out of 96 tax offices in the sample, the comparison includes 40; 21 tax offices in the lowest deciles and 19 in the highest deciles. For the comparison, I use a binary variable to indicate tax offices in the highest deciles (as 1) and in the lowest (as 0). The binary variable is interacted with 13 months in the sample, which provides estimates of the evolution of VAT revenue. Results from a linear regression with tax-office and time fixed-effects are presented in Column (1) of Table A.3 in Appendix A.2 and predicted values are plotted in Figure 4.1. VAT revenue are normalised to 1 with reference to August 2017, providing the log-point difference in every month.

**Fig. 4.1** Regional Tax Offices with Few/Many Winners (Linear Regression)




*Notes:* The figure presents monthly differences in VAT revenue (logarithmic form) between tax offices with many winners (solid line) and tax offices with few winners (dashed line). They are obtained from fitted values in a linear regression with interacted months and after controlling for month and tax offices fixed effects. Regression estimates are shown in Column (1) of Table A.3, taking as reference point December 2017, one period before receiving the lottery prize. The Few Winners(Many Winners) samples include tax offices with the 20 lowest(highest) winners as a percentage of the population. Monthly differences were normalised to 1 in August 2017. The y-axis shows the log-point increase/decrease with respect to that month. Confidence intervals are drawn at the 95% level with the only statistical significant difference being March 2018. The economic effects are higher for tax offices with many winners from January to June 2018. Robust standard errors are clustered at the tax office level.

Note, firstly, that the VAT revenue records distinct increases every quarter due to smaller firms declaring VAT on a quarterly basis. Secondly, the monthly evolution of VAT revenue between the two groups is identical. Thirdly, tax offices in the highest 2 deciles (solid line) record higher revenue for 5 months following the superdraw compared to tax offices in the 2 lowest deciles. By the sixth month, recorded revenue converge. Differences before the lottery remain small or negative, indicating that tax offices with many winners recorded comparatively less revenue that those with a few winners. Following the lottery, the difference is positive: economic differences in recorded revenue increase to 0.1% every month. The effect lasts from January 2018 to May 2018. Only one increase (in March 2018) is statistically significant at the 90% level, but significance might be affected by the relatively small sample of 40 tax offices. This comparison allows one to observe how the VAT revenue evolved in the highest and lowest winning regions (without taking into effect the entire variation of winners, which produces a more precise estimate).

Differences between tax offices remain when comparing unconditional mean and predicted values from a Poisson regression. Fitted values are illustrated in Figure A.4 and Figure A.5, respectively, in Appendix A.1. The results remain robust, with distinct increases ranging from 0.01% to 0.14% in 5 months following the superdraw, as documented in Column (2) of Table A.3 in Appendix A.2. In particular, a statistically significant difference for March 2018 at the 95% is observed. This corresponds to a 0.14% increase in VAT revenue for the particular month.

Overall, using the variation of winners in tax offices, the regression results in this section document a 0.01% increase in VAT revenue in an 8-month period following the superdraw. Monthly effects from the comparison of tax offices with high percentage of winners against tax offices with low percentage of winners suggest that the increase lasts for 5 months following the superdraw. What could explain the increase in VAT? One explanation could be that winners increase their electronic consumption after experiencing winning. This results in a higher volume of verifiable information in their local area, leading to firms reporting more revenue to their local tax offices. Alternatively, information about winning might be spreading to non-winners, who increase their electronic consumption in response. The remaining analysis investigates responses from winners and non-winners to understand the mechanisms by which winning the lottery increases VAT revenue.

# 5 Electronic Consumption of Winners

To explain the increase in VAT revenue, I examine how the payment behaviour of winners change once a prize is received. At the individual level, the sample includes monthly electronic consumption of winners and non-winners in 19 months; 12 before the superdraw and 7 after. This setup resembles a treatment group and a control group with a common treatment level (€1,000) and single timing (information on winning arriving in Christmas 2017 and prize money in early January). To ensure 'treatment' was random, one needs to control for spending, since higher spending increases the winning chances and determines the assignment in the treatment group.

A comparison of mean monthly electronic consumption between winners and non-winners is shown in Figure A.6. Winners exhibited higher mean electronic consumption by about €700 every month. Seasonality affects the winners' spending behaviour more than non-winners, as for instance during the end of the year, while mean electronic consumption is almost constant for non-winners. A histogram plotting the total annual electronic consumption for winners and non-winners can be seen in Figure A.7. About two-fifths of non-winners had annual electronic consumption below €1,000 and most individuals are concentrated on the left of the distribution. By contrast, winners exhibit more mass in the €3,000-7,000 area of annual electronic consumption. Spending more affects the probability of 'treatment', since it results in more tickets being awarded and higher chances of winning. Individual spending was a confounding factor in the winners' selection.[23]

Accounting for the probability of winning is necessary to create two comparable groups where a similar spending pattern during the lottery months occurred. This allows one to define a counterfactual of non-winners who exhibited similar payment behaviour to winners, but who did not experience winning. I proceed by, firstly, calculating propensity scores for the probability of winning the lottery, which generates a matching counterfactual for winners. Secondly, by calculating inverse probability weights that re-weigh winners and non-winners according to their electronic consumption. Weights are used in a difference-in-difference regression to control for the probability of assignment in the winners' group.

## 5.1 Propensity Score

The propensity score produces a metric for the probability of one individual being assigned to the winners' group during the superdraw. Let $w_n$ be a binary variable for individual $n$ with the value of 1 if winning (in any of the 9 draws) occurs and 0 otherwise. Let $T_t$ represent tickets received in months $t \in [1, 9]$ (January to September 2017).[24] The following logit model calculates the probabilities of winning:

$$P\left(w_n = 1\right) = \frac{1}{1 + \exp\left(-\left(\beta_0 + \sum_{t=1}^{9} \beta_t T_{n,t}\right)\right)} \tag{2}$$

The logistic regression is estimated using maximum likelihood with Firth's bias reduction (Firth, 1993; Heinze and Schemper, 2002). There are 1,000 winning tickets every month and about €10 billion in electronic consumption, which makes winning a rare event. The bias reduction ensures convergence of the maximum likelihood estimator. Results are presented in Table A.4 and show the increase in probability of winning for every ticket obtained in the months of January to September 2017. The effect of tickets on the probability of winning is positive for in all months and statistically significant at the 99% level.

Predicted values are plotted in a Kernel density function in Figure A.10. The graph shows the probability of winning in the superdraw for winners and non-winners. Note that the propensity scores differ for the two groups. The characteristics of the two functions are presented in Table A.5. The mean and median propensity scores for winners are 0.237 and 0.177 respectively, compared

[23]Differences in spending between winners and non-winners persist even when the lowest and highest 10% of the sample are truncated. Mean monthly electronic consumption is shown in graph Figure A.8 and the distribution of annual consumption in Figure A.9 in Appendix A.1. While seasonal spending differences are not as large as in the full sample, level differences of about €200-300 remain. The distributions of annual electronic consumption in 2017 for winners and non-winners remain similar to the full sample, with more mass for non-winners at the left of the distribution.

[24]Tickets were transformed from the electronic consumption of individuals, using the tickets-to-euros mapping as explained in Section 2.

to 0.133 and 0.103 for non-winners. That means, winners happened to be individuals with higher spending and about double the probability of winning than others. Whilst both density functions are skewed to the left, the non-winners' function exhibits a high mass below the 0.1 probability level. This means that large number of non-winners had a particularly low probability of treatment due to low electronic spending.

In order to make the two groups comparable, I limit the groups to ensure the probability of treatment, originating from the amount of tickets they received, was matching. I include individuals with propensity score below 0.17 and above 0.4. This corresponds to individuals in the 50th to the 90th percentile in the distribution of winners and from 80th to the 98th percentile in the distribution of non-winners. A comparison of the resulting Kernel density functions is shown in Figure A.10 with corresponding density function characteristics at the bottom part of Table A.5. As can be seen the two groups have a comparable probability of winning; they have an identical Kernel density function which translates to similar spending characteristics and tickets during the superdraw. The winners' sample includes 3,249 individuals and the non-winners' sample 8,144 individuals with similar probability of treatment and with the only difference that one group received a prize while the other one did not.

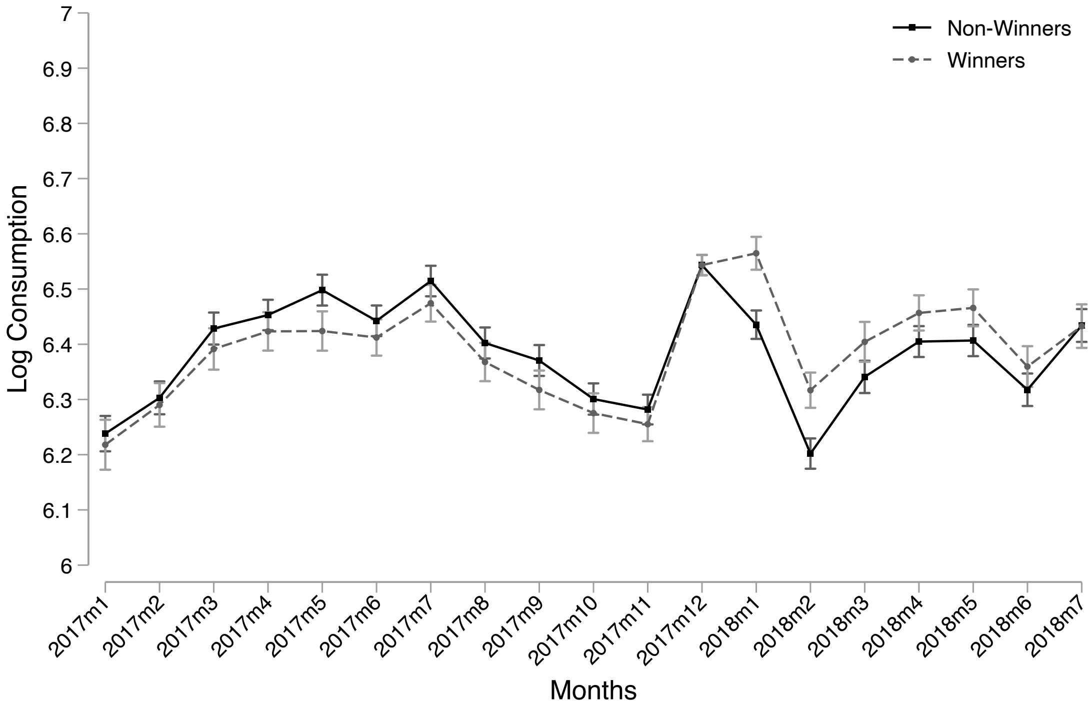
## 5.2 Estimation

The effect of winning on electronic consumption can be identified by comparing the two groups in a difference-in-difference setting with individual and time fixed effects. Given monthly electronic consumption $c^w_{n,t}$ for individual $n$ and indicator for winning $w_n$, I fit the following regression using the sub-sample of matching propensity scores:

$$\underbrace{c^w_{n,t}}_{Winners\, E-Consumption} = \alpha + \overbrace{\beta w_n \times Post_t}^{Winners\, indicator} + \chi_n + \lambda_t + \epsilon_{n,t} \tag{3}$$

Variables $\chi_n$ and $\lambda_t$ capture individual and time fixed effects respectively. In addition, inverse probability weights from the propensity score estimation are used to re-balance individuals and control for the probability of selection in the winners' group. This results in over-weighing individuals with high propensity score who have spent more and received more tickets and under-weighing individuals who have spent less and received less tickets. Results from Regression 3 are shown in Table A.6 and plotted in Figure 5.1.

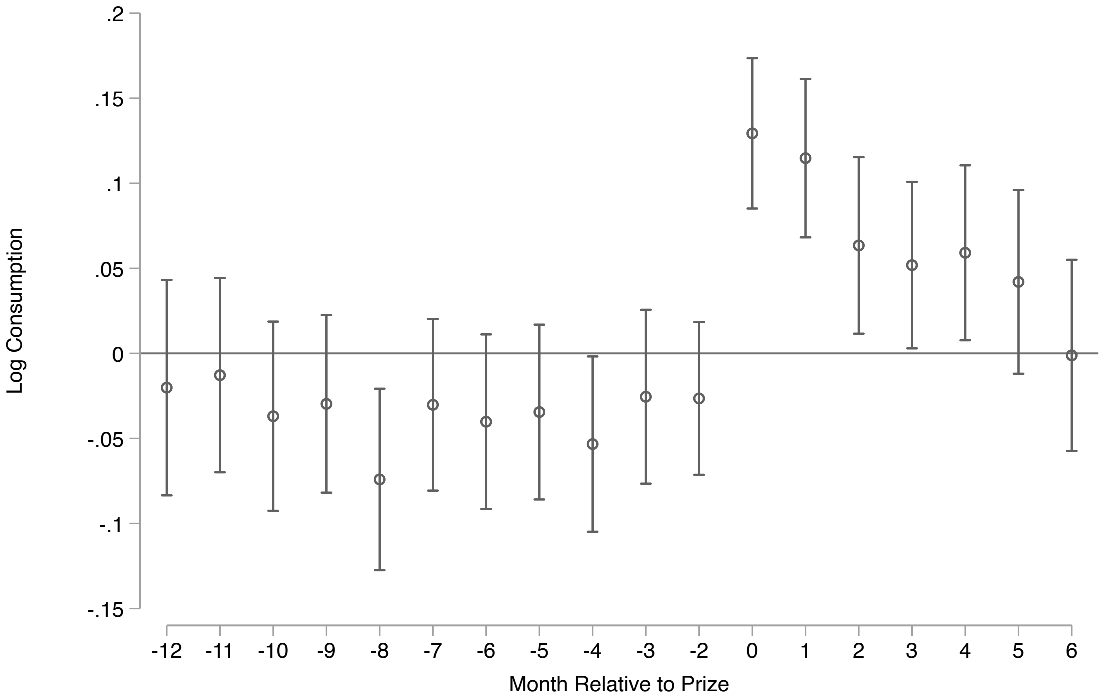
**Fig. 5.1** Effect of Winning on E-Consumption




*Notes:* The figure presents predicted values using Regression 3 for winners ($w_n = 1$) and non-winners ($w_n = 0$). Estimates are shown in Table A.6. The figure plots monthly differences in electronic consumption (logarithmic form) with reference to December 2017 (-1 period to receiving the lottery prize). The winners group (dashed line) includes 3,249 superdraw winners and the non-winners group (solid line) includes 8,144 non-winners. The groups were matched using propensity scores to ensure a similar spending pattern in the months of January to September 2017. Confidence intervals are at the 95% level. Robust standard errors are clustered at the individual level.

