



Essays on the Economics of Taxation

Berlin, June 2023

Panayiotis Nicolaides

Dissertation submitted to the Hertie School
in partial fulfilment of the requirements for the degree of

Doctor rerum politicarum (Dr. rer. pol.)

in the

Doctoral Programme in Governance

Advisors

First advisor

Prof. Dr. Christian Traxler

Hertie School

Second advisor

Prof. Gabriel Zucman, PhD

Paris School of Economics

Third Advisor

Prof. Niels Johannesen, PhD

University of Copenhagen

Fourth Advisor

Prof. Anne Brockmeyer, PhD

Institute for Fiscal Studies &
University College London

A Tax Lottery Story

(in place of acknowledgements)

In December 2013, during a meeting at the 10th floor of the Charlemagne Building in Brussels, I was introduced for the very first time to the idea of a tax lottery in Greece. Georges Siotis was on leave from Carlos III University of Madrid and involved with a technical assistance department of the European Commission - the Task Force for Greece - that was set up during the Greek economic crisis. Going through schemes that could provide positive incentives to improve tax compliance, he handed me a large printed file of a project in a blue European Commission folder, told me to have a look and asked me if I would be interested to work on the ground in Athens. Little did I know at the time that 10 years later I would be submitting a Ph.D thesis on this very topic.

The first tax lottery attempt in Greece was named "@podeixi" (the Greek for "receipt"), the inventor of which I met later in May 2014 in a pan-European meeting of tax lotteries, that took place at an old nuclear power plant in Ispra, Italy. The antithesis between the beauty of Lake Como on one hand and the industrial side of the workshop on the other, drew many parallels to the attempts of tax administrations to implement these schemes in the middle of the European debt crisis. Portuguese colleagues shared their success in introducing "Fatura da Sorte", and disappointed Greek colleagues discussed the problems faced in implementing "@podeixi" during a deep economic crisis.

Indeed, together with Georges, we would soon find out of the difficulties in convincing people about the lottery in the middle of the crisis in 2014: the Greek tax lottery could not gain traction during the government at the time. Its sole appearance in official documents was in a leaked infamous email by the then Minister of Finance Gkikas Hardouvelis, which included hard fiscal measures to bridge a gap in negotiations between Greece and its creditors. Admittedly, not the most charming place for a positive-incentives scheme! Contrary to our attempts for the lottery to gather pace, it gathered dust instead as it stayed on the shelves for years. In the meantime, I moved out of Greece with capital controls being implemented in the summer of 2015 and having lost all hope.

But since ideas die hard, so did the lottery. My phone rang in the autumn of 2015. At the other end was newly-appointed Alternate Minister of Finance, George Chouliarakis, who asked if I was interested in returning to Athens. "When?", I asked back. "Whenever you can" he said with a small pause and then continued "What about in a week?". With two suitcases packed, I set off to undust some old files. Luckily, on the reforms front things looked more positive: the tax authority became independent in 2017 and this led the way for its own budget and a tax lottery being implemented in the autumn of 2017.

It was only that autumn in 2017 that the idea to work on tax lottery *research* came by. Starting a Ph.D at the Hertie School that year, Christian Traxler planted a seed in my head: "Start from recent policies", he said, "what about this new tax lottery in Greece? I

have read that it has just started". Of course, being a project I knew too well, the idea spoke to my heart and a data request was soon filed.

For a year I waited patiently for the amazing staff of the Independent Authority for Public Revenue to complete the randomisation process, extract and anonymise data that would allow me to begin the analysis. Key to these developments were Ms Sofia Sechperidou, General Secretary for Public Revenue, and her colleagues, whom I cannot thank enough for the collaboration.

The data were finally received in October 2018 in an eventful way. On Friday, 4th of October 2018, I was expecting a call to receive them before flying out of Athens after a short visit. The call came around noon, mentioning that the offices were closing at 3pm. Suit and tie quickly on, I jumped in a taxi and headed to the tax authority, a few kilometers south of Athens, as if chased in a Hollywood movie. But to my luck, I got stuck midway in the most horrible of traffic jams. Desperate to reach the office before it was too late, I got out of the taxi, walked for 20 minutes under a hot Mediterranean sun and back in the first suburban train I could find. Not surprisingly, I arrived late and everyone (together with my hopes) were gone. Only a security guard alone in the entrance, and me, as if knowing the answer found the courage to ask:

- Are the offices closed?
- Yes, it's past 15pm, he replied. But who are you?
- I am here for some data. I told him.
- Wait a minute! He said.

And searching for some time, he pulled out a CD (double password-protected!) from under his desk and a paper for me to sign. Perplexed by what I experienced and holding the mysterious CD in my hands while having no CD-ROM in any known computer, I was in a desperate need to access the data. Salvation finally arrived in an internet cafe of a small Athenian neighbourhood nearby. One can only imagine the sight of a happy researcher in suit and tie shouting in joy while being watched by surprised gamers.

For these and for all, I am immensely and forever grateful to the main characters of this story, as well as to others who supported me along the way and who are too numerous to mention here. Georges Siotis, who introduced me to the Greek tax lottery and who shared his contacts generously in Greece and the EU, while supporting me despite being junior in a very challenging job. George Chouliarakis for giving me the chance to return to Greece, for closely collaborating in the economic stabilisation of the country, and for entrusting me with the most important files I have worked on in my life. Sofia Sechperidou and her colleagues at the tax authority for all the data work and cooperation. Christian Traxler for unlimited support, invaluable guidance and comments during the entire Ph.D time. Gabriel Zucman, for entrusting me with important tasks at the EU Tax Observatory, while providing the space and time to finish my research.

Subsequently, this thesis was written in Athens, Berlin, London and Paris; a truly European product, just as myself feel to have become. I am grateful to everyone who made this journey possible, enjoyable and eventful during the past 6 years.

Panayiotis Nicolaides
Paris, June 2023

Contents

1	The Political Economy of Tax Reforms during the Greek Economic Crisis	10
1.1	Economic Crisis and Adjustment Programmes	10
1.2	Tax Digitisation Reforms	12
1.3	A Tax Lottery on Electronic Payments	13
1.4	The Electronic Consumption Tax Discount	13
1.5	Conclusion	14
2	Are Tax Lotteries Regressive? Income, Consumption and, Occupational Characteristics of Winners	16
2.1	Introduction	16
2.2	The Greek Tax Lottery on Electronic Transactions	19
2.3	Data	23
2.4	Who Wins the lottery?	27
2.4.1	Descriptive Analysis	27
2.4.1.1	E-transaction and Income	27
2.4.1.2	Income Sources	30
2.4.1.3	Income Sources within Households	31
2.4.1.4	Descriptive Analysis Summary	32
2.4.2	Parametric Analysis	33
2.5	Lottery Reforms	35
2.5.1	Comparing Pre- and Post-Reform TAM	36
2.5.2	Other Reforms	39
2.6	Conclusion	40
2.7	Appendix	42

2.7.1	Complementary Figures and Tables	42
2.7.2	Construction of the Baseline Population of Tax Units	57
3	Income Tax Incentives for Electronic Payments: Evidence from Greece's Electronic Consumption Tax Discount	58
3.1	Introduction	58
3.2	Institutional Background	62
3.3	Conceptual Framework	64
3.3.1	Incentives for Electronic Consumption	64
3.3.2	Pre-filled Information and Adjustment Costs	66
3.4	Empirical Analysis	68
3.4.1	Data	68
3.4.2	Threshold Targeting	69
3.4.3	Responses in the Reporting Margin	71
3.4.4	Responses in the Electronic Consumption Margin	75
3.4.4.1	Empirical Strategy	75
3.4.4.2	Results	76
3.5	Interpretation	80
3.5.1	Marginal Cost Functions - Explicit Form	80
3.5.2	Diagrammatic Representation	82
3.6	Conclusion	84
3.7	Appendix	86
3.7.1	Additional Figures and Tables	86
3.7.2	ECTD Information	89
3.7.3	Proofs	90
3.7.3.1	Proposition 1	90
3.7.3.2	Proposition 2	90

4	Winning the Tax Lottery: Evidence from a Superdraw on Christmas Eve	91
4.1	Introduction	91
4.2	Institutional Background	94
4.3	Data	96
4.4	Effect of Winning on VAT Revenue	97
4.4.1	Identification Strategy	97
4.4.2	Results	98
4.5	Electronic Consumption of Winners	101
4.5.1	Propensity Score	102
4.5.2	Estimation	103
4.6	Spillovers to Non-Winners	105
4.6.1	Propensity Score	106
4.6.2	Estimation	106
4.7	Conclusion	110
4.8	Appendix: Figures and Tables	112
4.8.1	Figures	112
4.8.2	Tables	125
4.9	Appendix: Tax Lottery Information	132
4.9.1	Lottery	132
4.9.2	Sampling of Winners and Non-Winners	134
	Bibliography	135

List of Tables

2.1	Ticket-Awarding Mechanism (Jan 2017 – May 2019)	21
2.2	Summary Statistics: 2017- Winning Tax Units versus Population of Tax Unit	26
2.3	Estimation Results	34
2.4	Post-Reform Ticket-Awarding Mechanism (May 2019 onward)	35
2.5	Main Simulation Statistics	36
2.6	Main Simulation Statistics - Stricter Ticket Ceilings	40
2.7	Basic Summary Statistics	54
2.8	Non-parametric Estimates, by Primary Income Category	55
2.9	Estimates - Spouse's Primary Income Source	56
3.1	Threshold and ECTD	63
3.2	Income Tax in 2017	87
3.3	Sample Statistics	87
3.4	Event Studies - Monthly Consumption (% of Income)	88
4.1	Effect of Winning on VAT Revenue	99
4.2	Summary Statistics	125
4.3	Effect of Winning on VAT Revenue - Detailed	126
4.4	Few versus Many Winners (Bottom/Top 20% of Tax Offices)	127
4.5	Logistic Regression - Probability of Winning	128
4.6	Propensity Score - Kernel Density Functions	129
4.7	Winners E-Consumption	130
4.8	Non-winners E-Consumption	131

List of Figures

2.1	"Lottery" Google Trends in Greece	22
2.2	Distribution of Annual Electronic Transactions, 2017	28
2.3	Distribution of Declared Income 2017	29
2.4	Declared Annual Income and E-transaction by Income Sources	31
2.5	E-transaction Box Plots	37
2.6	E-transactions Distribution	38
2.7	Mean Annual E-transactions vs. Annual Declared Income	42
2.8	Declared Annual Income and E-transactions, Taxpayers with SB spouses	43
2.9	Declared Annual Income and E-transactions, Taxpayers with WG spouses	44
2.10	Declared Annual Income and E-transactions, Taxpayers with PE spouses	45
2.11	Declared Annual Income and E-transactions, Taxpayers with AG spouses	46
2.12	Income Distribution	47
2.13	E-transactions Distribution	48
2.14	Income Distribution for Stricter Ceilings	49
2.15	E-transactions Distribution for Stricter Ceilings	50
2.16	E-transactions Distribution (ranked by Income) for Stricter Ceilings	51
2.17	Winner's Distribution of Tickets in Simulations	52
2.18	Winner's Distribution of Tickets in Simulations	53
3.1	Threshold Targeting in Electronic Consumption	70
3.2	Within-Taxpayer Consumption Difference	72
3.3	Within-Taxpayer Difference - Below/Above Threshold	73
3.4	Within-Taxpayer Difference - Income Decomposition	74
3.5	Event Studies per Month of Reaching Threshold	78

3.6	Policy Attention without Liquidity Constraints	83
3.7	Policy Attention with Liquidity Constraints	84
3.8	Threshold Targeting - Truncation \pm €8000	86
3.9	Tax Filing - Codes 049 and 050	89
4.1	Regional Tax Offices with Few/Many Winners (Linear Regression)	100
4.2	Effect of Winning on E-Consumption	104
4.3	Event Study	105
4.4	Effect of Winning on Non-winners E-Consumption (10% Lowest/Highest)	107
4.5	Event Study (10% Lowest/Highest)	108
4.6	Effect of Winning on Non-winners E-Consumption (No winners)	109
4.7	Event Study (No winners)	110
4.8	Mean Tax Office Revenue	112
4.9	Variation of Winners in Tax Office Population	113
4.10	Variation of Winners in Postcode Population	114
4.11	Regional Tax Offices with Few/Many Winners (Unconditional Means)	115
4.12	Regional Tax Offices with Few/Many Winners (Poisson Regression)	116
4.13	Electronic Consumption (Unconditional Means)	117
4.14	Electronic Consumption in 2017	118
4.15	Electronic Consumption - Winsorized Sample (Unconditional Means)	119
4.16	Electronic Consumption in 2017 - Winsorized Sample	120
4.17	Propensity Score - Whole sample	121
4.18	Propensity Score - Matching e-consumption	122
4.19	Propensity Score - Non-winners with no winners in postcode	123
4.20	Propensity Score - Non-winners with 10% lowest winners in postcode	124
4.21	Euro-to-Ticket Scale	132
4.22	Superdraw Timeline	132
4.23	Tickets Example	133

Abbreviations

AADE	Independent Authority for Public Revenue (Greece)
AFM	Tax Identification Number (Greece)
AG	Agriculture (source of income)
DOY	Public Finance Service (tax collection office in Greece)
ECB	European Central Bank
ECTD	Electronic Consumption Tax Discount
EU	European Union
GDP	Gross Domestic Product
IBAN	International Bank Account Number
ID	Identification Document
IMF	International Monetary Fund
IRS	Internal Revenue Service (United States)
IT	Information Technology
NO	Null income
OLS	Ordinary Least Squares
PE	Pensions (source of income)
POS	Point of Sale
SB	Self-employed (source of income)
TAM	Ticket Awarding Mechanism (Greek tax lottery)
VAT	Value-added Tax
WG	Wages (source of income)

Chapter 1

The Political Economy of Tax Reforms during the Greek Economic Crisis

1.1 Economic Crisis and Adjustment Programmes

The Greek economic crisis unfolded in the aftermath of the global financial crisis and marked one of the most significant economic depressions faced by a country in recent economic history. GDP in current market prices collapsed from a peak of €242 billion in 2008 to a trough of €175 billion in 2016. A decrease of -28% such as this, was the largest recorded in a modern developed economy. The collapse in economic activity was reflected in the real economy with many insolvencies and a high unemployment rate, which increased from 8% in 2008 to a peak of 28% in 2013.

With public finance being pro-cyclical, an acute fiscal situation for the state followed. Total revenue fell from €98 billion in 2008 to €84 billion in 2015; a decrease of 8% as a percentage of GDP. At the same time, government spending increased as fiscal stabilisers worked in the opposite direction to support the reduction in household and business income.

The public debt situation was even more dramatic; Greece was shut off from financial markets as the cost of new debt was prohibitive for the government. News of tweaked budget statistics and a revision of the budget deficit (that was discovered to be greater than initially thought) at -15.4% of GDP, increased mistrust and interest rates for Greek Government debt jumped up to 36% for the 10-year bond.

Financial problems were present also in the banking sector, which held a large share of the Government's debt and faced the possibility of imminent default. The value of these bonds was later reduced in 2012 by a "private sector involvement" exercise, which imposed a haircut of 78% in face value (Cheng *et al.*, 2020).¹ In addition, increased insolvencies and high unemployment

¹The haircut was possible and was implemented by the Greek parliament since bonds were governed by Greek law.

rates meant that a high percentage of corporate loans and mortgages became non-performing, thus reducing the income stream for banks. Lending to the economy stopped and banks had to be recapitalised three times (Provopoulos, 2014).

Faced with a dire situation, in April 2010, Greece became the first country in the EU to seek financial assistance in the form of an economic adjustment programme from other European countries and with the contribution of the International Monetary Fund. The structural programme was negotiated by the European Commission, the European Central Bank and the IMF, which together formed the "Troika" of institutions. This first programme covered the period of 2010 - 2013 and foresaw bilateral loans to Greece in exchange for reforms. The extent of financial needs and the reforms that needed to be implemented led to a second programme being negotiated between the Troika and the Greek authorities in February 2012, covering the period 2012 to 2014. The combined financial assistance to Greece at the time was € 238 billion, equivalent to its entire pre-crisis GDP.

A third (and final) bailout programme was agreed in the summer of 2015. It made use of € 62 billion from the European Stability Mechanism, a new institution tasked with issuing syndicated debt and lending it to crisis countries in exchange for reforms. The programme was completed in August 2018, together with a comprehensive debt restructuring deal, that signaled the end of almost a decade of economic adjustment programmes for Greece.

During negotiations for all three programmes, taxation became a priority. The crisis exposed deep-rooted fiscal problems within Greece's economy, which included excessive public spending, unsustainable borrowing and weak tax collection. It became imperative that if fiscal discipline was to return at the end of the programmes and repayment of debts to EU countries and international institutions was to be ensured, more tax revenues would had to be generated. Being in a currency union left Greece with no options other than fiscal consolidation; a currency devaluation to regain competitiveness or high inflation to eliminate the real value of public debt could not be pursued.² Similarly, the crisis did not allow room for economic growth, that would have reduced the debt-to-GDP ratio. Thus, only high primary fiscal surpluses (more taxes than spending) could ensure debt sustainability in the long-term. For this reason, the three economic adjustment programmes, amongst a plethora of reforms, prioritised the improvement of tax collection.

Tax collection was seen as an important area for Greece to modernise, if sustainable annual tax revenues were to be generated. VAT evasion consistently appeared higher than other EU countries (Poniatowski *et al.*, 2021). Moreover, the structure of the economy with 99% of firms being micro-firms with less than 10 employees, facilitated evasion in all income streams; corporate, consumption and personal income (Kaplanoglou and Rapanos, 2015; Artavanis *et al.*,

²A devaluation would essentially mean a new drachma and leaving the euro area. This option was touted as Yianis Varoufakis (the Minister of Finance during the first Syriza government in 2015) plan B in case of failure to negotiate a third economic adjustment programme.

2016; Kourdoumpalou, 2016). Enforcement was not strong either, with the tax administration being understaffed, underfunded and without modern administrative systems, clear targets and incentives (Kaplanoglou and Rapanos, 2013). A strategy was soon drawn to give independence to the tax authority (similar to the IRS in the United States), with its own independent budget, which materialised in January 2017. Part of the reforms that the tax authority implemented after 2017, was the promotion of electronic payments as a way to generate third-party information, the details of which are described below.

1.2 Tax Digitisation Reforms

Tax digitalisation reforms in Greece were introduced in several stages and targeted the entire economy. Whilst cash was the predominant means of exchange, capital controls in the summer of 2015 saw a big increase in the number of bank cards being issued (Danchev *et al.*, 2020). In 2017, Greece implemented a gradual roll-out of compulsory use of Point-of-Sale (POS) machines in all professions, which facilitated the use of electronic payments.³ The obligation was extended to all professions by the end of August 2017.⁴ In addition, all firms and sole proprietors were obliged to hold a separate professional bank account linked directly with their POS machine.⁵

Importantly, banks were required since 2014 to send to the tax authority information about Greek taxpayers. This was made possible through an obligation to declare the personal tax number to the bank when opening a bank account. Information on the aggregate amount of payments, deposits and other financial assets held could be exchanged. Banks sent to the tax authority every month the aggregate amount of electronic transactions completed by each Greek taxpayer. Lastly, the income declaration process became completely electronic since 2014, with all relevant fields pre-filled. Individuals needed to declare their bank account's IBAN before submitting the form, which enabled a cross-check between bank accounts and personal tax numbers.

Building on these changes, the tax authority implemented two systems of incentives to encourage individuals to increase their electronic payments. A tax lottery on electronic payments and a requirement to use electronic payments during consumption in exchange of an income tax discount. A change from cash to electronic payments in the entire economy would facilitate third-party information and, thus, improve tax compliance (Kleven *et al.*, 2011).

³This was implemented with Law 4446/2016. Initially the following sectors were obliged to offer a POS until the end of June 2017: restaurants, bars and cafes; repairs by craftsmen; tuition for private schools, tutorial and foreign language institutes etc.; clothing and footwear shops; lease payments; beauty salons; gyms; legal and medical services; telecommunication services; utilities; and gaming.

⁴This was done with Law 4472/2017.

⁵The professional bank account was introduced into law in 2013, with Law 4172/2013. But actual enforcement steps took place only after the tax authority's independence in 2017: penalties and obligations were only announced in April 2017 (see AADE decision 45231/20-04-2017).

1.3 A Tax Lottery on Electronic Payments

In 2011, during the first years of the economic crisis, Greece embarked on its first attempt to establish a “random monetary reward” scheme, supported and promoted by technical assistance experts in the European Commission. The first scheme was called “@podeixi” (the Greek word for “receipt”), developed by the National Center of Scientific Research Demokritos in Athens.⁶ Mimicking tax lotteries from other countries, this first proposal sought to reward individuals who asked for a paper receipt after a transaction. It relied predominantly on mobile phone technology which was largely in place at the time. Individuals could input the information of a receipt and send them in a free text message. A random generator would award individuals with instant prizes. Due to lack of government support, however, this initiative never materialised.

As tax digitisation reforms progressed and the tax authority gained independence, a tax lottery was implemented in the autumn of 2017. The new form of a tax lottery promoted by AADE relied on information transmitted by banks to the tax authority and converted electronic payments into monthly tickets automatically. Prizes of €1 million were budgeted every month and the scheme has been running from October 2017 to this day. In this thesis, Chapter 2 and Chapter 4 provide an economic analysis of this scheme.

Chapter 2 analyses the characteristics of winners and documents an occupational premium and an income premium. In addition, using Monte Carlo simulations improvements in the policy are proposed to limit regressive effects that are inherent in the policy. This analysis enhances our understanding on how tax lotteries function in practice, how they can be improved and what the implications of the interlinkage with electronic payments are; an area that is becoming increasingly popular with tax authorities around the world.

Chapter 4 documents the VAT revenue effect of the tax lottery. Due to a delay in implementing the lottery in 2017, the tax authority decided to run 9 draws retroactively based on past spending. Using the winners’ variation across regional tax offices in Greece, this study identifies the winner’s effect on regional VAT revenue. Two mechanisms are documented. Firstly, idiosyncratic effects from winners themselves, by implementing a difference-in-difference comparison of their electronic consumption after winning against a random sample of non-winners. Secondly, spillover effects to non-winners. These findings add some unique evidence on tax lotteries, the literature of which remains particularly slim, with the exception of Naritomi (2019).

1.4 The Electronic Consumption Tax Discount

The second policy that incentivises electronic payments is analysed in Chapter 3. This is the first attempt by a tax authority to provide incentives through the income tax system. Individuals receive their annual tax allowance (about €10,000 taxed at 0% marginal tax rate) only if they

⁶The scheme was registered as a patented system in 2011 (Patent No. 1007355).

complete an annual amount of payments in electronic transactions. The amount they need to complete is a percentage of their income.

The policy evolved out of negotiations between Greece and its international lenders during the first economic adjustment programme in 2010. It aimed at incentivising taxpayers to demand paper receipts for their transactions, as a way to increase paper trail and fight tax evasion. It applied initially to paper receipts only: in 2010, individuals needed to keep receipts equivalent to 25% of annual income, reduced to 10% from 2011 onward but increasing marginally on income. Implementation of the policy relied on enforcement, as taxpayers had to present their paper receipts in case of an audit and were faced with an additional tax bill if they failed to produce the required amounts.

Paper receipt collection evolved drastically in 2017 to include only electronic payments. One can imagine that audits on individual paper receipts would have been particularly costly and difficult to implement. In addition, receipts could be easily transferable. However, an electronic record of payments automatised this procedure. The total amount of electronic payments completed by a taxpayer appeared automatically in the tax return of 2017, facilitated by the exchange of information between banks and the tax authority. Individuals had to check the amount and had the opportunity to modify it before filing taxes. If the taxpayer had spent less in electronic payments, an additional tax applied. The policy applies to this day and the spending limits have increased significantly (up to 30% of income has to be spent using electronic payments).

Chapter 3 utilises administrative tax data and evaluates how much taxpayers have declared versus how much they have spent in electronic transactions. It documents an increase in reported amounts, if the electronic payments are lower than what needed to be spent. In addition, spending increases during the end of the year, as the deadline to reach the spending limit approaches. These findings provide a comprehensive insight on how taxpayers behave in response to the introduction of this novel instrument.

1.5 Conclusion

The Greek economic crisis, 2009-2018, led to one of the largest contractions in GDP recorded in a modern economy. As public revenue collapsed and fiscal consolidation was the only way out of the crisis, it forced a rethink in policies to raise taxes. Tax reforms in this period were spearheaded by the use of third-party information to close a large VAT gap and broaden the tax basis.

Specifically, the Greek tax authorities decided to utilise digitalisation reforms, through the cross-checking of information and by providing incentives for electronic payments. Two of these innovative instruments form the focus of this thesis; a tax lottery on electronic payments in Chapters 2 and 4 and the electronic consumption tax discount in Chapter 3. Tax digitisation

reforms have been increasing in a number of countries, which adds to the significance of the findings presented here. In particular, economic, fiscal and welfare effects of these policies are not yet fully understood and the available evidence remain scarce. The data and analysis presented in this thesis provide important insides on the use of electronic payments to fight tax evasion.

Chapter 2

Are Tax Lotteries Regressive? Income, Consumption and, Occupational Characteristics of Winners

2.1 Introduction

Tax lotteries are schemes designed to increase the compliance of firms with the payment of value-added tax (VAT hereafter). They provide incentives to address VAT's "last-mile" problem: third-party reporting in business-to-business transactions deters tax evasion, since one firm's input serves as the other's output (Pomeranz, 2015), but in business-to-consumer transactions no incentive exists in asking and keeping a receipt of purchase. Unrecorded transactions result in higher tax evasion and lower public revenue. The monetary incentives are targeted to the final consumer – assigning to each receipt the expectation of winning a prize. The additional third-party information should then facilitate enforcement (Kleven *et al.*, 2011). While the effectiveness of tax lotteries has been the subject of study in economic literature – notably in Naritomi (2019) for the Brazilian lottery and in chapter 4 for the Greek lottery – less is known about the winners' characteristics. Who wins the lottery determines the policy's regressivity or progressivity.

Using administrative data from the Greek tax lottery, this chapter provides evidence on (i) the income, consumption and occupational characteristics of winners (ii) a quantification of the extent to which these characteristics affect their winning probabilities (iii) simulations on a ticket ceiling reform as a way to improve progressivity. The data allow for the reconstruction of a representative taxpayer population against which the population of tax lottery winners can be compared. They contain information on income, electronic consumption and occupation indicators for 68,897 tax units in the Greek population; 18,897 winning units in the Greek tax lottery during 19 monthly draws in 2017 and 2018, and a random sample of 50,000 non-winning tax units.

The comparison reveals an "income premium" in the lottery tickets. The chances of winning a prize increase proportionally to the level of income. The higher the income of a taxpayer, the higher the spending and the higher the chances of winning. Compared to the representative taxpayer population, the winners exhibit roughly seven times higher mean annual electronic transactions (€ 28,413 compared to € 3,931) and a higher mean income (€ 15,877 compared to € 9,403). In addition, annual electronic transaction differences are exacerbated between the two samples with extreme values in the winners' population; 369 winners exhibit more than € 1 million in annual e-transactions. By construction the scheme selects high income individuals, with a high propensity of payments, as winners.

Estimated elasticities from a linear probability model quantify the effect of income on (i) electronic transactions and (ii) the probability of winning. A 10% increase in income is associated with a 1.8% increase in the volume of electronic transactions and with a 0.11% increase in the probability of winning the tax lottery. Information from partners belonging to the same household and file taxes jointly, captures also an intra-household income effect. A 10% increase in spousal income is associated with 0.7% increase in electronic transactions and a 0.06% increase in the winning probability.

Turning to the income sources provides some striking results. I document a particularly large "occupational premium" for the self-employed in winning the lottery. Self-employed winners record annual mean electronic transactions of € 181,520 compared to € 14,626 for winners belonging to other occupational categories (wage-earners, pensioners and agricultural workers). They also record an economically large discrepancy between their electronic transactions and their declared income. Whilst for the representative population, spending by electronic means remains below income, for self-employed individuals who won the lottery, electronic transactions are 10 times higher than their income. This suggests that the self-employed might either under-declare their income or use business-transactions through their personal bank accounts. As a result, this group accounts for 8.3% of all winners compared to a population share of 4.1%. The lottery selects winners from the self-employed income category with a higher probability.

When quantifying this effect, I find that the occupational premium outweighs the income premium. Being self-employed increases electronic transactions by about 75% and the probability of winning by 18% compared to other income categories. Intra-household occupational effects are also significant; having a self-employed spouse increases the electronic transactions of their partner by 45% and their chances of winning the lottery by 3.5%. These result hold after controlling for a taxpayer's own, as well as, for spousal income.

The winning premiums documented in this chapter appear to be an inherent problem in tax lotteries with a direct effect on the policy's progressivity. By incentivising consumers to ask for receipts, it is by construction the high income individuals and high spenders who have higher winning chances and, thus, benefit the most from the scheme's prizes. In addition, the self-employed who might use their personal accounts for large business transactions have higher

chances of winning. Both income and occupational premiums might raise fairness concerns and ultimately undermine the effectiveness of tax lotteries.

To examine solutions that could make the scheme more progressive by construction, I perform monthly draws of the lotteries using Monte Carlo simulations under two scenarios. Firstly, a reformed ticket structure introduced in 2019 that became more concave, therefore awarding more tickets per euro spent in lower spending levels, combined with an upper limit of € 50,000 per month per individual after which no lottery tickets were awarded. Secondly, a stricter ticket limit in the monthly amount of tickets at € 1,000 and at € 5,000 per individual (instead of € 50,000). I find that the reformed ticket structure has only marginal effects, whilst the progressivity is improved partially through the upper limit of € 50,000. The annual mean electronic transactions halves, indicating that the ceiling becomes binding for extreme consumption values. The annual mean income of winners is reduced only marginally. Simulation results from the stricter limits of € 1,000 and € 5,000 per individual per month indicate a fairer distribution of prizes. However, the stricter the ceiling, the less the incentive to ask for receipts once the monthly limit is reached. In the Greek tax lottery, a ceiling of € 5,000 per individual per month strikes a good balance between the incentive to ask for receipts and a fair distribution of prizes.

The evidence presented in this chapter contribute to our understanding of tax lotteries. Other studies have documented the policy's effectiveness in raising tax revenue. Notably, Naritomi (2019) analyses the Brazilian tax lottery and identifies a 21% increase in reported sales and a lower, yet significant, increase of 9.3% in reported revenue for the state of Sao Paulo in Brazil. The study mentions whistle-blowing and collusion costs as potential mechanisms for the increase. The Greek tax lottery is analysed in chapter 4. In contrast to the Brazilian tax lottery, tickets are awarded only when taxpayers complete payments through electronic payments, and in addition, no registration of payments is necessary; the tickets are awarded automatically to the entire population proportional to their amount of electronic payments. In chapter 4, I document an increase in regional VAT by 0.01% per additional winner. The main mechanism is idiosyncratic effects from winners, who spent more in electronic payments after winning, as well as, spillover effects in electronic consumption from winners to non-winners.

Beyond the policy's effectiveness in raising revenue, none of the previous studies have considered so far the income, consumption and occupational characteristics of winners. These characteristics are important in assessing how regressive or progressive the policy can be. Evidence in this chapter contribute to the literature by bringing forth an inherent regressivity in the design of tax lotteries. Firstly, since tickets increase proportional to spending, the higher the income of individuals, the higher the chances of winning. Secondly, winning is not occupationally-neutral; some income categories who might utilise consumption for professional purposes gain an advantage. These lottery characteristics result in high-income earners, high-spenders and the self-employed/business-owners being selected more frequently as winners.

This chapter contributes also to a vast literature that studies the efficiency of intra-household allocations of income. The income pooling hypothesis predicts that only household (or joint) income should matter for allocation decisions, and not who receives it. Browning *et al.* (1994) using Canadian household expenditure data show that who gets what depends on the income of each individual. Additional evidence are presented in Lundberg *et al.* (1997) using a natural experiment in the United Kingdom with child benefits being allocated to wives increasing women and children clothing expenditure. A survey of results in developed countries is presented in Chiappori *et al.* (2020). These findings are particularly applicable in developing countries, since household decisions are key for development; evidence from a number of countries suggest imperfect income sharing within a household (Quisumbing and Maluccio, 2003). Adding to these evidence, I quantify the elasticity of electronic consumption with respect to a change in spousal income. A 10% increase in spousal income increases electronic consumption of their partner by 0.76%, lending support to evidence of imperfect income sharing.

Finally, the findings add to existing empirical evidence on the marginal propensity to consume. Carroll *et al.* (2014) estimate the aggregate marginal propensity to consume for Greece to range between 0.10, when fitting a net wealth distribution, and 0.35, when fitting a liquid assets distribution. I estimate a marginal propensity to consume of 0.18, which falls within this range. The estimate in this study is important as it is derived directly from the real electronic transactions of individuals and declared income in tax returns.

The remaining chapter is organised as follows. Section 2.2 provides institutional information on the Greek tax lottery. Section 2.3 describes the data. A descriptive analysis of the winners' characteristics, followed by a parametric analysis that quantifies these effects is provided in Section 2.4 . Simulations of reforms are shown in Section 2.5. Lastly, Section 2.6 concludes.

2.2 The Greek Tax Lottery on Electronic Transactions

Tax lotteries have become a common tool to mobilise consumers as a source of third-party reporting, to expand the tax base and ultimately to increase tax revenues (Naritomi, 2019). This trend is also visible in Europe, where several countries introduced tax lotteries during the European debt crisis (Fookan *et al.*, 2015).¹ In 2017, when the Greek Tax Authority was granted a strong institutional and financial independence, a tax lottery scheme gained traction.

¹In 2011, during the first years of the economic crisis, Greece embarked on its first attempt to establish a “random monetary reward” scheme. The first project was called “@podeixi” (the Greek word for “receipt” in Latin characters), developed by the National Center of Scientific Research Demokritos in Athens. It received a Greek patent in 2011 (Patent No. 1007355). For more information see <http://www.obi.gr/obi/Default.aspx?tabid=127&idappli=X410275>. Mimicking tax lotteries from other countries, this first proposal sought to reward individuals who asked for paper receipts (regardless of the payment method). Due to lack of government support, however, this initiative never materialised. The tax lottery appeared only once in an official document; a 2014 leaked email of the Minister of Finance (also known as “Hardouvelis Email”) to the heads of monitoring institutions known as the “Troika” during Greece’s second economic adjustment programme. It was intended to serve as an eleventh hour proposal to bridge fiscal differences before the structural programme expiry in early 2015.