The graph plots the electronic consumption in logarithmic form from January 2017 to July 2018. Firstly, note that consumption is parallel between winners and non-winners; the comparison takes place between groups that exhibited a matching spending pattern and therefore similar chances of winning. Seasonality is matching and monthly fluctuations are identical. Secondly, a parallel trend is maintained throughout the pre-winning period, from January 2017 to December 2017. The trend fluctuates monthly and differences are statistically insignificant, with the exception of April 2017. The non-winners group provides a valid counterfactual of how the winners' electronic consumption would have evolved had they not experienced winning in the superdraw. Thirdly, as can be seen, the electronic consumption of winners diverges temporarily after receiving the prize, but reverts back to the non-winners consumption level after 6 months.

**Fig. 5.2** Event Study




*Notes:* The figure presents results from an event study that correspond to the results in Figure 5.1 and for estimates are Table A.6. It quantifies the increase in electronic consumption by winners after receiving the lottery prize. Month and individual fixed effects were used in the estimation. Monthly differences in electronic consumption (logarithmic form) are drawn with reference to December 2017 (-1 period to receiving the lottery prize) and with the 0-horizontal line representing the electronic consumption of non-winners. The winners group includes 3,249 superdraw winners and the non-winners group includes 8,144 non-winners. The groups were matched using propensity scores to ensure a similar spending pattern in the months of January to September 2017. Confidence intervals are at the 95% level. Robust standard errors are clustered at the individual level.

Figure 5.2 presents the effect on winner's electronic consumption in an event study. Winners increased their electronic consumption by 13.8% and 12.1% in the first two months after winning. The increase is reduced to 8.9%, 6.8% and 8% in the third to fifth month, before subsiding to pre-winning consumption levels by month 6 and 7. Overall, winning produced an economically large short-term response in winners. However, it did not cause a permanent change in their payment habits (from cash to electronic payments).

As was documented in Table A.1, the average monthly electronic consumption for winners ranged from €1,021 to 1,370. This implies that the 5-month electronic consumption increases was about €589 per winner or € 5.3 million for 9,000 winners in the superdraw. Conditional on all revenue being recorder, a 24% VAT rate on the additional amount spent implies a €1.3 million revenue for the state. The increase in electronic consumption identified in winners can provide one explanation of the increased VAT revenue documented in Section 4.

# 6 Spillovers to Non-Winners

A complementary effect that could explain the increase in VAT revenue, pertains to non-winners increasing their electronic consumption. Whilst the data do not allow for a direct observation of information exchange, one can examine if non-winners alter their payment behaviour (increasing electronic consumption) based on the number of winners in their area. This would indicate the presence (or absence) of spillover effects from winning. Spillover effects could arise either from other individuals sharing information of their winning experience or from firms adjusting to a more widespread use of electronic payments.

I investigate the presence of spillovers by utilising the variation of winners at the *postcode level*. Figure A.3 shows a distribution of the percentage of winners over the population in each postcode. As can be observed, some postcodes did not receive any winners whilst other postcodes experienced up to 1% of winners in their population. A 'treatment' group consists of *non-winners* with many winners in their postcode, whilst a 'control' group with *non-winners* with few or no winners in their postcode.

There are 1,099 postcodes in total in our sample. Ranking postcodes by the percentage of winners in their population, a first comparison group is made of individuals in the highest 10% versus the lowest 10% (*Group 1*). A second comparison group is made of a tighter sample of 39 postcodes, which received no or very few winners (from 0 to 0.03% of the postcode population), against non-winners from postcodes who experienced higher than 0.3% of winners in their population (*Group 2*). There are 756 non-winners in the former and ?? non-winners in the latter. I compare the electronic consumption of non-winners in low against high percentages of winners, controlling for postcode, individual and time fixed effects.

## 6.1 Propensity Score

To produce a meaningful comparison between non-winners, I follow the same approach as in Section 5.1. Firstly, I calculate propensity scores based on spending during January to September 2017 using Regression 2. This is necessary in order to ensure that individuals in different postcodes had a similar spending pattern and for assessing what their electronic consumption would had been absent of winners in their area. Recall that one should consider treatment in this context as assigning a high number of winners in a non-winner's location. The propensity score controls for confounding between higher spending in the area, which increases the probability of treatment. That is, non-winners in postcodes with high spending have a higher chance of being 'treated' with more winners in their area.

The propensity scores produce samples with comparable spending levels, but who happened to reside in areas with many versus few winners. Kernel density functions of propensity scores for the Group 1 and Group 2 are shown in Figure A.13 and Figure A.12 respectively. Similar to Section 5.1, I limit the samples to propensity scores below 0.17 and above 0.4 to produce comparable samples in terms of spending.

### 6.2 Estimation

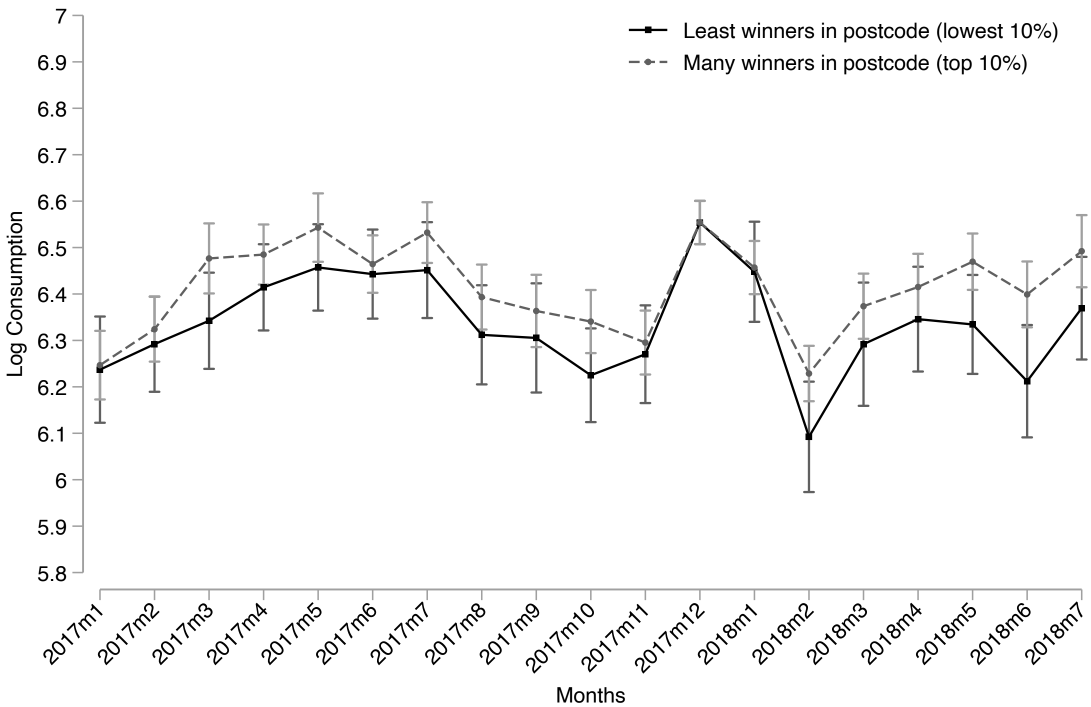
Spillover effects in the electronic consumption of non-winners are estimated using difference-in-difference with postcode, individual and time fixed effects. Let $c^{nw}_{p,n,t}$ denote the electronic consumption at time $t$ of a non-winning individual $n$ residing at postcode $p$. Let $w_p$ be a binary variable with the value of 1 if the non-winner belonged to a postcode of many winners and 0 to a postcode of few or no winners. The regression takes the following form:

$$\underbrace{c^{nw}_{p,n,t}}_{Non-winners\, consumption} = \alpha + \overbrace{\beta w_p \times Post_t}^{Many/few\, winners\, indicator} + \delta_p + \chi_n + \lambda_t + \epsilon_{p,n,t} \tag{4}$$

Variables $\delta_p$, $\chi_n$ and $\lambda_t$ capture postcode, individual and time fixed effects respectively. Inverse probability weights, calculated from the propensity scores using Equation 2. The weights re-balance the non-winners based on their electronic consumption; a non-winner with high electronic consumption is over-weighted and therefore more likely to have winners in the area of residence. Similarly, non-winners with low levels of electronic consumption are down-weighted.

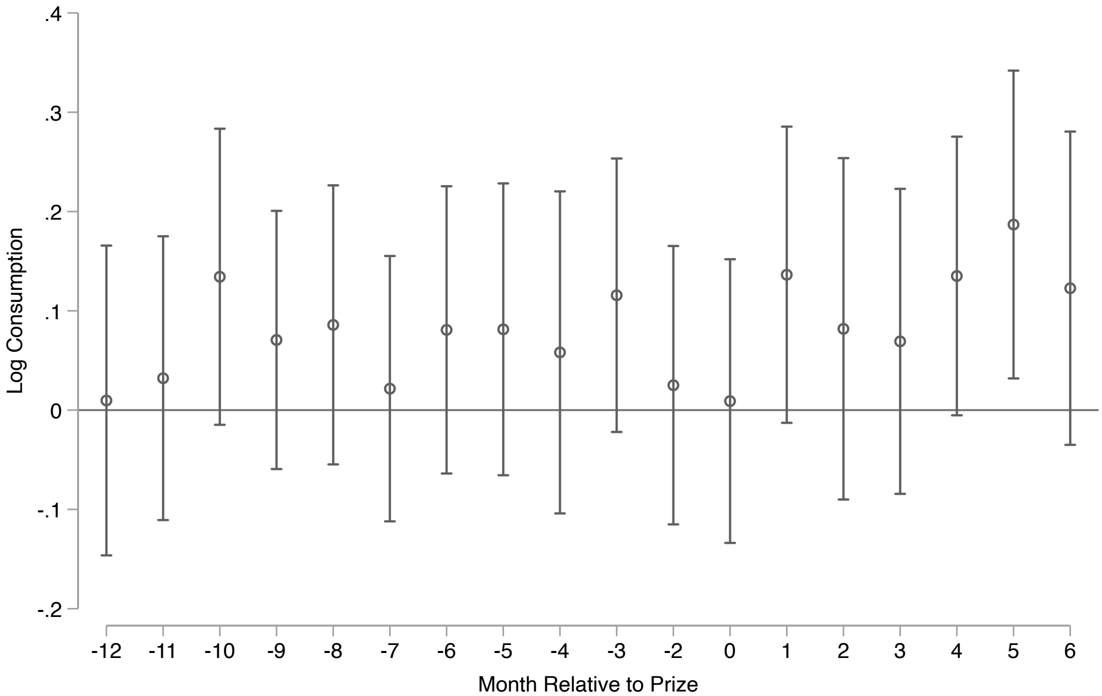
Results from Group 1 are shown in Table A.7 and monthly differences from Regression 4 are presented in Figure 6.1. Non-winners in postcodes with the highest/lowest percentage of winners exhibit a similar electronic consumption pattern prior to the lottery with matching seasonal fluctuations. Monthly differences are not statistically significant for the two samples prior the lottery. The two samples diverge slightly following the superdraw, with two monthly differences being statistically significant at the 10% level and one at the 95% level. An event study in Figure 6.2 presents the differences between the two samples, relative to the month winners received the prize.

**Fig. 6.1** Effect of Winning on Non-winners E-Consumption (10% Lowest/Highest)




*Notes:* The figure presents monthly differences in electronic consumption (logarithmic form) between non-winners in postcodes with many winners (dashed line) and non-winners in postcodes with few winners (solid line). It shows spillovers in electronic consumption from winners to non-winners. The estimates are obtained from fitted values in a linear regression with interacted months and after controlling for individual, month and postcode fixed effects. Estimates are shown in Table A.7, taking as reference point December 2017, one period before winners received the lottery prize. The two groups were formed by ranking postcodes by the percentage of winners in their population and taking non-winners from the lowest/highest 10%. Non-winners were matched using propensity scores to ensure a similar pattern of spending prior to the lottery. Confidence intervals are drawn at the 95% level. Robust standard errors are clustered at the postcode level.