Technical advice towards the Greek Tax Authority supported and widely promoted the use of electronic payments in the economy, as a way to fight tax evasion.² The design of the lottery was thus incorporated in a broader strategy that aimed at curbing tax evasion by encouraging the use of electronic payments over cash payments. For this reason, the Greek tax lottery focuses specifically on incentivising electronic transactions over cash payments.

Electronic Transactions. For a long time, cash has been the prevailing payment method in Greece. However, the imposition of capital controls in July 2015 led to a substantial increase in electronic payments.³ From 2015 onward, Greece recorded a massive increase of debit cards issuance, which have been associated with increased tax buoyancy in the years that followed, despite a large negative economic shock (Hondroyiannis and Papaoikonomou, 2017). Additional incentives to promote electronic payments were introduced in 2016 (Law No. 4446/2016). Among others, they included an annual tax credit to all taxpayers who spent (from 2017 onward) a certain fraction of their (reported) incomes in electronic payments; this policy is analysed in chapter 3.⁴ At the same time, acceptance of electronic payments and introduction of Point-Of-Sale (POS) terminals became mandatory (in gradual roll-out phases starting 2017, depending on the profession). Overall, the legislative measures had a strong, positive impact on card use (Danchev *et al.*, 2020).

These reforms were complemented by the introduction of a comprehensive IT system. Starting from January 2017, banks were required to automatically report the total volume of electronic transactions per individual to the tax authority. This information, on the one hand, enabled the tax authority to assess whether a taxpayer would pass the minimum threshold of electronic payments (see chapter 3). On the other hand, the reporting system also served as key building block for a further incentive for electronic payments: a lottery that rewards electronic transactions.

Electronic Transactions Lottery. At the end of each month, the *aggregated volume* of all eligible electronic transactions (but not each single transaction) of each bank client are submitted to the authority. All electronic payments – online banking transactions as well as debit, prepaid, and credit cards payments – by Greek taxpayers to businesses with presence in Greece or in other

²A comprehensive strategy for the promotion of electronic payments to tackle tax evasion, appeared as a key deliverable by Greece's creditors in the summer of 2015, the implementation of which was linked to financial disbursements in the Memorandum of Understanding of the European Stability Mechanism programme to Greece. See p. 9 in https://ec.europa.eu/info/sites/info/files/01_mou_20150811_en1.pdf.

³Cash withdrawals were limited to € 60 per day per individual, but electronic payments remained unlimited.

⁴For 2017, the cutoffs were 10% for declared annual income up to € 10,000, increasing to 15% for any income between € 10,000 to € 30,000 and 20% for incomes exceeding € 30,000 (with an upper limit for very high incomes). (Such a tax credit already existed since 2011. Instead of being tied to electronic payments, however, individuals had to collect and keep paper receipts in case of a tax audit.) If electronic payments would fall below the cutoff, individuals would face higher tax obligations (which are automatically calculated by the tax authority and saliently reported in tax returns and online bank accounts). The tax credit thus provided a clear incentive for consumers to use electronic payments in everyday transactions.

EU countries are eligible.⁵ Based on the unique tax identification number of each individual, the tax authority adds up the total volume of completed electronic transactions in a given month for each taxpayer.⁶ The resulting sum serves as the key input for the lottery.⁷

The monthly volume of electronic transactions are converted into tax lottery tickets according to a given *ticket-awarding mechanism* (TAM hereafter). In an attempt to increase the relative chances of low-consumption groups to win the lottery, the TAM has a concave structure: at higher levels of electronic transaction volumes, an additional euro would translate into fewer tickets. This point is documented in Table 2.1, which presents the detailed structure of TAM.⁸ While the first euro of monthly electronic consumption would translate into one ticket, the € 1,001st would yield only 0.25 tickets. Note further that the TAM does not contain any upper bound.

Table 2.1 Ticket-Awarding Mechanism (Jan 2017 – May 2019)

Total amount of monthly e-transactions	Tickets awarded	Maximum number of tickets
€ 1 – 100	1 ticket per € 1	100
€ 101 – 500	1 ticket per € 2	300
€ 501 – 1,000	1 ticket per € 3	466
> € 1,000	1 ticket per € 4	No limit

Applying the TAM on the volume of electronic transaction data from a given month m yields the total number of tickets that enter a draw in month $m + 1$. A random draw then picks 1,000 winning tickets (with the constraint that an individuals can only win once in a given month).⁹ Each winner receives € 1,000. The prize is tax exempt and cannot be confiscated.

⁵Eligible electronic payments include, for instance, paying an airline ticket, buying a product on online, or paying for online subscription services. Eligible credit and debit card payments include paying for any product or services at any shop within the EU. Non-eligible payments include paying for rent and any government related payments such as taxes and fines. Transactions with firms from non-EU countries are excluded since these cannot be fully identified and classified by the electronic systems of banks.

⁶Banks request the tax identification number when opening a bank account and, thus, have information on each individual. Multiple accounts and bank cards registered on an individual are aggregated using the their tax numbers. Joint bank accounts include the tax identification numbers of multiple individuals. However, the means of a payment (a credit or a debit card, for instance) is always registered on the name of one individual. Banks take into account this differentiation when submitting data to authorities.

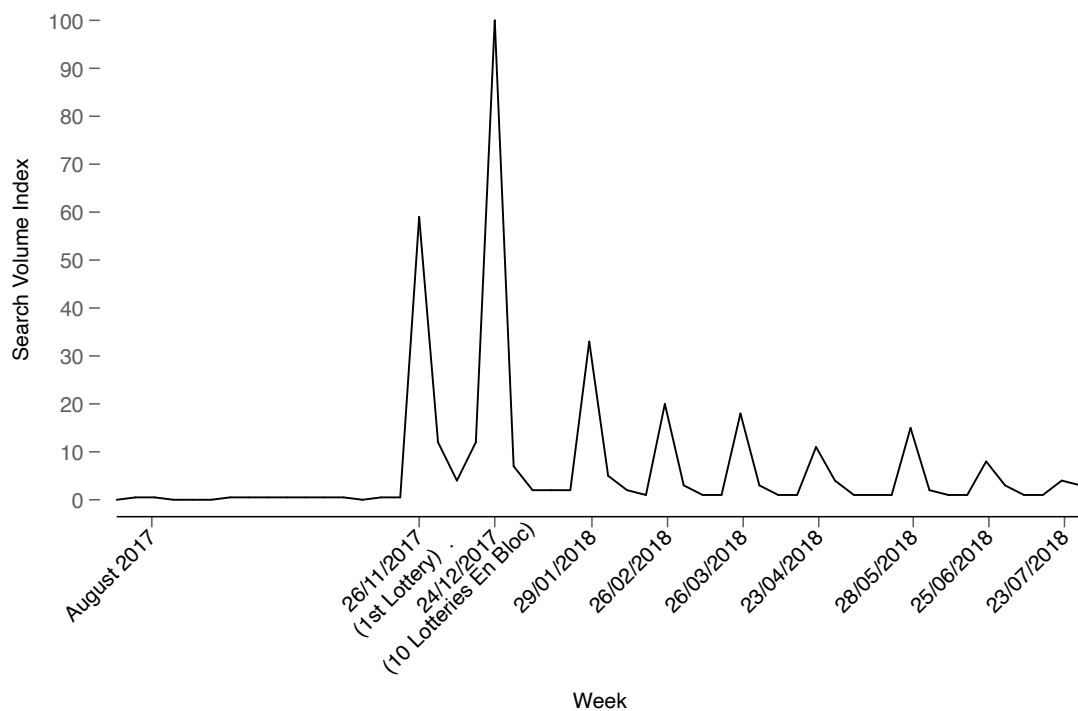
⁷Details on the functioning of the lottery can be found in Law No. 4446/2017, and Ministerial Decision 1161/17-10-2017. Communication by the Tax Authority on matters regarding the lottery can be found in <https://www.aade.gr/menoy/miniaies-synallages-kai-lahnoi>.

⁸This scheme was in place until May 2019. In Section 2.5 I discuss a reform and compare the initial with the post-reform TAM.

⁹To prevent that the lottery gets rigged, the tax authority follows a two-step process. First, the random draw is outsourced to a university. Based on the total number of monthly tickets, the university generates a series of random numbers which are submitted to the tax authority. Second, the authority modifies these numbers using a pre-determined formula, which is unknown to the university.

The initial plan was to introduce the first tax lottery draw in January 2017. While the bank reporting system started operating in January, there were several legislative and technical issues that resulted in a delay. In chapter 4, I take advantage of this delay to identify the VAT revenue effect of the tax lottery. A public announcement of the lottery took place in early October 2017. At that point in time, the TAM and the prize structure were made public.¹⁰ Subsequently, the first draw took place in November 2017, based on electronic transactions from October. In December 2017, draws from the previous months in 2017 took place en bloc, based on electronic transactions from January to September 2017. The tax authority announced 10,000 winners in December; corresponding to 1,000 winners for each month from January to September 2017 and 1,000 winners for December 2017. Search volumes in Google recorded in Greece at the time for the word “Lottery” in Greek are shown in Figure 2.1. While the search volume is close to zero in the months prior to the first lottery, the volume spikes at the end of November (1st lottery), while the highest volume was recorded at the end of December (at the time of the 10 draws), indicating increasing public awareness.¹¹

Fig. 2.1 "Lottery" Google Trends in Greece



Notes: The figure plots the Google search volumes (indexed from 0-100 on the y-axis) for the word “lottery” in Greek. The geographical area is constraint to Greece alone. The timeline is shown on the x-axis, containing weekly trends for every week starting with the first week of August 2017 and ending in the last week of July 2018.

¹⁰The TAM was more widely discussed in the press when the first lottery draw took place.

¹¹The search volume index records increases at the end of each month thereafter, in line with the time of monthly draws.

Since November 2017, the lottery is running on a monthly basis, each time based on the previous month’s transaction data. On the day of the draw (usually the 24th day of a month), the authority publishes the tax IDs of all winners on a dedicated website. Thereafter, winners are also contacted by emails or via text messages on their phone. The prizes are transferred directly to winners’ bank accounts about a week after winning.

The Greek tax lottery is unique and differs from other tax lotteries in numerous ways. Firstly, and different from all other tax lotteries, the Greek scheme is hardly based on self-selection. Typically, consumers must register in a system and collect receipts and in order to participate in a tax lottery (see, e.g., Naritomi, 2019). Instead, the Greek lottery is the first to almost automatically include the vast majority of taxpayer: everybody who (i) holds a bank account and (ii) makes an electronic transactions in a given month takes part in the lottery.¹² Secondly, tickets are exclusively tied to electronic payments. Hence, the objective of the Greek lottery is to incentivise consumers to switch from cash to e-payments – which tend to be more difficult to conceal and should thus facilitate monitoring and tax enforcement. In fact, as discussed above, the lottery emerged jointly with a system that provides the tax authority with full information on an individual taxpayer’s level of monthly electronic payments.

One caveat concerns the use of private bank accounts for business purposes. In an attempt to separate individual from business transactions, the tax authority obliged all firms to use separate business bank accounts (Joint Ministerial Decision No. 45231/2017). However, the implementation of this measure was significantly delayed and the enforcement of the requirement was gradually pushed from summer 2017 to spring 2019.¹³ Hence, while the intention of the tax authority was to include only private electronic transactions (‘consumption’) by individuals, the lottery initially (i.e., before 2019) included a non-trivial volume of business transactions. The line separating business and individual transactions will be, as a result, particularly blurry for self-employed individuals and owners of small firms, since business transactions can be made through their personal bank accounts (which are tied to their unique tax identification number).

2.3 Data

The data contain (a) information on the monthly level of electronic transactions for the period from January 2017 to July 2018 as well as (b) matched information from tax returns in 2017.¹⁴ I observe the annual declared pre-tax income for submitted tax returns of tax units, which

¹²According to the World Bank’s Global Findex database, 85% of individuals in Greece above the age of 15 had a bank account in 2017. As some of these are joint ‘family’ accounts (see fn. 6), the formal banking system includes almost the entire population.

¹³The initial ministerial decision was published in April 2017 giving business owners 3 months to comply with the decision. Renewed decisions were issued 3 times with deadlines being pushed to 15/1/2019, 28/2/2019 and 30/4/2019.

¹⁴The last day of tax return submission for the tax year of 2017 was July 30, 2018. The tax returns underwent a basic plausibility check and tax payment statements were issued by the tax authority in August 2018. The data in this paper were received in October 2018.

include either a single or a joint filing (the latter consists of the main taxpayer's and the spouse's income) from economic activities in 2017. In case of joint filings, the monthly level of electronic transactions correspond to only one of the two individuals in the tax unit. The data provided by the tax authority were anonymous, while monthly electronic transactions were rounded to the nearest € 10 and annual declared income information to the nearest € 5.¹⁵ For joint filings, I observe income values for both partners, enabling the calculation of the tax unit's declared income.¹⁶ For single filings I observe the declared income of the single person in the filing, which is also a single-household income.¹⁷ Lastly, it is compulsory to file tax returns even if an individual has exactly zero income. Thus, the data include many students above the age of 18 (in tertiary education) as well as the unemployed. With an unemployment rate of 21.5% in 2017, the latter group constituted a significant proportion of the working population in post-crisis Greece. The zero income group, however, might also include tax units who conceal all of their incomes.

In addition to the declared amount, the data indicates the source(s) of income from five different categories: income from wages (subsequently *WG*); self-employed/business income (*SB*); agricultural income (*AG*), income from pensions (*PE*), or zero-declared income (*NO*).¹⁸ *WG* includes income received from salaried activities. Hence, I observe a tax unit's reported annual gross salary. The *SB* category includes sole proprietorships, such as the self-employed, sole traders and small firms. This is the most common legal form of business activities in Greece.¹⁹ The data contain annual net profit from business activities (but would not report losses). *AG* contains declared annual income from agricultural activities, such as for farm owners, agricultural workers and small cultivations. *PE* includes all individuals who receive pensionable income from main or auxiliary pensions. The data report the (pre-tax) annual pension income. *NO* contains individuals who have reported zero income in 2017. This category includes individuals who are obliged to submit tax returns, even if their income is zero, such as tertiary education students and the unemployed. However, it might also contain individuals from the *SB* and *AG* income categories, who report zero income. The percentages of single filings, joint filings for households and income categories, included in the sample, are shown in Table 2.7.

Note that a given tax unit might of course declare incomes from multiple sources (categories). Below, I will use indicators that define the primary source of reported incomes. These dummies also serves as a proxy for the (primary) occupational activity.

¹⁵The declared income does not include any income received from the government as a subsidy to the household, such as social welfare transfers for poor households, nor any tax credits added before the final tax calculation.

¹⁶For the 2017 tax returns, joint filing was mandatory for married couples. Law no. 4172/2013 provided that the main taxpayer of the household is the husband, responsible for submitting the tax return, while the wife must sign-in before finalising the submission and give consent to the declared amounts.

¹⁷However, I cannot distinguish individuals and households in the case were the main taxpayer has declared some level of income, and the spouse has declared 0 income.

¹⁸These income categories in Greece corresponded to different pension insurance funds contributions that existed in the past.

¹⁹In 2017, 1.2 million sole proprietorships existed out of a total of 1.4 million firms. These correspond to the *SB* category. See the Statistical Business Register in 2017 available from the Hellenic Statistical Authority at: <https://www.statistics.gr/en/statistics/-/publication/SBR01/->.

Due to between category variation in third-party reporting (Kleven *et al.*, 2011), there are major differences in the opportunities to under-report incomes. For WG and PE income, the income values (as reported by employers or pension funds) appear automatically in individuals' tax returns. SB and AG incomes, on the contrary, are self-reported. Hence, as noted above, some individuals with non-zero incomes from these sources might not report any income and thus end up in the NO category.

Sample. I obtained data for two different samples. First, the universe of 18,897 tax units with members that have won the lottery during the first 19 consecutive lottery draws (based on the e-transactions for the months from January 2017 to July 2018);²⁰ second, a randomly drawn sample of 50,000 tax units that did *not* win the tax lottery in any of these 19 draws. For the winners' sample, and for all 19 months covered, I observe the monthly volume of electronic transactions for the winning individual, while for the randomly-drawn sample I observe electronic transactions for one individual in the tax unit (as reported from banks to the tax authority). Among the first sample, I also observe the month of winning. For both samples, I observe the tax return filing, which contains one annual income value for single member tax units and two annual income values for joint filing households. In the case of joint filings, monthly e-transactions correspond to one of the two persons in the household. Table 2.7 presents basic summary statistics for the two samples.

Based on the information from these data, the analysis will focus on comparing the characteristics of winners against a reconstructed baseline population of tax units. Specifically, the winners from the 2017 draws will be used, since this has two advantages. Firstly, one can compare the annual declared income from 2017 with the monthly e-transactions $z_{i,m}$ for an entire calendar year, i.e., I compute $Z_i = \sum_{m=1}^{12} z_{i,m}$ for each taxpayer i covered by in the data. A second advantage is that the e-transactions values from this year are, to a large extent, not influenced by the lottery itself. To see this, recall that the tax lottery was announced in early October 2017 and the first draw took place in late November (see Section 2.2). Note further that the broader public only took notice of the lottery after the first draw and, in particular, after the draws that took place en bloc at the end of December 2017. Hence, electronic transaction values for the months January to September 2017 were recorded before the announcement of the lottery; and even for the remaining months the tax lottery might play a minor role in shaping annual e-transactions.

To allow for a meaningful comparison of winners relative to the baseline population, one has to account for the different sampling of the two samples. The non-winners sample was drawn randomly from the population of taxpayers – conditional on not having won (given the TAM and the taxpayers monthly levels of e-transactions). The other sample contains the universe of

²⁰The sample covers less than 19,000 winners (1,000 winners for each of the 19 lotteries) since some of the prizes were refused or not claimed by those who were drawn. This happened in 103 out of 19,000 cases, of which 40 from draws in 2017. In informal conversation with the tax authority it has been reported that because essentially all taxpayers are included by default in the lottery (and do not actively opt-in), some winners are unaware of its existence and regard the winning message by the tax authority as fraud.

winner (which were drawn under the same TAM conditional on the taxpayers’ e-transaction pattern). To arrive at a sample that represents the population of taxpayers, I expand (or re-weight) the non-winner population such that they match the overall number of ‘lottery tickets’ (i.e., the aggregated amount of e-transactions) observed in 2017. As a plausibility check, I compare the obtained number of tax units with in the corresponding number in the population of taxpayers. The details of this approach are described in Appendix 2.7.2.

Table 2.2 Summary Statistics: 2017- Winning Tax Units versus Population of Tax Unit

	<i>Winners</i>			<i>Population</i>		
	Obs.	Income	E-Trans.	Obs.	Income	E-Trans.
<i>by Primary Income Category:</i>						
SB	988	20,753	181,520	266,317	12,120	11,420
Self-Employed/Business inc.	8.3%	(32,955)	(695,170)	4.1%	(25,891)	(60,163)
WG	5,773	18,357	10,857	2,890,322	11,418	4,064
Wage-Earner	48.3%	(38,738)	(45,598)	44.7%	(13,941)	(6,138)
PE	2,704	14,631	10,964	1,573,228	11,875	3,322
Pensions	22.6%	(6,347)	(67,821)	24.3%	(6,046)	(5,350)
AG	503	47,423	15,532	340,746	17,582	3,817
Agriculture	4.2%	(106,648)	(33,355)	5.27%	(38,113)	(6,627)
NO	1,818	0	27,618	1,397,996	0	2,935
Zero-declared inc.	15.2%	(0)	(197,309)	21.6%	(0)	(15,109)
No Filing	174	-	37,119	-	-	-
(Tax return not submitted)	1.45%	-	(342,630)	-	-	-
Total	11,960	15,877	28,413	6,468,897	9,403	3,931
	100%	(37,277)	(229,919)	100%	(15,036)	(15,243)

Notes: The table presents the number of observations, the mean declared income and the mean e-transactions Z_i in 2017 (nominal €values) for individuals in tax units with matched e-transactions information. They are presented by primary income source as has been declared in their tax returns. The groups pair taxpayers according to their main source of reported incomes coming from wages (WG), self-employed/business incomes (SB), agricultural income (AG) or income from pensions (PE). An additional category indicates zero-declared income (NO). Standard deviations are in parenthesis. The ‘Winners’ sample includes all individual winners from the monthly draws based on 2017 e-transactions. For joint-filings income and e-transactions information are matched for one individual, either main taxpayer or spouse. The ‘Population’ sample is a reconstructed sample of the tax unit population (see Appendix 2.7.2).

Basic summary statistics from the resulting, expanded sample are presented in Table 2.2. The table compares the baseline tax unit population with the tax units of winners from 2017. Comparison takes place at the individual level within tax units, since I have e-transaction

information only for one individual in each unit. The tabulation reveals several peculiar observations. Firstly, the mean e-transactions and mean income of winners is much higher than the rest of the population. Secondly, the SB income category is over-represented and exhibits very high levels (and variance) of e-transaction. Thirdly, winners in the NO category, had a particularly high level of e-transactions. These discrepancies form the centre of the analysis in the following section.

2.4 Who Wins the lottery?

This section examines the characteristics of lottery winners highlighted in Table 2.2. I first provide descriptive evidence on the selection implied by the TAM (considering the e-transaction volume, income, and the primary income source at the individual and tax unit level). Second, I parametrically quantify the selection along multiple dimensions.

2.4.1 Descriptive Analysis

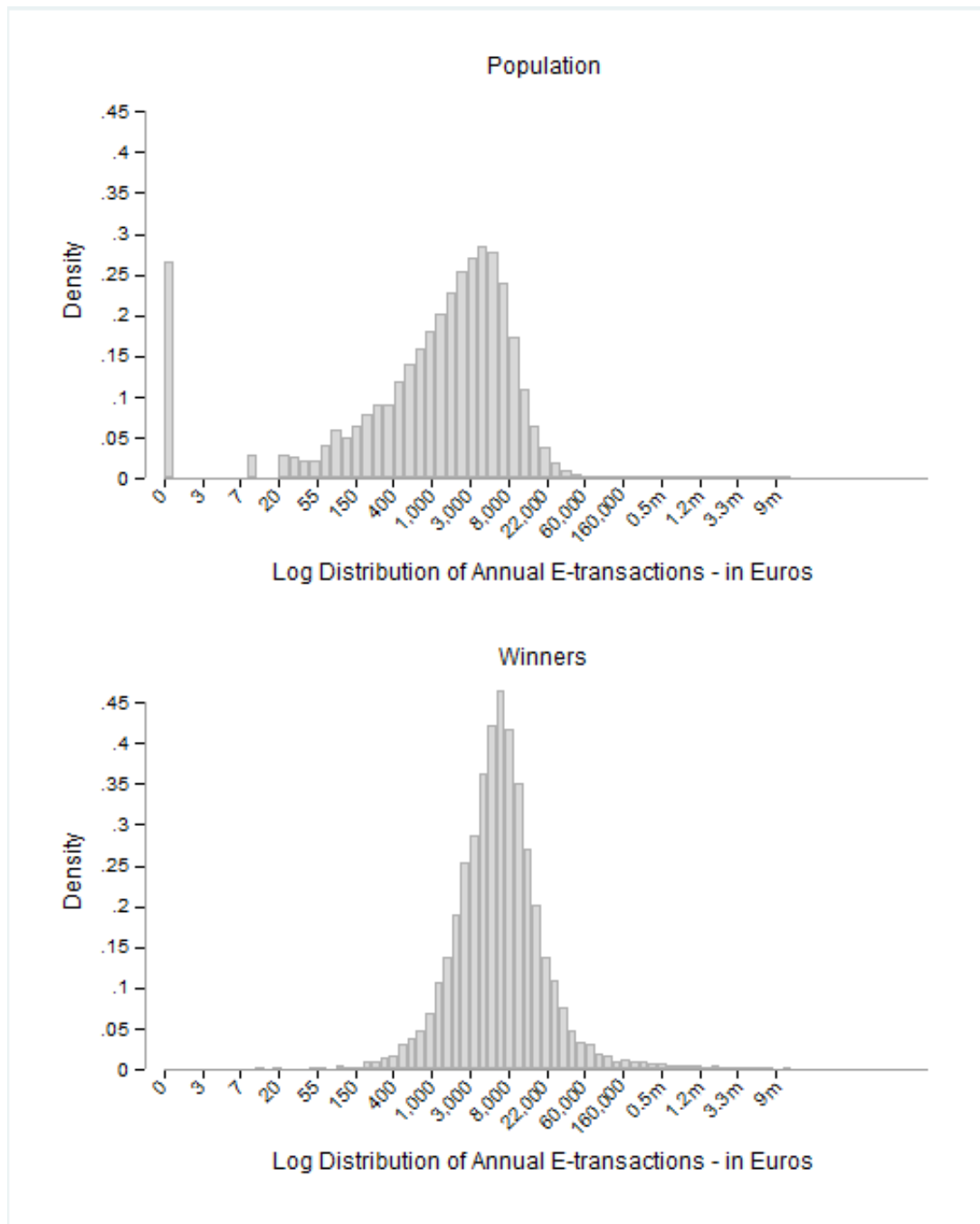
2.4.1.1 E-transaction and Income

A comparison of the mean annual e-transaction is shown in Table 2.2. Individual winners exhibit roughly seven times as high mean annual e-transaction compared to the representative population. This reflects the basic property of the lottery: the chances of winning increase proportionally to the level of e-transaction. Since the TAM contains no upper bound (there is no maximum number of assigned tickets in a given month), the probability of winning *cet.par.* approaches unity if $z_{i,m} \rightarrow \infty$. While this holds for a given month m , as e-transactions fluctuate between months, the annual level Z_i is only an indirect indicator for the selection implied by the TAM.²¹

The difference in e-transaction (more specifically, in $\log(Z_i)$) can also be seen in Figure 2.2. Among the total taxpayer population, the distribution is right skewed and bi-modal: almost 25% of taxpayers have basically zero e-transactions. The mean annual e-transaction volume is € 3,931 (median € 1,940). For winners, log-e-transactions are symmetrical and normally-distributed, with a significantly higher mean of € 28,413 (median € 6,400). In contrast to the general population, where there is hardly any mass in the range $Z_i > € 22,000$, there is a heavier right-hand tail for winners, with a non-trivial share of annual e-transactions volumes well above € 60,000. For 2017, I observe 334 winners with more than € 1 million annual e-transactions, 34 with more than € 2 million and one extreme value of more than € 9 million e-transactions (who has won twice in 2017).

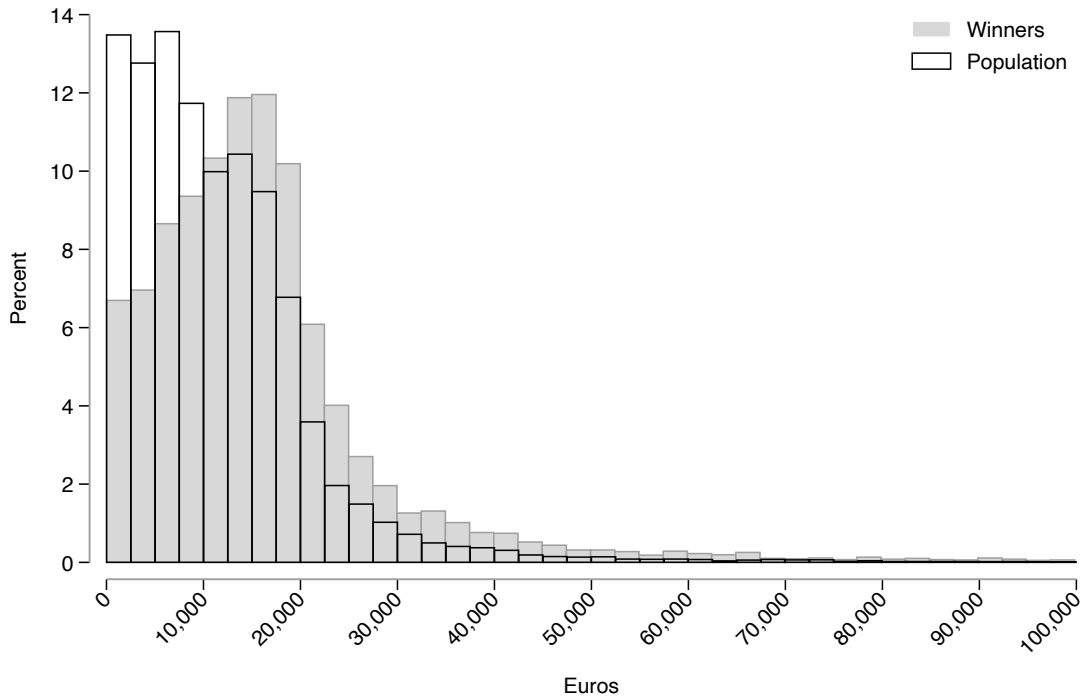
²¹The difference between mean e-transactions for winners relative to the population is even more evident in a monthly comparison. The taxpayer population's monthly mean e-transactions followed an upward trend in 2017, fluctuating between € 278 (in the beginning of the year) to € 445 (at the end of the year). The mean e-transactions of those who have won in a particular month fluctuated around € 4,000 (without observing any upward trend).

Fig. 2.2 Distribution of Annual Electronic Transactions, 2017



Notes: The figures present the log of annual e-transactions distributions in 2017 between individuals in the tax unit population and individuals who have won the lottery in 2017. The x-axis is a log scale representing the equivalent values in €. Tickers are rounded to the nearest thousand in €. The population distribution includes the individuals from 6.4 million tax units. The winners distribution includes 11,960 winners in tax lotteries that took place in 2017. Monthly e-transactions of individuals were summed up over the 12 months to create the annual e-transactions. Monthly values in the data were rounded to the nearest € 10 by the tax authority. The distributions are drawn on the same scale, sharing the same y- and x- axes.

Fig. 2.3 Distribution of Declared Income 2017



Notes: The figure compares the distribution of declared income in 2017 of lottery winners with the corresponding distribution of individuals in the tax unit population. The population has been reconstructed from a random sample of 50,000 tax units. The graph is truncated at € 100,000, as right-tails diminish quickly in the distribution.

The differences in e-transaction are mirrored in the levels of reported income. The mean declared income among winners is € 15,877, whereas it is € 9,403 among the tax unit population (median values are € 12,113 and € 6,850, respectively). These differences are economically and statistically significant. They are also illustrated in Figure 2.3. As can be seen a substantial proportion of individuals in the tax unit population report income below € 10,000. Among lottery winners, there is less mass in this income range but over-proportionately many cases with higher income levels. Overall, the figure clearly reflects that the lottery winners are (judged against the general population) higher-income taxpayers.²²

I then compare the gap between annual e-transactions and incomes within winners and the tax unit population. Among the latter group, the log values of the two variables display a positive correlation of 0.29. For winners, the mean e-transactions are about twice as high as the mean declared income (see Figure 2.7 in the Appendix). This difference, however, is largely driven by some outliers (with very high e-transactions values). The correlation between the log values of the two variables in the winners' sample remains positive but falls to 0.11, indicating that e-transactions become weakly associated to income. Lastly, the e-transaction/income ratio is

²²For illustration purposes, Figure 2.3 is truncated at € 100,000. However, there are a few observations of winners with very high incomes well above € 100,000.

1.79 among winners and 0.42 among the general population. Winners spent almost two times their income on e-transactions, while the general population spent less than half. For every third winner (33.5%) I observe a ratio above unity, i.e., an e-transaction volume that is above the declared income.

2.4.1.2 Income Sources

This section examines the extent to which the observed patterns are driven by differences in the (primary) income source. Firstly, note that Table 2.2 documents significant differences in mean incomes between winners and the general population within each group with a given primary income. The table further documents that individuals from the SB category are massively over-represented among lottery winners. Relative to a population share of 4.1% , this group accounts for 8.3% of all winners.²³ While WG, PE and AG individuals are observed at similar percentages as in the population of tax units, taxpayers in the zero declared income (NO) group are under-represented.²⁴

Comparing the groups with different primary income sources, Table 2.2 reports very high levels of e-transactions among the SB category: the mean among winners is € 181,520 (median of €17,565), with a high standard deviation. Among the total population, individuals with SB incomes have a mean (median) e-transaction volume of € 11,420 (€ 4,410). Hence, in addition to the fact that individuals from the SB group win the lottery more often than others, the lottery also selects (within the SB group) winners with unusually high e-transactions volumes.

The discrepancy of e-transaction for SB individuals is depicted in Figure 2.4, where declared income and e-transactions of SB are compared against the pooled groups WG, PE and AG.²⁵ For the latter income categories, mean annual e-transactions are about one-third of mean declared incomes in the tax unit population. Among the SB population, the corresponding share is around 90%. Hence, these income categories display a different pattern. As shown in Figure 2.4 (b), these differences are even more pronounced among winners: the mean e-transactions of SB winners are ten times as high as their mean declared income. Among the winners from the other income categories, the e-transaction/income ratio is below one.²⁶

The stark difference between e-transactions and declared income among winners from the SB group suggests that these individuals are using electronic means and their private bank accounts when paying for business-purpose expenses. The flow of business transactions results

²³See Section 2.3 for more information on the SB category.