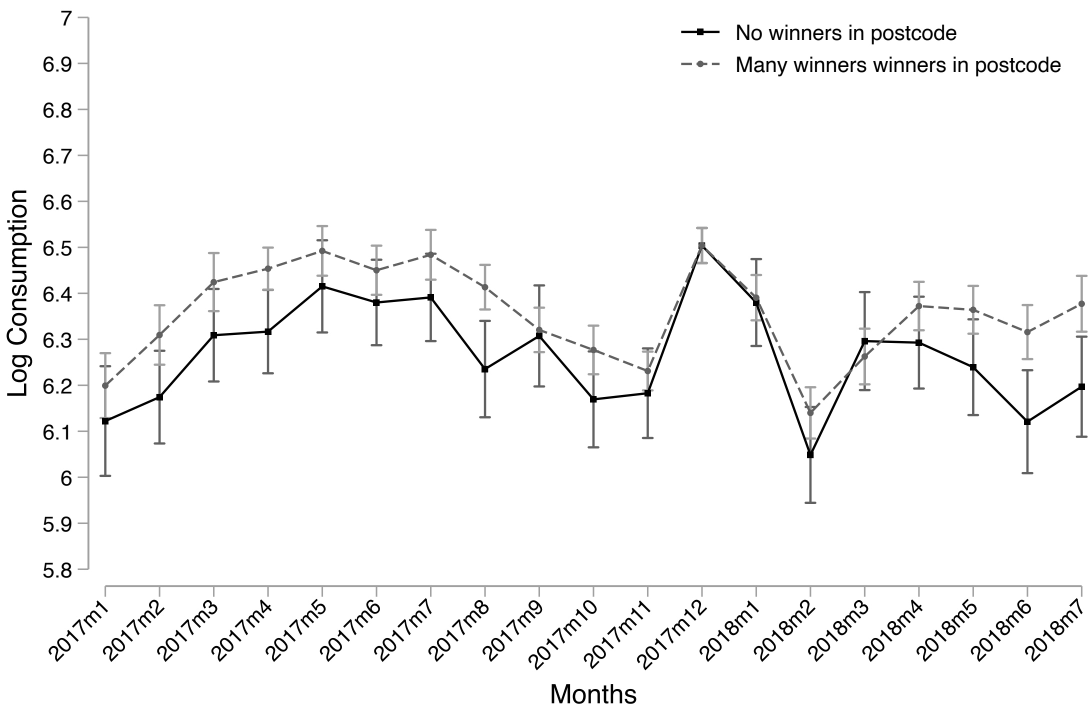
**Fig. 6.2** Event Study (10% Lowest/Highest)




*Notes:* The figure plots estimates from an event study that correspond to the results in Figure 6.1 and Table A.7. It quantifies the changes in electronic consumption by non-winners after winners in their postcodes receive the lottery prize. Individual, month and postcode fixed effects were used in the estimation. Monthly differences in electronic consumption (logarithmic form) are drawn with reference to December 2017 (-1 period to winners receiving the lottery prize) and with the 0-horizontal line representing the electronic consumption of non-winners in postcodes with the 10% lowest percentage of winners. The groups were matched using propensity scores to ensure a similar spending pattern in the months of January to September 2017. Confidence intervals are at the 95% level. Robust standard errors are clustered at the postcode level.

A similar effect can observed when a tighter sample is used in Group 2. This sample includes non-winners with no or very few winners in their postcode (less than 0.3% of the postcode population). Their electronic consumption is compared against non-winners in postcodes with many winners (more than 0.3% of the postcode population). As in the analysis above, I re-balance the sample using inverse probability weights to ensure similar electronic consumption pattern between the two samples. Similar to Group 1, the electronic consumption pattern follows a matching trend without statistically significant differences prior to the superdraw, as can be seen in Figure 6.3. The only statistically significant difference prior to the lottery is observed August 2018, at the 95%-level. This is due to individuals with no winners in their postcode having a lower electronic consumption in August 2018, before compensating with higher electronic consumption in the following month. Electronic consumption differences begin to diverge with differences 5 to 7 months after the superdraw becoming statistically and economically significant.

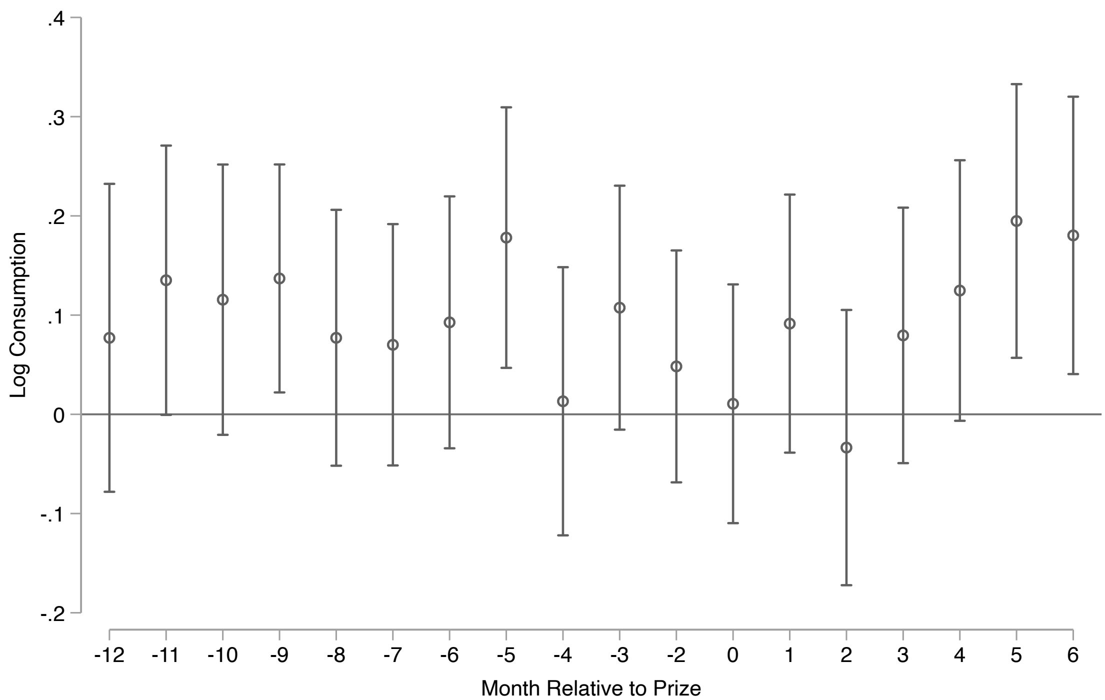
**Fig. 6.3** Effect of Winning on Non-winners E-Consumption (No winners)




*Notes:* The figure presents monthly differences in electronic consumption (logarithmic form) between non-winners in postcodes with many winners (dashed line) and non-winners in postcodes with no winners (solid line). It shows spillovers in electronic consumption from winners to non-winners. The estimates are obtained from fitted values in a linear regression with interacted months and after controlling for individual, month and postcode fixed effects. Estimates are shown in Table A.7, taking as reference point December 2017, one period before winners received the lottery prize. The two groups were formed by ranking postcodes by the percentage of winners in their population and taking non-winners from postcodes which exhibited more than 0.3% of winners in their population (many winners) and less than 0.03% (no winners). Non-winners were matched using propensity scores to ensure a similar pattern of spending prior to the lottery. Confidence intervals are drawn at the 95% level. Robust standard errors are clustered at the postcode level.

The differences between the two groups with reference to the month winners received prizes are depicted in an event study in Figure 6.4. In month 5 after the draw (May 2018) the electronic consumption of non-winners in postcodes with many winners increases by 13.3%, statistically significant at the 10% level (p-value 0.06), compared to non-winners in postcodes with no winners. Month 6 and 7 (June and July 2018) exhibit increases of 21.5% and 19.8% respectively, statistically significant at the 99% level (p-values 0.00 and 0.01). This provides evidence of a delayed effect in the electronic consumption of non-winners which is limited to some of the months.

**Fig. 6.4** Event Study (No winners)




*Notes:* The figure plots estimates from an event study that correspond to the results in Figure 6.3 and Table A.7. It quantifies the changes in electronic consumption by non-winners after winners in their postcodes receive the lottery prize. Individual, month and postcode fixed effects were used in the estimation. Monthly differences in electronic consumption (logarithmic form) are drawn with reference to December 2017 (-1 period to winners receiving the lottery prize) and with the 0-horizontal line representing the electronic consumption of non-winners in postcodes with less than 0.03% in percentage of winners over the postcodes population. The groups were matched using propensity scores to ensure a similar spending pattern in the months of January to September 2017. Confidence intervals are at the 95% level. Robust standard errors are clustered at the postcode level.

Overall, the findings from spillover effects provide a mixed picture. There is no evidence of an immediate effect in the months after the superdraw, but electronic consumption of non-winners with many winners in their postcodes increases in later months (from May to July 2018). These monthly increases are economically large ranging from 13.3% to 21.5%. This serves as limited evidence of a delayed effect and should be interpreted with caution given the short-term focus of the study, which includes 7 months of electronic consumption observations after the superdraw.

# 7 Conclusion

This paper estimates the effect of winning the Greek Electronic Payments Tax Lottery on VAT revenue and identifies short-term changes in the payments behaviour of individuals. An unexpected superdraw on Christmas Eve in 2017 generated 9,000 winners and allocated €9m in prizes. Using the variation of winners in tax offices, this paper documents a 0.01% increase in VAT revenue per additional winner. The effect can be decomposed in an idiosyncratic effect from winners and in spillover effects from winners to non-winners. Winners increase their electronic consumption for five months after winning, but by the sixth month they revert back to pre-winning electronic consumption levels. Spillover effects appear from the fifth month following the superdraw. Initially no response is recorded in non-winners during the first four months.

The results have a number of implications for third-party reporting policies and tax lotteries in particular. Firstly, in line with the findings of Naritomi (2019) in the Brazilian tax lottery, the conclusions confirm the positive effect tax lotteries can have in increasing VAT revenue through additional verifiable information. The analysis sheds light on the winners' channel; experiencing winning incentivises higher electronic consumption (in the short-term), which increases third-party information and VAT revenue. In addition, it provides evidence on the existence of spillovers effects in third-party reporting; winning can have a reinforcing effect through non-winners.

Secondly, considerable government innovation in the use of big data and digitalisation has taken place in later years. Tax administrations have been exploring ways of utilising the latest advancements in information and communications technology (Gupta *et al.*, 2017). Empirical evidence point to a positive effect of digitalisation policies in facilitating formality in firms (Ali *et al.*, 2021; Lovics *et al.*, 2019; Bellon *et al.*, 2019; Okunogbe and Pouliquen, 2022) and in incentivising individuals to use electronic payments (Brockmeyer and Somarriba, 2022). Results from this tax lottery corroborate with similar evidence on the success of electronic payments to incentivise individuals, albeit in the short-term. While the setting of the policy did not allow for an assessment of tax compliance, the evidence suggests the policy was fiscally positive with limited risk for the government and positive revenue potential.

Thirdly, the temporary effect of winning on electronic consumption indicates the limitations of the policy in facilitating a switch from cash to electronic payments. The tax lottery does not succeed in changing payment habits permanently; the latter appear to be persistent, given that winners revert back to their initial electronic consumption levels. Similar evidence on the limitations of third-party reporting policies have been documented in firms' responses in Carrillo *et al.* (2017) and Bjørneby *et al.* (2021).

Lastly, design characteristics in tax lotteries play an important role in their success or failure. Over the years the main mechanisms of tax lotteries remained in principle the same, but their characteristics became more diversified as more countries began to adopt them. A variety of ticket structures, prizes, participation criteria and information technology systems currently exists. In the absence of a common best-practices approach, tax authorities often optimise by trial and error, relying on small fine-tuning interventions following a lottery's introduction. Analysing different versions of tax lotteries is necessary to enhance our understanding of good policy practices in the future.

## References

Adhikari, B., Alm, J., Collins, B., Sebastiani, M. and Wilking, E. (2022). Using a natural experiment in the taxicab industry to analyze the effects of third-party income reporting. *Journal of Economic Behavior  Organization*, **193**, 312–333.

—, — and Harris, T. F. (2021). Small business tax compliance under third-party reporting. *Journal of Public Economics*, **203**, 104514.

Ali, M., Shifa, A. B., Shimeles, A. and Woldeyes, F. (2021). Building fiscal capacity in developing countries: Evidence on the role of information technology. *National Tax Journal*, **74** (3), 591–620.

Almunia, M. and Rodriguez, D. L. (2014). Heterogeneous responses to effective tax enforcement: Evidence from spanish firms, Banco de Espana Working Paper No. 1419.

Bellon, M., Chang, J., Dabla-Norris, E., Khalid, S., Lima, F., Rojas, E. and Villena, P. (2019). Digitalization to improve tax compliance: Evidence from vat e-invoicing in peru, IMF Working Paper No. 19/231.

Bjørneby, M., Alstadsæter, A. and Telle, K. (2021). Limits to third-party reporting: Evidence from a randomized field experiment in norway. *Journal of Public Economics*, **203**, 104512.

Brockmeyer, A. and Somarriba, M. S. (2022). Electronic Payment Technology and Tax Compliance: Evidence from Uruguay's Financial Inclusion Reform, CEPR Discussion Paper No. DP17097, Available at SSRN: https://ssrn.com/abstract=4069905.

Carrillo, P., Pomeranz, D. and Singhal, M. (2017). Dodging the taxman: Firm misreporting and limits to tax enforcement. *American Economic Journal: Applied Economics*, **9** (2), 144–64.

Danchev, S., Gatopoulos, G. and Vettas, N. (2020). Penetration of Digital Payments in Greece after Capital Controls: Determinants and Impact on VAT Revenues. *CESifo Economic Studies*.

Das, S., Gadenne, L., Nandi, T. K. and Warwick, R. (2022). Does Going Cashless Make You Tax-Rich? Evidence from India's Demonetization Experiment, CEPR Discussion Paper No. DP16891, Available at SSRN: https://ssrn.com/abstract=4026778.

Dzansi, J., Jensen, A., Lagakos, D. and Telli, H. (2022). *Technology and Tax Capacity: Evidence from Local Governments in Ghana*. Working Paper 29923, National Bureau of Economic Research.

ECB (2020). Study on the payment attitudes of consumers in the euro area (SPACE).

Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika*, **80** (1), 27–38.

Fooken, J., Hemmelgarn, T. and Herrmann, B. (2015). Improving VAT Compliance – Random Awards for Tax Compliance, Directorate General Taxation and Customs Union, European Commission, Taxation Paper No. 51.

Gupta, S., Keen, M., Shah, A. and Verdier, G. (2017). *Digital Revolutions in Public Finance.* International Monetary Fund.

Heinze, G. and Schemper, M. (2002). A solution to the problem of separation in logistic regression. *Statistics in Medicine*, **21** (16), 2409–2419.

Kleven, H. J., Knudsen, M., Kreiner, C., Pedersen, S. and Saez, E. (2011). Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark. *Econometrica*, **79** (3), 651–692.

Lovics, G., Szîoke, K., Tóth, C. G. and Ván, B. (2019). *The effect of the introduction of online cash registers on reported turnover in Hungary.* MNB Occasional Papers 137, Budapest.

Naritomi, J. (2019). Consumers as Tax Auditors. *American Economic Review*, **109** (9), 3031–3072.

Nicolaides, P. (2022). Income Tax Incentives for Electronic Payments: Evidence from Greece's Electronic Consumption Tax Discount, Working Paper.

Okunogbe, O. and Pouliquen, V. (2022). Technology, taxation, and corruption: Evidence from the introduction of electronic tax filing. *American Economic Journal: Economic Policy*, **14** (1), 341–72.

Pomeranz, D. (2015). No taxation without information: Deterrence and self-enforcement in the value added tax. *American Economic Review*, **105** (8), 2539–69.

Poniatowski, G., Bonch-Osmolovskiy, M. and Śmietanka, A. (2021). VAT gap in the EU : report 2021, Directorate General Taxation and Customs Union, European Commission.

Slemrod, J., Collins, B., Hoopes, J. L., Reck, D. and Sebastiani, M. (2017). Does credit-card information reporting improve small-business tax compliance? *Journal of Public Economics*, **149**, 1–19.

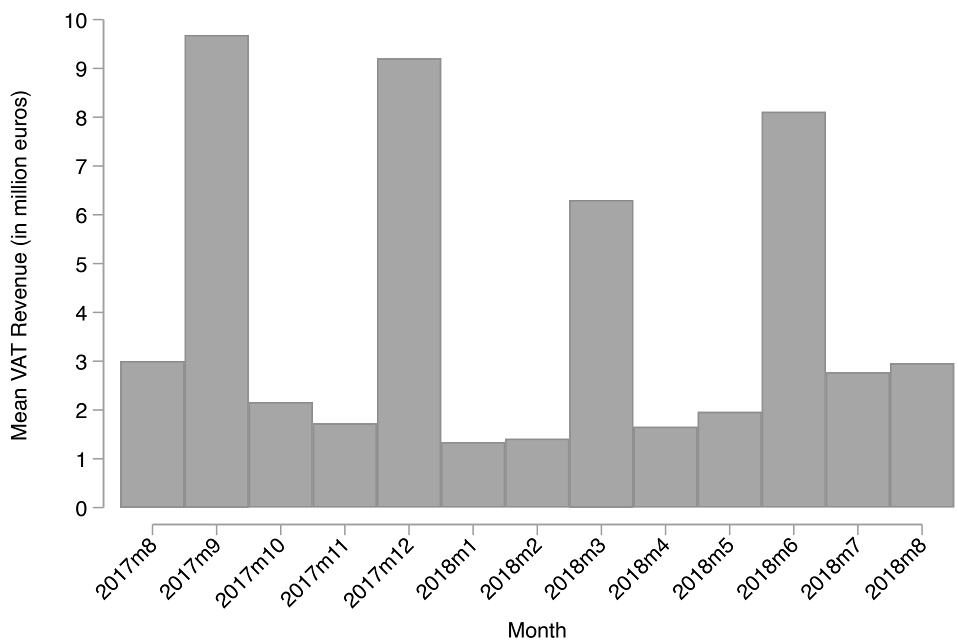
Wan, J. (2010). The incentive to declare taxes and tax revenue: The lottery receipt experiment in china. *Review of Development Economics*, **14** (3), 611–624.

Waseem, M. (2020). The Role of Withholding in the Self-Enforcement of a Value-Added Tax: Evidence from Pakistan. *The Review of Economics and Statistics*, pp. 1–44.

# A Appendix: Figures and Tables

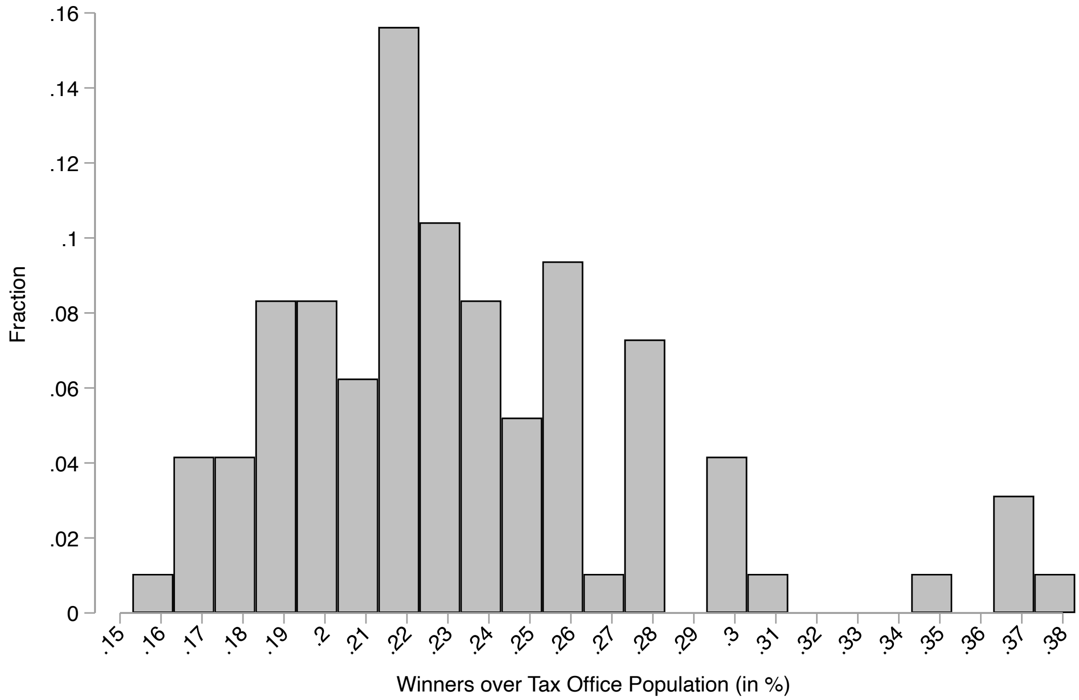
## A.1 Figures

**Fig. A.1** Mean Tax Office Revenue



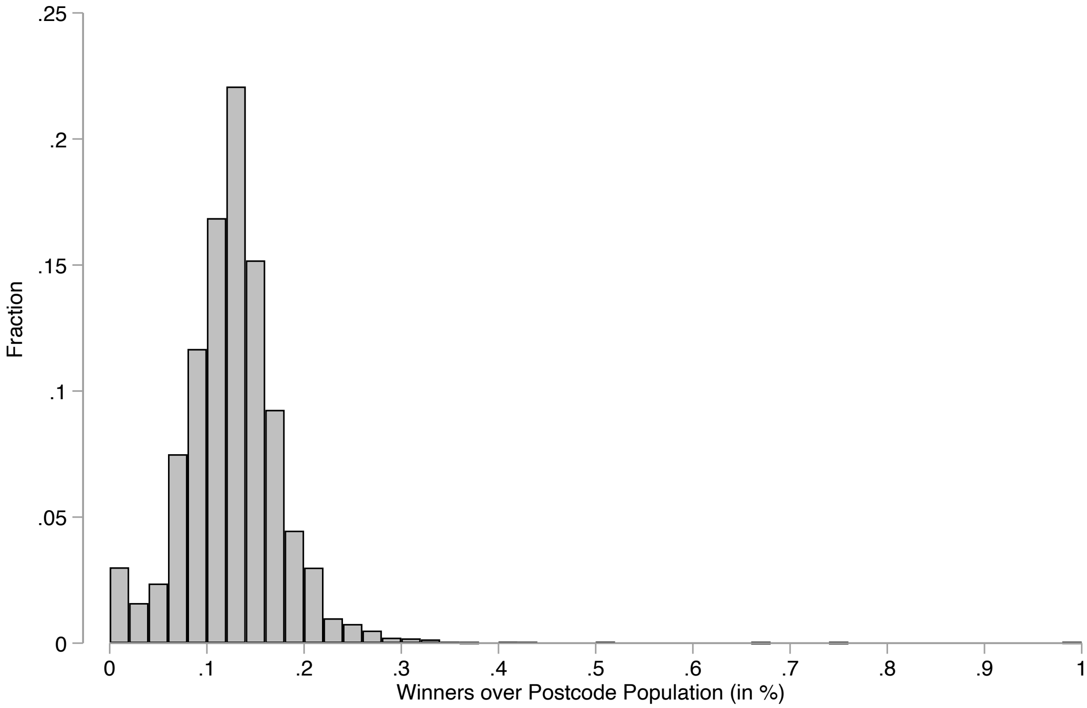

*Notes:* The graph plots mean VAT revenue in million euros, collected in 96 tax offices in Greece from August 2017 to August 2018. VAT revenue correspond to accounting code 1119 in the Greek public finance system, which corresponds to tax revenue from activities other than building, tobacco, fuel and alcohol products. There were 101 tax offices in Greece in 2017. The graph excludes 3 tax offices, which include listed companies and do not deal with companies based on a geographical basis and 2 tax offices for which data are incomplete. Tax revenue are reported by firms monthly or quarterly based on their size and type. Smaller firms report quarterly leading to the spikes observed in the graph in September, December, March and June.

**Fig. A.2** Variation of Winners in Tax Office Population



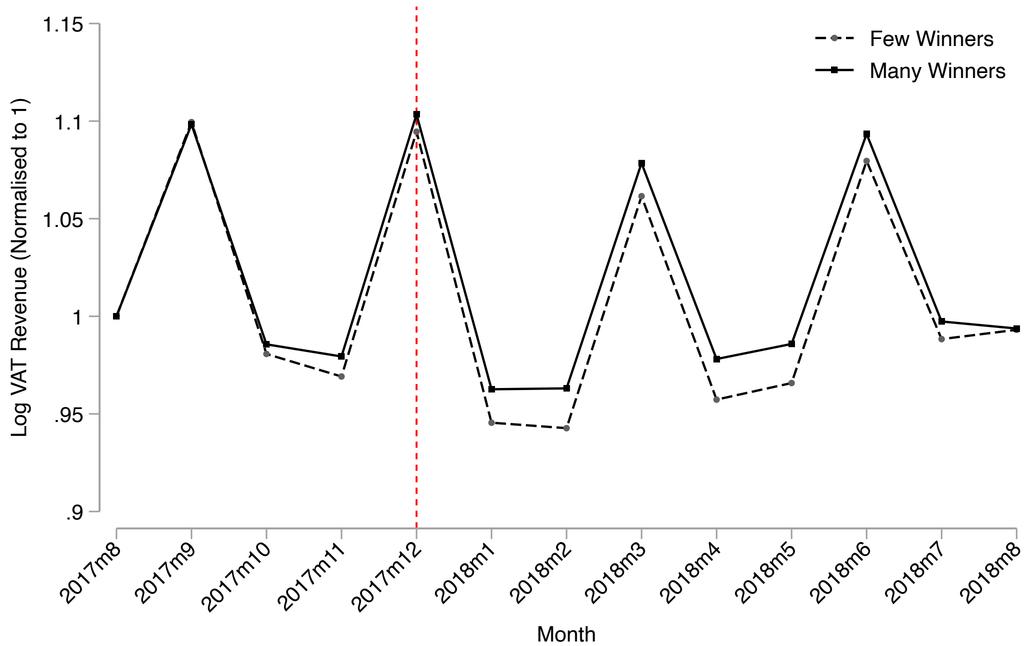

*Notes:* The figure presents the distribution of winners as a percentage of the population in each tax office. To construct this, winners in each tax office are divided by the tax office population and multiplied by 100. This gives the percentage of winners (x-axis). For example taking the most frequent observation, 0.22% of the tax office's population have won in the superdraw. The population is constructed using the randomly-drawn sample of non-winners as described in Appendix B.2. The distribution includes 96 tax offices.