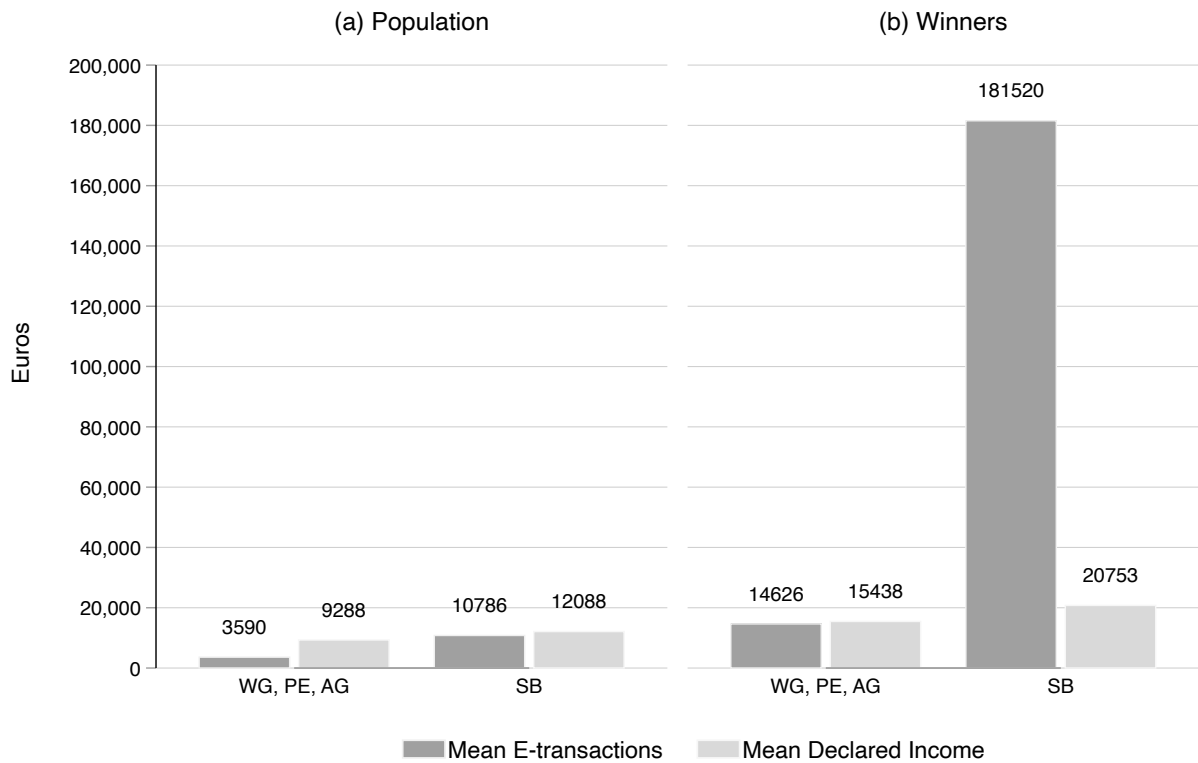
²⁴The latter group, which still accounts for around 15% of all winners, can be composed by heterogeneous types: (a) students or unemployed individuals who have non-zero e-transactions; (b) individuals with self-employed or business income that report zero or negative incomes (losses).

²⁵To allow for a meaningful comparison, the NO category individuals are excluded.

²⁶Out of 988 winners from the SB category, 64% exhibit e-transactions higher than their income. Among the SB group in the general population, the equivalent percentage is 39% – which is still much high compared to taxpayers with other income sources. For example, in the WG group, about 9% of the population (16% of winners) have annual e-transactions levels higher than their incomes.

in particularly high e-transaction levels compared to both, their own declared income as well as the e-transaction volumes of other income groups. As a result, one can observe disproportionately more tax units with primary SB income among winners. Recall from Section 2.2, that the use of private bank accounts for business purposes was prohibited in 2019. Hence, it was still legal in 2017 for some business expenses to be channeled through private bank accounts.²⁷

Fig. 2.4 Declared Annual Income and E-transaction by Income Sources



Notes: The figures compares mean e-transactions and mean declared income for groups with different primary income sources: self-employed/business income (SB) vs other non-zero incomes from wages, pensions and agricultural activities (WG, PE, and AG). Zero-declared income individuals are excluded from this comparison. Figure (a) is based on the population of tax units and figure (b) presents the lottery winners from 2017.

2.4.1.3 Income Sources within Households

Another characteristic that affects winning chances is the household composition and their income source types. I first examine the influence of having a spouse with SB income. As long as some couples share their (private) bank account, and if partners with SB income use the accounts for business (e-)transactions (see above), one should expect to observe higher e-transactions levels for individuals jointly filing with a SB (rather than a non-SB) spouse. To

²⁷The observed pattern might also originate from the illegal underreporting of incomes: since the SB group has (relative to third-party reported incomes) more opportunities to conceal income (Kleven *et al.*, 2011), the vast e-transaction/income gap may therefore – at least in parts – reflect income tax evasion. The data do not allow us to quantify this channel.

assess this case, I focus on individuals who filed jointly in 2017. Overall, I observe that 37% tax units in the data file jointly, which is very close to the official percentage of 40% for 2017.²⁸

Figure 2.8 in the Appendix compares e-transaction and income levels of individuals who jointly file with spouses that either receive their primary income from SB or with spouses in the WG/PE/AG income group.²⁹ (To facilitate interpretation, the sample underlying this graph excludes individuals from the SB and the NO income group). Having a partner in the SB income group is associated with higher levels of e-transactions. This holds for the general population (Panel (a) of Figure 2.8) but, more strongly among the group of winners (Panel b).³⁰ At the same time, the partner's income source does not make much of a difference for the reported income. Overall, the data indicates that (many) jointly filing couples seem to share private bank accounts and that SB partners seem to use these accounts for business transactions.³¹

2.4.1.4 Descriptive Analysis Summary

The descriptive analysis documents several key implications of the lottery's design. First, as the chance of winning is proportional to the volume of e-transactions, the lottery selects (more) high e-transaction taxpayers as winners. Since e-transactions positively correlate with incomes, I document an over-representation of higher-income taxpayers among winners. Whilst the concave structure of the TAM was introduced in the scheme to ensure better winning chances for the low-income taxpayers, the results indicate that this was not sufficient from ensuring equal chances. A sizable gap between the income of lottery winners relative to the population exists, suggesting a regressive policy.

Second, e-transaction patterns appear to be strongly affected by the use of private bank accounts for business purposes. High e-transactions associated with business activities result in more SB individuals being selected as winners. In 2017, the TAM thus rewarded a subgroup of individuals from occupations (with SB incomes) that resulted in very high e-transaction levels, well above their declared income. Third (and consistent with the previous point), I observe spillovers within jointly filing couples: a taxpayer, whose spouse receives SB income, has a higher amount of e-transactions. This increases the chances of winning the lottery.

²⁸This information is included in the annual statistics published by the tax authority in <https://www.aade.gr/menoy/statistika-deiktes/eisodima/etisia-statistika-deltia>.

²⁹Appendix Figures 2.9, 2.10 and 2.11 illustrate the same type of sample split for individuals with WG, PE and AG spouses, respectively.

³⁰The differences are even greater if the spouse receives any part of income from SB activities (instead of having SB as a primary income source; column 1). Table 2.9, column (2), documents that this difference in annual e-transactions is economically and statistically highly significant. The difference is hardly affected by controlling for annual declared income in 2017, as shown in column (3).

³¹The pattern might also be shaped by individuals who record certain private, household expenses (such as the purchase of a personal computer) as business input costs in order to exempt these costs from VAT.

2.4.2 Parametric Analysis

This section quantifies the pattern documented in the descriptive analysis. In particular, I explore the role of individual and spousal income level and income sources for (i) the level of e-transactions and (ii) the probability of winning the lottery. First, I consider models of the structure:

$$\log(Z_i) = \beta_0 + \beta_1 \log(Y_i) + \beta_2 \log(Y_{j|i}) + \beta_3 SB_i + \beta_4 SB_{j|i} + \beta_5 Joint_i + \varepsilon_i \quad (2.1)$$

where Y_i indicates the reported income and SB_i is a dummy variable indicating business or self-employed being the primary income source. The sub-index $j|i$ measures these variables for i 's spouse j . $Joint_i$ is a dummy variable indicating an individual who has filed jointly with a spouse. Note that β_1 and β_2 capture taxpayer i 's elasticity of e-transactions with respect to their own and their spouse's income, respectively. Perfect income sharing within a household (plus equal propensities to spend money electronically) would imply $\beta_1 = \beta_2$.

Columns (1)–(3) in Table 2.3 reports OLS estimates that follow the structure of equation (2.1). The estimated β_1 suggests that a 10% higher income correlates with a 1.8% increase in e-transactions. This measure is similar to a marginal propensity to consume estimate for electronic consumption only. It captures how much electronic consumption changes to a change in income. Carroll *et al* estimate the aggregate marginal propensity to consume for Greece to range between 0.10, when fitting a net wealth distribution, and 0.35, when fitting a liquid assets distribution (Carroll *et al.*, 2014). The estimate in Table 2.3 falls within this range. The coefficient hardly changes in Column (2), when spousal income is controlled for. The correlation with the partner's income is significantly lower: a 10% higher income of the spouse is associated with a 0.7% increase in individual i 's e-transactions. F-test reject the null $\beta_1 = \beta_2$ with $p < 0.001$. The results are quantitatively similar in column (3), where I focus on taxpayers who filed jointly. The findings indicate imperfect income sharing within a household (Browning *et al.*, 1994; Lundberg *et al.*, 1997) or differential propensities to engage in e-transactions.

The estimates further document that, consistent with the descriptive evidence from above, income from self-employment or business activities is associated with significantly higher levels of e-transactions. The estimated semi-elasticities imply that receiving the primary income from this source (SB) is associated with an approximately 75% higher level of e-transactions. A similar estimate is obtained for a spouse with primary income from SB: the corresponding semi-elasticity for the $SB_{j|i}$ dummy is around 45% (see Column 2). It is worth stressing that this holds while controlling for the taxpayer's own, as well as, for the spousal income. Hence, the pattern reflects an occupational rather than a mere income correlation.

Table 2.3 Estimation Results

Dependent var:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Log(e-transactions)			P(winning)			
	joint-filers			joint-filers			
Log-Income (β_1)	0.181*** (0.003)	0.179*** (0.003)	0.205*** (0.008)	0.011*** (0.000)	0.011*** (0.001)	0.016*** (0.001)	0.002 (0.002)
Log-Income/Spouse (β_2)		0.073*** (0.005)	0.076*** (0.006)		0.006*** (0.001)	0.008*** (0.001)	0.001 (0.001)
Primary Income: SB (β_3)	0.773*** (0.050)	0.745*** (0.050)	0.519*** (0.074)	0.179*** (0.014)	0.177*** (0.014)	0.186*** (0.022)	0.021 (0.025)
Primary Income/Spouse: SB (β_4)		0.445*** (0.063)	0.466*** (0.063)		0.036** (0.017)	0.035** (0.017)	-0.005 (0.021)
Joint Filing		-0.130*** (0.044)			-0.010 (0.007)		-0.008 (0.008)
Tickets 2017							0.000*** (0.000)
Constant	5.582*** (0.026)	5.371*** (0.028)	5.002*** (0.090)	0.098*** (0.003)	0.080*** (0.004)	0.009 (0.014)	0.019 (0.012)
<i>F-Tests</i> (p-values):							
$\beta_1 = \beta_2$		0.000	0.000		0.000	0.000	0.461
$\beta_3 = \beta_4$		0.000	0.594		0.000	0.000	0.338

Notes: The table presents estimation results following the structure of equation (2.1). The dependent variable in columns (1)–(3) are e-transactions ($\log(Z_i)$) and, in columns (4)–(7), an indicator for winning the lottery. Coefficients and standard errors in columns (4) – (7) are multiplied by 100. The sample is $N = 6,468,609$ observations, except for columns (3) and (6), where the sample is constrained to 2,406,683 jointly filing taxpayers. Robust standard errors (clustered at the level of 50,000 + 11,960 unique taxpayers) in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

As a second step, I replace the dependent variable from equation (2.1) with a dummy W_i which indicates that the taxpayer has won in the lottery during 2017. The results from linear probability model estimates are presented in Columns (4)–(7) in Table 2.3 and they document a positive correlation between the own and the partner’s income with the chance of winning, which in itself reflects the positive correlation between income and e-transactions (see columns 1 – 3). It is striking to observe that individuals with primary income from SB or (jointly filing) taxpayers with a spouse that receives SB income have a significantly higher chance to win in the lottery. The point estimates from column (5) indicate that an individual with SB income has a 0.00177 log-point higher chance of winning the lottery – which is non-trivial given that the baseline chance of winning are the same magnitude (roughly 12,000 tax units over 6.5 million tax units). Put differently, having an SB income is *cet.par.* associated with a twice as high chance of winning the lottery. The association with the partner’s primary income source being SB is smaller but still statistically significant at the 5% level. Results are qualitatively unchanged when I estimate only in the sample of jointly filing taxpayers.

Finally, it is worth noting specification (7) of Table 2.3 which adds the annual number of tickets assigned to an individual in 2017. It is reassuring to observe that, controlling for the total number of tickets, renders all other variables statistically insignificant. This suggests that the lottery is not rigged. The correlations between winning and the level and source of incomes are merely shaped by the association of these variables with the amount of e-transactions. It is

the latter which then translates into tickets and, ultimately, the probability of winning the tax lottery.

2.5 Lottery Reforms

The evidence from above indicates that individuals (and households) reporting incomes from business activities or self-employed labor, benefit from the lottery design; private bank accounts are used for business purposes, which inflates the (seemingly private) e-transaction volumes of these taxpayers. This ‘occupational premium’, however, applied in the beginning of the lottery. Since May 2019, all firms, self-employed and sole proprietorships were required to switch all business activities to separate professional bank accounts. While the level of enforcement of (and compliance with) this rule remains unclear, there is at least an appropriate regulatory framework in place.

The analysis further documents that the lottery – by design (i.e., via the TAM) – tends to select higher-income individuals as winners. While the primary objective of the tax lottery is not to engage in income redistribution (but rather to incentivise the transition from cash to electronic payments), the tax authority and the Ministry of Finance (who together are formally in charge of the lottery design) nevertheless responded to this observation.³² In an attempt to generate more lower-income winners, they reformed the ticket-awarding mechanism in May 2019. The initial TAM (from Table 2.1) was replaced by a more concave ticket scheme, displayed in Table 2.4. Monthly e-transactions up to € 200 (pre-reform: € 100) would be now one-to-one converted into tickets. For e-transaction volumes above € 1,000, an addition Euro would only yield 1/6 of a ticket (pre-reform: 1/4). Moreover, the new TAM introduced an upper bound with a maximum of 8,682 tickets in a given month (which is, in the new scheme, equivalent to e-transactions of € 50,000). Below I explore the potential impact of this reform for different taxpayer’s propensity to win the lottery.

Table 2.4 Post-Reform Ticket-Awarding Mechanism (May 2019 onward)

Total amount of monthly e-transactions	Tickets awarded	Maximum number of tickets
€ 1 – 200	1 ticket per € 1	200
€ 201 – 500	1 ticket per € 2	350
€ 501 – 1,000	1 ticket per € 3	516
> € 1,000	1 ticket per € 6	8,682

³²Following en bloc draws in December 2017 (see Section 2.2), some Media reported on cases where individuals won more than once. See for example, <https://www.dimokratianews.gr/dimokratia/forolotaria-me-symptoseis-adianoites/> (in Greek).

2.5.1 Comparing Pre- and Post-Reform TAM

Based on the distribution of 2017 e-transactions in the general population, I simulate who would have won under the new TAM.³³ By using the 2017 distribution of e-transactions, this approach ignores potential behavioral responses to the lottery's post-reform TAM. In a non-static framework behavioural adjustments could have a non-trivial impact on the post-reform distribution of e-transactions and on the results presented below.³⁴ Starting from the data, I first transform the monthly e-transactions of the population (from 2017) into lottery tickets – using both pre- and post-reform TAMs. I then simulate (based on 100 iterations) the 1,000 lottery winners of the 12 lottery draws in a calendar year (this ensures that the simulation results take into consideration all months and are not affected by specific spending patterns in particular month). For each of the 1,200 iterations, I record the winners' characteristics and then compare the 1.2 million simulated (12 months \times 100 iterations \times 1,000) winners under the initial TAM with the winners from the post-reform TAM.

Table 2.5 Main Simulation Statistics

	Pre-reform TAM				Post-reform TAM			
	Mean	p10	p50	p90	Mean	p10	p50	p90
Annual E-transactions	20,794 (227,467)	1,000	4,890	16,730	10,480 (69,923)	990	4,720	14,980
Annual Income	13,870 (23,226)	0	11,225	25,370	13,476 (21,495)	0	11,190	24,695

Notes: The table presents the main statistics from lottery simulations using the pre-reform TAM (left-hand side) and post-reform TAM (right-hand side). Each simulation aggregates 1,200,000 observations of winners, drawn based tickets assigned from each structure (100 lottery iterations, drawing 1,000 winners in each iteration, for each of the 12 months in 2017). The first column in each TAM presents the mean values and standard deviation in parentheses. The median values are presented in the "p50" columns, together with the lowest and highest percentiles in "p10" and "p90" respectively.

The main results from the simulations are shown in Table 2.5. Winners' selection in pre-reform TAM is comparable to winners' selection in the actual lottery; both the mean and standard deviation values of annual e-transactions and income are similar to the respective values in Table 2.2.³⁵ In the simulations one can observe the results of the occupational premium of

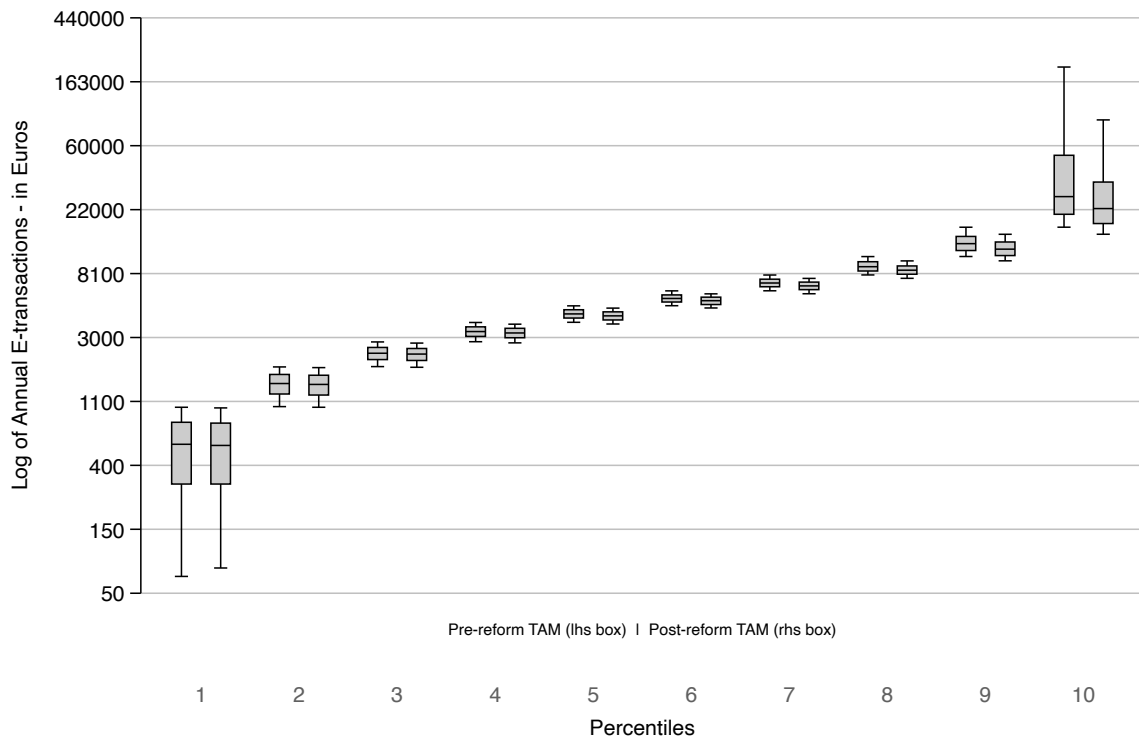
³³Recall that the data available do not no cover the post-reform period.

³⁴The upper ceiling of monthly tickets, for instance, creates an incentive to 'smooth' e-transactions above € 50,000 over different months. However, the data from the pre-reform period do not indicate any bunching of monthly e-transactions at the kink-points of the initial TAM. This null-observation might be due to the limited salience of monthly e-payment volumes.

³⁵These values are not precisely the same because the actual lottery is similar to only one iteration (1000 winners) for each month in 2017. The precision in the simulation increases, since the lottery is performed 100 times each month, and thus, converging to more precise values.

the policy; higher mean e-transactions than mean income prevails among winners, € 20,794 against € 13,870; lower median than mean values in both income and e-transactions; and a very high standard deviation value for annual e-transactions at € 227,467. The latter indicates the existence of extreme values, confirming that the pre-reform TAM selects individuals with high e-transaction volumes in the population.

Fig. 2.5 E-transaction Box Plots



Notes: The figure presents box plots of the log-value of annual e-transactions for each decile in the pre- and post-reform TAM simulations. The pre-reform TAM are presented on the left-hand side of each decile and the post-reform TAM on the right-hand side. Each simulation contains 1.2 million observations of winners. The plots exclude extreme values, which are present in 1st and, in particular, the 10th deciles. The y-axis is drawn in the log value of annual e-transactions but has the values have been computed to Euros, rounded up to the nearest tenth for values of up to € 8100, and to the nearest thousand for higher values.

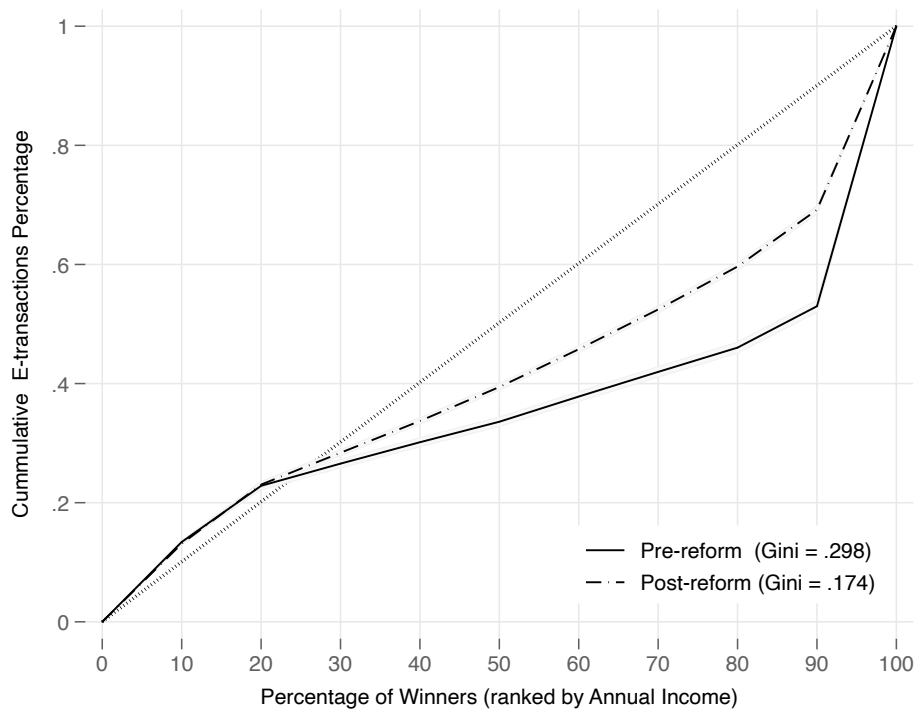
Importantly, the results in Table 2.5 indicate a change in the winners population once the post-reform TAM is applied. This change works predominantly through the € 50,000 limit in monthly e-transactions. The annual mean (standard deviation) e-transactions for winners drops from € 20,794 (227,467) to € 10,480 (69,923), while the mean declared income is reduced only marginally, indicating that the upper ceiling can limit the distortion of winners exhibiting higher e-transactions than their declared income. Mean annual e-transactions for the 9th decile reduce accordingly from € 16,730 to € 14,980.³⁶ Moreover, while decreasing e-transactions in the highest decile are non-trivial, the post-reform TAM produces a lower e-transactions level for low-to-middle deciles. This point is illustrated in Figure 2.5, which presents box plots of

³⁶Notice also that the lowest decile exhibits 0 income, while having positive annual e-transactions. This group consists of individuals who declare zero income and happen to win the lottery.

the log of e-transactions for every decile in pre- and post-reform TAM. Deciles in the middle of the distribution present small reductions in log e-transactions, while the largest reduction happens at the highest decile. An overall comparison between the income and e-transactions distributions can be seen in Figures 2.12 and 2.13, respectively.

Finally, I provide a measure to assess the extent to which the reform can limit the distortion of winners exhibiting higher e-transactions volumes than their declared income. I plot the e-transactions distribution curves for each TAM (on the y-axis), ranking individuals by their annual reported incomes (on the x-axis) in Figure 2.6. This produces an e-transactions distribution of winners, where if the slope is higher [lower] than the 45-degree line, winners at a particular decile exhibit proportionally more [less] e-transactions than income.³⁷

Fig. 2.6 E-transactions Distribution



Notes: The figure plots e-transaction distribution curves for 2×1.2 million winners in Monte Carlo Simulations of the lottery. The x-axis represents the population percentiles of winners, ranked by annual income in 2017. The y-axis plots the cumulative percentage of e-transactions resulting from simulated lottery winners in pre- and post-reform TAM. The dotted line is a 45-degree line, at which the winners' income percentage equals the winners' e-transactions percentage. For the pre- [post-]reform curve, tickets were assigned based on the initial [post-reform] TAM presented in Table 2.1 [Table 2.4].

First, note that the lowest quintile of the population (according to their annual e-transactions) covers more than 20% of winners. This property, which holds for both the initial and the new TAM, seems to reflect the TAM's concavity, favoring lower income households who exhibit proportionately higher monthly e-transactions than their declared income (for example, by being

³⁷If all slopes at all deciles are equal to 45-degree, this serves as a point of equality, where the percentage of e-transactions for each decile equals their percentage in the income distribution.

net welfare benefit receivers). The slope is higher than the 45-degree line in the first decile, while it is close to 45-degrees for the second decile, indicating that the lowest benefits the most in both TAMs. Second, for the pre-reform TAM, the top decile in the income distribution accounted for 47% of all e-transactions. With the post-reform TAM, this percentage drops to 34%. The observed decline originates from a decreased slope in the highest decile, as well as, increased slopes in lower deciles. The change suggests that the post-reform TAM – in particular, its upper ticket limit – is partly effective in reducing the winning chances associated with extreme e-transactions volumes. Yet, high income individuals still exhibit higher chances of winning. Third, the reform-beneficiaries are located between the 2nd and the 9th decile of the e-transaction distribution (higher slopes in the post-reform, than in the pre-reform TAM). Overall, the Gini coefficient associated with the e-transaction curve, when ranked by income, can serve as measure of this distortion; it falls from 0.30 to 0.17. Hence, the reform had a non-trivial effect on achieving a more equal distribution of prizes.

2.5.2 Other Reforms

Following the simulation approach from above, I also consider the impact of more radical upper limits in the monthly amount of tickets; at € 1,000 and € 5,000, respectively.³⁸ I find that both ceilings are effective in limiting the winning chances associated with very high monthly e-transactions. As shown in Table 2.6, the e-transactions of the highest decile fall to € 15,350 at the € 5,000 ceiling and to € 13,550 at the € 1,000 ceiling. As in the pre-/post-reform TAM the annual income distribution of winners changes only marginally. Changes to the distribution of income and e-transactions are illustrated over the whole population of winners in Figures 2.14 and 2.15 respectively.

By drawing e-transaction distribution curves over the income distribution of winners in Figure 2.16, one can investigate the beneficiaries of such reforms. The stricter ceilings reduce even further the chances of winning for the highest decile of the winners' income distribution (the slope at the top decile approaches the slope of the 45-degree line), while individuals in the 2nd to the 9th decile stand to benefit. The Gini coefficients fall further to 0.162 for the € 1,000 ceiling and to 0.148 for the € 5,000 ceiling. The latter records a more equal distribution than the former, because, when the € 1,000 ceiling is used, a much higher fraction of individuals ends up receiving the maximum amount of tickets in several months. (see Figures 2.17 and 2.18). As the very strict ceiling of € 1,000 becomes more binding in the population, the chances of winning become more detached from the individuals' monthly level of e-transactions. Overall, the € 5,000 monthly ceiling seems to strike a better balance between, on one hand, limiting the distortion caused by high e-transaction volumes and, on the other, achieving a fairer distribution of prizes in the population.

³⁸The simulations adopt the euro-to-ticket structure from the pre-reform TAM. The ceilings thus limit the maximum number of monthly tickets at 467 and 1,467, respectively.

Table 2.6 Main Simulation Statistics - Stricter Ticket Ceilings

	Ceiling € 1,000				Ceiling € 5,000			
	Mean	p10	p50	p90	Mean	p10	p50	p90
Annual E-transactions	6,994 (24,783)	970	4,630	13,550	8,316 (33,949)	990	4,800	15,350
Annual Income	13,088 (19,307)	0	11,210	24,160	13,553 (21,259)	0	11,225	24,950

Notes: The table presents the main statistics from lottery simulations using a € 1,000 monthly ticket ceiling per individual (left-hand side) and a corresponding € 5,000 ceiling (right-hand side). Each simulation aggregates 1,200,000 observations of winners, drawn based tickets assigned from each structure (100 lottery iterations, drawing 1,000 winners in each iteration, for each of the 12 months in 2017). The first column in each TAM presents the mean values and standard deviation in parentheses. The median values are presented in the "p50" columns, together with the lowest and highest percentiles in "p10" and "p90" respectively.

2.6 Conclusion

This chapter documents how certain design characteristics of a tax lottery (a scheme that provides monetary rewards to consumers when they ask for payment receipts) assign higher winning probabilities to certain groups of individuals against others. Using a reconstructed taxpayer population in Greece and by comparing their income and annual electronic spending against those of winners, I document an "income premium" and an "occupational premium" in lottery winners.

When it was introduced in 2017, the Greek tax lottery selected winners from high income categories, who spent more in electronic consumption and were awarded with more tickets than lower income categories. The probability of winning increased by 0.11% in response to a 10% increase in income. The lottery exhibited also an "occupational premium", selecting winners from the self-employed/business-income category more frequently. Being self-employed increased the chances of winning by 18% compared to wage-earners, pensioners and agricultural workers, after controlling for individual and household income. This might have resulted from business transactions channeled through personal bank accounts, or from under-declaration of self-employed income. In addition, intra-household income and occupational effects are economically and statistically significant. The probability of winning increases by 0.7% in response to a spousal income increase by 10%; it increases further by 3.5% when the spouse is self-employed.

These results have important implications for tax lotteries. Since tax lotteries provide incentives to the final consumer to ask for a receipt at the point of purchase, it is high-income/high-spending

consumers who stand to benefit the most from the lottery's monetary rewards. In addition, spending for business purposes cannot be distinguished easily from personal spending, resulting in some occupational categories winning more frequently than others. These effects are amplified by intra-household effects, which move in the same direction as idiosyncratic effects. Both occupational and income premiums might dampen the effectiveness of the policy. Firstly, if the policy is perceived as unfair it might discourage individuals from participating. Secondly, by awarding winning prizes to high income individuals makes the policy regressive by construction. Prizes in low income individuals might act as a stronger incentive mechanism due to a higher prize-to-income ratio. Thirdly, the policy becomes less salient in low income individuals who experience winning less often than high income individual, but who might be using electronic payments less frequently than high-income individuals.

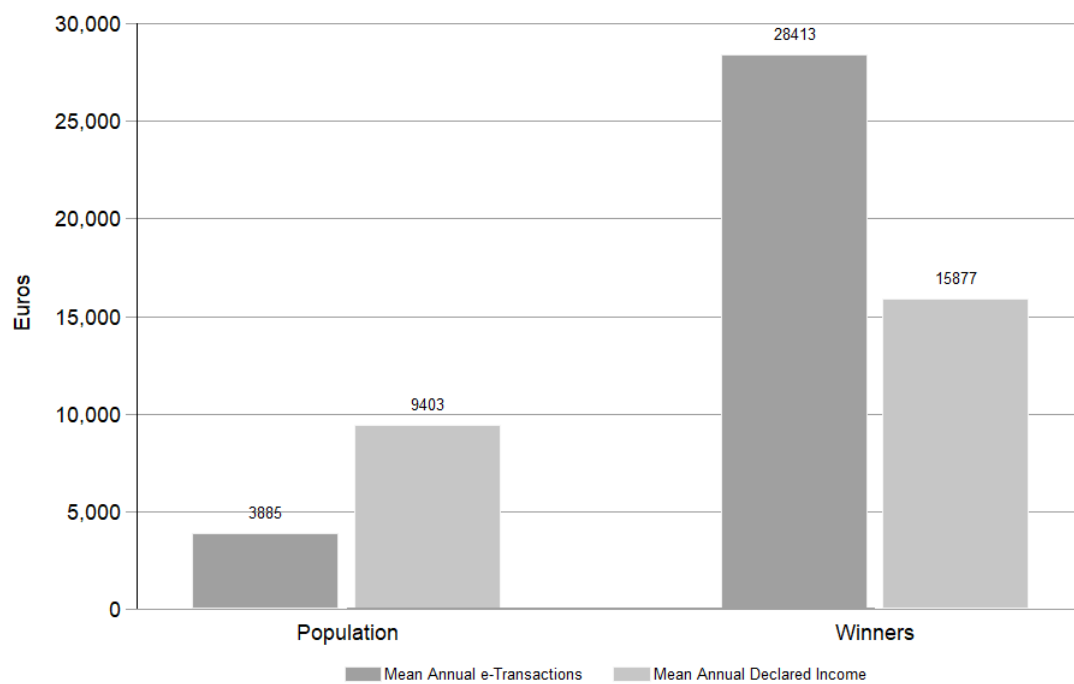
To mitigate the lottery's shortcomings and limit ticket premiums in certain groups of the population, the analysis in this paper considered a reform of the lottery in 2019. A more concave ticket-to-euro structure and a limit of € 50,000 per individual per month in the amount of tickets awarded were introduced. Using Monte Carlo simulations in a static framework, tax lottery draws were performed to determine the policy's effect. The results establish that the € 50,000 ticket ceiling is effective in improving fairness by limiting extreme consumption values for high spending individuals, whilst the more concave ticket-to-euro structure has only a marginal effect. In the absence of a ceiling high-income/high-spending individuals would have had higher chances of winning. In a second step, I consider stricter limits and I find that a € 5,000 limit in the amount of tickets awarded, per month per individual, would have resulted in an even fairer distribution of prizes, without placing excessive limits on the incentivisation of taxpayers to ask for receipts.

The evidence in this paper provide guidance for tax lottery design. As regressivity seems to be a characteristic of this policy, inherent by construction and difficult to predict prior to the policy's implementation, placing a ticket ceiling will benefit a tax lottery's effectiveness. The ceiling can limit excess ticket premiums to certain groups of individuals and thereby improve a tax lottery effectiveness.

2.7 Appendix

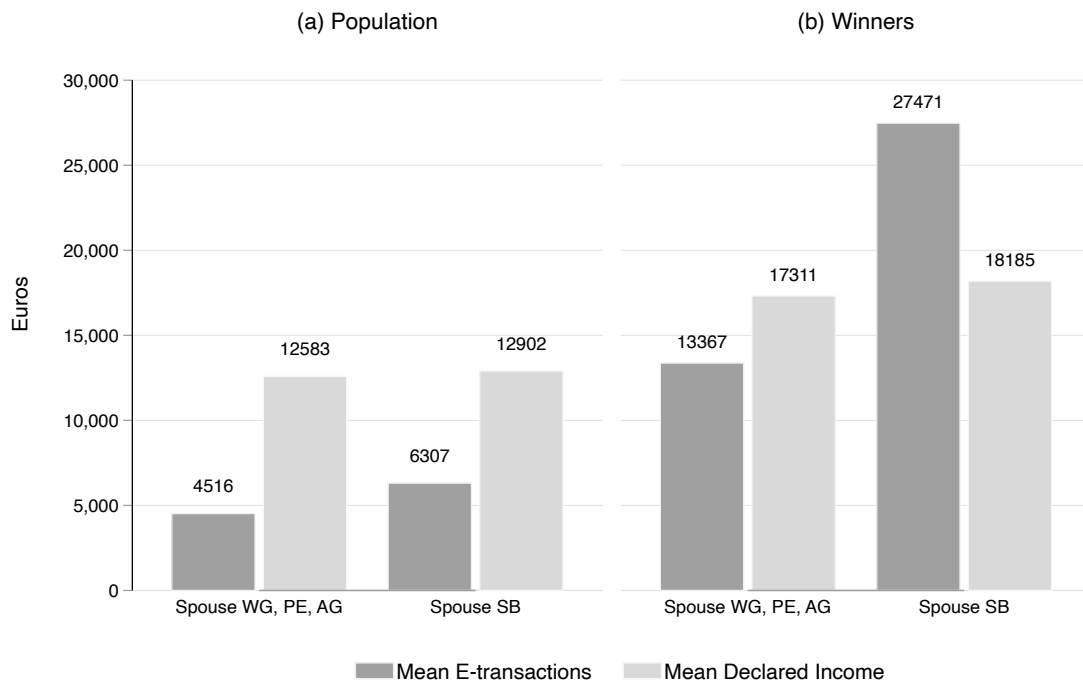
2.7.1 Complementary Figures and Tables

Fig. 2.7 Mean Annual E-transactions vs. Annual Declared Income



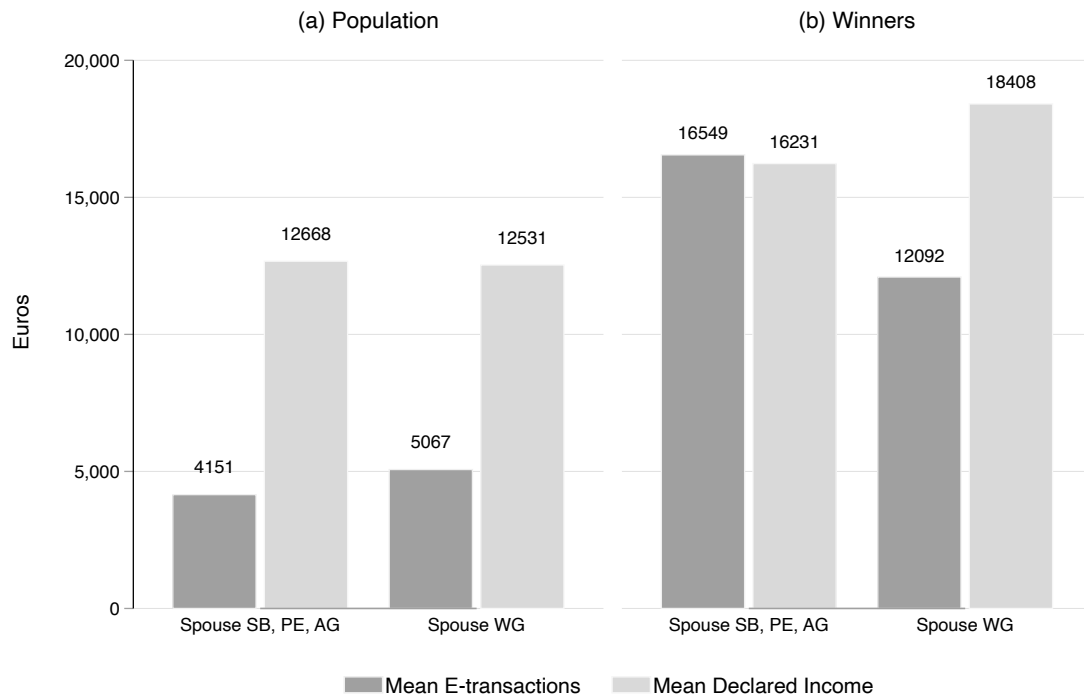
Notes: The figure compares the annual mean electronic transactions against the mean annual declared income of winners and of the representative taxpayer population. Only winners from lotteries in 2017 are included in the winners' sample. Non-parametric estimates of the differences are provided in Table 2.8

Fig. 2.8 Declared Annual Income and E-transactions, Taxpayers with SB spouses



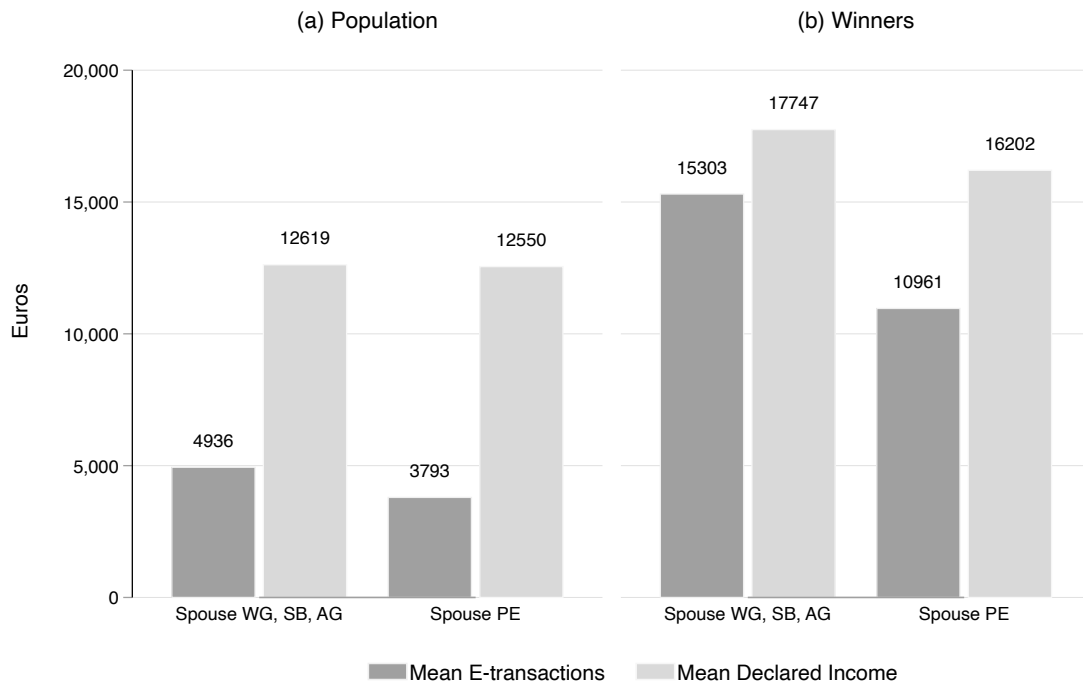
Notes: The figure compares the mean annual declared income and mean annual e-transactions of the taxpayer population, figure (a) on the left-hand side and for winners, figure (b) on the right-hand side. The left-hand side columns of each figure include individuals who have a spouse with primary income from WG, PE and AG, against individuals who have a spouse with primary income from SB. Individuals with primary SB income and NO income are excluded from the sample.

Fig. 2.9 Declared Annual Income and E-transactions, Taxpayers with WG spouses



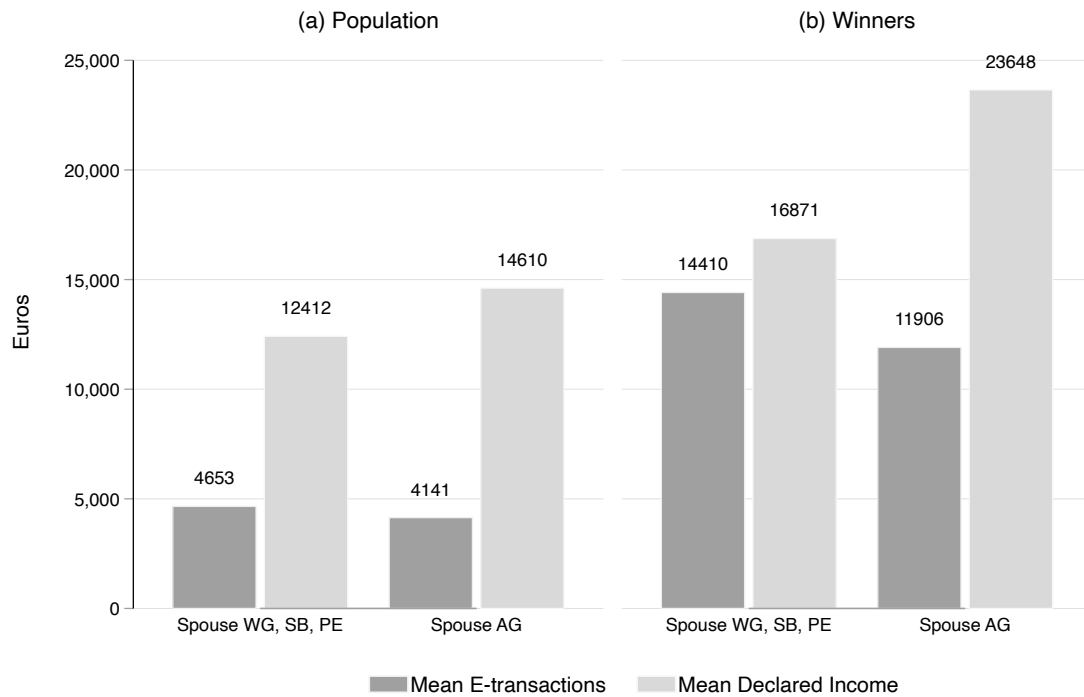
Notes: The figure compares the mean annual declared income and mean annual e-transactions of the taxpayer population, figure (a) on the left-hand side and for winners, figure (b) on the right-hand side. The left-hand side columns of each figure include individuals who have a spouse with primary income from SB, PE and AG, against individuals who have a spouse with primary income from WG. Individuals with primary SB income and NO income are excluded from the sample.

Fig. 2.10 Declared Annual Income and E-transactions, Taxpayers with PE spouses



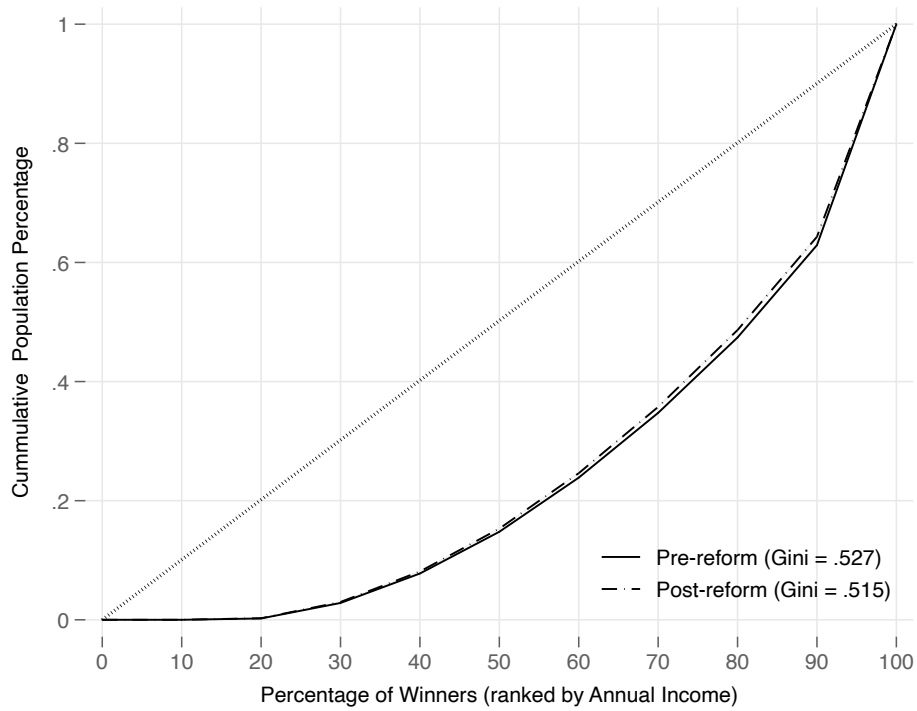
Notes: The figure compares the mean annual declared income and mean annual e-transactions of the taxpayer population, figure (a) on the left-hand side and for winners, figure (b) on the right-hand side. The left-hand side columns of each figure include individuals who have a spouse with primary income from WG, SB and AG, against individuals who have a spouse with primary income from PE. Individuals with primary SB income and NO income are excluded from the sample.

Fig. 2.11 Declared Annual Income and E-transactions, Taxpayers with AG spouses



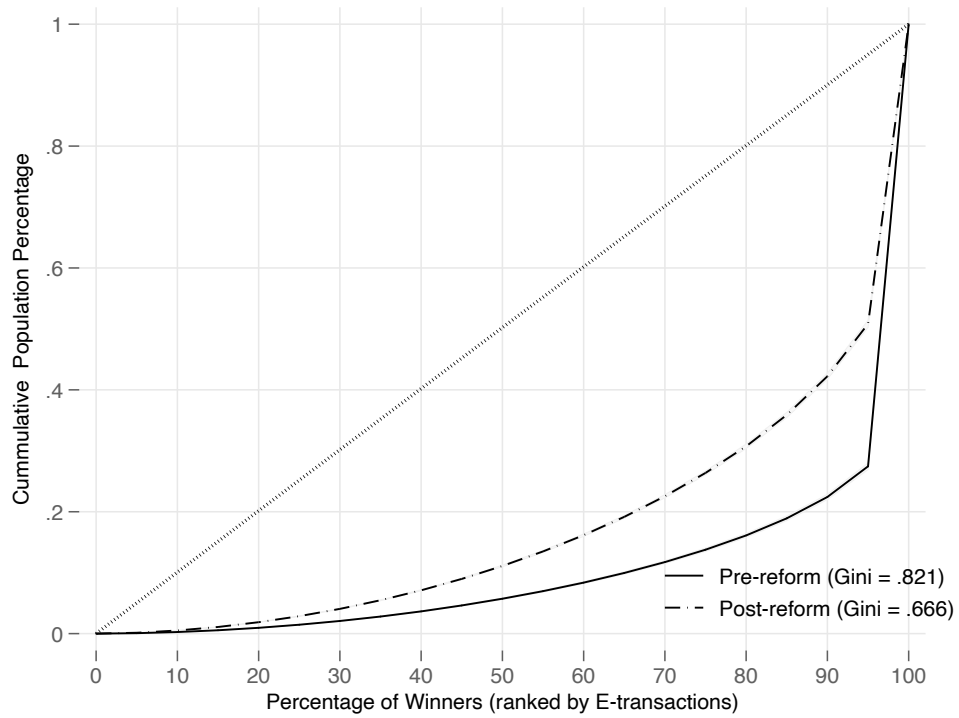
Notes: The figure compares the mean annual declared income and mean annual e-transactions of the taxpayer population, figure (a) on the left-hand side and for winners, figure (b) on the right-hand side. The left-hand side columns of each figure include individuals who have a spouse with primary income from SB, PE and WG, against individuals who have a spouse with primary income from AG. To allow for a meaningful comparison, SB individuals are excluded from the sample since these have exhibited a very high volume of e-transactions as shown in Fig. 2.4. NO income category and single filings are excluded from the sample.

Fig. 2.12 Income Distribution



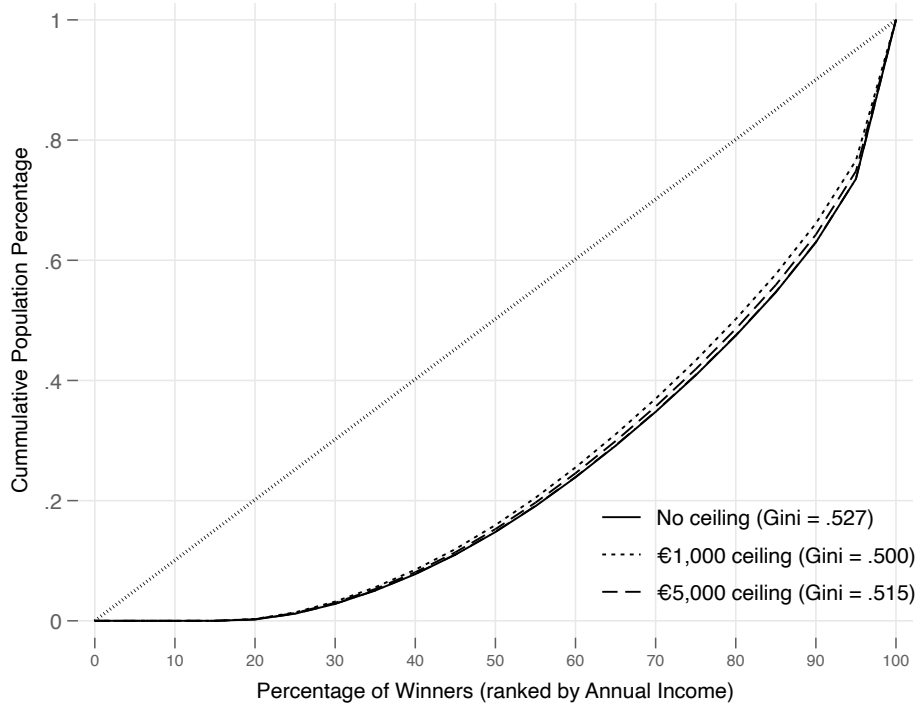
Notes: The figure plots e-transaction distribution curves for 2×1.2 million winners in Monte Carlo Simulations of the lottery. The x-axis represents the population percentiles of winners, ranked by annual income in 2017. The y-axis shows the percentage of individuals who have won the lottery in the simulations. The dotted line is a 45-degree line, at which the population percentage equals the winners percentage in the distribution. For the pre-[post-]reform curve, tickets were assigned based on the initial [post-reform] TAM presented in Table 2.1 [Table 2.4].

Fig. 2.13 E-transactions Distribution



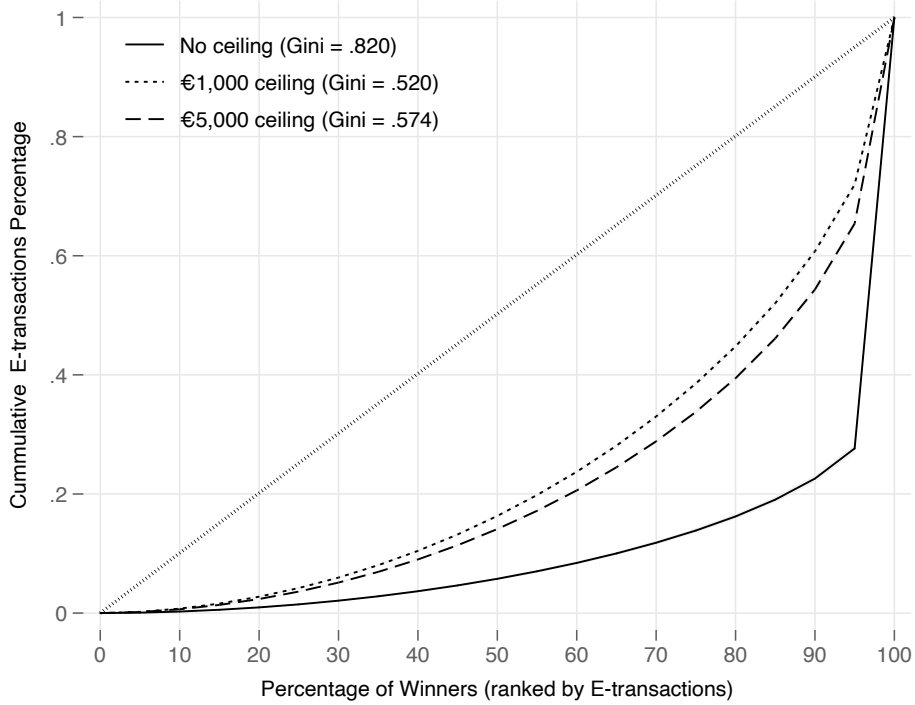
Notes: The figure plots e-transaction distribution curves for 2×1.2 million winners in Monte Carlo Simulations of the lottery. The x-axis represents the population percentiles of winners, ranked by annual e-transactions in 2017. The y-axis shows the percentage of individuals who have won the lottery in the simulations. The dotted line is a 45-degree line, at which the population percentage equals the winners percentage in the distribution. For the pre- [post-]reform curve, tickets were assigned based on the initial [post-reform] TAM presented in Table 2.1 [Table 2.4].

Fig. 2.14 Income Distribution for Stricter Ceilings



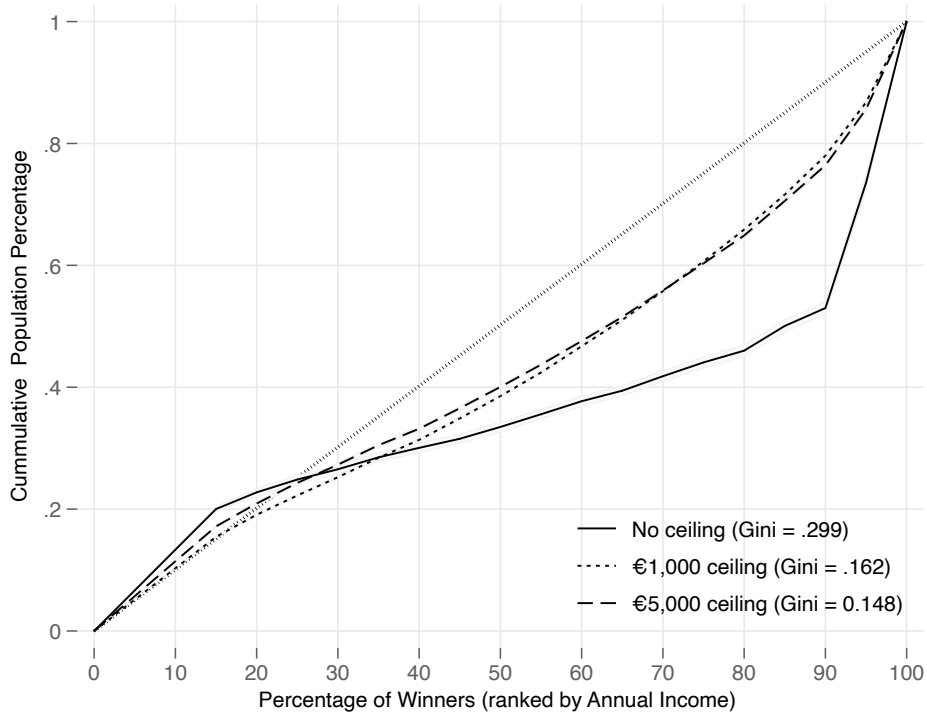
Notes: The figure plots income distribution curves for 1.2 million winners in Monte Carlo Simulations of the lottery. The x-axis represents the population percentiles of winners, ranked by their declared annual income in 2017. The y-axis shows the percentage of individuals who have won the lottery in the simulations. The dotted line is a 45-degree line, at which the population percentage equals the winners percentage in the distribution. The "no ceiling" curve is a simulation of the lottery assigning tickets using the pre-reform TAM in Table 2.1. The "€ 1,000 ceiling" curve retains the pre-reform TAM characteristics, but introduces a maximum ceiling in monthly tickets. For e-transactions beyond € 1,000 per month no more tickets are awarded to individuals. Similarly, the € 5,000 curve, retains the characteristics of pre-reform TAM, but introduces a ceiling at the € 5,000 monthly e-transaction level.

Fig. 2.15 E-transactions Distribution for Stricter Ceilings



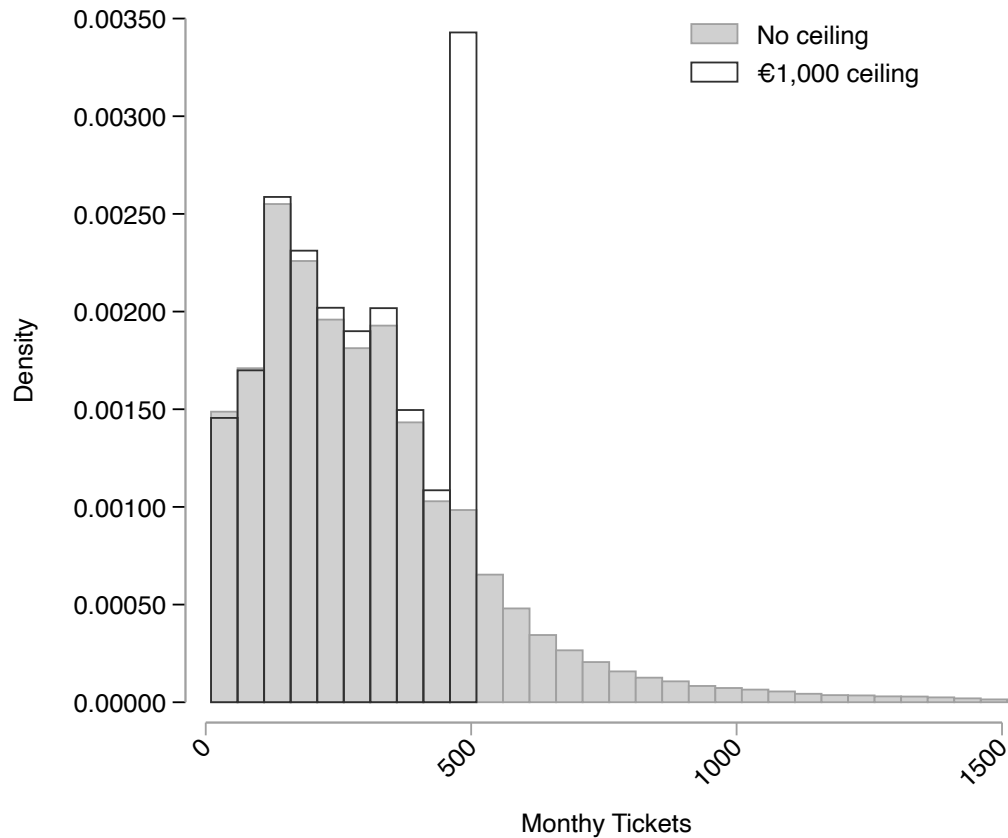
Notes: The figure plots e-transactions distribution curves for 1.2 million winners in Monte Carlo Simulations of the lottery. The x-axis represents the population percentiles of winners, ranked by e-transactions in 2017. The y-axis plots the cumulative percentage of e-transactions resulting from simulated lottery winners in pre- and post-reform TAM. The y-axis shows the percentage of individuals who have won the lottery in the simulations. The dotted line is a 45-degree line, at which the population percentage equals the winners percentage in the distribution. The "no ceiling" curve is a simulation of the lottery assigning tickets using the pre-reform TAM in Table 2.1. The "€ 1,000 ceiling" curve retains the pre-reform TAM characteristics, but introduces a maximum ceiling in monthly tickets. For e-transactions beyond € 1,000 per month no more tickets are awarded to individuals. Similarly, the € 5,000 curve, retains the characteristics of pre-reform TAM, but introduces a ceiling at the € 5,000 monthly e-transaction level.

Fig. 2.16 E-transactions Distribution (ranked by Income) for Stricter Ceilings



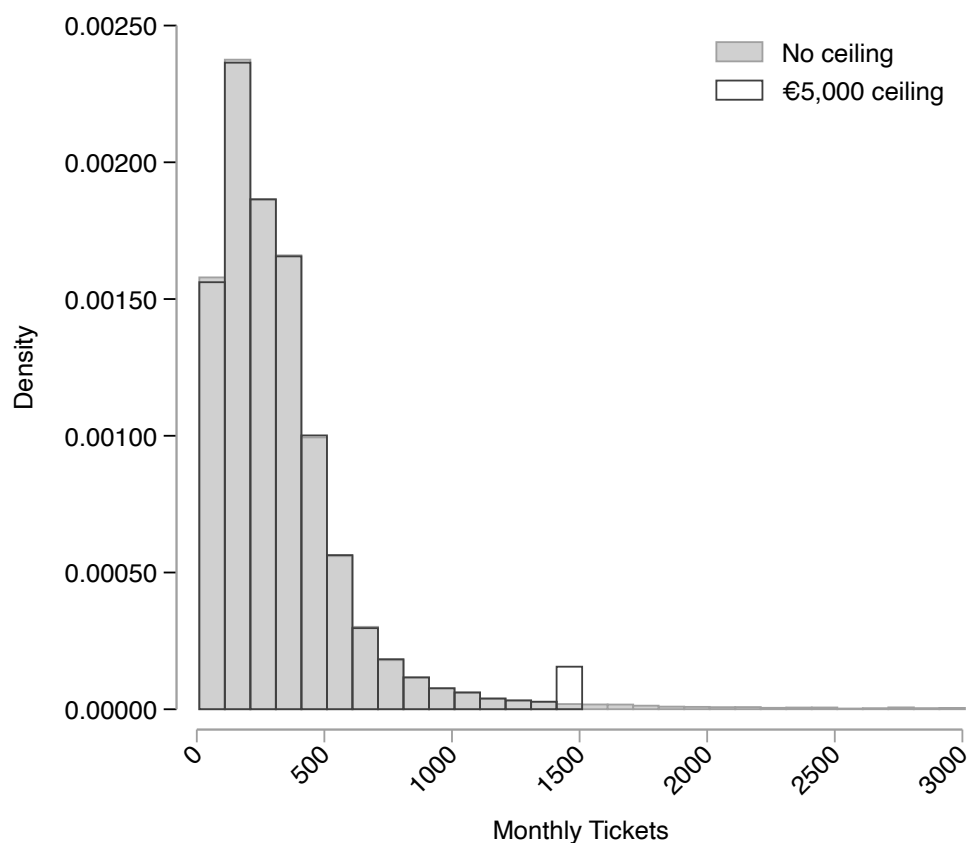
Notes: The figure plots e-transactions distribution curves for 1.2 million winners in Monte Carlo Simulations of the lottery. The x-axis represents the population percentiles of winners, ranked by annual income in 2017. The y-axis plots the cumulative percentage of e-transactions resulting from simulated lottery winners in pre- and post-reform TAM. The dotted line is a 45-degree line, at which the winners' income percentage equals the winners' e-transactions percentage. The "no ceiling" curve is a simulation of the lottery assigning tickets using the pre-reform TAM in Table 2.1. The "€ 1,000 ceiling" curve retains the pre-reform TAM characteristics, but introduces a maximum ceiling in monthly tickets. For e-transactions beyond € 1,000 per month no more tickets are awarded to individuals. Similarly, the € 5,000 curve, retains the characteristics of pre-reform TAM, but introduces a ceiling at the € 5,000 monthly e-transaction level.

Fig. 2.17 Winner's Distribution of Tickets in Simulations



Notes: The figure compares the distribution of winners' tickets assigned in Monte Carlo Simulations and the effect on tickets by placing a maximum ticket ceiling at € 1,000. This translates to a maximum number of 467 monthly tickets. Both simulations contain 1,200 iterations of the lottery (100 for each month in 2017), drawing 1,000 winners in each iteration. Both distributions contain 1.2 million winners. For the "No ceiling" distribution tickets were assigned using the the pre-reform TAM in Table 2.1. The "€ 1,000 ceiling" distribution retains the pre-reform TAM characteristics, but introduces a maximum ceiling in monthly tickets. For e-transactions beyond € 1,000 per month no more tickets are awarded to individuals. The distributions is truncated at 1,500 tickets, as right-tails diminish quickly in the distribution beyond this point.

Fig. 2.18 Winner's Distribution of Tickets in Simulations



Notes: The figure compares the distribution of winners' tickets assigned in Monte Carlo Simulations and the effect on tickets by placing a maximum ticket ceiling at € 5,000. This translates to a maximum number of 1,467 monthly tickets. Both simulations contain 1,200 iterations of the lottery (100 for each month in 2017), drawing 1,000 winners in each iteration. Both distributions contain 1.2 million winners. For the "No ceiling" distribution tickets were assigned using the the pre-reform TAM in Table 2.1. The "€ 5,000 ceiling" distribution retains the pre-reform TAM characteristics, but introduces a maximum ceiling in monthly tickets. For e-transactions beyond € 1,000 per month no more tickets are awarded to individuals. The distributions are truncated at 3,000 tickets, as right-tails diminish quickly in the distribution beyond this point.

Table 2.7 Basic Summary Statistics

	Samples:		Single/Joint Filing:	
	Non-Winners	Winners	Single Filers	Joint Filers
	Freq (Percent)	Freq (Percent)	Freq (Percent)	Freq (Percent)
<i>By Primary Income Source:</i>				
SB : Self-Employed/Business Income	2,052 (4.10)	1,609 (8.52)	1,855 (4.46)	1,806 (6.61)
WG : Wage-Earner	22,335 (44.67)	9,107 (48.19)	17,205 (41.38)	14,237 (52.11)
PE : Pensions (Main and Auxiliary)	12,163 (24.33)	4,201 (22.23)	8,979 (21.60)	7,385 (27.03)
AG : Agriculture	2,635 (5.27)	831 (4.40)	1,463 (3.52)	2,003 (7.33)
NO : Zero-declared Income	10,815 (21.63)	2,861 (15.14)	12,072 (29.04)	1,604 (5.87)
No Filing : Tax return not submitted	-	288 (1.52)	-	288 (1.05)
Total	50000	18897	41574	27323

Notes: The table presents basic summary statistics for the tax unit samples, per income source category. The left-hand side columns present the number of observations and percentages (in parentheses), of the non-winners and winners samples in the tax lottery. The winners sample includes winners in 19 consecutive months, from January 2017 to July 2018. The non-winners sample has been randomly drawn from the population of tax units. The right-hand side columns present the frequencies and percentages of single and joint-filing in each primary income source category. Joint-filers can be indirectly deduced from the sample. based on annual declared income from both spouses in a household. The case where the main taxpayer declares positive income and the spouse zero income cannot be identified in the sample.