**Fig. A.3** Variation of Winners in Postcode Population



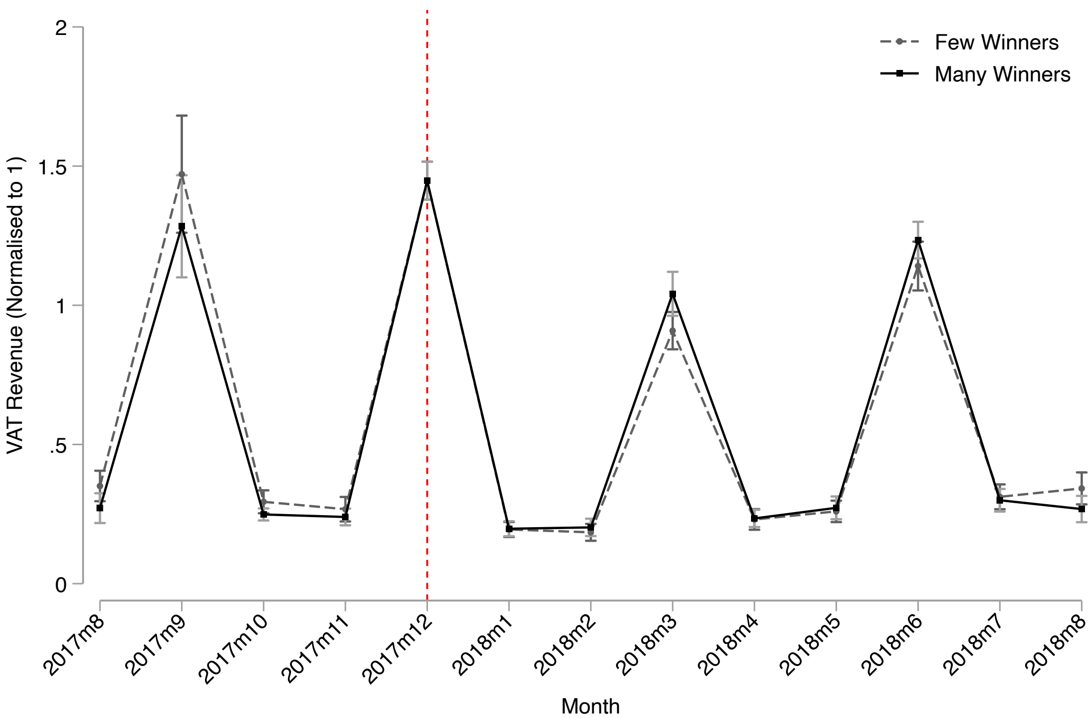

*Notes:* The figure presents the distribution of winners as a percentage of the population in each postcode. To construct this, winners in each tax office are divided by the postcode population and multiplied by 100. This gives the percentage of winners (x-axis). For example taking the highest observation, 1% of that postcodes' population have won in the superdraw. The population is constructed using the randomly-drawn sample of non-winners as described in Appendix B.2. The distribution includes 1,099 postcodes.

**Fig. A.4** Regional Tax Offices with Few/Many Winners (Unconditional Means)



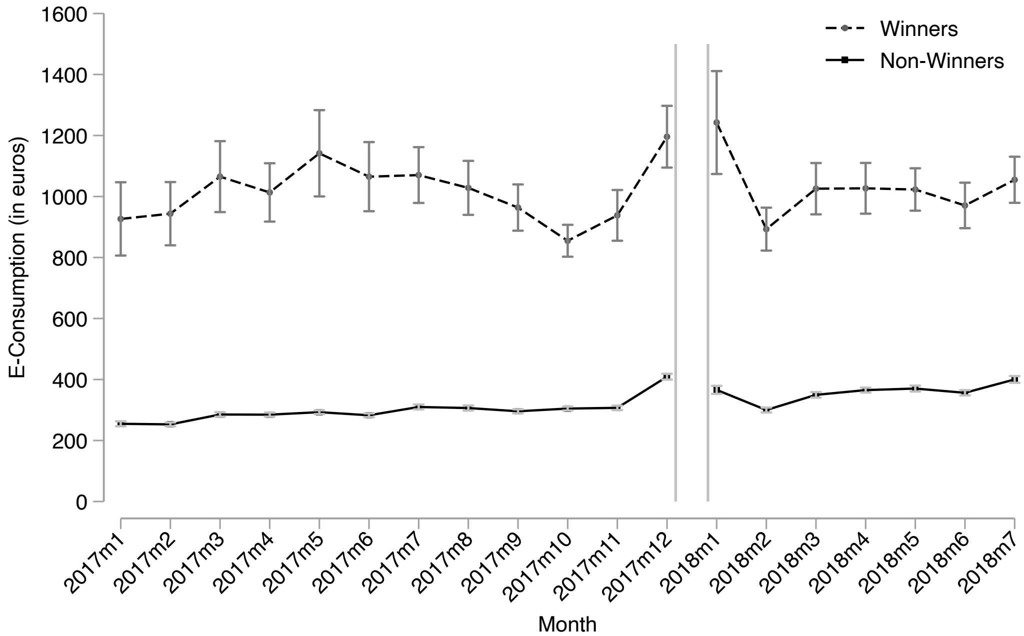

*Notes:* The figure presents unconditional monthly differences in VAT revenue (logarithmic form) between tax offices with many winners (solid line) and tax offices with few winners (dashed line). The red line is a reference point to December 2017, the month of the lottery and before winners received the lottery prize. The Few Winners(Many Winners) samples include tax offices with the 20 lowest(highest) winners as a percentage of the population. Monthly differences were normalised to 1 in August 2017. The y-axis shows the log-point increase/decrease with respect to that month.

**Fig. A.5** Regional Tax Offices with Few/Many Winners (Poisson Regression)



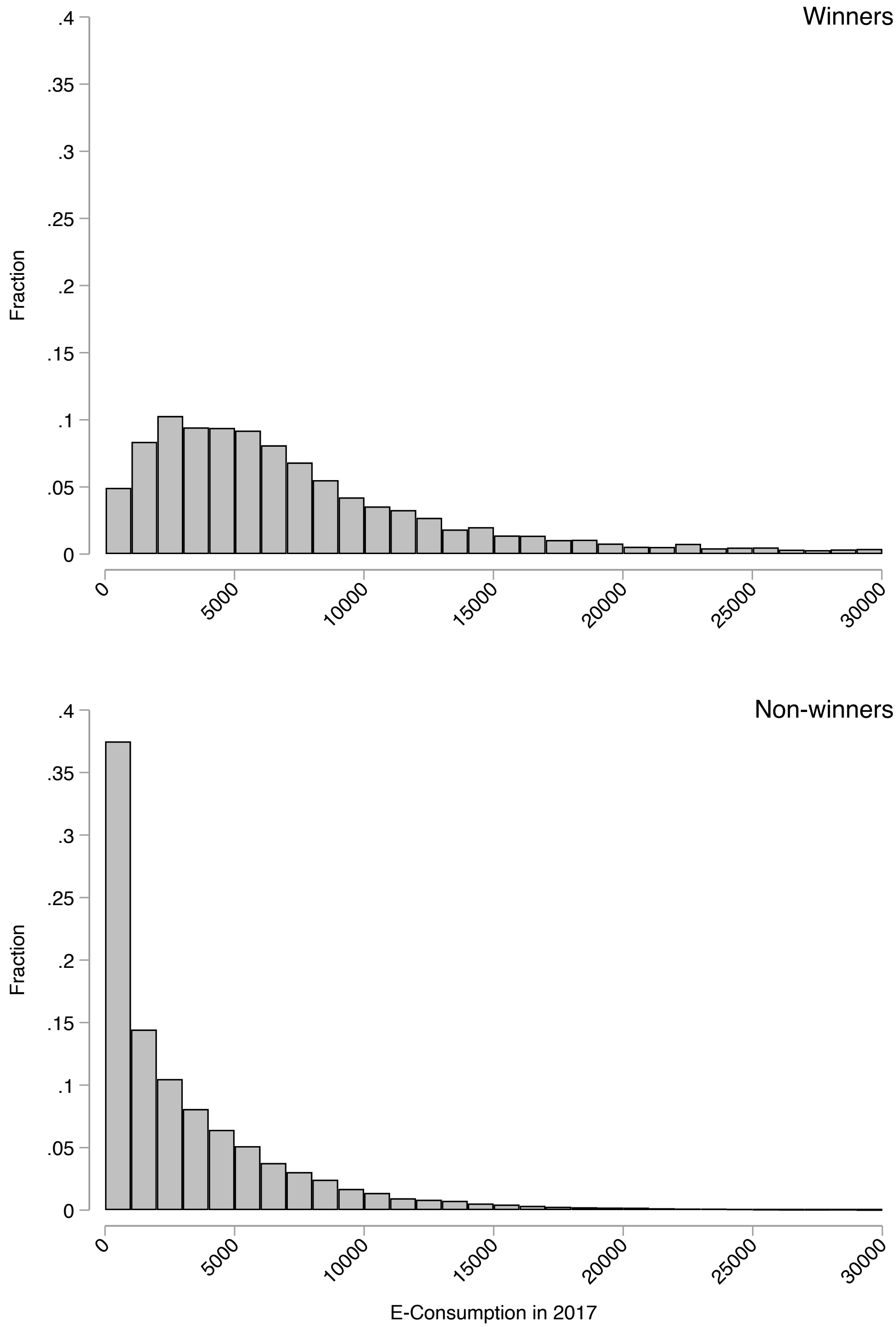

*Notes:* The figure presents monthly differences in VAT revenue between tax offices with many winners (solid line) and tax offices with few winners (dashed line). They are obtained from fitted values in a Poisson regression with interacted months (using absolute values) and after controlling for month and tax offices fixed effects. Regression estimates are shown in Table A.3, taking as reference point December 2017, one period before receiving the lottery prize. The Few Winners(Many Winners) samples include tax offices with the 20 lowest(highest) winners as a percentage of the population. Monthly differences were normalised to 1 in August 2017. The y-axis shows the log-point increase/decrease with respect to that month. Confidence intervals are drawn at the 95% level with the only statistical significant difference being March 2018. Robust standard errors are clustered at the tax office level.

**Fig. A.6** Electronic Consumption (Unconditional Means)



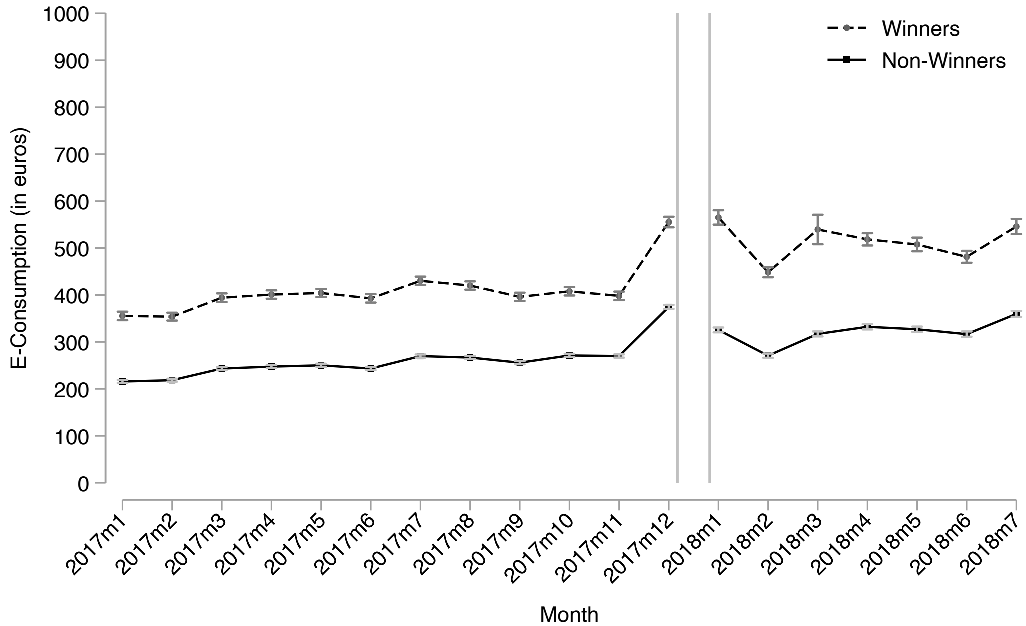

*Notes:* The figure shows unconditional mean electronic consumption for winners (dashed line) and non-winners (solid line). Winners include the 9,000 superdraw winners and non-winners the 50,000 randomly drawn sample. Electronic consumption is shown in the y-axis and months in the x-axis. The sample includes 19 months in total, where their electronic consumption is observed. The vertical lines indicate the period of the superdraw (24th of December 2017).

**Fig. A.7** Electronic Consumption in 2017



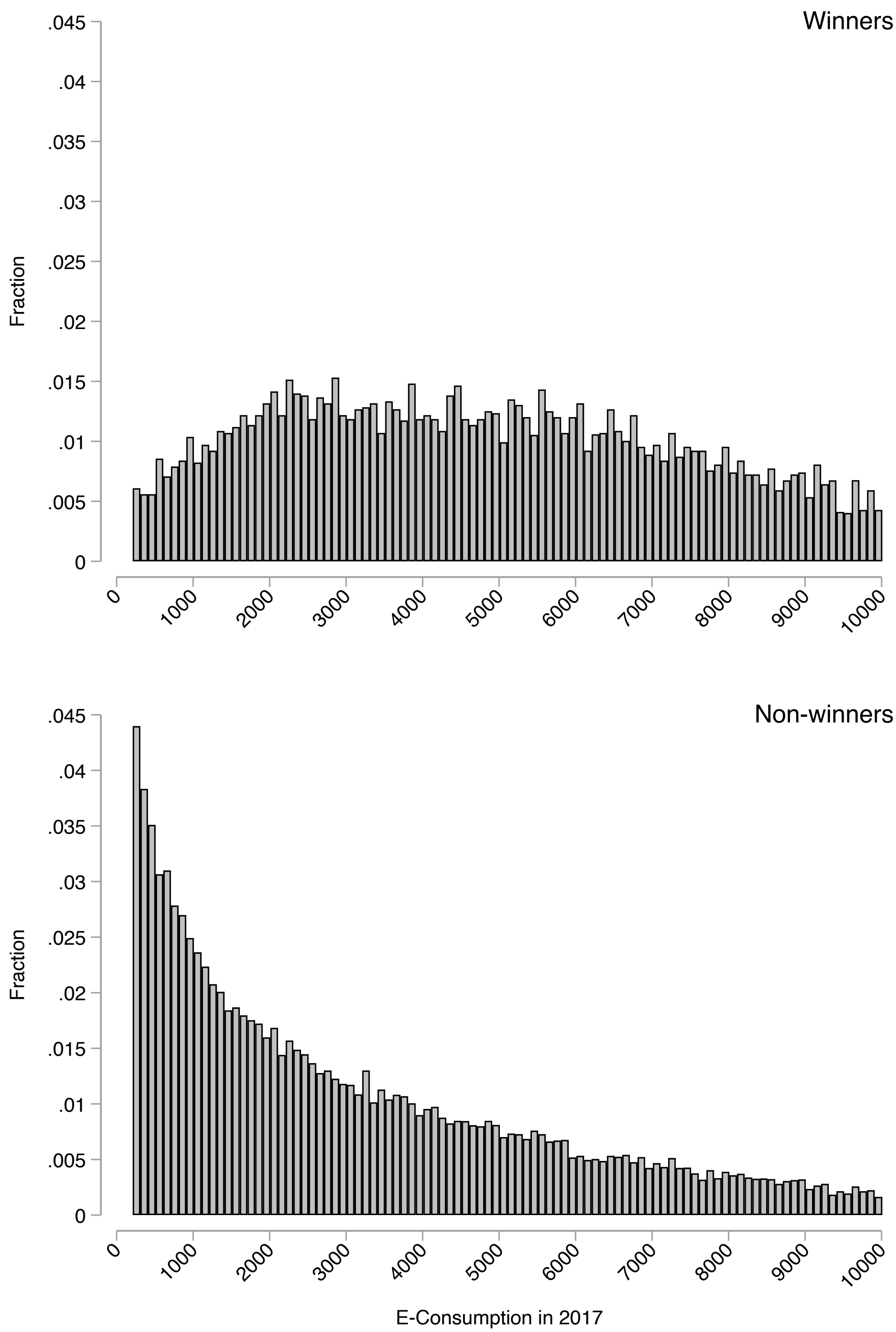

*Notes:* The figures plot the distributions of annual electronic consumption in 2017 for winners (top panel) and non-winners (bottom panel). The annual electronic consumption is shown in the x-axis. The sample includes 9,000 winners and 50,000 non-winners. The bins of the distributions are at €1,000 and both are trancated at €30,000 for illustration purposes.