Table 2.8 Non-parametric Estimates, by Primary Income Category

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Annual Income</i>	SB	WG	PE	AG	NO	No Filing
Winner in 2017	8,665*** (1,192)	6,952*** (518)	2,760*** (134)	29,884*** (4,807)	0 (0)	0 (0)
Constant	12,088*** (568)	11,404*** (92)	11,870*** (55)	17,538*** (733)	0 (0)	0 (0)
<i>Annual E-transactions</i>						
Winner in 2017	170,733*** (22,125)	6,807*** (601)	7,655*** (1,305)	11,732*** (1,491)	24,715*** (4,628)	53,286 (33,389)
Constant	10,786*** (860)	4,050*** (37)	3,309*** (41)	3,800*** (126)	2,903*** (125)	4,925*** (1,137)
Observations	266,317	2,890,322	1,573,228	340,746	1,397,996	288

Notes: The table presents estimation results per primary income source. Results on top of the table use annual declared income as independent variable and at the bottom, annual electronic transactions. The NO and No Filing categories in columns (5) and (6) do not record results for annual income regressions, since no income was declared. Annual e-transactions are observed for the latter categories. Robust standard errors (clustered at the individual level, depending on the number of individuals in each income source category) in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.9 Estimates - Spouse's Primary Income Source

	(1)	(2)	(3)	(4)	(5)	(6)
	e-trans	e-trans	e-trans	e-trans	e-trans	e-trans
Spouse's Primary Income SB	2,006*** (317)		1,778*** (320)	1,781*** (256)	1,791*** (268)	1,747*** (256)
Spouse's Any Income from SB		2,321*** (357)				
Annual Income in 2017			0.153*** (0.027)	0.139*** (0.021)		0.138*** (0.021)
Winner in 2017					8,851*** (1,369)	8,196*** (1,374)
Winner in 2017 & Spouse SB					12,313** (6,088)	12,236** (6,081)
Constant	4,970*** (119)	4,943*** (119)	3,049*** (328)	2,788*** (266)	4,516*** (79)	2,774*** (265)

Notes: The table presents estimation results for individuals with SB spouses. The sample is restricted to 2,406,971 individuals in the population who filed jointly in regressions (1) - (3). Observations are restricted to 2,279,469 in (4) - (6), which include joint-filers, but exclude individuals who declared SB as their primary income source. Robust standard errors (clustered at the level of 27,323 and 25,517 unique taxpayers for regressions (1) - (3) and (4) - (6) respectively) in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

2.7.2 Construction of the Baseline Population of Tax Units

In addition to the information from two samples, I observe the total number of lottery tickets issued in each calendar month, \bar{T}_m . Given that lottery tickets are derived from monthly e-transactions via the TAM described in Table 2.1, I 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, I utilise $\hat{T}_{i,m,1}$.

Given this, the following identity must hold:

$$\sum_{m=1}^{12} \bar{T}_m = \sum_{m=1}^{12} \sum_{i=1}^{N_1} T_{i,m,1} + \sum_{m=1}^{12} \sum_{i=1}^{N_1} \hat{T}_{i,m,1} + \omega \sum_{m=1}^{12} \sum_{i=1}^{N_2} T_{i,m,2} \quad (2.2)$$

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 ω , the weight or expansion factor that I have to use to arrive at a sample that matches the population in terms of lottery tickets, since it is the only unknown. One can observe the total number of tickets in 2017, $\sum_{m=1}^{12} \bar{T}_m$ and the total number of tickets in the samples.

A further plausibility check is that $N_1 + \omega N_2 \cong N$. The calculation derives ω to be 129. Expanding the random sample gives a total tax unit population of 6.45 million ($50,000 \times 129$), 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 2017.³⁹

In addition to expanding the population, I have explored other ways that would allow a direct comparison and solve the pre-selection of winners problem. One of this has been to obtain an additional random sample without pre-selecting winners from the population. That would had produced a random sample over the entire population. However, since the number of winners is very small, only very few winners would have been selected, therefore inhibiting a comparison of their characteristics. The random sample drawn in this case would have required to be particularly large in order to reach a point where a large number of winners are selected, to allow for a meaningful comparison. Thus, pre-selecting the winners was essential for providing insights to the lottery. Lastly, the random sample of 50,000 non-winners is large enough as to allow for a good approximation of characteristics to the actual population of taxpayers.

³⁹Annual statistics for the 2017 filing are published by the Tax Authority at <https://www.aade.gr/menoy/statistika-deiktes/eisodima/etisia-statistika-deltia>.

Chapter 3

Income Tax Incentives for Electronic Payments: Evidence from Greece's Electronic Consumption Tax Discount

3.1 Introduction

The taxation literature is abundant with examples of how taxes on income are being used to encourage or discourage behaviour. For instance, income tax deductions offer incentives for saving, investment and charitable giving; child credit supports population growth and tackles income disparities; earned-income tax credit incentivises work and reduces poverty. Can income tax incentives be extended to address other objectives? And how effective are features of the income tax system in incentivising behavioural change?

This chapter studies an extension of such incentives to the payments behaviour, using the introduction of the Electronic Consumption Tax Discount in Greece (or ECTD hereafter).¹ The policy attempts to induce third-party reporting information on a large scale, by conditioning the annual personal tax allowance on electronic consumption. Incentives encourage the use of electronic payments over cash to improve tax compliance. The chapter presents evidence of how taxpayers respond; either by increasing electronic consumption or by reporting increased amounts during tax filing to gain the full tax allowance. Subsequently, the chapter proposes an explanation of how adjustment costs, in the form of policy inattention, liquidity constraints and perceived costs of audit, shape the observed responses.

The ECTD forms part of third-party reporting policies, the role of which in improving tax compliance has generated a substantial body of empirical and theoretical work (Kopczuk and Slemrod, 2006; Kleven *et al.*, 2011; Kleven, 2016). In value-added taxes, most of these schemes attempt to address the "last mile" problem. Business-to-business transactions are self-enforcing since inputs to production can be tax exempt and firms have a direct economic benefit in requesting transaction receipts. The third-party information generated has been shown to

¹In Greece this policy became known as "electronic receipts". It was introduced in December 2016 with L.4446/2016 and implemented as of January 2017.

facilitate enforcement (Pomeranz, 2015). However, the self-enforcement element breaks down in retail sales to the final consumer, as there are no incentives to report transactions (Slemrod, 2007).

Hitherto, governments have devised a variety of schemes to enhance incentives and create value in retail transactions, all of which have relied on active consumer participation. Tax lotteries have been a widely-used instrument in recent years as surveyed in Fookien *et al.* (2015). However, only a few schemes have been documented in economic literature. For instance, the Brazilian lottery in Naritomi (2019) has been shown to have increased reported sales in firms by 21%, resulting in an increase of 9% in tax revenue. Similarly, consumer VAT rebates in Uruguay increase card transactions as shown in Brockmeyer and Sáenz Somarriba (2022).

A noteworthy feature of ECTD, which distinguishes it from other third-party reporting policies, is the use of the income tax system as a means for incentivisation. Instead of relying on active participation, taxpayers are included by default through their personal tax allowance which is granted conditional on spending by electronic means (credit cards or electronic transfers). The scheme is also unique in the information technology system used for monitoring transactions. Aggregate amounts of electronic payments completed by all Greek taxpayers are transferred automatically every month from financial institutions to the tax authority. Tax ID and IBAN are then matched to provide an annual aggregation of electronic consumption, which appears directly in the tax returns of the following year. Using self-assessment, the taxpayers can accept or modify the pre-filled amounts and the reported electronic consumption amount is used to calculate the final tax bill.

To analyse how the scheme's incentives work and to motivate the empirical analysis, I embed the taxpayers' choice in a stylised model. Taxpayers can either increase electronic consumption or change the pre-filled information during tax filing, a process closely resembling how ECTD works. The tax discount they receive increases conditional on the level of electronic consumption during the financial year, up to a maximum amount (or *threshold* hereafter). The threshold is determined by the combination of taxable income and an exogenous percentage set by the government. Since thresholds are based on taxable income, they are taxpayer-specific and their task is to place an upper limit on the amount of tax discount that can be granted.

The model produces three testable predictions. Firstly, that reporting electronic consumption on or beyond the threshold (or threshold-targeting hereafter) is optimal. Through threshold-targeting taxpayers experience a positive income effect from maximising their tax discount, thereby avoiding an additional tax liability. The second and third predictions relate to how taxpayers choose to engage in threshold-targeting. For this the model defines a margin of responses; either increasing their electronic consumption during the year or changing the pre-filled information during tax filing.

All predictions are tested using a randomly-drawn administrative sample of 50,000 Greek taxpayers. The data include information on annual reported electronic consumption and monthly volumes of actual electronic consumption in 2017 and 2018, as transferred by financial institutions to the tax authority at the end of each month. By aggregating the monthly electronic consumption volumes, one can retrieve the annual aggregate amount for each taxpayer and

evaluate the difference between what has been pre-filled in their tax returns versus what has been reported during tax filing. The data are complemented by tax returns information, including the taxable income declared by taxpayers, which allows for the calculation of their threshold values.

Consistent with the first prediction, I find evidence of threshold-targeting. A substantial visual mass of taxpayers (8%) report electronic consumption on or close to their threshold and the vast majority (84%) reports beyond their threshold, with only the remaining (8%) reporting less. Driven by tax discount maximisation, the distinct cutoff on and beyond threshold provides evidence of strong responses to the policy. At the same time, I document a mismatch between reported and pre-filled amounts, with less taxpayers concentrating on threshold (4%) and beyond threshold (69%), while a higher mass (27%) exhibits pre-filled amounts less than their threshold. Individuals with pre-filled amounts less than the threshold, increase their reported values during tax filing.

Further evidence of changes to pre-filled amounts are documented by (a) constructing within-taxpayer differences, which capture the amounts taxpayers actually spent versus the amounts reported in tax returns, and (b) by decomposing the sample to those below and above threshold at the end of the financial year. This analysis reveals the presence of rounding effects above threshold when below-threshold taxpayers report consumption. Visual excess mass is concentrated on multiples of fifty and hundred euros, indicating buoyancy of the pre-filled amounts.

Reporting higher amounts to gain the full tax discount is an unintended consequence of the policy. The implications of this result are even more important given the occupational type of taxpayers, who are predominantly wage-earners and pensioners, having traditionally lower evasion opportunities than the self-employed (Engström and Holmlund, 2009; Hurst *et al.*, 2014). Evidence of increases in the reported amounts by these occupational groups suggests that taxpayers react rationally to these opportunities, regardless of their occupational type.

To assess if taxpayers increase their electronic consumption in response to the policy, I use monthly event studies with the end-of-year deadline as a cutoff. The data include all months of 2017, in which the threshold is built, and 7 months of 2018. Taxpayers are then grouped in cohorts, based on the month their threshold was reached and an additional group with individuals who did not reach threshold. The evolution of electronic consumption in relation to the end-of-year deadline is assessed for each of these cohorts.

The ECTD affects electronic consumption in the following ways. Firstly, about half of the taxpayers in the regression sample reach their threshold by August 2017, 8 months after the policy's introduction. These individuals exhibit a moderate to high consumption pattern during the year, which enables them to achieve and surpass their threshold. Secondly, a significant mass of taxpayers who reach threshold in September, October, November and December exhibit end-of-year spikes in electronic consumption, up to 13% of annual income, as the deadline to build the threshold approaches. These cohorts have particularly low propensity of electronic consumption during the year, while the end of year spikes are inconsistent with their annual spending behaviour and seasonal spending fluctuations in other cohorts. Thirdly, there are

evidence of spending becoming entrenched in the beginning of the new financial year in cohorts who experienced consumption spikes. Their electronic consumption increases and stabilises at about 5% of annual income every month as the new threshold-building attempt begins.

A final result in consumption responses regards individuals who remain below threshold and form about one-third of the regression sample. These individuals exhibit minimal consumption during the financial year and end up responding by changing their pre-filled information during tax filing. As the new financial year begins, their electronic consumption increases incrementally from 1% to 3.6% of annual income.

Overall, these results suggest a mixture of responses in both the reporting margin and electronic consumption margin. These responses have implications for the policy's effectiveness. Changing pre-filled amounts to gain the discount implies limited responses in electronic consumption and no generation of third-party reporting information. In addition, end-of-year consumption spikes might have welfare implications for liquidity constraint taxpayers. What determines the mixture of responses we observe? And why do they differ?

I propose an explanation for the observed pattern through adjustment costs in the form of policy inattention, liquidity constraints and perceived costs of audit, which seem to shape taxpayers' responses. Due to policy inattention, taxpayers spend lower amounts at the beginning of the year. As the end of year approaches, higher policy attention induces individuals to spend more to reach their threshold before the deadline. Taxpayers begin their threshold building attempt earlier during the new financial year after their last end-of-year experience. Additionally, responses might also depend on liquidity constraints: despite a higher policy attention, liquidity constraints might prevent some taxpayers from reaching their threshold. Lastly, for taxpayers who remain below threshold, changing their pre-filled amounts might be explained by relatively low perceived costs of audit and penalty. This seems to be a reasonably fitting explanation of the various adjustment costs that shape taxpayers' responses to the ECTD. While features of the income tax can exacerbate a change in behaviour, the overall result depends on taxpayers' adjustment costs.

This chapter relates to the third-party reporting literature at the consumer level. A handful of schemes have been analysed so far, notably Naritomi (2019) on the Brazilian lottery and Brockmeyer and Sáenz Somarriba (2022) on Uruguay's VAT rebates. The majority of studies in third-party reporting focuses on the role of information trail on firm behaviour (Almunia and Rodriguez, 2014; Pomeranz, 2015; Slemrod *et al.*, 2017; Waseem, 2022) and on reporting frictions that may inhibit the effectiveness of VAT policies (Carrillo *et al.*, 2017). This chapter is the first to present consumer level third-party reporting incentives through the income tax system.

The chapter also contributes to a growing literature on the use of digitalisation to fight tax evasion (Gupta *et al.*, 2017; Haichao Fan and Wen, 2018; Bellon *et al.*, 2019). In particular, a number of studies have taken advantage of Greece as a testing ground, given the fast pace of reforms during the years of economic adjustment and an increasing digitalisation trend. Danchev *et al.* (2020) study the penetration of electronic payments following capital controls in 2015 and Hondroyiannis and Papaoikonomou (2017) using macroeconomic data, estimate that 1% increase

in card payments leads to 1% increase in VAT tax revenue. These estimates are in line with other studies at the EU level (Madzharova, 2020; Immordino and Russo, 2018).

Lastly, the analysis builds on closely-related studies on bunching (Kleven, 2016), that have documented responses at points of interest such as kinks (Saez, 2010; Chetty *et al.*, 2011; Bastani and Selin, 2014) and notches (Kleven and Waseem, 2013). In particular, tax reforms combined with the use of consumption expenditure provide a fertile ground to document such responses and related costs, as shown in Gorodnichenko *et al.* (2009). Also, recent studies have focused on how adjustment costs shape responses at these points (Chetty *et al.*, 2011; Adam *et al.*, 2021; Gelber *et al.*, 2020). I build on these findings, to analyse responses to thresholds and the role of adjustment costs.

The chapter is organised as follows. Section 3.2 introduces the ECTD and provides a detailed description of its main elements. Section 3.3 develops a conceptual framework to analyse how it functions and derives testable predictions. Section 3.4 provides an empirical analysis, documenting threshold-targeting and responses in the reporting and electronic consumption margins. Section 3.5 interprets the findings and Section 3.6 concludes.

3.2 Institutional Background

Greece's attempt to increase tax revenue during the economic crisis in 2010 gave rise to a number of novel policies one of which was a requirement for taxpayers to collect a percentage of their total annual consumption in paper receipts in exchange for a personal tax allowance.² The policy aimed at incentivising taxpayers to demand paper receipts for their transactions, as a way to increase paper trail and fight tax evasion. Implementation of the policy relied on enforcement, as taxpayers had to present their paper receipts in case of an audit and were faced with additional tax bill if they failed to produce the required amounts.³

Paper receipt collection evolved drastically in 2017 to include only electronic payments, aided by the introduction of an information technology system by the tax authority, which linked individual taxpayer information to their bank accounts.⁴ This formed part of a strategy to incentivise the use of electronic payments against the use of cash in the economy, and thereby

²The amount of paper receipts to be collected was set initially at 25% of annual income, and changed in the years that followed to 10%, increasing marginally on income. The policy evolved out of negotiations between Greece and its international lenders during the first economic adjustment programme in 2010. The IMF developed a policy position of expanding the Greek tax base to generate revenue, by lowering the tax allowance limit which stood at € 12,000 at the time, the existence of which resulted in a large number of taxpayers not having any income tax obligations. Given the political cost of lowering the limit, the counter-proposal of the Greek delegation was to link the annual income tax allowance to a level of paper receipts, thereby expanding the tax base by fighting tax evasion.

³Evaluation of the measure is absent in economic literature and in government's reports. However, it must have been very difficult to apply effectively on a large scale. Firstly, auditors had to check receipts one-by-one for millions of transactions, costing them time from other audits. Secondly, receipts could be transferable to other individuals. For example, tax accountants could keep a single file of receipts that could be used for any of their clients in case of an audit.

⁴Individuals declare their tax ID when opening a bank account and their IBAN number during tax filing, facilitating automatic matching.

generate third-party information on a large scale.⁵ More specifically, financial institutions were obliged to send to the tax authority the aggregate amount of electronic transactions (but not single transactions) for all taxpayers at the end of the month. The tax authority could then aggregate the amount of electronic consumption completed by every taxpayer in the economy and check automatically if the required level was reached. Two new tax codes were introduced in tax returns (049 and 050 for taxpayer and spouse, respectively) with the pre-filled amounts of payments completed and allowing individuals to either confirm or modify the final amounts.

The policy links the personal tax allowance to electronic consumption as follows. Income is taxed progressively according to the tax brackets in Table 3.2 in Appendix 3.7.1, and taxpayers receive a 22% tax discount on each euro spent by electronic consumption, up to a maximum amount (or *threshold*). Note that the discount is equivalent to the tax rate of the first income bracket, essentially making the amount of income spent up to the threshold, income tax free. This is equivalent to receiving a personal tax allowance, as long as the threshold amount is reached. The threshold and ECTD are taxpayer-specific, calculated as a percentage of declared taxable income, using Table 3.1.⁶ It increases marginally on income up to € 160,000, thus higher incomes up to this level are incentivised to spend more by electronic means. This has been the main form of personal tax allowance in Greece since 2017.

Table 3.1 Threshold and ECTD

Income Bracket €	Marginal Rate %	Threshold Bracket €	ECTD €
0 - 10,000	10	0 - 1,000	0 - 220
10,001 - 30,000	15	1,000 - 4,000	220 - 880
30,001 - 160,000	20	4,000 - 30,000	880 - 6,600
160,000 <	0	30,000	6,600

ECTD-eligible taxpayers, are those receiving income from wages, pensions and agriculture, constituting the majority of the taxpayer population in Greece. The main exemptions are the self-employed, sole proprietors and business owners, taxpayers above the age of 70, individuals declaring null income, guaranteed minimum income receivers and residents in villages with less than 500 residents and islands with less than 3,100 residents (but which are not touristic locations). In addition, a special provision applies to joint-filing households.⁷

⁵Other measures in the strategy included the compulsory introduction of point-of-sale machines in all firms (gradual roll-out from mid-2017 on-wards, based on occupation) and the introduction of a tax lottery using electronic transactions to generate eligible tickets (monthly draws began in October 2017)

⁶For example, consider a taxpayer with taxable income of € 5,000 in 2017. The threshold corresponds to € 500 (10% of income) in electronic consumption, giving the taxpayer a tax discount of 22% per euro spent up to € 500. The maximum tax discount for the individual is € 110 (500x0.22). If the taxpayer spends less, the discount is reduced accordingly; for example if only € 250 is spent, the discount is € 55, and the taxpayer pays € 55 more in tax than if having received the full discount. Electronic consumption in excess of the threshold is not rewarded.

⁷All married individuals were obliged to file jointly in 2017. Each electronic consumption amount is declared individually in the tax returns (049 for husband and 050 for wife), but if the amount of one person is higher than the personal threshold while for the other person's amount is lower, the extra amount is taken into account when calculating the other person's tax obligation. For example, if an individual is 500 EUR below the personal threshold and the partner 500 EUR above the threshold, then the full tax discount applies.

Consumption expenditures that count towards ECTD are those completed by electronic means (debit or credit cards and online banking) in Greece and in other EU countries. Non-consumption expenditure such as rents and mortgage payments, government payments (fines and taxes), buying vehicles (motorbikes, cars, boats), any kind of investment products (bonds or stocks) do not count towards the threshold. By excluding these categories, the policy aims at increasing electronic payments in everyday consumption transactions.

Reporting electronic consumption takes place in the annual tax returns, the filing of which is compulsory in Greece for all adults (even if declaring null income). Aggregation of payments happens during the financial year (January to December), followed by tax filing starting in April and ending in July the following year. During the tax year, financial institutions report monthly to the tax authority on the amount of electronic payments completed per taxpayer and the latter links the taxpayer's IBAN to their tax ID. The total annual amount of eligible electronic payments appears in dedicated fields during filing (codes 049-050) as shown in Figure 3.9 in Appendix 3.7.2. Before submitting the form, the taxpayers can observe what the tax authority has registered as electronic payments for the year, and they can edit the fields and declare a different amount, higher or lower, or accept the amount shown. Tax obligations are calculated once filing is completed, based, amongst others, on the declared amount of electronic payments, which is used to determine the final tax discount. Thus, the final amount of electronic payments declared has a direct and immediate effect on the taxpayer's overall tax obligation.

3.3 Conceptual Framework

3.3.1 Incentives for Electronic Consumption

In order to analyse the taxpayer behaviour, consider a static labour supply model with a choice for cash and electronic consumption.⁸ Assume a quasi-linear, iso-elastic utility function, U , which depends positively on after-tax income c and negatively on before-tax income z , the former representing utility from consumption and the latter the effort from earning income. To analyse the choice between cash and electronic consumption, assume that labour income z is exogenous.

Utility from after-tax income c , is derived either from cash consumption, c_c , or electronic consumption, c_e ; the two being perfect substitutes and utility strictly increasing in both c_c and c_e . In this benchmark model, switching between c_c and c_e is costless and either cash or electronic consumption yield the same level of utility.

The taxpayer chooses c_c , c_e to maximise U , subject to a linear budget constrain:

$$U(c_c, c_e, z) = c_c + c_e - z \tag{3.1}$$

$$s.t \ c_c + c_e = (1 - \tilde{t})z - tF(z, a) + tc_e \tag{3.2}$$

⁸This modeling choice follows closely the elasticity of taxable income literature, as used in Feldstein (1999), Saez (2010) and as surveyed in Saez *et al.* (2012).

where $(1-\tilde{t})z$ represents net-of-tax income, with \tilde{t} income tax rate. An additional tax $t \in (0, 1)$, is linked to electronic consumption conditional on function $F(z, a)$, which represents the taxpayer's threshold, determined by labour income z and an exogenous parameter $a \in (0, 1)$. Initially the threshold implies an additional tax obligation corresponding to a percentage of income, that can be nullified by increasing electronic consumption c_e (equivalent to receiving a tax discount).⁹ In its simplest form, the threshold $F(z, a)$ is a percentage of labour income:

$$F(z, a) = za \quad (3.3)$$

Equation 3.3 defines the maximum tax discount that can be reached given income z . To avoid granting the tax discount beyond $F(z, a)$, I posit the following condition on t , where the marginal tax discount stops being granted when the threshold is reached:

$$t = 0 \quad \text{if} \quad c_e \geq F(z, a) \quad (3.4)$$

The model resembles how the ECTD works. The government sets exogenous policy parameters \tilde{t} , t , rates a ; i.e. the income tax rate, the tax discount rate and the percentage of electronic consumption conditional on income that determines the taxpayer's maximum discount threshold. Given an income z , the taxpayers choose electronic consumption, c_e , to maximise U . The tax discount increases on electronic consumption c_e , until the threshold $F(z, a)$ is reached. Any increase of c_e beyond $F(z, a)$ does not carry any additional discount.

Formally, the taxpayer's maximisation problem and first order condition become:

$$\underset{c_e}{\operatorname{argmax}} U(c_e, z) = (1 - \tilde{t})z - tF(z, a) + tc_e - z \quad (3.5)$$

$$FOC : \frac{\partial U}{\partial c_e} = 0 \implies t = 0 \quad (3.6)$$

Equation 3.6 implies that the maximising response of electronic consumption would be one that nullifies the additional tax. According to Equation 3.4, this takes place when the electronic consumption is equal or exceeds the threshold. In other words, the marginal benefit of increasing electronic consumption, which is the tax discount t , increases up to the point where additional electronic consumption yields no more benefits.

Proposition 1 (Threshold Targeting). *Assume cash consumption and electronic consumption are perfect substitutes. For a given level of labour income $z > 0$, tax discount t and threshold defined by $F(z, a)$: $U(c_e, z)$ is maximised at $c_e \geq F(z, a)$.¹⁰*

Overall, the model predicts that taxpayers respond by shifting electronic consumption on or above their personal threshold. Increasing electronic consumption implies a positive income effect derived from maximising the tax discount.

⁹In the ECTD, $F(z, a)$ is a piece-wise linear function, increasing on income as shown in Table 3.1. For simplicity, I model $F(z, a)$ to correspond to a mapping of electronic consumption that can be reached at every point of the before-tax income z .

¹⁰Proof provided in Appendix 3.7.3.

3.3.2 Pre-filled Information and Adjustment Costs

The benchmark model can be extended to include the choice of changing pre-filled information during tax filing and adjustment costs. This choice resembles closely the ECTD's self-assessment process during tax returns; the tax authority pre-fills the amounts spent during the financial year and the taxpayer can either accept the amount or modify it before the tax obligations are determined.

Define c_r as any excess electronic consumption reported beyond the pre-filled electronic consumption, c_e . Changing the amounts may result in some costs, captured by function ψ , associated with a higher audit probability and penalty as is standard in the tax compliance literature (Allingham and Sandmo, 1972). Assume also some costs of electronic consumption, captured by function ξ . Costs of electronic consumption include adjustment costs of behaviour from cash to electronic means.¹¹

The utility function and budget constraint (in Equations 3.1 and 3.2) are then modified accordingly:

$$U(c_c, c_e, c_r, z) = c_c + c_e - \xi(c_e) - \psi(c_r) - z \quad (3.7)$$

$$s.t \ c_c + c_e = (1 - \tilde{t})z - tF(z, a) + t(c_e + c_r) \quad (3.8)$$

Firstly, note that c_r does not enter the utility function directly, in contrast to c_c and c_e which represent actual consumption expenditure. Secondly, c_r can increase utility through the budget constraint by increasing the tax discount awarded in the case where electronic consumption, c_e , falls short of the threshold $F(z, a)$. As in the benchmark model, to avoid a tax subsidy if electronic consumption increases beyond the threshold, the condition of Equation 3.4 is modified accordingly to include the choice of increasing reported amounts:

$$t = 0 \quad \text{if} \quad c_e + c_r \geq F(z, a) \quad (3.9)$$

Equation 3.9 implies that a tax discount is awarded by either increasing electronic consumption or by increasing the reported amounts on or beyond $F(z, a)$ during tax filing. Any excess amount spent or reported is not rewarded beyond the threshold. The maximisation problem for the individual and the first order conditions become:

$$\underset{c_e, c_r}{\operatorname{argmax}} U(c_e, c_r, z) = (1 - \tilde{t})z - tF(z, a) + t(c_e + c_r) - \xi(c_e) - z \quad (3.10)$$

$$FOC \ 1: \frac{\partial U}{\partial c_e} = 0 \implies t = \xi'(c_e) \quad (3.11)$$

¹¹Both cost functions are assumed to be well-behaved. An explicit functional form of ψ and ξ is shown in Section 3.5.

$$FOC\ 2 : \frac{\partial U}{\partial c_r} = 0 \implies t = \psi'(c_r) \quad (3.12)$$

The first order conditions are similar to the benchmark model in Section 3.3.1, with the difference that a margin of responses is introduced between increasing electronic consumption or changing the pre-filled amounts when reporting. Firstly, note that in this extension of the model, it is still optimal for individuals to increase (electronic or reported) consumption on or above their threshold, by setting $t = 0$. By increasing c_e or c_r , the individual receives marginal benefit t (the tax discount) up to the threshold, $F(z, a)$. At maximum, one can combine Equations 3.11 and 3.12; the choice to either increasing consumption or changing pre-filled amounts depends on their marginal costs, $\psi'(c_r)$ and $\xi'(c_e)$.

Let electronic consumption costs be given by $\xi(c_e)$ and costs of changing pre-filled information by $\psi(c_r)$. Assume further a level of labour income $z > 0$, tax discount t and threshold defined by $F(z, a)$. Then, it can be shown:

Proposition 2 (Response Margin). *$U(c_e, c_r, z)$ is maximised at $c_e + c_r \geq F(z, a)$, with the response margin depending on the marginal costs of c_e and c_r , satisfying $\xi'(c_e) = \psi'(c_r)$ at maximum.*¹²

Intuitively, Proposition 2 implies a response margin for individuals to reach their threshold $F(z, a)$ and maximise their tax discount. They can respond through the *electronic consumption margin*, by increasing electronic consumption during the financial year. If electronic consumption is below the threshold, they can respond through the *reporting margin* by changing their pre-filled information during tax filing.

The theoretical analysis provides a clear framework to explain how the ECTD works and to assess the margin of responses by taxpayers. Theoretical predictions can be summarised as follows:

1. *Threshold Targeting.* Taxpayers maximise utility by either increasing electronic consumption or reporting it on or above their threshold.
2. *Responses in the Reporting Margin.* If electronic consumption is less than the threshold, utility is maximised by changing pre-filled amounts up to the threshold, subject to marginal costs of doing so.
3. *Responses in the Electronic Consumption Margin.* Electronic consumption increases on or beyond the threshold, subject to the marginal costs of doing so.

¹²Proof provided in Appendix 3.7.3.

3.4 Empirical Analysis

3.4.1 Data

I utilise a unique administrative dataset of 50,000 randomly-drawn taxpayers to study responses to the ECTD following the theoretical predictions of Section 3.3.¹³ The data include aggregate monthly amounts of electronic consumption per taxpayer from all 12 months of 2017 and from the first 7 months of 2018, in the form that these were transferred from financial institutions to the tax authority (rounded to the nearest € 10). This enables us to examine monthly electronic spending that counts towards the taxpayers' threshold and which pre-fills the tax returns. Monthly electronic consumption data is matched at the taxpayer's level with information from tax returns (rounded to the nearest € 5).

The data include information for either a single individual or two individuals who file jointly in a household, corresponding to 31,409 and 18,591 observations, respectively. For each tax unit with a single individual, the tax return data contain the declared annual income, the declared amount of electronic payments and the postcode. In addition, I observe the occupational income source, being (a) from wages (b) from pensions (c) from self-employed or business activity and (d) from agricultural income. For joint-filing tax units, the data contain almost the same information as single-filers. I observe the declared annual income and declared electronic consumption for both individuals as well as their occupational income sources. However, I observe monthly electronic consumption payments for only one of the two individuals in a household, thus, limiting the analysis of any intra-household effects on electronic consumption. Sample statistics by primary income source are shown in Table 3.3 in Appendix 3.7.1.

A special category is also included in the sample, shown in the last row of Table 3.3; individuals who have declared null income, representing 21.6% of the sample (10,815 individuals). This arises from a legal requirement of compulsory tax filing in Greece for all adults, even if the income is null. Thus, this group contains students above the age of 18 (for instance, in tertiary education), unemployed individuals, as well as tax units who conceal all of their income.¹⁴ Declaring null income implies that these individuals do not face a requirement to reach a specific threshold of electronic payments (since the income is null, so is the threshold).

The sample utilised to examine taxpayer responses is restricted to include those eligible for ECTD. These are individuals who earn up to €160,000 and receive their primary income source from wages, pensions or agriculture (see Section 3.2). Business owners and the self-employed are exempt from the policy. To limit intra-household effects (transfers of income, electronic consumption and sharing of thresholds between spouses) the analysis excludes joint-filer. The

¹³The data were provided by the Greek tax authority in October 2018 based on tax filings completed by July 30, 2018 (the last day of submission). The tax returns underwent a basic plausibility check and tax payment statements were issued by the tax authority in August 2018.

¹⁴Individuals in this group occupy a large proportion of the overall sample due the post-crisis recovery phase of the Greek economy in 2017. Firstly, annual youth unemployment stood at 43.5%, indicating a slow absorption rate of youth in the labour market and a large number of graduates declaring null income. Secondly, Greece experienced a very high unemployment rate of 21.5% in 2017. Thirdly, it is likely that more firms and workers shifted to the shadow economy during the crisis. These reasons can explain the high number of null-income declarations in the taxpayer population.

null income category is also excluded as individuals do not abide to a specific threshold. Summary statistics are shown in the "ECTD" column of Table 3.3. Overall, this subsample includes 20,676 taxpayers, 12,685 of which are wage-earners, 6,880 pensioners and 1,111 have income from agriculture.

3.4.2 Threshold Targeting

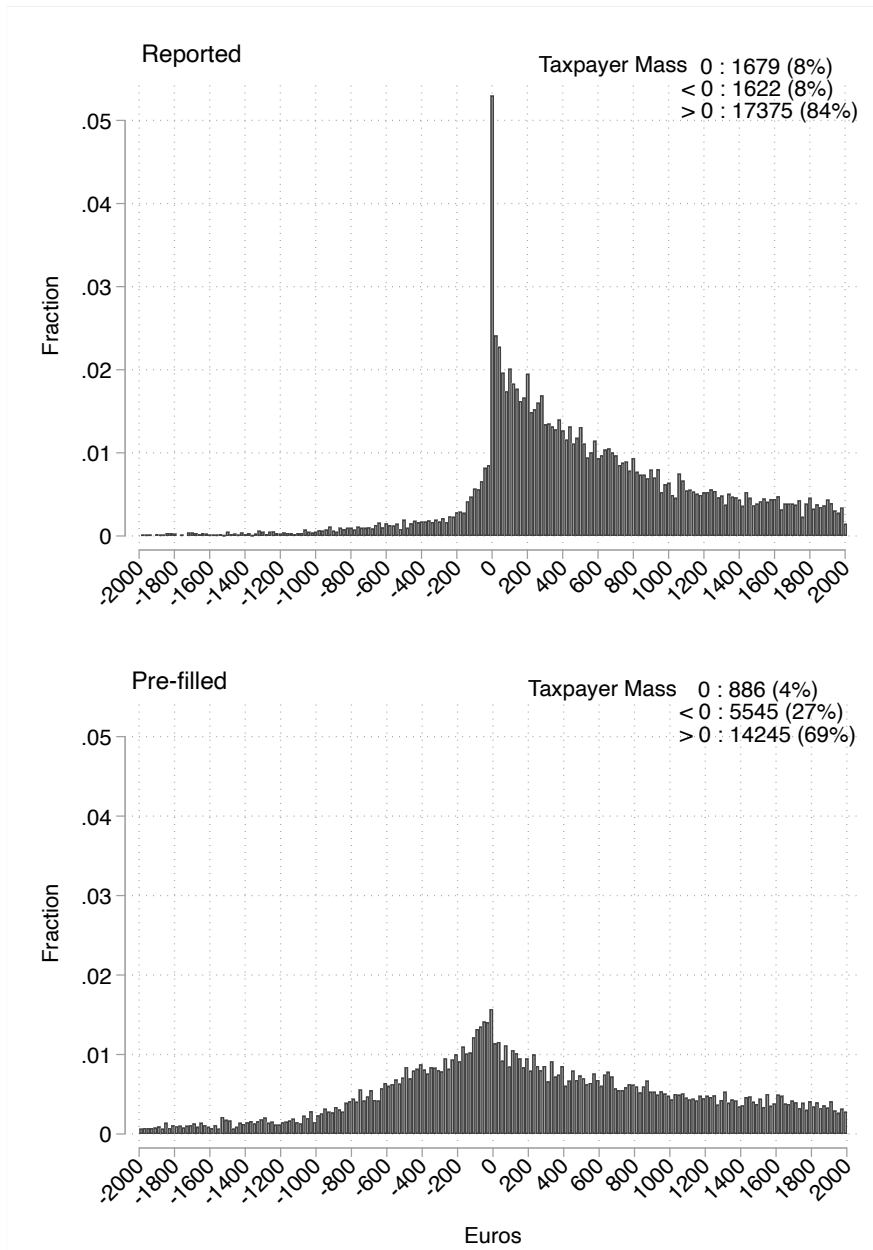
The data allow for the examination of threshold targeting and for a comparison between reported and pre-filled consumption information. Figure 3.1 depicts distributions of within-taxpayer threshold differences *vis-à-vis* reported electronic consumption (top graph) and pre-filled electronic consumption (bottom graph). The reference values at 0, indicate reported (top graph) and pre-filled (bottom graph) electronic consumption corresponding to the taxpayer's threshold, conditional on income. Values to the right of 0 represent over-reporting beyond threshold (top graph) and over-spending beyond threshold (bottom graph).

Firstly, threshold targeting is evident by the higher visual mass at the 0 value in both distributions. Figure 3.1 shows a spike in the reported electronic consumption distribution (top graph) and a smaller, yet significant, spike in the pre-filled electronic consumption distribution (bottom graph). The mass of the distribution around 0 ($\pm \text{€ } 50$), can provide a measure of taxpayer responsiveness to the threshold. Concentrated mass at 0 indicates that taxpayers know exactly what their threshold value is and respond precisely to match that value. When reporting electronic consumption, 8% of the sample report the exact required amount. The pre-filled electronic consumption amount is targeted by 4% of the taxpayers (using their monthly spending).

Secondly, the vast majority of taxpayers (84%) report more than their threshold. This indicates a strong response to the policy. When combined with taxpayers who concentrate responses at 0, almost all individuals (92%) declare electronic consumption on or above their threshold. A small number of individuals (8%) report less than their threshold, possibly due to policy inattention, honest reporting or being exempt from the threshold requirement (see Section 3.2 for exemptions).¹⁵

¹⁵This group is diverse and we cannot determine the precise reasons for responding below the threshold. It includes 753 wage-earners, 515 pensioners and 354 with agricultural income. Based on their postcode, they exhibit geographical dispersion, some of them reside in rural others in urban areas. Also, they do not belong to a particular decile of the income distribution.

Fig. 3.1 Threshold Targeting in Electronic Consumption



Notes: The figures present distributions of the cardinal difference of reported (top) and pre-filled (bottom) electronic consumption from taxpayers' threshold. The threshold value is derived from taxpayers declared income, by applying the scale in Table 3.1. The distance is measured in euros and both distributions are truncated at $\pm \text{€}2000$ with bin width $\text{€}20$. The top graph shows reporting of electronic consumption with reference to each taxpayer's threshold. The 0-value indicates that electronic consumption reported in tax returns matches the threshold value. Positive (negative) differences indicate reporting above (below) threshold. The bottom graph shows the pre-filled consumption with reference to the taxpayer's threshold. The 0-value indicates that pre-filled electronic consumption matches the threshold value. Positive (negative) differences indicate over(under)-spending in reference to the threshold. The taxpayer mass at 0 measures the frequency (and percentage) of taxpayers at $\pm \text{€} 50$ around 0 and the mass below (above) 0 is calculated from $\pm \text{€} 50$. The sample used corresponds to the ECTD column in Table 3.3.

Thirdly, an important insight of how taxpayers respond is given by the difference in the skewness of the two distributions. A broader version of the graphs (at \pm €8,000 differences and €100 bin width) is shown in Figure 3.8 in Appendix 3.7.1. The taxpayers' pre-filled electronic consumption distribution (bottom graph) is more dispersed compared to the reported electronic consumption distribution (top graph). The latter displays a visible cutoff at 0 and with significant skewness at positive values to the right. While the distribution of pre-filled electronic consumption (bottom graph) is still skewed to the right with 73% of individuals spending on or above their threshold, this percentage increases to 92% when taxpayers report their electronic consumption. At the same time the mass below 0 falls from 27% in pre-filled consumption to 8% when reporting it. This provides some initial evidence of individuals increasing their pre-filled information during tax filing, which I explore in detail in Section 3.4.3.

Overall, the two distributions show that taxpayers respond strongly to the ECTD, targeting their threshold level or reporting above it. Spikes at thresholds of both graphs suggest that taxpayers use both margins of responses; either maximising their tax discount using electronic consumption or increasing the reported values tax filing. The pattern of behaviour provides support to the first theoretical prediction, showing clear evidence of threshold-targeting. In addition, taxpayers seem to behave rationally by maximising their utility through reporting electronic consumption on or above their thresholds. Decomposing further the responses in the two response margins, provides additional insights into how taxpayers respond to the policy.

3.4.3 Responses in the Reporting Margin

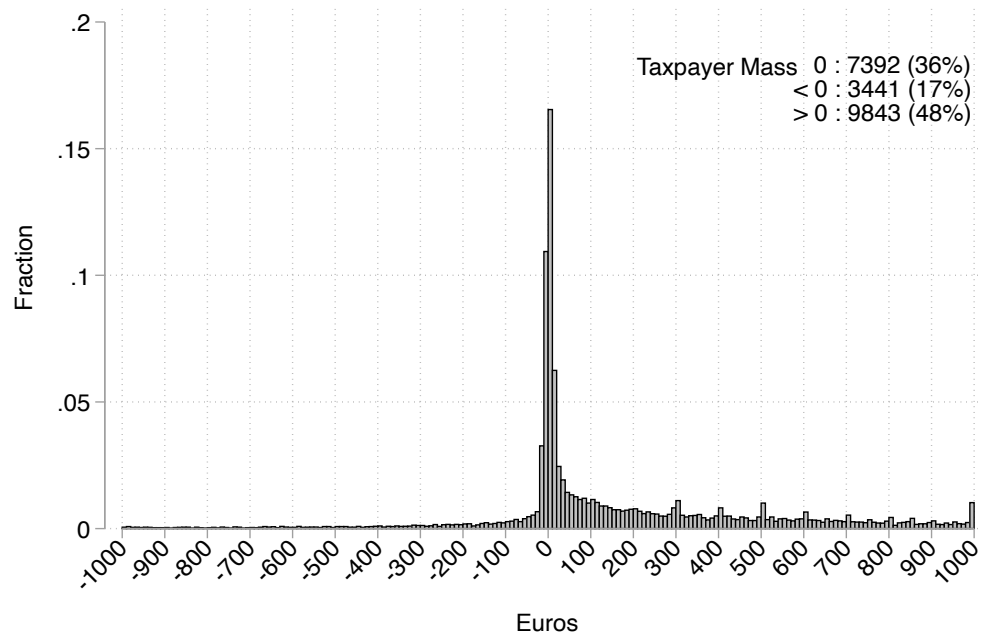
I examine changes to pre-filled electronic consumption information during tax filing, using within-taxpayer differences. These are calculated by subtracting the pre-filled consumption from the reported consumption for each individual, which reveals c_r ; what has been reported by the individual over and above the pre-filled information indicated by the tax authority in the tax returns. A positive (negative) difference indicates over(under)-reporting of electronic consumption during tax filing. Conditional on their income and threshold, the value at 0 represent no change between the pre-filled and reported amounts. Thus, it serves as a measure of equivalence between the two.

The distribution of within-taxpayer differences is shown in Figure 3.2. The highest single-value mass (36%) is observed at or near 0 (\pm € 50), corresponding to pre-filled and reported being equivalent, whilst the majority of taxpayers are situated to the right of the distribution, indicating over-reporting of electronic consumption during tax filing. Over-reporting taxpayers constitute almost half (48%) of the sample.

Adding to evidence of increased reporting during filing, Figure 3.2 documents spikes at positive round numbers of the distribution. A visually higher mass of responses is observed at every hundred euros to the right of the distribution. Spikes indicate that over-reporting does not occur as a result of reporting omissions in electronic consumption by financial institutions which the taxpayers attempt to amend, as this would had produced a smoother distribution. Instead, taxpayers over-report in multiples of hundreds beyond their pre-filled electronic consumption.

To investigate the mechanism that leads taxpayers to report more, consider the responses of those who ended up spending below (*below-threshold* hereafter) versus those spending above their threshold (*above-threshold* hereafter).¹⁶ Below-threshold taxpayers stand to gain from increasing their tax discount when they over-report on their pre-filled amounts, whilst for above-threshold taxpayers over-reporting carries no additional benefit. Figure 3.3 shows distributions of within-individual differences for below-threshold (top graph) versus above-threshold taxpayers (bottom graph), composed of 6,051 and 14,625 taxpayers respectively.

Fig. 3.2 Within-Taxpayer Consumption Difference



Notes: The figure presents the distribution of differences between reported and pre-filled consumption. The difference is calculated by subtracting pre-filled electronic consumption amounts from reported electronic consumption amounts for each taxpayer. A 0-value serves as a measure of no change, where reported consumption corresponds precisely to pre-filled consumption. Positive (negative) values indicate over(under)-reporting of pre-filled consumption in tax returns. The distribution is truncated at $\pm \text{€} 1000$ and with bin width $\text{€} 10$. The taxpayer mass at 0 measures the frequency (and percentage) of taxpayers at $\pm \text{€} 50$ around 0 and the mass below (above) 0 is calculated from $\pm \text{€} 50$. The sample used corresponds to the ECTD column in Table 3.3.

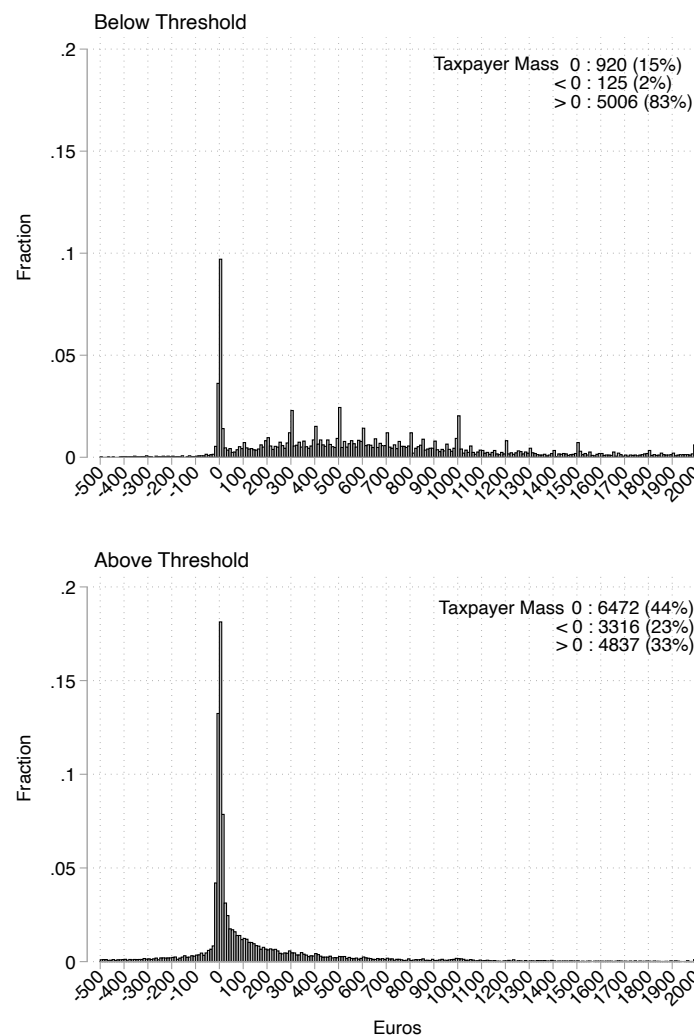
Consider the below-threshold taxpayers (top graph). Conditional on their declared income and threshold, reporting equivalence between pre-filled and reported amounts (at the 0-value) is limited to 15% of taxpayers in this group, while 83% report a larger electronic consumption amount. By contrast, reporting equivalence for above-threshold taxpayers (bottom graph) stands at 44%. In the overall sample, 7,392 individuals report equal amounts and out of these, 6,472 belong to the above-threshold group and 920 to the below-threshold group. Similarly, there were 9,843 taxpayers reporting consumption above their threshold, 5,006 of which belong to the below-threshold group. The above-threshold distribution exhibits a skewness to the right

¹⁶Note that these are endogenously determined groups, making results from this comparison difficult to generalise.

indicating an overall tendency to over-report, but to a lesser extent when compared to the below-threshold distribution. The sub-samples are split to the left and right of 0, by 23% and 33% respectively in the above-threshold group and to 15% and 83% respectively in the below-threshold group.

Importantly, excess visual mass at round numbers (multiples of 50 and 100 euros) are present only in below-threshold responses, which can explain a similar pattern at round numbers observed in Figure 3.1. By contrast, mass at round numbers is absent from the above-threshold taxpayers' distribution.

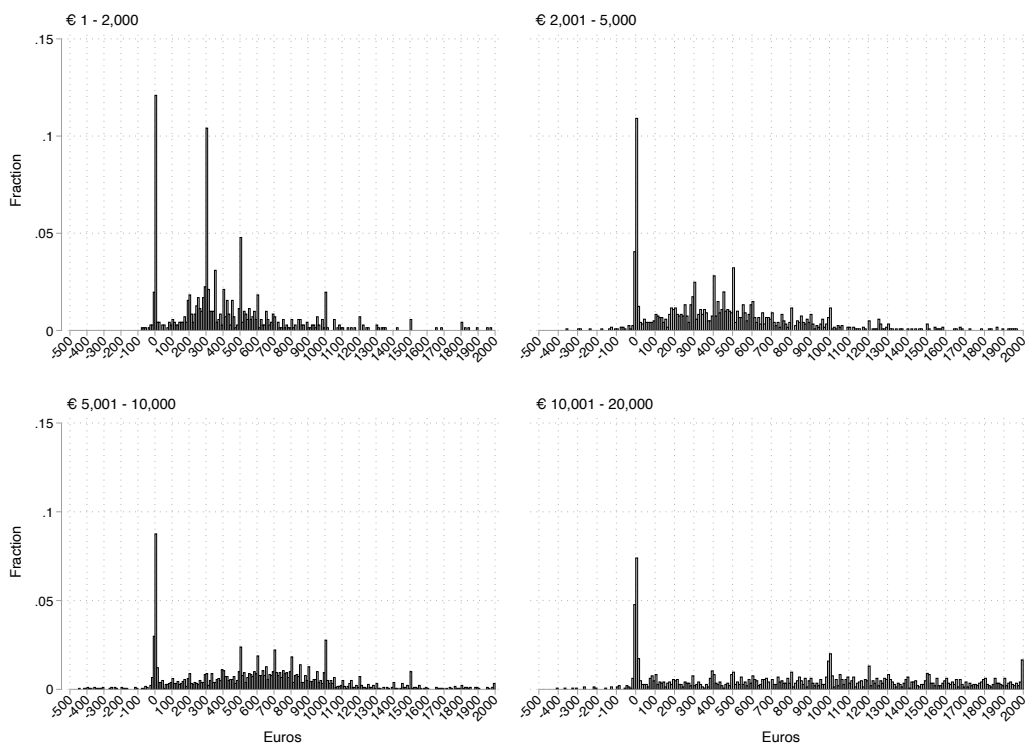
Fig. 3.3 Within-Taxpayer Difference - Below/Above Threshold



Notes: The figures presents the distribution of differences between reported and pre-filled consumption for individuals who spend less than their threshold (top graph) and more than their threshold (bottom graph). The difference is calculated by subtracting pre-filled electronic consumption from reported electronic consumption for each taxpayer. A 0-value serves as a measure no change, where reported consumption corresponds precisely to pre-filled consumption. Positive (negative) values indicate over(under)-reporting of pre-filled consumption in tax returns. The distributions are truncated at $-\text{€ } 500$ and $+\text{€ } 2000$ with bin width $\text{€ } 10$. The taxpayer mass at 0, measures the frequency (and percentage) of taxpayers at $\pm \text{€ } 50$ around 0 and the mass below (above) 0 is calculated from $\pm \text{€ } 50$.

Decomposing the below-threshold distribution in distinct income bands, provides further evidence of increases relative to the taxpayer’s threshold. Figure 3.4 shows a breakdown in 4 income bands; € 1-2,000 (Group 1), € 2,001-5,000 (Group 2), € 5,001-10,000 (Group 3) and € 10,001-20,000 (Group 4). Out of 6,051 below-threshold taxpayers, these income groups include 5,749 observations with the rest being spread in higher income groups. Since the threshold is increasing in income (see Table 3.1), so is the excess visual mass in the 4 groups. For Group 1, that must satisfy a 10% of income threshold, over-reporting is more prominent at € 300. Such an amount satisfies their threshold level, even if they choose not to engage at all in electronic spending. As income increases in Group 2, so do threshold spikes at € 300, 400 and 500, corresponding to 10% threshold conditional on income. A similar pattern is observed also for Group 3 and 4, with the latter exhibiting the highest spikes at € 1,000, 1,200 and 2,000.

Fig. 3.4 Within-Taxpayer Difference - Income Decomposition



Notes: The figures present distributions of differences between reported and pre-filled consumption for individuals who spend less than their threshold decomposed by income. The difference is calculated by subtracting pre-filled electronic consumption from reported electronic consumption for each taxpayer. A 0-value serves as a measure of no change, where reported consumption corresponds precisely to pre-filled consumption. Positive (negative) values indicate over(under)-reporting of pre-filled consumption in tax returns. The distributions are truncated at $-\text{€ } 500$ and $+\text{€ } 2000$ with bin width $\text{€ } 10$. The taxpayer mass at 0 measures the frequency (and percentage) of taxpayers at $\pm \text{€ } 50$ around 0 and the mass below (above) 0 is calculated from $\pm \text{€ } 50$.

Overall, the within-taxpayer distributions provide strong evidence of responses in the reporting margin. The decomposition for above- and below-threshold individuals, identifies taxpayers who are more likely to increase their reported amounts beyond the pre-filled amounts. Reporting more is prominent in individuals who spent less than their threshold in electronic consumption. Such an increase would result in gaining the full tax discount and avoiding an additional tax liability.

Excess mass at round numbers in distinct income groups, corresponding to their threshold values, provides evidence of the mechanism by which taxpayers respond: those who have not increased electronic consumption to reach their threshold, choose to increase their reported electronic consumption during tax filing.

3.4.4 Responses in the Electronic Consumption Margin

Having established that taxpayers respond strongly to the policy during tax filing, I turn to responses in their monthly electronic consumption. This is an important policy parameter, since the ECTD aims at increasing electronic spending and thereby generating third-party information. Recall that payments counting towards the threshold must take place during the financial year, which runs from January to December. Using monthly event studies with the end-of-year deadline as a before/after cut-off in 2017, this section documents increases in electronic consumption as the deadline approaches. The end-of-year increases appear to stabilise and become entrenched in the first half of 2018.

3.4.4.1 Empirical Strategy

To identify increases in consumption, I utilise variation in individuals who reach their thresholds in different months of the year. Overall, the dataset includes information on spending in 19 months; for 12 months in 2017 and for the first 7 months of 2018. Taxpayers are grouped in 12 cohorts based on the month their threshold was reached in 2017 (from February to December, including a group that did not reach threshold).¹⁷ I analyse how their electronic consumption evolves in all months of the year, relative to the end-of-year deadline. Recall that once a taxpayer's electronic consumption has reached threshold, any additional consumption does not contribute to a tax discount. Reaching the threshold early in the year eliminates incentives to increase electronic consumption until the end of year. By contrast, for taxpayers who spent small amounts early in the year, incentives for spending are higher as the deadline approaches. Therefore, the initial hypothesis is that conditional on spending less than the threshold, the closer taxpayers are to the deadline, the larger the electronic consumption responses will be.

Importantly, note that monthly cohorts are not varied experimentally and are determined endogenously instead, based on the spending of individuals. A number of factors can affect the probability of a taxpayers belonging to a monthly cohort. Firstly, spending habits in electronic payments. Some individuals might have a higher propensity to use electronic payments than others. Those having high propensity can reach the threshold early in the year. Secondly, a large payment in a particular month, such a utility bill, which can eliminate the threshold requirement. Thirdly, other factors affecting electronic spending, such as the availability of point of sale machines and the social norms regarding electronic payments in the place of residence. Lack of exogenous variation limits our understanding on the precise mechanisms that might cause increases in consumption. However, the evidence remains suggestive on how electronic

¹⁷A group for January is not included since two months are required to assess the monthly increase in electronic consumption.

spending evolves in relation to the ECTD, given an end of year cut-off that has to be met for the tax discount to apply.

For estimating monthly event studies I follow three steps. Firstly, cohorts are established by identifying individuals above threshold in month m , who were below threshold in month $m - 1$. This indicates the month at which an individual has passed the required threshold. Secondly, monthly consumption $C_{i,m}$ is parameterised by the taxpayer's annual declared income, Y_i , using:

$$\tilde{C}_{i,m} = \frac{C_{i,m}}{Y_i} \times 100 \quad (3.13)$$

Transforming monthly electronic consumption in Equation 3.13 enables comparability across individuals. The dependent variable becomes monthly electronic consumption as a percentage of declared income. Thirdly, I consider changes in consumption before and after the end-of-year deadline using monthly event studies for each m -cohort. These are specified as follows:

$$\tilde{C}_{i,m} = \alpha + \sum_{k=1}^{11} \beta_k (Lag\ k)_{i,m} + \sum_{j=1}^7 \gamma_j (Lead\ j)_{i,m} + \varepsilon_{i,m} \quad (3.14)$$

where for i individual in month m , $Lag\ k$ are binary variables indicating consumption being k -months away from end-of-year deadline and $Lead\ j$ are binary variables indicating consumption j -months after the deadline. The event is considered to be end of December 2017.

3.4.4.2 Results

Results are presented in Figure 3.5 with regression estimates in Table 3.4 in Appendix 3.7.1. They show the evolution of consumption for individuals who have reached their threshold in each month of 2017 and, in addition, for those who have not reached their threshold. Months before or after, are in reference to December 2017 (the deadline for reaching the threshold) with the base month being January 2017. Spending in 2017 is represented by indicators -11 to -1 and in the first months of 2018 by indicators +1 to +6.

The following patterns can be observed. Firstly, note that about half of taxpayers reach their threshold in the first 8 months of the year (10,166 out of 19,648 in our sample). In September, October, November and December about 1,000 individuals each month reach their threshold, while 6,051 individuals do not reach threshold. Secondly, in most cohorts electronic consumption as a percentage of income fluctuates around 5% each month, with the exception of the February cohort, which exhibits a higher propensity to spend by electronic means. Thirdly, cohorts that reach threshold early in the year exhibit seasonal fluctuation with a small upward trend in electronic consumption.

Responses in electronic consumption are prominent in end-of-year cohorts. Taxpayers who reach their threshold in September, October, November and December, exhibit minimal spending during the year, whilst at the end of the year, these groups exhibit statistically and economically higher percentages in electronic consumption. In the November and December cohorts (with

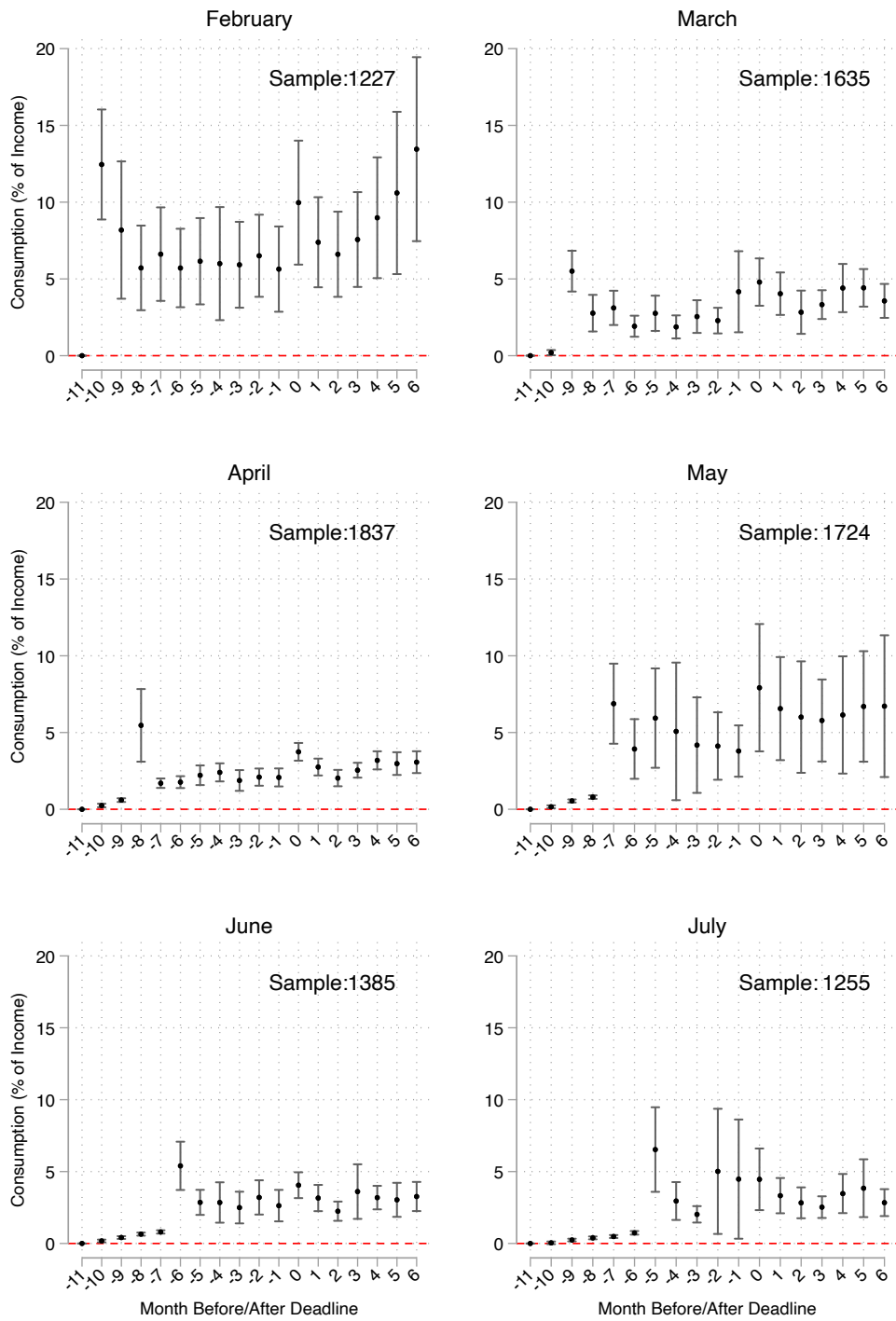
only two and one month respectively to reach threshold) monthly spending increases to 13% and 12.5% of annual income respectively, which is an economically large amount of spending relative to their income. This is in contrast to previous months of the year, in which consumption was close to 0% and exceeded 1% only once during the year. The increase goes beyond the seasonal increase exhibited by other cohorts, which is in the range of 0% to 6%, even with a higher propensity in using electronic payments. The large increases in end-of-year cohorts provide suggestive evidence of taxpayers increasing electronic spending to reach their threshold before the deadline.

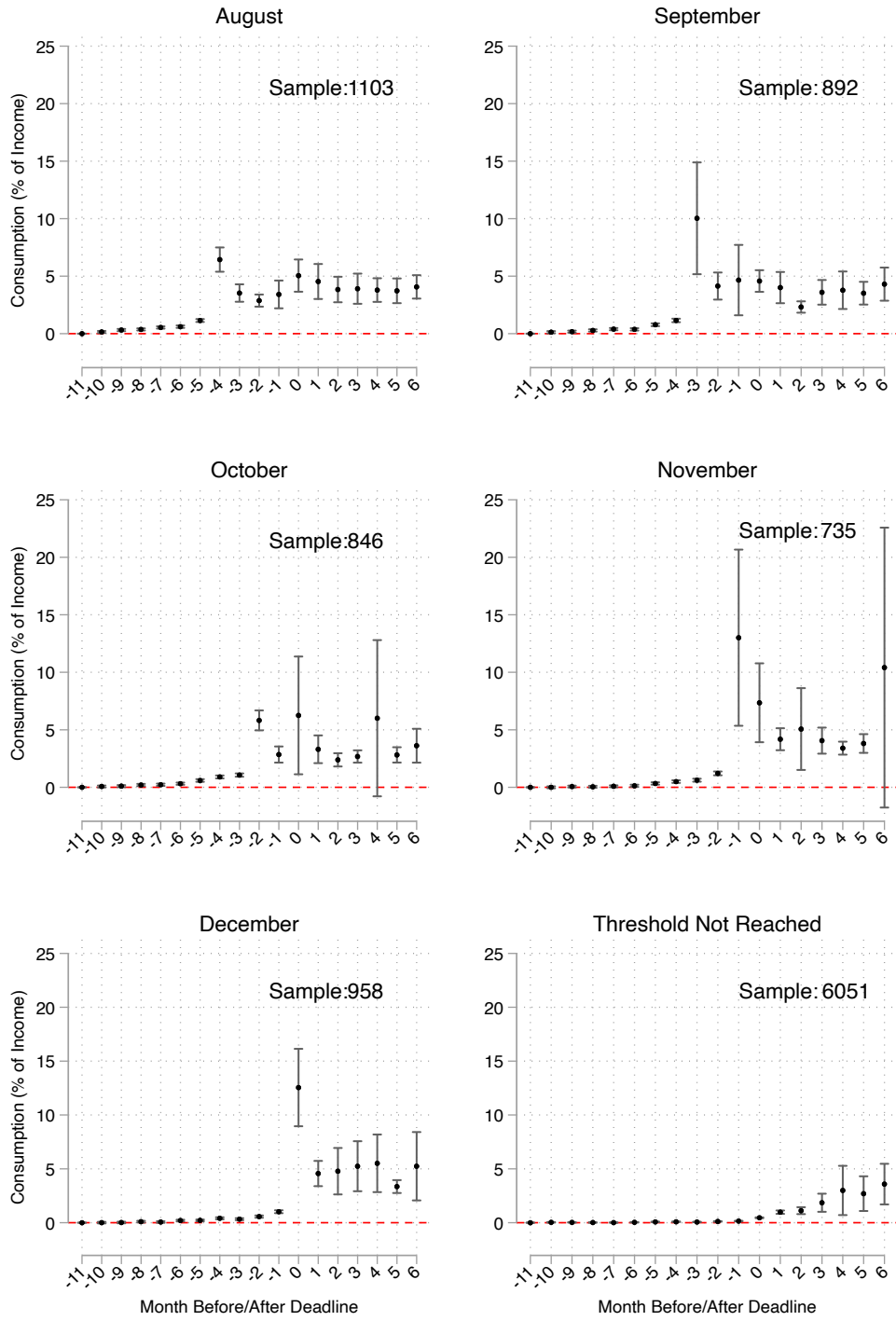
Importantly, about a quarter of individuals in our sample that are subject to the policy do not reach their threshold. These individuals are more likely to increase their electronic consumption when reporting as was shown in Section 3.4.3. As can be observed in the "Threshold Not Reached" graph of Figure 3.5, responses in electronic consumption remain close to 0 during the year, with only marginal increases are observed at the end of year. However, the group exhibits an increase in electronic consumption during the new financial year starting from 1% in January 2018 and slowly increasing to 3.6% by June. This indicates that taxpayers who do not reach their threshold in 2017, increase payments gradually in 2018.

A similar behaviour during the new financial year can be observed in end-of-year cohorts who exhibit strong consumption responses at the end-of-year deadline in 2017. At the start of the new financial year electronic spending increases and remains stable at about 5% of income every month. This finding is shown in Figure 3.5 and provides evidence of electronic spending becoming entrenched once individuals experience an end-of-year deadline spike in consumption. Habit formation in the population or strategically spreading consumption earlier in the year to avoid an end-of-year spike can explain this pattern. Delayed responses in 2017, combined with higher consumption percentages in 2018 can also be a sign of policy inattention in 2017: individuals who noticed the policy late during the year responded by increasing consumption, whilst as the policy became salient, taxpayers increased monthly electronic consumption earlier in the year.