**Fig. A.8** Electronic Consumption - Winsorized Sample (Unconditional Means)



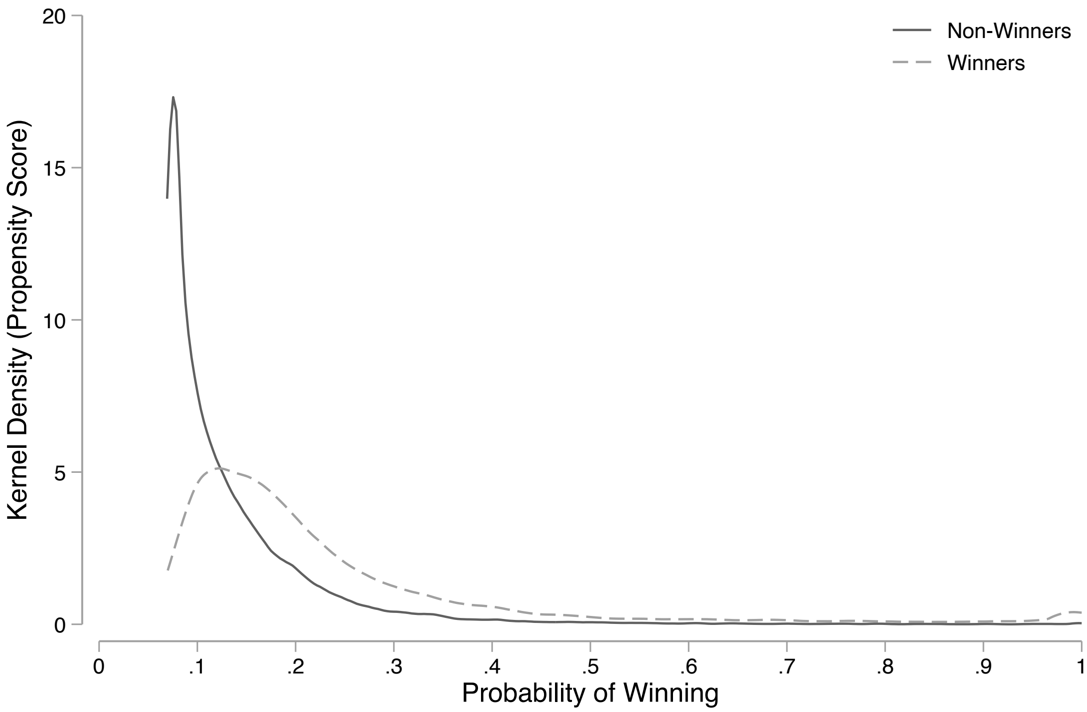

*Notes:* The figure shows unconditional mean electronic consumption for winners (dashed line) and non-winners (solid line). The sample is winsorized at top/bottom 10% based annual electronic consumption. Electronic consumption is shown in the y-axis and months in the x-axis. The sample includes 19 months in total, where their electronic consumption is observed. The vertical lines indicate the period of the superdraw (24th of December 2017).

**Fig. A.9** Electronic Consumption in 2017 - Winsorized Sample



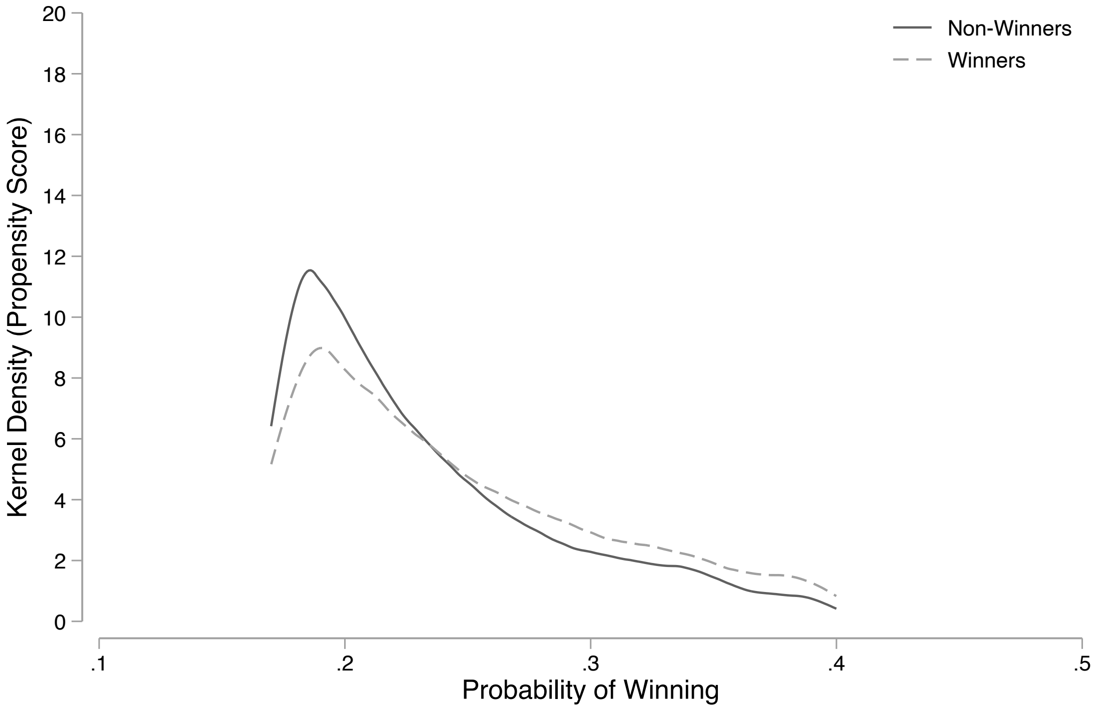

*Notes:* The figures plot the distributions of annual electronic consumption in 2017 for winners (top panel) and non-winners (bottom panel). The sample is winsorized at top/bottom 10% based annual electronic consumption. The annual electronic consumption is shown in the x-axis. The bins of the distributions are at €100 and both are trancated at €10,000 for illustration purposes.

**Fig. A.10** Propensity Score - Whole sample



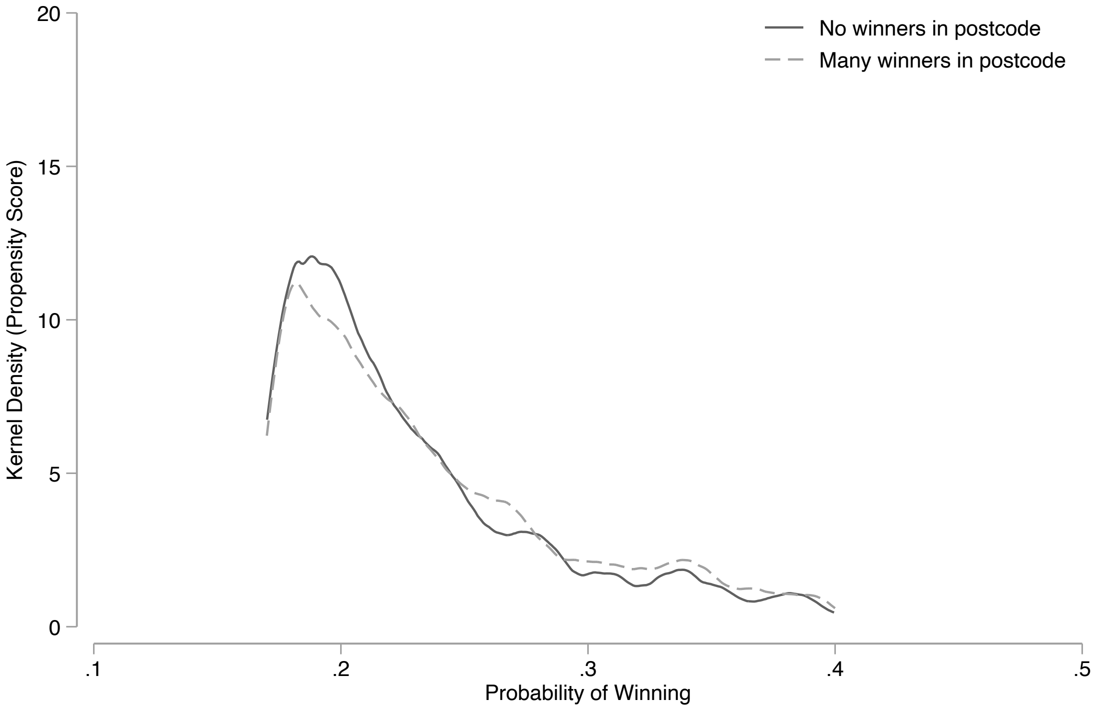

*Notes:* The graph plots kernel density functions of the propensity scores generated by Equation 2 for winners (dashed line) and non-winners (solid line). The propensity scores indicate the probability of individuals in winning in the superdraw based on the tickets generated in the months of January to September 2017 (which corresponded to the lottery draws).

**Fig. A.11** Propensity Score - Matching e-consumption



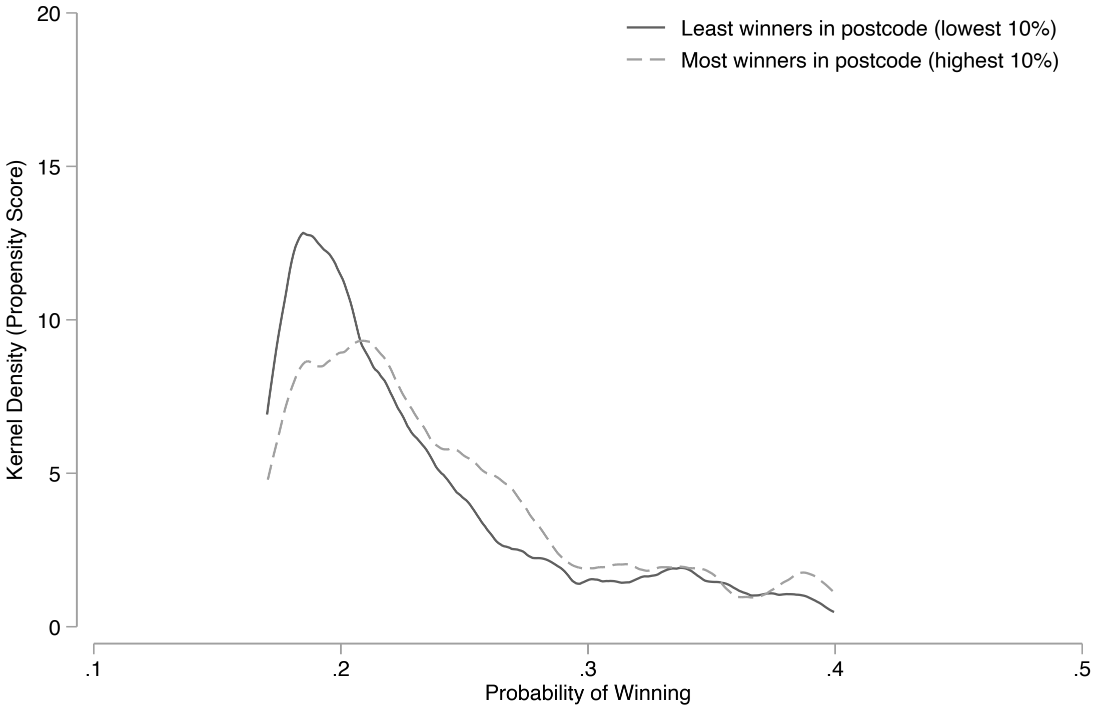

*Notes:* The graph plots kernel density functions of the propensity scores generated by Equation 2 for winners (dashed line) and non-winners (solid line). The propensity scores indicate the probability of individuals in winning in the superdraw based on the tickets generated in the months of January to September 2017 (which corresponded to the lottery draws). The sample includes individuals with propensity score between 0.17 and 0.4, to create samples of winners and non-winners with the same probability of winning.

**Fig. A.12** Propensity Score - Non-winners with no winners in postcode




*Notes:* The graph plots kernel density functions of the propensity scores generated by Equation 2 for non-winners with many winners in their postcode (dashed line) and non-winners with no (or very few) winners in their postcode (solid line). The first group is generated using non-winners residing in postcodes with a percentage of winners in their population of 0.3% or higher. The second group is generated using non-winners residing in postcodes with a percentage of winners in their population of 0.03% or lower. The propensity scores indicate the probability of individuals in winning in the superdraw based on the tickets generated in the months of January to September 2017 (which corresponded to the lottery draws). The samples of winners and non-winners are truncated at 0.17 and 0.4 to create comparable samples.

**Fig. A.13** Propensity Score - Non-winners with 10% lowest winners in postcode




*Notes:* The graph plots kernel density functions of the propensity scores generated by Equation 2 for non-winners with the most winners in their postcode (dashed line) and non-winners with the least winners in their postcode (solid line). Firstly, the postcodes are ranked by the percentage of winners over the postcodes' population. Then, the first group is generated using non-winners residing in postcodes with the lowest 10%. The second group is generated using non-winners residing in postcodes with the highest 10%. The propensity scores indicate the probability of individuals in winning in the superdraw based on the tickets generated in the months of January to September 2017 (which corresponded to the lottery draws). The samples of winners and non-winners are truncated at 0.17 and 0.4 to create comparable samples.

## A.2 Tables

**Table A.1** Summary Statistics

| | Winners | | | Non-winners | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | N | Mean | St.Dev. | N | Mean | St.Dev. |
| E-Consumption | | | | | | |
| Jan 2017 | 7,748 | 1,088 | 8,265 | 44,383 | 275 | 925 |
| Feb 2017 | 7,748 | 1,088 | 7,708 | 44,383 | 273 | 712 |
| Mar 2017 | 7,748 | 1,257 | 8,966 | 44,383 | 308 | 838 |
| Apr 2017 | 7,748 | 1,197 | 8,020 | 44,383 | 307 | 754 |
| May 2017 | 7,748 | 1,370 | 11,428 | 44,383 | 316 | 702 |
| Jun 2017 | 7,748 | 1,246 | 9,053 | 44,383 | 305 | 714 |
| Jul 2017 | 7,748 | 1,271 | 8,572 | 44,383 | 334 | 922 |
| Aug 2017 | 7,748 | 1,224 | 8,987 | 44,383 | 331 | 945 |
| Sept 2017 | 7,748 | 1,152 | 8,673 | 44,383 | 319 | 750 |
| Oct 2017 | 7,748 | 1,021 | 8,319 | 44,383 | 329 | 741 |
| Nov 2017 | 7,748 | 1,144 | 10,414 | 44,383 | 331 | 817 |
| Dec 2017 | 7,748 | 1,349 | 11,135 | 44,383 | 441 | 1,086 |
| Jan 2018 | 7,748 | 1,357 | 11,153 | 44,383 | 392 | 1,494 |
| Feb 2018 | 7,748 | 1,004 | 8,659 | 44,383 | 321 | 913 |
| Mar 2018 | 7,748 | 1,184 | 10,138 | 44,383 | 373 | 1,041 |
| Apr 2018 | 7,748 | 1,188 | 10,593 | 44,383 | 389 | 899 |
| May 2018 | 7,748 | 1,189 | 10,064 | 44,383 | 393 | 1,089 |
| Jun 2018 | 7,748 | 1,151 | 10,298 | 44,383 | 377 | 990 |
| Jul 2018 | 7,748 | 1,242 | 10,883 | 44,383 | 422 | 1,253 |

*Notes:* The table presents summary statistics for the two samples of winners (Columns (1) - (3)) and non-winners (Columns (4) - (5)). Statistics are shown per month of electronic consumption. Overall we observe the electronic consumption in 19 months (12 before the superdraw and 7 after). Columns (1) and (4) present the number of individuals in the sample. The initial sample included 9,000 winners and 50,000 individuals. From these, I exclude winners and non-winners with 0 consumption in 2017 (not participating in the lottery) and individuals with income from business. The latter oftentimes use their personal bank accounts for professional purposes, thus generating large amounts of electronic consumption, which are not comparable to other individuals. Columns (2) and (5) present mean electronic consumption values and Columns (3) and (6) to standard deviations for winners and non-winners respectively. These correspond to the plot in Figure A.6.

**Table A.2** Effect of Winning on VAT Revenue - Detailed

| | Monthly | | | Quarterly | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | Log Revenue | Log Revenue | Log Revenue | Log Revenue | Log Revenue | Log Revenue |
| Superdraw Winners | 0.0017*** | 0.0012*** | 0.0010** | 0.0023*** | 0.0012*** | 0.0012*** |
| | (0.0004) | (0.0004) | (0.0004) | (0.0005) | (0.0004) | (0.0004) |
| Lagged E-Consumption: | | | | | | |
| -4 months | | | -0.1895 | | | |
| | | | (0.1196) | | | |
| -5 months | | | -0.0902 | | | |
| | | | (0.1132) | | | |
| -6 months | | | -0.0382 | | | |
| | | | (0.1291) | | | |
| -7 months | | | 0.0497 | | | |
| | | | (0.1280) | | | |
| -8 months | | | 0.1922 | | | |
| | | | (0.1754) | | | |
| -9 months | | | 0.4956*** | | | |
| | | | (0.1651) | | | |
| -10 months | | | 0.3780** | | | |
| | | | (0.1804) | | | |
| -11 months | | | 0.4253** | | | |
| | | | (0.1906) | | | |
| -12 months | | | 0.2366 | | | |
| | | | (0.1802) | | | |
| -2 quarters | | | | | | -0.3712* |
| | | | | | | (0.2138) |
| -3 quarters | | | | | | 0.2114 |
| | | | | | | (0.1506) |
| -4 quarters | | | | | | -0.1922 |
| | | | | | | (0.1947) |
| Constant | 14.4859*** | 14.3048*** | -10.0551 | 15.8568*** | 15.7121*** | 22.0163*** |
| | (0.0031) | (0.0045) | (8.6603) | (0.0092) | (0.0143) | (7.8097) |
| Tax Office FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Time FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1248 | 768 | 768 | 384 | 192 | 192 |
| Tax Offices | 96 | 96 | 96 | 96 | 96 | 96 |
| Adj. R-squared | 0.90 | 0.90 | 0.90 | 0.89 | 0.96 | 0.96 |

*Notes:* The table presents estimates from Regression 1. The table is similar to Table 4.1, with the addition of controls for past values, that are shown here in detail. The variable "Superdraw Winners" corresponds to the variation of winners in tax offices following the superdraw. For all regressions tax office fixed effects, time fixed and robust standard errors clustered at the tax office level are used. Columns (1), (2) and (3) use 13 months of VAT observations in 96 tax offices. Columns (4), (5) and (6) use four quarterly observations. All regressions present estimates of the association of winners and VAT revenue in logarithmic form. Column (1) and (4) include regressions without lagged electronic consumption values (no controls). Columns (2) and (5) include time observations *only* after the superdraw (i.e., in the months or quarters in 2018). Columns (3) and (6) correspond to the full specification of Regression 1 at the monthly and quarterly level respectively. These include lagged e-consumption values, resulting in the same observations as Columns (2) and (5) in the monthly and quarterly regression respectively.

**Table A.3** Few versus Many Winners (Bottom/Top 20% of Tax Offices)

| | Linear<br>(1)<br>Log VAT Revenue | Poisson<br>(2)<br>VAT Revenue |
|---|---|---|
| Many Winners Interaction with: | | |
| August 2017 | -0.0088 | -0.2571* |
| | (0.0145) | (0.1442) |
| Septempter 2017 | -0.0100 | -0.1361 |
| | (0.0096) | (0.1146) |
| October 2017 | -0.0039 | -0.1682* |
| | (0.0077) | (0.0893) |
| November 2017 | 0.0014 | -0.1103 |
| | (0.0092) | (0.1116) |
| January 2017 | 0.0083 | 0.0141 |
| | (0.0105) | (0.1047) |
| February 2018 | 0.0115 | 0.0919 |
| | (0.0114) | (0.1250) |
| March 2018 | 0.0081* | 0.1360** |
| | (0.0045) | (0.0561) |
| April 2018 | 0.0119 | 0.0157 |
| | (0.0106) | (0.1167) |
| May 2018 | 0.0112 | 0.0481 |
| | (0.0104) | (0.1166) |
| June 2018 | 0.0050 | 0.0786 |
| | (0.0037) | (0.0544) |
| July 2018 | 0.0002 | -0.0406 |
| | (0.0103) | (0.1153) |
| August 2018 | -0.0082 | -0.2445* |
| | (0.0133) | (0.1418) |
| Constant | 1.0988*** | 1.6200*** |
| | (0.0028) | (0.0240) |
| Tax Office FE | Yes | Yes |
| Time FE | Yes | Yes |
| Number of Observations | 520 | 520 |
| Number of Tax Offices | 40 | 40 |

*Notes:* The table presents estimates that correspond to Figure 4.1 in Column (1) and Figure A.5 in Column (2). These are generated from Regression 1 using samples of tax offices with the least and the most winners in their population. To generate the samples, the tax offices are ranked according to the percentage of winners in their population. The bottom 20% form the group with least winners and the top 20% form the group with most winners. The first group includes 21 tax offices and the second group 19 tax offices. Column (1) presents estimates generated from a linear regression of VAT revenue in logarithmic form. Column (2) presents estimates from a Poisson regression using the absolute values of VAT revenue. The coefficients of the former are in log-points. For the latter, they can be interpreted as percentages. The regressions include 12 periods of VAT revenue from August 2017 to August 2018, with December 2017 dropped, since the superdraw took place at the 24th of December 2017 with the awarded in early January 2018. Robust standard errors are clustered at the tax office level.

**Table A.4** Logistic Regression - Probability of Winning

| | $P(W_n = 1)$ |
|---|---|
| Tickets in: | |
| January | 0.0003631*** |
| | (0.0000755) |
| February | 0.0005676*** |
| | (0.0000800) |
| March | 0.0002974*** |
| | (0.0000670) |
| April | 0.0005647*** |
| | (0.0000810) |
| May | 0.0004566*** |
| | (0.0000745) |
| June | 0.0004548*** |
| | (0.0000841) |
| July | 0.0004820*** |
| | (0.0000808) |
| August | 0.0002367*** |
| | (0.0000780) |
| September | 0.0003536*** |
| | (0.0000738) |
| Constant | -2.5983843*** |
| | (0.0210077) |
| Number of Individuals | 52,131 |

*Notes:* The table presents estimates from the logistic regression in Equation 2. This is used to generate the propensity score of winning the lottery. The months used correspond to the months that generated the tickets for the superdraw, from January to September 2017. These are regressed on the sample of 7,748 winners (assigned the value of 1) and 44,383 non-winners (assigned the value of 0). The total sample is 52,131. Winning was a rare event, hence to ensure convergence of the maximum-likelihood function, a Firth logistic regression is used. The positive values indicate the percentage increase in the probability of winning of one extra ticket in each of the months. The regression produces propensity scores, which are plotted for winners and non-winners in Figure A.10.

**Table A.5** Propensity Score - Kernel Density Functions

Whole Sample

| | N | Mean | p10 | p25 | p50 | p75 | p90 |
|---|---|---|---|---|---|---|---|
| Non-winners | 44,383 | 0.133 | 0.071 | 0.078 | 0.103 | 0.154 | 0.226 |
| Winners | 7,748 | 0.237 | 0.097 | 0.126 | 0.177 | 0.265 | 0.431 |

Matching Sample

| | N | Mean | p10 | p25 | p50 | p75 | p90 |
|---|---|---|---|---|---|---|---|
| Non-winners | 8,144 | 0.233 | 0.177 | 0.190 | 0.217 | 0.263 | 0.321 |
| Winners | 3,249 | 0.245 | 0.179 | 0.195 | 0.229 | 0.284 | 0.341 |

*Notes:* The tables presents comparisons of the kernel density functions produced by Equation 2 and plotted in Figure A.10 for the top "Whole Sample" panel and in Figure A.11 in the bottom "Matching Sample" panel. The top panel includes the whole sample and the bottom one is truncated at 0.17 and 0.4 of propensity scores. This produces a matching sample. The columns present the number of individuals in the samples, the means and values at different percentiles.