Overall, the results provide evidence that the policy affects the electronic consumption of about half of the taxpayers in our sample who are subject to the policy. Taxpayers who notice the policy close to the deadline and those with low propensity of electronic consumption, increase their spending at the end of year to about 13% of annual income and maintain a stable percentage around 5% every month during the new financial year. About one-third of taxpayers in our sample that are subject to the policy do not respond during the year and their electronic consumption remains below threshold. This group knows about the policy at the time of tax filing and resorts to changing the pre-filled information, increasing their reported electronic payments to gain the full tax discount. Finally, the remaining individuals reach their thresholds in the beginning to middle of the year, indicating that for part of the taxpayer population the thresholds are reachable and the policy is salient.

Fig. 3.5 Event Studies per Month of Reaching Threshold





Notes: The figures present monthly event studies for groups of taxpayers who have reached their threshold in particular months. The last graph includes individuals who have not reached their threshold. The dependent variable is monthly electronic consumption as a percentage of annual declared income. Point estimates in graphs follow the specification in Equation 3.14, which includes binary variables before and after the deadline in December 2017 and, in addition, individual fixed effects. Results are shown relative to -11 month from the deadline (February 2017). Detailed regression estimates are shown in Table 3.4 in Appendix 3.7.1. The sample number in each graph represents the number of individuals in our sample who have reached threshold on the specified month in 2017.

3.5 Interpretation

This chapter has established three findings. First, the ECTD triggers strong responses in the taxpayer population, evidenced by their threshold-targeting behaviour. Taxpayers report on or beyond their threshold during tax filing. Second, about one-third of individuals subject to the policy do not increase their electronic consumption and they respond by changing the pre-filled amounts during tax filing. Responses in the reporting margin are strongest for individuals who failed to reach their threshold at the end of the year. Third, some taxpayers who exhibit low electronic spending during the year respond by increasing their electronic consumption sharply before the end-of-year deadline. Electronic consumption for these individuals becomes entrenched and increases during the new financial year. These evidence are consistent with the three theoretical predictions of Section 3.3; it is rational for taxpayers to target their threshold and they do so by either increasing electronic consumption or by changing the pre-filled information.

What could explain the pattern of behaviour we observe in the two margins of responses? For taxpayers with sharp consumption responses at the end of year one reason is policy inattention. Low salience of the policy during the first year of introduction might have resulted in late consumption responses. At the start of the new financial year, taxpayers increase consumption earlier, spreading payments to avoid end-of-year spikes. For below-threshold taxpayers the lack of consumption responses can be explained by a combination of factors. Policy inattention in their case can be extreme, hence noticing the policy after the financial year. However, the policy is salient at the time of tax filing (taxpayers respond to threshold when reporting takes place). Alternatively, the policy may be salient but these individuals may face liquidity constraints in increasing electronic consumption. Lastly, higher reported amounts can also signify perceived low costs of audit and penalties. If the tax authority's audit strategy is not credible enough, or if the expected costs of evading are low, taxpayers might prefer to report higher amounts than increase their electronic consumption. An interplay of these factors can provide a fitting explanation of the observed pattern.

3.5.1 Marginal Cost Functions - Explicit Form

The taxpayers' choice between increasing electronic consumption or the reported amounts, can be expressed formally using the following explicit functional form for ψ and ξ , which can account for the facts that were documented in the empirical analysis. Firstly, consider the costs of increased reporting, $\psi(c_r)$, that were defined implicitly in Section 3.3.2. In the subjective sense of the individual, an increase of the reported amount of electronic consumption beyond the pre-filled amount is associated with increased risks of tax audit and penalty. Consider a function that captures these expected audit and penalty costs in the following form:

$$\psi(c_r) = \pi \rho c_r \tag{3.15}$$

where $\pi \in (0, 1)$, represents a financial penalty proportional to the increase, c_r , and $\rho \in (0, 1)$, a perceived probability of audit. Note that reporting the pre-filled amount ($c_r = 0$), does not imply any additional costs $\psi(c_r) = 0$, while increasing the reported amounts, increases the expected costs of audit and penalty

Secondly, consider a possible function for $\xi(c_e)$, arising from policy inattention as in Taubinsky and Rees-Jones (2017) and liquidity constraints. Suppose that electronic consumption c_e can be separated in 12 months represented by m , and define total electronic consumption in the year as the sum of these months $\sum_{m=1}^{12} c_{e,m}$. Policy inattention varies from month to month and can be defined in the cost function by a parameter $\theta_m \in (0, \infty)$, capturing policy inattention ($\theta_m > 1$) and excess attention ($\theta_m < 1$). The policy attention benchmark is given by $\theta_m = 1$.¹⁸ In a similar fashion, to allow for liquidity constraints, the functional form should capture increasing costs of delaying electronic consumption at the end of year. I posit a factor $\delta^m \in (0, \infty)$ to account for these costs, with liquidity constraints in reaching the threshold as months progress ($\delta^m > 1$) and excess liquidity at the end of the year ($\delta^m < 1$). At $\delta^m = 1$ month-to-month liquidity remains the same. The cost function can be written as:

$$\xi(c_e) = \sum_{m=1}^{12} \theta_m \delta^m c_{e,m} \quad (3.16)$$

Equation 3.16 can determine the timing of increased spending during the year. Higher attention(inattention) results in low(high) costs in a given month, and a higher(lower) preference to spend by electronic means. Similarly, higher(lower) liquidity constraints at the end of year, increase(decrease) the costs of electronic consumption.

Recall that the decision between increasing electronic consumption or the reported amounts depends on marginal costs (Proposition 2). Substituting the explicit form of the cost functions yields:

$$\pi \rho = \theta_m \delta^m |_{m} \quad (3.17)$$

Equation 3.17 provides an explicit form for the margin of responses, which maximises utility. In every month, m , the taxpayer weighs the marginal costs of increasing electronic consumption to the marginal cost of reporting higher amounts during tax filing.

This offers a concise framework to interpret the observed empirical evidence. Firstly, the expected costs of reporting higher amounts are fixed for all months and depend on the penalty and the perceived probability of audit. Secondly, policy inattention varies from month to month, starting high at the beginning of the year and decreasing at the end of the year, as the deadline approaches. Higher policy attention, implies lower costs of electronic consumption, leading to increased levels of electronic spending at the end of year. Thirdly, as policy attention increases, higher liquidity constraints at the end of year might prevent some taxpayers from reaching their

¹⁸Note that policy inattention in a particular month results in high costs of electronic consumption, implying that taxpayers would prefer paying in cash. Increased policy attention in a particular month, lowers the costs of electronic consumption, making individuals more inclined to spend by electronic means (to reach their threshold).

threshold. Alternatively, below-threshold behaviour accompanied by changing of the pre-filled amounts during tax filing, can be explained by perceived audit and penalty costs being too low.

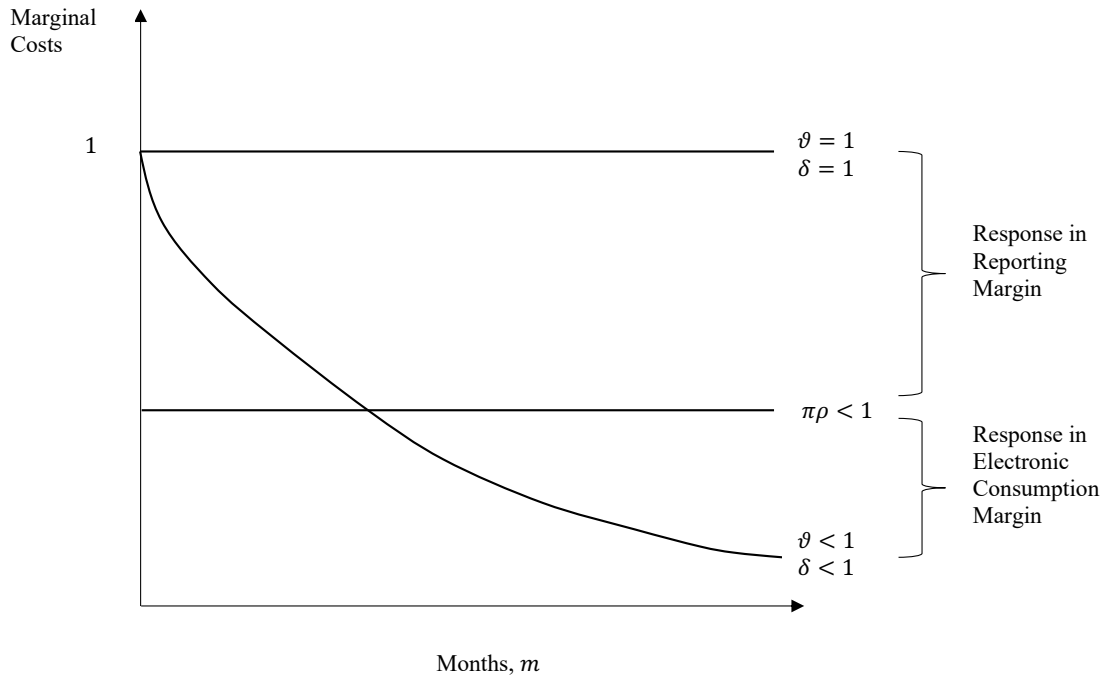
3.5.2 Diagrammatic Representation

Our understanding can be aided by a diagrammatic representation of the different cases. Consider the case where policy inattention is high in the beginning of the year ($\theta > 1$) and falls slowly as the deadline approaches ($\theta < 1$). In this first case, individuals do not face liquidity constraints in reaching their threshold ($\delta < 1$). The taxpayer's choice is illustrated in Figure 3.6, which graphs the marginal costs of increased reporting and electronic consumption against months, m .

Notice that the benchmark case for costs of electronic consumption is a horizontal line with $\theta = 1$ and $\delta = 1$, in all months of the year. The marginal costs of increased reporting are also represented by the straight $\pi\rho$ -line, being fixed in all months. Since $\pi, \rho \in (0, 1)$, the marginal costs of reporting line is lower than the benchmark case of marginal costs of electronic consumption line. This means that without a higher policy attention, individuals will not increase electronic consumption and will prefer to increase the reported amounts instead. Higher policy attention is shown in the downward slopping curve, where θ decreases as months progress. At the point where the marginal costs of electronic consumption fall below the marginal costs of increased reporting, higher policy attention leads individuals to respond in the electronic consumption margin to reach their threshold. Compared to the empirical evidence, this corresponds to the sharp increase in electronic consumption observed in some cohorts as the end-of-year deadline approaches.

Consider now the case where policy inattention falls as months progress but individuals face liquidity constraints. This is shown in Figure 3.7, where the electronic consumption cost curve is still downward slopping ($\theta < 1$), but at a slower pace, indicated by the slope of the curve. High liquidity constraints ($\delta > 1$) increase the costs for taxpayers to respond in the electronic consumption margin. These costs are exacerbated at the end of the year and especially for the case where policy inattention is still high. The combination of high liquidity constraints and late policy attention, prevents some taxpayers from reaching their threshold using electronic consumption responses. These individuals resort to increasing their pre-filled electronic consumption amounts during tax filing.

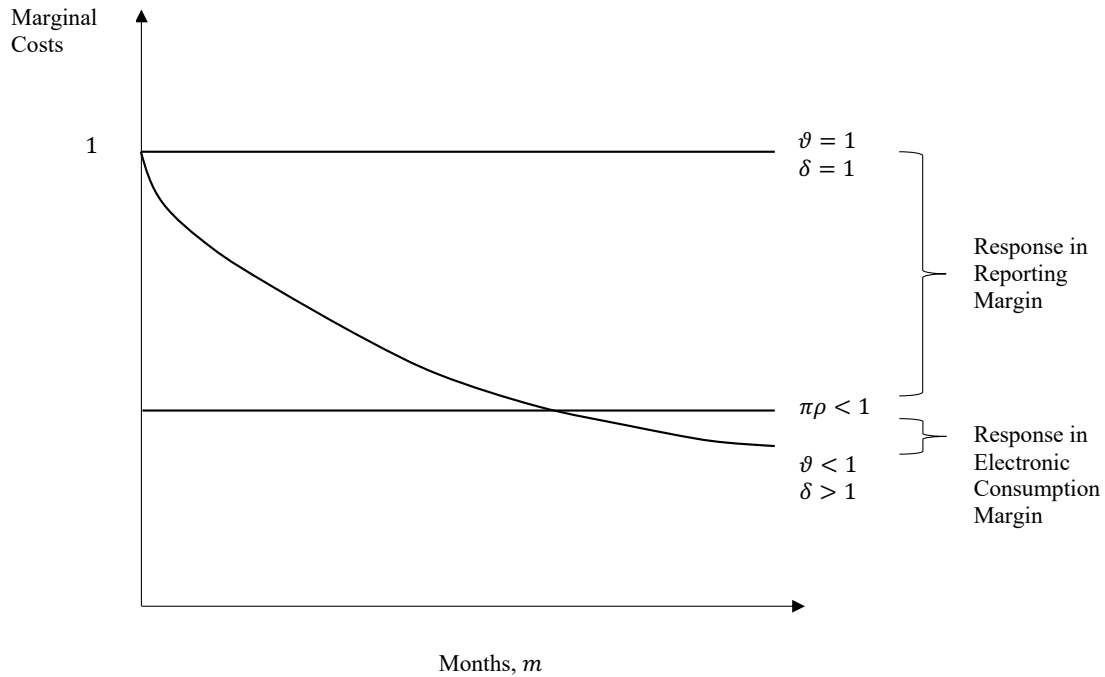
Fig. 3.6 Policy Attention without Liquidity Constraints



Notes: The diagram presents the taxpayer’s choice of increasing electronic consumption or change their pre-filled amounts during reporting, based on the explicit definition of marginal costs in Equation 3.17. Marginal costs are drawn against months of the year to build the threshold (beginning of year on the left, to end of year on the right). The straight line at the top of the graph represents the benchmark case of responses in the electronic consumption margin. Policy inattention and liquidity constraints take the value of 1 and are therefore the same in all months. The downward slopping curve represents decreasing costs of electronic consumption, due to increasing policy attention $\theta < 1$ and decreasing liquidity constraints $\delta < 1$. The $\pi\rho$ -line represents the marginal costs of reporting, being fixed in all months. In this case, lower marginal costs of electronic consumption than cost of reporting in particular months of the year, lead taxpayers to increase responses in the electronic consumption margin to reach their threshold. This is shown on the diagram, using the (relatively) large area of responses in electronic consumption margin.

Lastly, notice how perceived audit and penalty costs determine the choice of increasing the reported amounts. If the perceived probability of audit, ρ , is very low, due to the audit strategy being non-credible, or the penalty, π , is very small, the cost of reporting higher amounts falls, eliminating the need for electronic consumption responses to reach the threshold. This case can still occur if policy attention is high and liquidity constraints are low. Low perceived costs of audit and penalties, might explain the lack of consumption responses and the subsequent changes in the pre-filled information during tax filing. The different cases can explain the variety of taxpayer behaviour we observe in response to the policy.

Fig. 3.7 Policy Attention with Liquidity Constraints



Notes: The diagram presents the taxpayer’s choice of increasing electronic consumption or increasing amounts during reporting, based on the explicit definition of marginal costs in Equation 3.17, when the liquidity constraints are high. Marginal costs are drawn against months of the year to build the threshold (beginning of year on the left, to end of year on the right). The straight line at the top of the graph represents the benchmark case of responses in the electronic consumption margin. Policy inattention and liquidity constraints take the value of 1 and are therefore the same in all months. The downward sloping curve represents decreasing costs of electronic consumption, due to increasing policy attention $\theta < 1$, which is offset by increasing liquidity constraints $\delta > 1$. The $\pi\rho$ -line represents the marginal costs of reporting, being fixed in all months. Lower marginal costs of electronic consumption than cost of reporting in particular months of the year, lead taxpayers to increase responses in the electronic consumption margin to reach their threshold. In this case, responses in the electronic consumption margin are small, due to a combination of late policy attention during the year and high liquidity constraints. Higher marginal costs of electronic consumption than marginal costs of reporting, lead taxpayers to increase their pre-filled amounts during tax filing.

3.6 Conclusion

Features of the income tax system have been widely used as means for incentivising behavioural change. Since they apply to the majority of the taxpayer population, incentivisation is commonly broad, covering a large number of individuals. This chapter has analysed a third-party reporting policy that conditions personal tax allowance on consumption by electronic means, thereby using a feature of the income tax system to incentivise a widespread change in payments behaviour. The analysis provided insights on how taxpayers change their behaviour when the policy is introduced and an explanation of how adjustment costs shape the variety of taxpayer responses.

By examining within-taxpayer differences between reported and pre-filled consumption, the chapter produced evidence of threshold-targeting, indicating strong responses to the policy. About 92% of the sample report on and beyond the required amounts to gain the tax discount.

The findings support similar evidence of bunching at kink-points and cutoffs of the tax schedule (Kleven, 2016), and extend these to personal thresholds.

In addition, the analysis has documented evidence of increased reported amounts, in cases where electronic consumption is lower than required. Taxpayers who exhibit low propensity of electronic consumption during the financial year, end up changing their pre-filled information. This is an unintended consequence of the policy, which might be linked to tax evasion behaviour. The fact that these taxpayers are predominantly wage-earners and pensioners, suggests that individuals with traditionally low opportunities to evade, might not hesitate to do so when granted the opportunity. The implications for the evasion elasticity are likely to be large in this case, conditional on the opportunities granted by the institutional framework.

Inducing electronic consumption responses is the ultimate aim of the policy. The chapter has provided evidence of how electronic consumption evolves during the financial year. Firstly, about half of the eligible population reaches threshold using electronic consumption by the third quarter of the financial year. Secondly, some taxpayers who exhibit low propensity of electronic consumption during the year, increase their responses as the end-of-year deadline approaches. Using monthly event studies, the chapter has documented increases of up to 13% of annual income in the final months of the year. Thirdly, about a third of the eligible sample exhibits low electronic consumption and does not reach threshold. These individuals are more prone to increasing their electronic consumption amounts during tax filing. Fourthly, there are evidence of electronic consumption behaviour becoming entrenched as the new financial year begins, indicating an attempt by taxpayers to build their threshold earlier in the year.

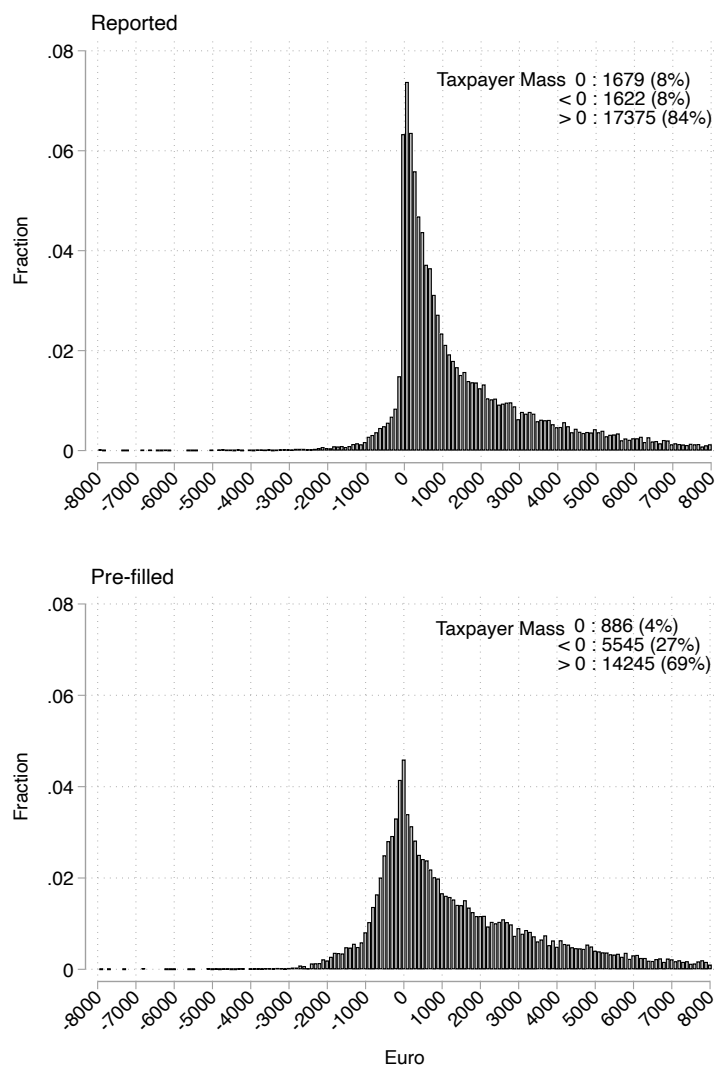
The variety of responses observed in the analysis can be explained through adjustment costs, which seem to affect the policy's objectives. Similar studies on bunching document how adjustment costs determine responses (Chetty *et al.*, 2011; Adam *et al.*, 2021; Gelber *et al.*, 2020), and how these often related to the various elasticities (Piketty *et al.*, 2014). In the ECTD, the interplay of policy inattention, liquidity constraints and the perceived costs of audit are frictions that shape the final outcome and might dampen the policy's effect.

The scheme constitutes one of the first attempts by a tax administration to use features of the income tax as a way to change payments behaviour, and thereby induce third-party information at the final consumer level. Recent advances in information technology and an increasing digitalisation trend, are steering tax administrations towards using data that can monitor the entire volume of transactions, without the need for consumers to register their receipts (Gupta *et al.*, 2017). Nevertheless, the economic and welfare implications of this shift are yet to be fully understood. Evidence in this chapter, from threshold-targeting behaviour and related adjustment costs, suggest that when incentives for increasing electronic consumption are combined with key elements of the income tax system, they result in strong responses in the taxpayer population.

3.7 Appendix

3.7.1 Additional Figures and Tables

Fig. 3.8 Threshold Targeting - Truncation $\pm \text{€}8000$



Notes: The figures present distributions of the cardinal difference of reported (top graph) and pre-filled (bottom graph) electronic consumption from taxpayers' threshold. The threshold value is derived from taxpayers' declared income, by applying the scale in Table 3.1. The distance is measured in euros and both distributions are truncated at $\pm \text{€}8000$ with bin width $\text{€}100$. The top graph shows reporting of electronic consumption with reference to each taxpayer's threshold. The 0-value indicates that electronic consumption is reported in tax returns matches the threshold value. Positive(negative) differences indicate reporting above(below) threshold. The bottom graph shows the pre-filled consumption with reference to the taxpayer's threshold. The 0-value indicates that pre-filled electronic consumption matches the threshold value. Positive (negative) differences indicate over(under)-spending in reference to the threshold. The taxpayer mass at 0, measures the frequency (and percentage) of taxpayers at $\pm \text{€} 50$ around 0 and the mass below (above) 0 is calculated from $\pm \text{€} 50$. The sample used corresponds to the ECTD column in Table 3.3.

Table 3.2 Income Tax in 2017

Income Bracket €	Marginal Tax %
0 - 20,000	22
20,001 - 30,000	29
30,001 - 40,000	37
40,000 <	45

Table 3.3 Sample Statistics

	Sample	Single Filers	Joint Filers	ECTD
	Freq (Percent)	Freq (Percent)	Freq (Percent)	Freq (Percent)
<i>Primary Income Source:</i>				
Self-Employed/Business	2,052 (4.1)	1,070 (3.4)	982 (5.3)	- -
Wage-Earner	22,335 (44.7)	12,691 (40.4)	9,644 (51.9)	12,685 (25.4)
Pensioner	12,163 (24.3)	6,880 (21.9)	5,283 (28.4)	6,880 (13.8)
Agricultural Income	2,635 (5.3)	1,123 (3.6)	1,512 (8.1)	1,111 (2.2)
Null Income	10,815 (21.6)	9,645 (30.7)	1,170 (6.3)	- -
Total	50,000	31,409	18,591	20,676

Notes: The table presents the number of observations in the randomly-drawn sample from the Greek taxpayer population in 2017, decomposed by primary income source. There are four income categories and, in addition, a null income category. The null income includes individuals who declared null in their tax returns. Primary income is defined as the highest declared among the four income categories. The first column presents the overall sample. The second column presents single-individual tax units. The third column presents individuals who belong to a household comprising from two individuals and, thus file jointly (note that all such households were required by law to file jointly in 2017). Lastly, the fourth column includes single-filers only, who are eligible for ECTD. This category excludes null income taxpayers and those that earning more than € 160,000, as well as, the self-employed and business income category.

Table 3.4 Event Studies - Monthly Consumption (% of Income)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Threshold Reached in:	Not Reached	February	March	April	May	June	July	August	September	October	November	December
Month to Deadline:												
-10	0.037*** (0.012)	12.454*** (1.827)	0.197** (0.086)	0.247*** (0.059)	0.162*** (0.048)	0.175*** (0.048)	0.043 (0.047)	0.149*** (0.045)	0.129*** (0.048)	0.070 (0.044)	-0.000 (0.043)	0.007 (0.037)
-9	0.027** (0.011)	8.181*** (2.280)	5.505*** (0.677)	0.598*** (0.061)	0.542*** (0.052)	0.412*** (0.049)	0.240*** (0.047)	0.315*** (0.048)	0.175*** (0.046)	0.101** (0.045)	0.062 (0.043)	0.020 (0.037)
-8	0.018 (0.011)	5.717*** (1.405)	2.768*** (0.608)	5.467*** (1.206)	0.792*** (0.058)	0.651*** (0.054)	0.389*** (0.051)	0.370*** (0.047)	0.281*** (0.051)	0.199*** (0.049)	0.043 (0.043)	0.097** (0.040)
-7	0.010 (0.011)	6.613*** (1.553)	3.113*** (0.569)	1.699*** (0.157)	6.878*** (1.330)	0.803*** (0.058)	0.484*** (0.052)	0.539*** (0.051)	0.394*** (0.050)	0.222*** (0.050)	0.086** (0.043)	0.066* (0.038)
-6	0.036*** (0.012)	5.710*** (1.302)	1.919*** (0.350)	1.770*** (0.197)	3.928*** (0.991)	5.400*** (0.855)	0.739*** (0.063)	0.608*** (0.053)	0.375*** (0.053)	0.314*** (0.050)	0.129*** (0.049)	0.200*** (0.041)
-5	0.076*** (0.012)	6.151*** (1.433)	2.759*** (0.586)	2.216*** (0.327)	5.939*** (1.651)	2.858*** (0.444)	6.532*** (1.499)	1.144*** (0.064)	0.777*** (0.061)	0.590*** (0.058)	0.331*** (0.052)	0.216*** (0.041)
-4	0.088*** (0.013)	5.996*** (1.877)	1.873*** (0.382)	2.405*** (0.301)	5.074** (2.284)	2.850*** (0.715)	2.952*** (0.670)	6.442*** (0.538)	1.151*** (0.077)	0.901*** (0.064)	0.495*** (0.059)	0.410*** (0.047)
-3	0.067*** (0.012)	5.921*** (1.425)	2.546*** (0.543)	1.877*** (0.344)	4.181*** (1.587)	2.498*** (0.562)	2.024*** (0.291)	3.532*** (0.387)	10.031*** (2.477)	1.061*** (0.072)	0.617*** (0.066)	0.331*** (0.046)
-2	0.117*** (0.014)	6.508*** (1.363)	2.283*** (0.426)	2.096*** (0.285)	4.120*** (1.120)	3.203*** (0.608)	5.015** (2.219)	2.875*** (0.269)	4.139*** (0.600)	5.815*** (0.441)	1.209*** (0.085)	0.568*** (0.053)
-1	0.151*** (0.014)	5.639*** (1.416)	4.163*** (1.346)	2.075*** (0.299)	3.794*** (0.852)	2.632*** (0.557)	4.477** (2.113)	3.411*** (0.615)	4.659*** (1.559)	2.848*** (0.357)	13.008*** (3.901)	1.020*** (0.069)
Deadline - Dec 2017	0.457*** (0.019)	9.967*** (2.058)	4.794*** (0.786)	3.743*** (0.294)	7.921*** (2.114)	4.054*** (0.458)	4.460*** (1.092)	5.048*** (0.714)	4.577*** (0.480)	6.251** (2.608)	7.348*** (1.748)	12.546*** (1.831)
+1	0.978*** (0.080)	7.387*** (1.495)	4.037*** (0.706)	2.753*** (0.280)	6.559*** (1.712)	3.161*** (0.465)	3.326*** (0.625)	4.536*** (0.776)	4.005*** (0.691)	3.304*** (0.614)	4.179*** (0.486)	4.565*** (0.597)
+2	1.118*** (0.163)	6.605*** (1.414)	2.829*** (0.718)	2.032*** (0.272)	6.005*** (1.851)	2.244*** (0.341)	2.827*** (0.548)	3.839*** (0.566)	2.318*** (0.251)	2.385*** (0.293)	5.064*** (1.812)	4.781*** (1.098)
+3	1.854*** (0.429)	7.565*** (1.574)	3.325*** (0.478)	2.546*** (0.246)	5.784*** (1.363)	3.608*** (0.968)	2.526*** (0.384)	3.909*** (0.669)	3.595*** (0.546)	2.679*** (0.272)	4.059*** (0.577)	5.244*** (1.182)
+4	2.994*** (1.166)	8.983*** (2.005)	4.404*** (0.801)	3.187*** (0.301)	6.143*** (1.947)	3.189*** (0.418)	3.471*** (0.693)	3.791*** (0.524)	3.777*** (0.831)	6.006* (3.460)	3.404*** (0.287)	5.516*** (1.363)
+5	2.697*** (0.821)	10.598*** (2.694)	4.418*** (0.625)	2.976*** (0.378)	6.699*** (1.835)	3.034*** (0.603)	3.844*** (1.024)	3.727*** (0.547)	3.520*** (0.504)	2.816*** (0.339)	3.809*** (0.415)	3.357*** (0.305)
+6	3.584*** (0.962)	13.451*** (3.054)	3.566*** (0.563)	3.069*** (0.363)	6.717*** (2.354)	3.266*** (0.516)	2.841*** (0.477)	4.068*** (0.518)	4.307*** (0.734)	3.617*** (0.749)	10.409* (6.199)	5.242*** (1.616)
Constant	0.220 (0.185)	5.380*** (1.349)	3.565*** (0.385)	2.484*** (0.169)	1.825* (1.021)	1.408*** (0.340)	1.190*** (0.390)	0.795*** (0.246)	0.640** (0.299)	0.561 (0.448)	0.515 (0.559)	0.423 (0.290)
Observations ($N \times T$)	108,918	22,086	29,430	33,066	31,032	24,930	22,590	19,854	16,056	15,228	13,230	17,244
Taxpayers (N)	6,051	1,227	1,635	1,837	1,724	1,385	1,255	1,103	892	846	735	958

Notes: The table presents monthly event studies for groups of taxpayers who have reached their threshold in particular months. Column (1) includes individuals who have not reached their threshold. The dependent variable is monthly electronic consumption as a percentage of annual declared income. The income has been declared by taxpayers in 2018 for the financial year 2017. Both the financial year and the deadline to build the threshold run from January to December every year. The regressions follow the specification in Equation 3.14, which includes binary variables before and after the deadline in December 2017 and, in addition, individual fixed effects. Results are shown relative to -11 month from the deadline (February 2017). January 2017 is excluded from the regression since two months are needed for assessing monthly spending. The sample excludes (a) joint-filers (to limit any intra-household consumption effects), (b) individuals who declare income from business and self-employed (c) individuals who declared 0 income (no threshold to follow). The results correspond to the graphs presented in Figure 3.5. Robust standard errors presented in parentheses, clustered at the individual's level.

3.7.2 ECTD Information

Fig. 3.9 Tax Filing - Codes 049 and 050

ΠΙΝΑΚΑΣ 7. ΠΟΣΑ ΔΑΠΑΝΩΝ ΠΟΥ ΑΦΑΙΡΟΥΝΤΑΙ ΑΠΟ ΤΟ ΣΥΝΟΛΙΚΟ ΕΙΣΟΔΗΜΑ Ή ΑΠΟ ΤΟ ΦΟΡΟ					
1. Δαπάνη αγοράς αγαθών και παροχής υπηρεσιών	049	5.773,99	050		
2. Δωρεές χρημ. ποσών άρθρ19 ΚΦΕ κτλ: Στο εξωτερικό 031	032			αι συνολικά	059
3. Δωρεές χρημ. ποσών στο λογ/σμό αλληλοβοήθειας για την απόσβεση του Δημόσιου χρέους	075				060
4. Πολιτιστικές χορηγίες του ν.3525/2007	061				076
5. Μισθώματα επιχειρήσεων τριτογενούς τομέα (παρ. Β2 άρθρ.43 και 44 ν.4030/2011)	077				062
6. Ποσό επένδυσης για κινηματογραφικές ταινίες (παρ. 9 άρθρ. 73 ν. 3842/2010)	663				078
7. Εμπίπτειτε στις διατάξεις των περιπτώσεων α ή β της παρ. 9 του άρθρου 73 ν. 3842/2010;		033	α	035	β
8. Ποσό δαπάνης για την ιδιωτική χρηματ. πολιτικ. κόμματος ή συνασπ. κομμάτων (άρθρ. 8 ν.3023/2002)	055				664
9. Ποσό δαπάνης για την ιδιωτ. χρηματ. υποψ. Βουλής Ελλήνων και Ευρωπ. Κοινοβ. (άρθρ. 8 ν.3023/2002)	057				056
					058

Notes: The picture presents an example of how the aggregate amount of electronic consumption appeared in tax returns in 2017. Code 049 corresponds to the pre-filled information of annual electronic consumption, as sent by financial institutions to the tax authority. This is an example of a single-filing tax unit. For a joint-filing, an additional amount would appear in Code 050, for the spouses' spending. The individual can modify this amount before filing taxes, but the amount reported to the tax authority by financial institutions is salient to the taxpayer. The code name translates to "Consumption expenditure for goods and services".