## Table A.6 Winners E-Consumption

| | Linear (1) Log E-Consumption | Poisson (2) E-Consumption |
|---|---|---|
| Winners Interaction with: | | |
| January 2017 | -0.0201 | 0.00285 |
| | (0.0323) | (0.0233) |
| February 2017 | -0.0128 | -0.00975 |
| | (0.0291) | (0.0220) |
| March 2017 | -0.0369 | -0.00468 |
| | (0.0284) | (0.0235) |
| April 2017 | -0.0297 | -0.0145 |
| | (0.0266) | (0.0211) |
| May 2017 | -0.0741*** | -0.0506** |
| | (0.0272) | (0.0217) |
| June 2017 | -0.0302 | -0.0210 |
| | (0.0257) | (0.0206) |
| July 2017 | -0.0402 | -0.0307 |
| | (0.0262) | (0.0208) |
| August 2017 | -0.0345 | -0.0260 |
| | (0.0262) | (0.0206) |
| September 2017 | -0.0533** | -0.0431** |
| | (0.0263) | (0.0208) |
| October 2017 | -0.0255 | -0.0223 |
| | (0.0261) | (0.0220) |
| November 2017 | -0.0265 | -0.0416** |
| | (0.0229) | (0.0211) |
| January 2018 | 0.129*** | 0.101*** |
| | (0.0225) | (0.0224) |
| February 2018 | 0.115*** | 0.0862*** |
| | (0.0238) | (0.0255) |
| March 2018 | 0.0635** | 0.0562 |
| | (0.0265) | (0.0368) |
| April 2018 | 0.0519** | -0.00235 |
| | (0.0250) | (0.0220) |
| May 2018 | 0.0591** | 0.0242 |
| | (0.0262) | (0.0282) |
| June 2018 | 0.0420 | 0.0142 |
| | (0.0275) | (0.0229) |
| July 2018 | -0.00115 | -0.00386 |
| | (0.0287) | (0.0291) |
| Constant | 6.543*** | 6.927*** |
| | (0.00938) | (0.00795) |
| Tax Office FE | Yes | Yes |
| Month FE | Yes | Yes |
| Number of obs. | 212243 | 212243 |
| Number of Individuals | 11174 | 11174 |

*Notes:* The table presents monthly difference-in-difference estimates from Regression 3. Column (1) presents estimates of a linear regression on electronic consumption in logarithmic form and Column (2) of a Poisson regression on electronic consumption (using absolute values). The estimates of the former can be interpreted as log-point differences and of the latter as percentage differences. Results are plotted in Figure 5.1. The regressions use inverse probability weights generated from the propensity scores to control for the level of electronic spending, which determines the individuals' probability of winning. For both samples the propensity scores that generate the inverse probability weights correspond to those illustrated in Figure A.11. Robust standard errors are clustered at the individual level.

## Table A.7 Non-winners E-Consumption

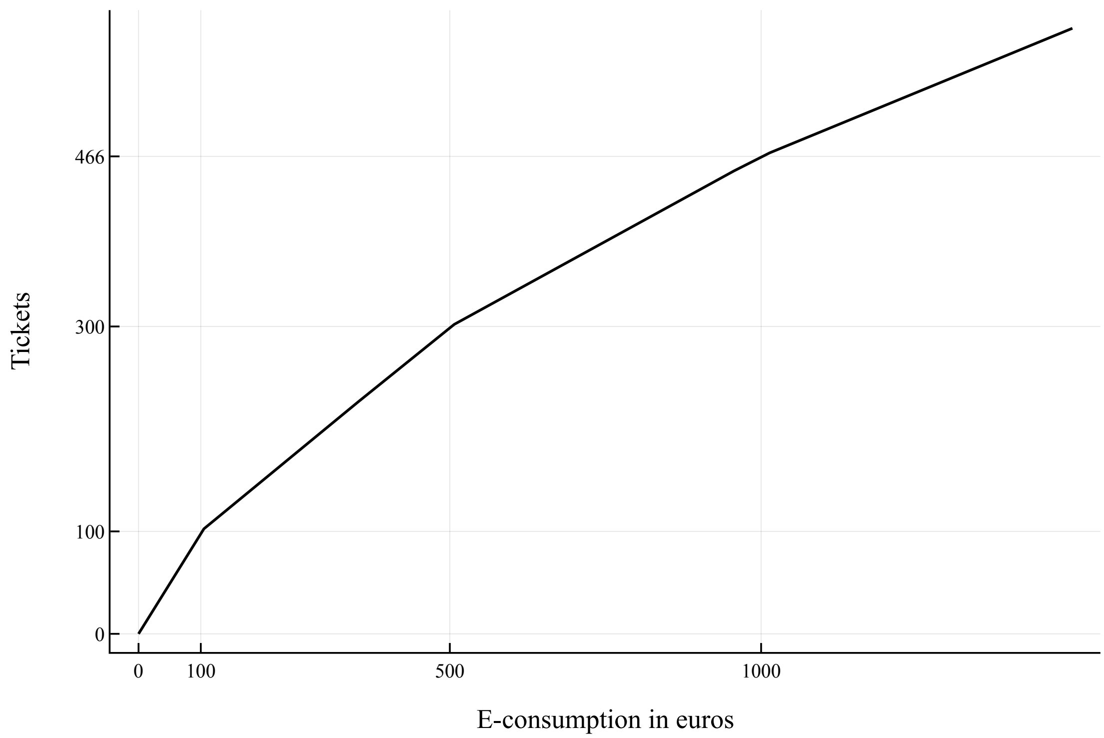
| | Lowest/Highest 10% (1) Log E-Consumption | No/Many Winners (2) Log E-Consumption |
|---|---|---|
| Non-winners with Many Winners in Postcode Interaction with: | | |
| January 2017 | 0.00972 | 0.0771 |
| | (0.0796) | (0.0792) |
| February 2017 | 0.0322 | 0.135* |
| | (0.0729) | (0.0692) |
| March 2017 | 0.134* | 0.116* |
| | (0.0761) | (0.0695) |
| April 2017 | 0.0706 | 0.137** |
| | (0.0663) | (0.0586) |
| May 2017 | 0.0858 | 0.0771 |
| | (0.0717) | (0.0658) |
| June 2017 | 0.0216 | 0.0701 |
| | (0.0682) | (0.0620) |
| July 2017 | 0.0808 | 0.0927 |
| | (0.0738) | (0.0648) |
| August 2017 | 0.0814 | 0.178*** |
| | (0.0750) | (0.0670) |
| September 2017 | 0.0581 | 0.0131 |
| | (0.0827) | (0.0689) |
| October 2017 | 0.116 | 0.107* |
| | (0.0703) | (0.0627) |
| November 2017 | 0.0251 | 0.0483 |
| | (0.0715) | (0.0596) |
| January 2018 | 0.00910 | 0.0106 |
| | (0.0729) | (0.0614) |
| February 2018 | 0.136* | 0.0914 |
| | (0.0761) | (0.0664) |
| March 2018 | 0.0819 | -0.0335 |
| | (0.0877) | (0.0708) |
| April 2018 | 0.0692 | 0.0796 |
| | (0.0784) | (0.0657) |
| May 2017 | 0.135* | 0.125* |
| | (0.0716) | (0.0670) |
| June 2017 | 0.187** | 0.195*** |
| | (0.0791) | (0.0704) |
| July 2017 | 0.123 | 0.180** |
| | (0.0805) | (0.0713) |
| Constant | 6.554*** | 6.504*** |
| | (0.0238) | (0.0194) |
| Tax Office FE | Yes | Yes |
| Month FE | Yes | Yes |
| Postcode FE | Yes | Yes |
| Number of Observations | 29507 | 43206 |

*Notes:* The table presents monthly difference-in-difference estimates from Regression 4. Column (1) includes a sample with the lowest/highest percentage of winners in their population as illustrated in Figure 6.1. Column (2) includes a sample with no winners and many winners. The former includes postcodes that experienced no winners at all, and the latter postcodes that experienced more than 0.3% of winners in their population. Point estimates correspond to Figure 6.3. The regressions use inverse probability weights generated from the propensity scores to control for the level of electronic spending, which determines the individuals' probability of winning. For the sample in Column (1) the propensity scores correspond to those illustrated in Figure A.13. For Column (2) the propensity scores correspond to Figure A.12. The depended variable is electronic consumption of individuals in logarithmic form. Robust standard errors are clustered at the postcode level.

# B Tax Lottery Information

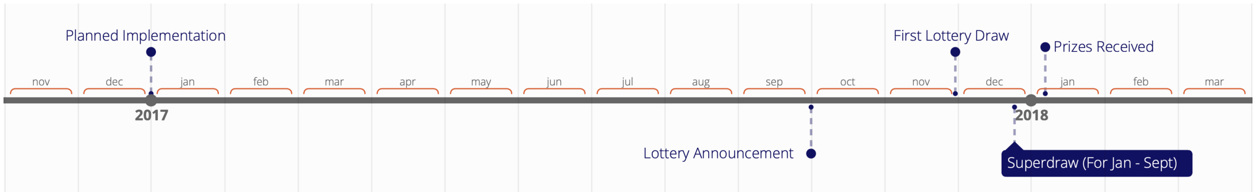
## B.1 Lottery

**Fig. B.1** Euro-to-Ticket Scale




*Notes:* The graph illustrates the scale used to convert the aggregate level of monthly electronic consumption to eligible tickets in the lottery. Banks sent the aggregate level of electronic consumption completed by each individual and this is converted to ticket using the following scale. At €1-100, tickets correspond at 1 for every €1. At €101-500, tickets correspond at 1 for every €2. At €501-1,000, tickets correspond at 1 for every €3. For over 1,000, tickets correspond at 1 for every €4. There was no upper limit in tickets. Details about eligible payments and additional information on the institutional structure are explained in Section 2.

**Fig. B.2** Superdraw Timeline




*Notes:* The figure shows an indicative timeline of the superdraw that took place on Christmas Eve 2017. The planned implementation was January 2017. The lottery announcement took place in October 2017 with the first draw taking place at the end of November 2017 for payments completed in October. The superdraw took place on the 24th of December 2017, for payments corresponding to months of January to September 2017. Prizes were handed out directly to the individuals' bank accounts in early January 2017.

**Fig. B.3** Tickets Example

| α/α Κληρ. | Ημ/νία Κλήρωσης | Περίοδος Συν/γών | Ποσό Συναλλαγών | Πλήθος Λαχνών | από ... λαχνό | έως ... λαχνό |
|---|---|---|---|---|---|---|
| 9 | 24/12/2017 | 2017-02 | 330,72 | 216 | 7572 3176 | 7572 3391 |
| | | | α/α και λαχνός που κερδίζει: | | 0 | 0 |
| 8 | 24/12/2017 | 2017-03 | 289,79 | 195 | 1 6336 2988 | 1 6336 3182 |
| | | | α/α και λαχνός που κερδίζει: | | 0 | 0 |
| 7 | 24/12/2017 | 2017-04 | 1.185,64 | 514 | 5239 1026 | 5239 1539 |
| | | | α/α και λαχνός που κερδίζει: | | 0 | 0 |
| 6 | 24/12/2017 | 2017-05 | 210,89 | 156 | 6 3365 9750 | 6 3365 9905 |
| | | | α/α και λαχνός που κερδίζει: | | 0 | 0 |
| 5 | 24/12/2017 | 2017-06 | 447,73 | 274 | 1 4227 5137 | 1 4227 5410 |
| | | | α/α και λαχνός που κερδίζει: | | 0 | 0 |
| 4 | 24/12/2017 | 2017-07 | 934,60 | 445 | 6136 0963 | 6136 1407 |
| | | | α/α και λαχνός που κερδίζει: | | 0 | 0 |
| 3 | 24/12/2017 | 2017-08 | 944,54 | 449 | 10 6711 7986 | 10 6711 8434 |
| | | | α/α και λαχνός που κερδίζει: | | 0 | 0 |
| 2 | 24/12/2017 | 2017-09 | 236,41 | 169 | 10 3466 3194 | 10 3466 3362 |
| | | | α/α και λαχνός που κερδίζει: | | 0 | 0 |
| 1 | 30/11/2017 | 2017-10 | 50,54 | 51 | 8 4499 7607 | 8 4499 7657 |
| | | | α/α και λαχνός που κερδίζει: | | 0 | 0 |
| Number of Draw | Draw Date | E-Consumption Period | E-Consumption Amount | Number of Tickets | Starting ticket number | ... to ticket number |

*Notes:* The picture shows a real example of a Greek taxpayer who took part in the lottery. The first column shows the number of draws ranked by the date that these took place. The draw date indicates the exact date of the draw and the corresponding consumption period in which the tickets were generated. Notice the superdraw taking place on Christmas Eve for transactions that took place in previous months. The 4th column shows the electronic payments transfered from the banks to the tax authority and the 5th column the awarded tickets after the euro-to-ticket conversion is applied as illustrated in Figure B.1. The last two columns indicate the corresponding ticket numbers and the red number shows the winning tickets (0 in this case). This information is accessible to each individual via a dedicate website.

## B.2 Sampling of Winners and Non-Winners

In order to make the two samples comparable I utilise the total number of lottery tickets issued in each calendar month, $\bar{T}_m$. Given that lottery tickets are derived from monthly e-transactions, one can compute $T_{i,m,s}$, the number of tickets from individual $i$ in month $m$ in sample $s$, where $s \in \{1, 2\}$ indicates the winner and non-winner sample, respectively. In addition, non-winners in 2017, who were winners in 2018 must be added in the expansion. To avoid a different subscript for the year, $\hat{T}_{i,m,1}$ is used instead.

Given this, the following identity must hold:

$$\sum_{m=1}^{12} \bar{T}_m = \sum_{m=1}^{12} \sum_{i=1}^{N_1} T_{i,m,1} + \sum_{m=1}^{12} \sum_{i=1}^{N_1} \hat{T}_{i,m,1} + \omega \sum_{m=1}^{12} \sum_{i=1}^{N_2} T_{i,m,2} \tag{5}$$

where $N_s$ indicates the size of the samples $s$ (with $N_1 = 18,897$ and $N_2 = 50,000$).

From this, it is straightforward to derive $\omega$, the *weight or expansion factor* used to arrive at a sample that matches the population in terms of lottery tickets, since it is the only unknown. Both the total number of tickets in 2017, $\sum_{m=1}^{12} \bar{T}_m$ and the total number of tickets in the two samples of winners and non-winners are known.

A further plausibility check is that $N_1 + \omega N_2 \cong N$. The calculation derives $\omega$ to be 129. Expanding the random sample gives a total taxpayer population of 6.45 million (50,000×129), to which 18,897 winners are added. This is very close to official statistics from the tax authority, indicating 6.37 million tax returns being filed for 2017.[25]

[25] Annual statistics for the 2017 filing are published by the Greek tax authority at https://www.aade.gr/menoy/statistika-deiktes/eisodima/etisia-statistika-deltia.