3.7.3 Proofs

3.7.3.1 Proposition 1

The first part involves amounts of electronic consumption below threshold $F(z, a)$. Proof by contradiction. Suppose c_e^* is consumption below threshold, such that $c_e^* < F(z, a)$, and is a utility maximising point, such that $U(c_e^*, z) > U(c_e, z)$. Consider another consumption amount $c_e' > c_e^*$. Since $U(c_e, z)$ is increasing on c_e below the threshold, it must hold that $U(c_e', z) > U(c_e^*, z)$.

The second part involves amounts of consumption more than or equal to $F(z, a)$. Consider two consumption amounts $\tilde{c}_e > \hat{c}_e \geq F(z, a)$. Recall from Equation 3.4 that $t = 0$ if $c_e \geq F(z, a)$. Hence, $U(\tilde{c}_e, z) = U(\hat{c}_e, z)$. For any two electronic consumption points above threshold, the utility is the same. This completes the proof.

3.7.3.2 Proposition 2

The maximised response is proved in the same way as Proposition 1, for reporting and electronic consumption margins. What remains is to prove that the margin of responses depends on marginal costs.

Suppose $\xi'(c_e) > \psi'(c_r)$. Increasing c_r , increases utility U , until $t = 0$. At $t = 0$, $\psi'(c_r) = 0$. By Equation 3.9, $t = 0$ if $c_e + c_r \geq F(z, a)$, which is the utility maximising point.¹⁹

By symmetry, the same applies at $\xi'(c_e) < \psi'(c_r)$. Increasing c_r , increases U . At $t = 0$, $\xi'(c_r) = 0$. Therefore $t = 0$ is a utility maximising point, satisfying $\xi'(c_e) = \psi'(c_r)$. This completes the proof.

¹⁹In terms of the policy this means the following. If the marginal costs of changing the pre-filled amounts during reporting are lower than the costs of increasing electronic consumption to reach the threshold, the individual responds by increasing the reported values up to the threshold point, which gives maximum tax discount, and thus, maximum utility.

Chapter 4

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

4.1 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 chapter 3) 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 chapter 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

¹According 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).

²The 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.

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.

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 onward. 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 onward.

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,

³Tax 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 4.2.

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 Fooker *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 chapter 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 chapter 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, 2022), 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 Sáenz 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 chapter confirms the positive effect of electronic payments on tax revenue. Whilst a similar tax revenue effect is not present in the Uruguayan rebate programme

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

in Brockmeyer and Sáenz 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 4.2 describes the lottery and Section 4.3 the data. Section 4.4 documents the effect of winning on VAT revenue. Section 4.5 and Section 4.6 investigate changes to the payment behaviour of winners and spillovers to non-winners, respectively. Finally, Section 4.7 concludes.

4.2 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 4.21.⁸ 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

⁵The 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").

⁶It 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.

⁷Individuals 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.

⁸Note 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 €50,000 per month. However, for this chapter only the initial scale applies.

$m + 1$.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.

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 4.23 in Appendix 4.9.1. 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.

⁹For 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.

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

¹¹The 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>.

¹²This 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.

¹³A visual illustration of the lottery's timeline in 2017 is shown in Figure 4.22 in Appendix 4.9.1

4.3 Data

Winners Three administrative datasets are used in this chapter. 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 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 4.9.2.

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

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

¹⁵Note, 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.

¹⁶The 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 4.2.

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 4.8 in Appendix 4.8.1 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 4.2.

4.4 Effect of Winning on VAT Revenue

4.4.1 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 4.9 in Appendix 4.8.1. 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 4.2. 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

¹⁷Smaller 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.

¹⁸Note that the overall VAT amount was €12 billion. Listed companies belonged to 3 dedicated tax offices as mentioned in Section 4.2, recorded the remaining €7 billion.

and non-winners sample by multiplying (or expanding) the non-winners population by a factor $\omega = 129$, as explained Appendix 4.9.2. 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 tax office with winners W and non-winners NW the monthly aggregate consumption per tax office becomes:

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

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:

$$\underbrace{R_{i,t}}_{VAT\ Revenue} = \alpha + \underbrace{\beta W_i \times Post_t}_{Winners\ variation} + \underbrace{\sum_{\ell}^{\mathcal{L}} \gamma_{t-\ell} C_{i,t-\ell}}_{Consumption\ Period} + \delta_i + \lambda_t + \epsilon_{i,t} \quad (4.1)$$

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 δ_i and λ_t , respectively.

4.4.2 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)

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

²⁰Detailed estimates for the controls are shown in Table 4.3.

²¹A back-of-the-envelope 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.

²²Quarterly observations avoid large spikes every quarter originating from the reporting requirements of smaller firms, as was discussed in Section 4.2.

and (5) include only 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 4.1, 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.

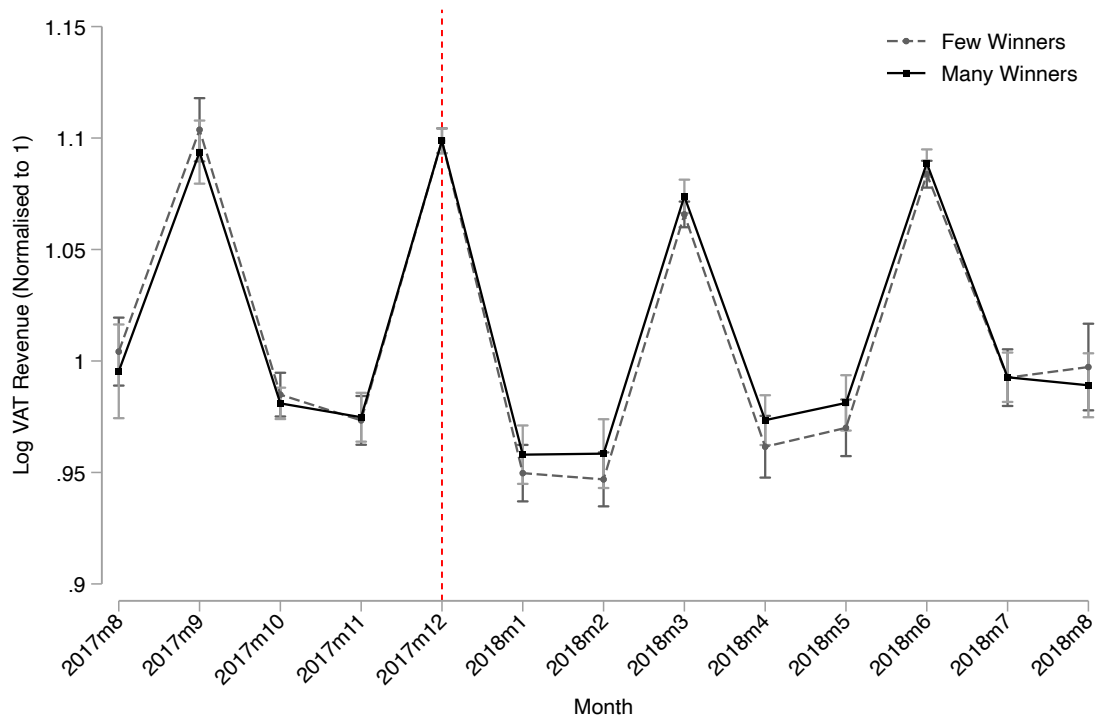
Table 4.1 Effect of Winning on VAT Revenue

	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*** (0.0004)	0.0012*** (0.0004)	0.0010** (0.0004)	0.0023*** (0.0005)	0.0012*** (0.0004)	0.0012*** (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 4.1. "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 4.1 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 4.4 in Appendix 4.8.2 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 4.4, 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 than 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 4.11 and Figure 4.12,

respectively, in Appendix 4.8.1. 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 4.4 in Appendix 4.8.2. 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.

4.5 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 4.13. 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 4.14. 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

²³Differences 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 4.15 and the distribution of annual consumption in Figure 4.16 in Appendix 4.8.1. 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.

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.

4.5.1 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:

$$P(w_n = 1) = \frac{1}{1 + \exp\left(-\left(\beta_0 + \sum_{t=1}^9 \beta_t T_{n,t}\right)\right)} \quad (4.2)$$

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 4.5 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 4.17. 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 4.6. The mean and median propensity scores for winners are 0.237 and 0.177 respectively, compared 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

²⁴Tickets were transformed from the electronic consumption of individuals, using the tickets-to-euros mapping as explained in Section 4.2.

in the distribution of non-winners. A comparison of the resulting Kernel density functions is shown in Figure 4.17 with corresponding density function characteristics at the bottom part of Table 4.6. 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.

4.5.2 Estimation

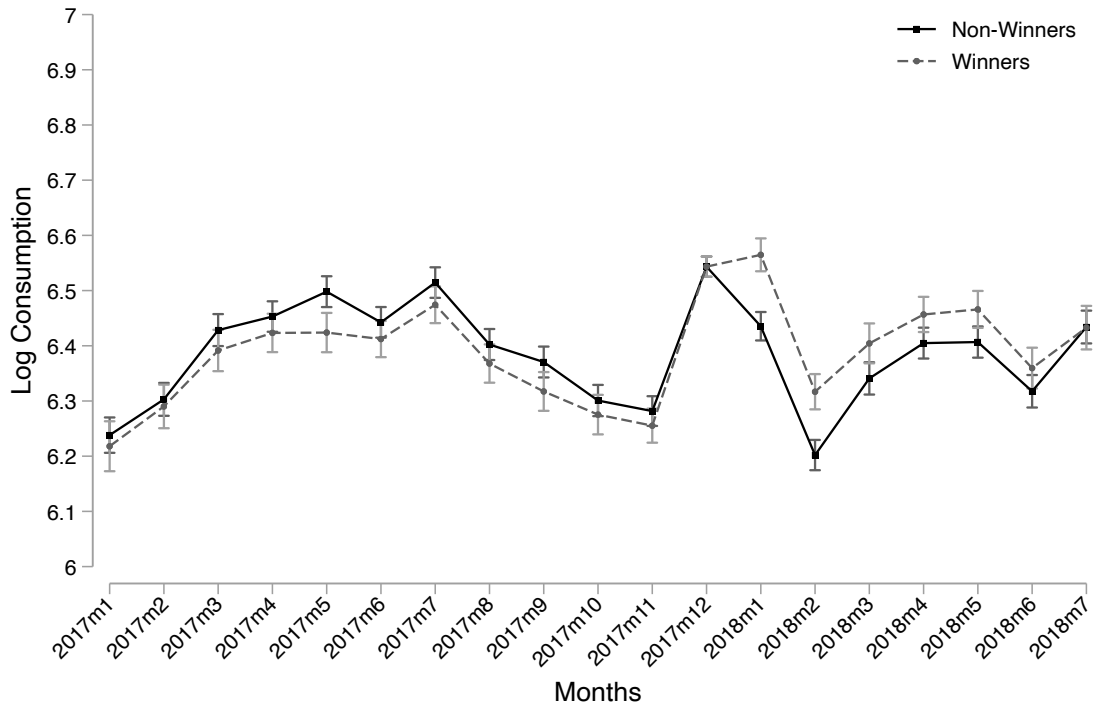
The effect of winning on electronic consumption can be identified by comparing the two groups in a difference-in-difference setting with individual and time fixed effects. Given monthly electronic consumption $c_{n,t}^w$ for individual n and indicator for winning w_n , I fit the following regression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

$$\underbrace{c_{n,t}^w}_{\text{Winners } E\text{-Consumption}} = \alpha + \overbrace{\beta w_n \times Post_t}^{\text{Winners indicator}} + \chi_n + \lambda_t + \epsilon_{n,t} \quad (4.3)$$

Variables χ_n and λ_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-weighting individuals with high propensity score who have spent more and received more tickets and under-weighting individuals who have spent less and received less tickets. Results from Regression 4.3 are shown in Table 4.7 and plotted in Figure 4.2.

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. 4.2 Effect of Winning on E-Consumption

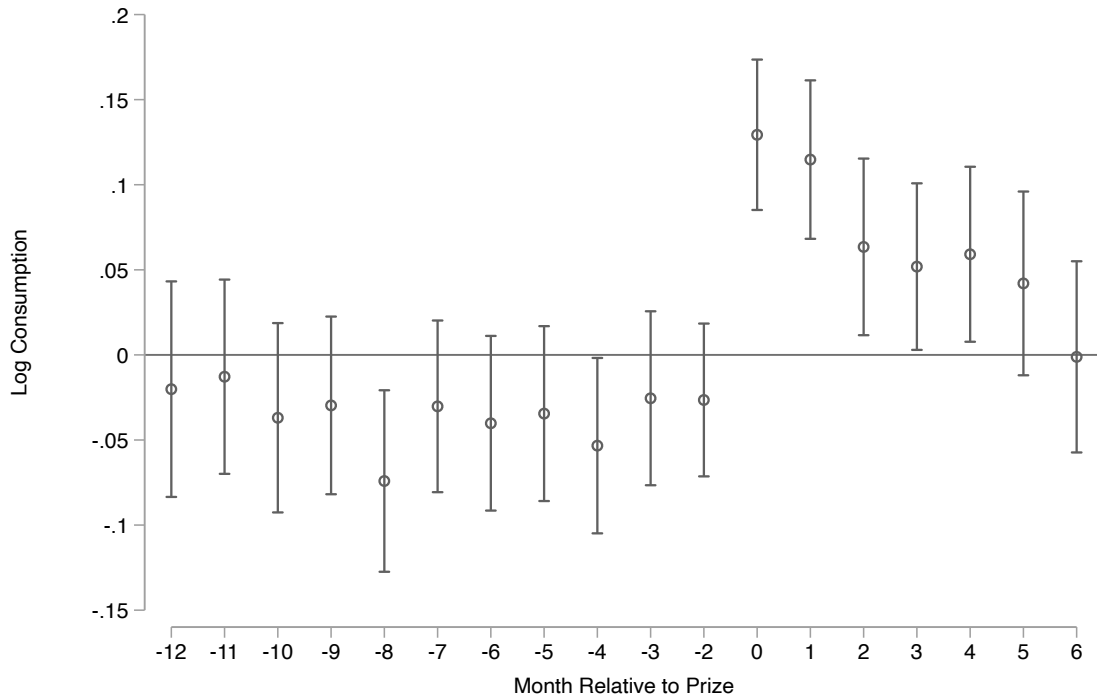


Notes: The figure presents predicted values using Regression 4.3 for winners ($w_n = 1$) and non-winners ($w_n = 0$). Estimates are shown in Table 4.7. The figure plots monthly differences in electronic consumption (logarithmic form) with reference to December 2017 (-1 period to receiving the lottery 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.

Figure 4.3 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 4.2, 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 4.4.

Fig. 4.3 Event Study



Notes: The figure presents results from an event study that correspond to the results in Figure 4.2 and for estimates are Table 4.7. 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.

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 4.10 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 the 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*). I compare the electronic consumption of non-winners in low against high percentages of winners, controlling for postcode, individual and time fixed effects.

4.6.1 Propensity Score

To produce a meaningful comparison between non-winners, I follow the same approach as in Section 4.5.1. Firstly, I calculate propensity scores based on spending during January to September 2017 using Regression 4.2. 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 4.20 and Figure 4.19 respectively. Similar to Section 4.5.1, I limit the samples to propensity scores below 0.17 and above 0.4 to produce comparable samples in terms of spending.

4.6.2 Estimation

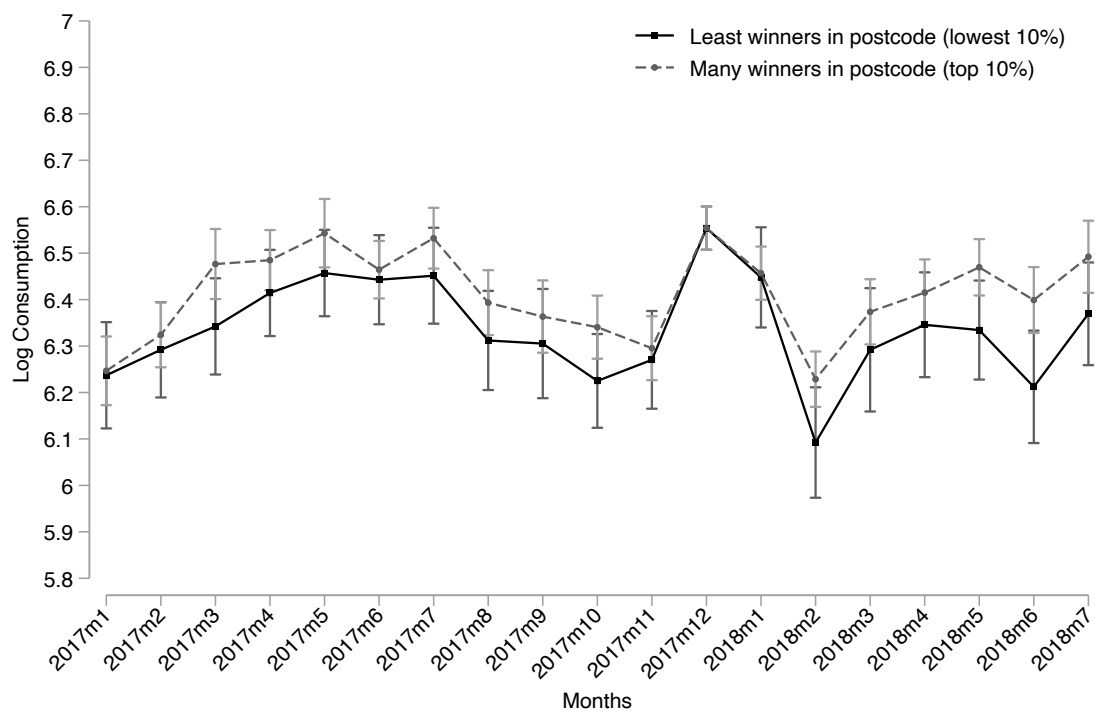
Spillover effects in the electronic consumption of non-winners are estimated using difference-in-difference with postcode, individual and time fixed effects. Let $c_{p,n,t}^{nw}$ denote the electronic consumption at time t of a non-winning individual n residing at postcode p . Let w_p be a binary variable 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:

$$\underbrace{c_{p,n,t}^{nw}}_{\text{Non-winners consumption}} = \alpha + \overbrace{\beta w_p \times Post_t}^{\text{Many/few winners indicator}} + \delta_p + \chi_n + \lambda_t + \epsilon_{p,n,t} \quad (4.4)$$

Variables δ_p , χ_n and λ_t capture postcode, individual and time fixed effects respectively. Inverse probability weights, calculated from the propensity scores using Equation 4.2. 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.

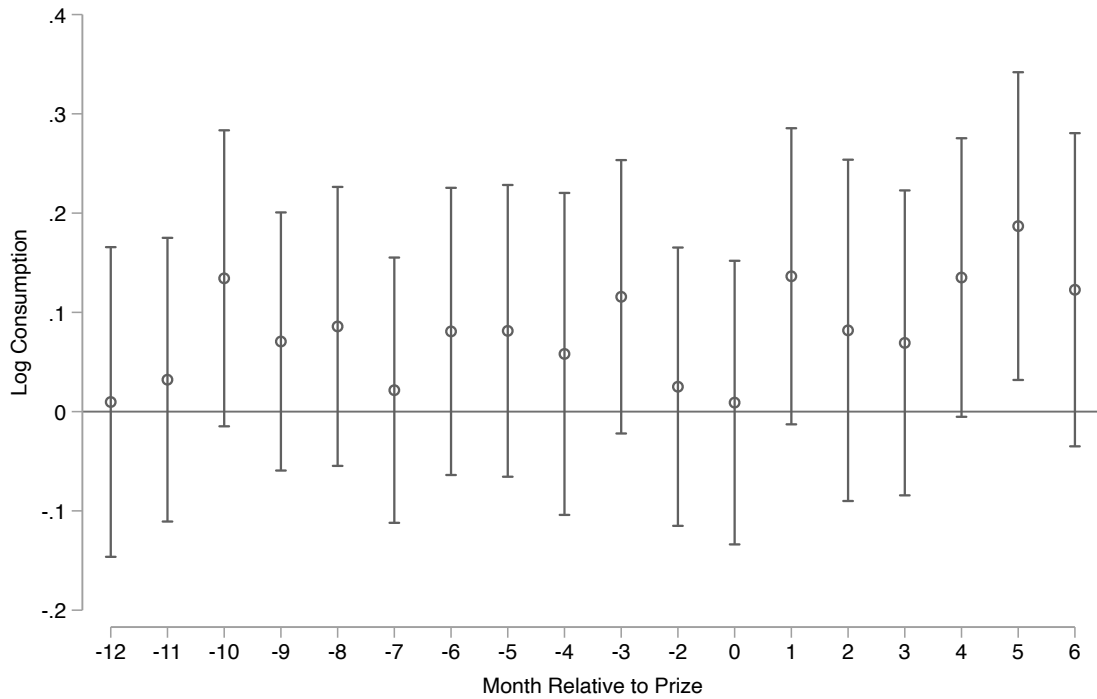
Results from Group 1 are shown in Table 4.8 and monthly differences from Regression 4.4 are presented in Figure 4.4. Non-winners in postcodes with the highest/lowest percentage of winners exhibit a similar electronic consumption pattern prior to the lottery with matching seasonal fluctuations. Monthly 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 4.5 presents the differences between the two samples, relative to the month winners received the prize.

Fig. 4.4 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 4.8, 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.

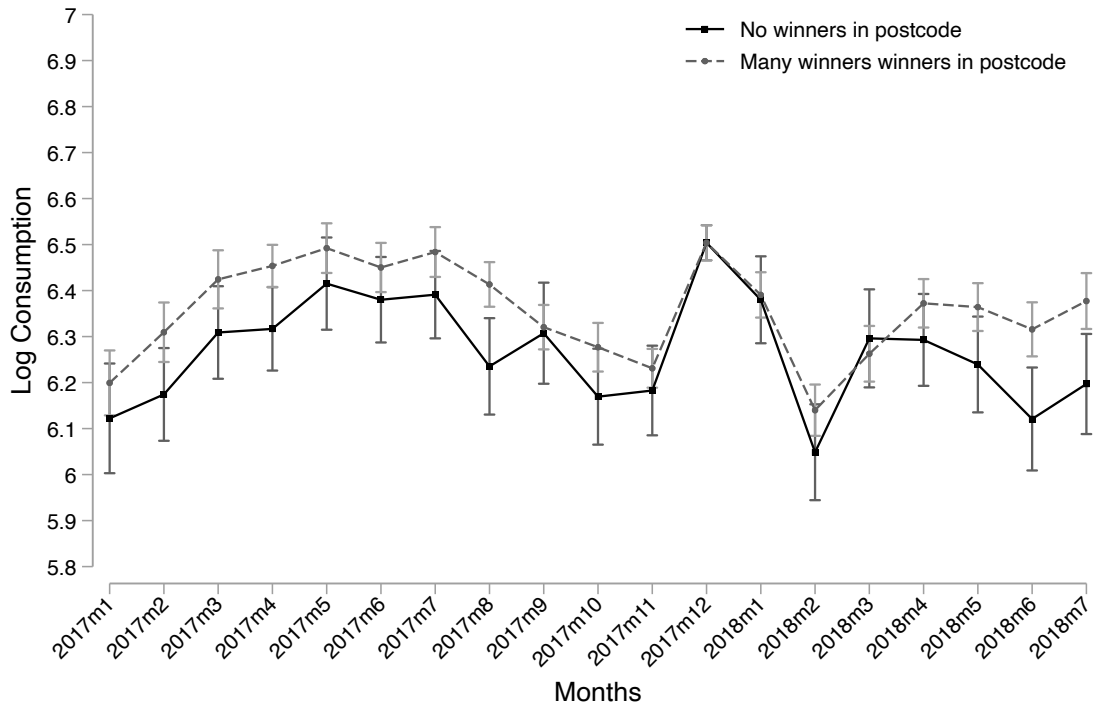
Fig. 4.5 Event Study (10% Lowest/Highest)



Notes: The figure plots estimates from an event study that correspond to the results in Figure 4.4 and Table 4.8. 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 be 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 patterns 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 4.6. The only statistically significant difference prior to the lottery is observed in 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.

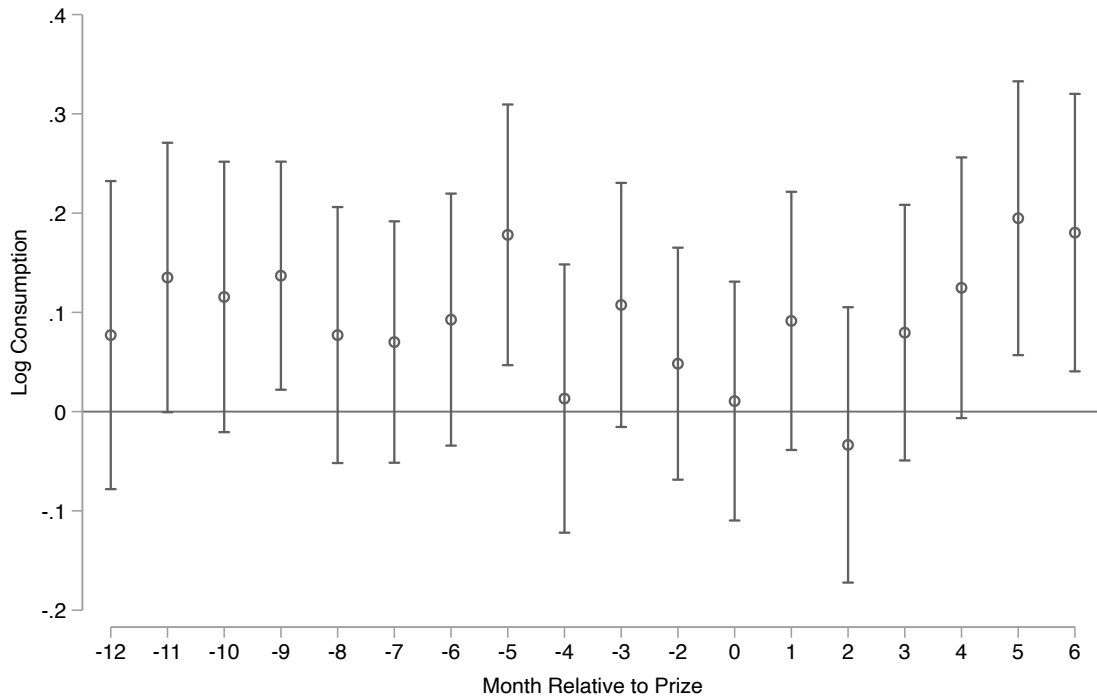
Fig. 4.6 Effect of Winning on Non-winners E-Consumption (No winners)



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 4.8, 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 4.7. 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. 4.7 Event Study (No winners)



Notes: The figure plots estimates from an event study that correspond to the results in Figure 4.6 and Table 4.8. 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.

4.7 Conclusion

This chapter 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 chapter 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 Sáenz 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.

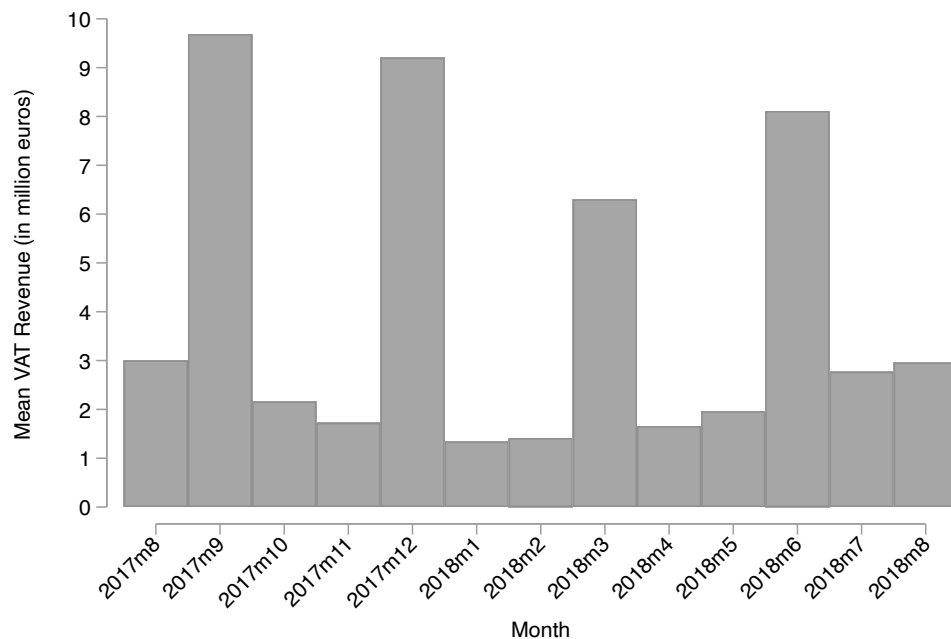
Thirdly, the temporary effect of winning on electronic consumption indicates the limitations of the policy in facilitating a switch from cash to electronic payments. The tax lottery does not succeed in changing payment habits permanently; the latter appear to be persistent, given that winners revert 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.

4.8 Appendix: Figures and Tables

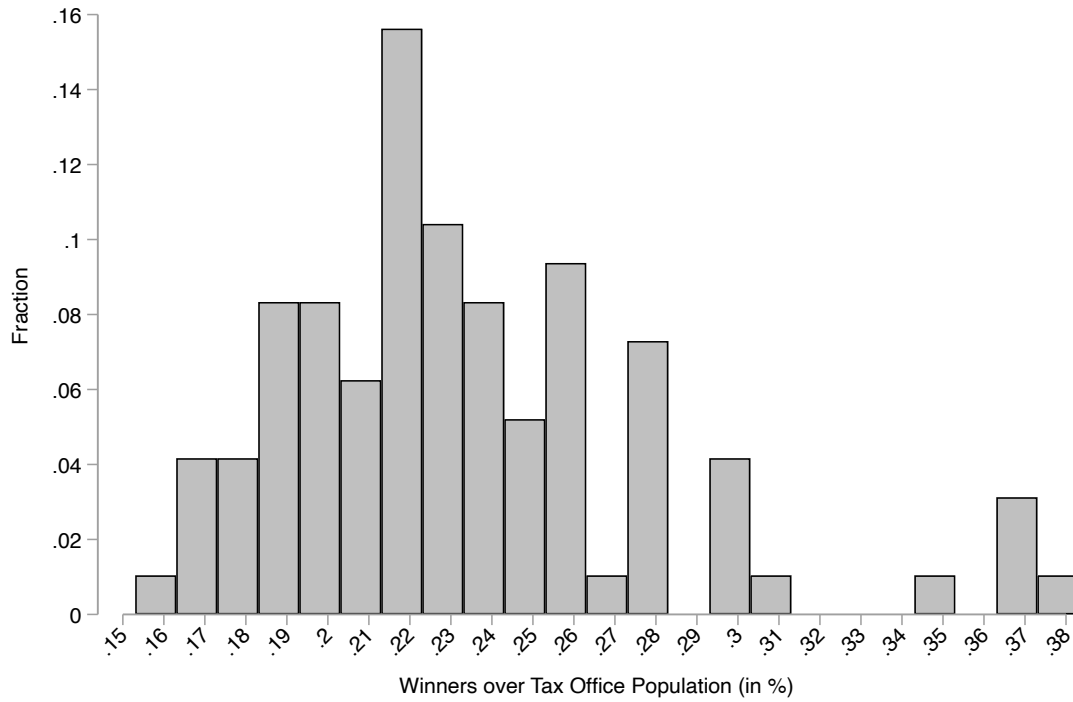
4.8.1 Figures

Fig. 4.8 Mean Tax Office Revenue



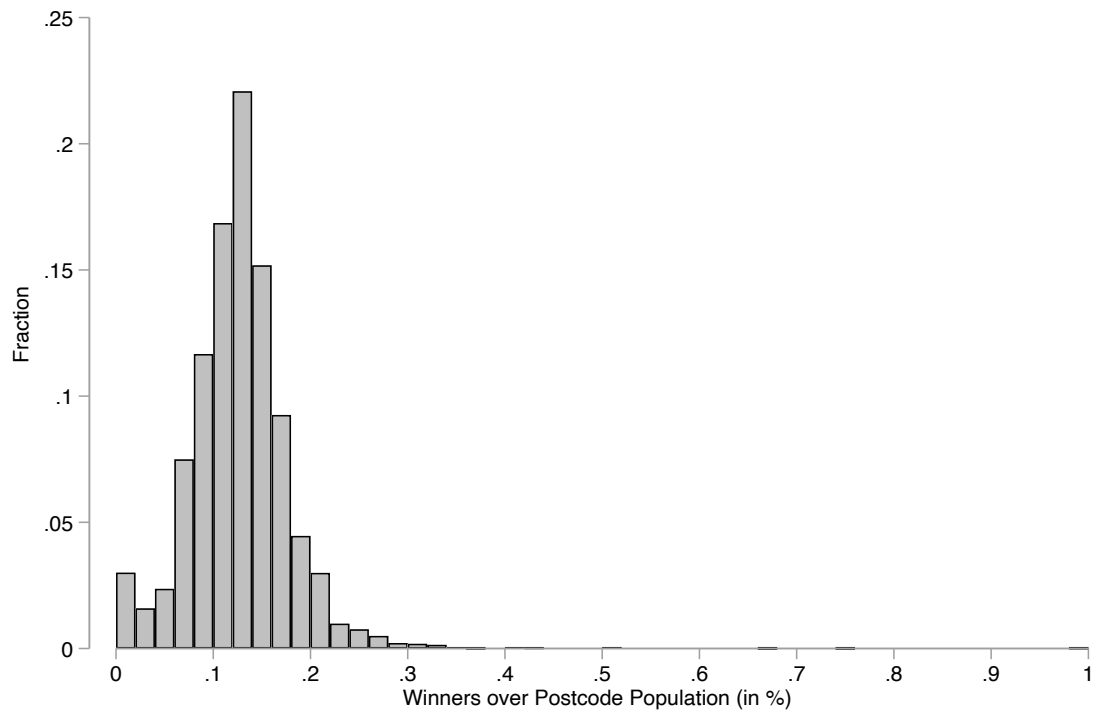
Notes: The graph plots mean VAT revenue in million euros, collected in 96 tax offices in Greece from August 2017 to August 2018. VAT revenue correspond to accounting code 1119 in the Greek public finance system, which corresponds to tax revenue from activities other than building, tobacco, fuel and alcohol 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. 4.9 Variation of Winners in Tax Office Population



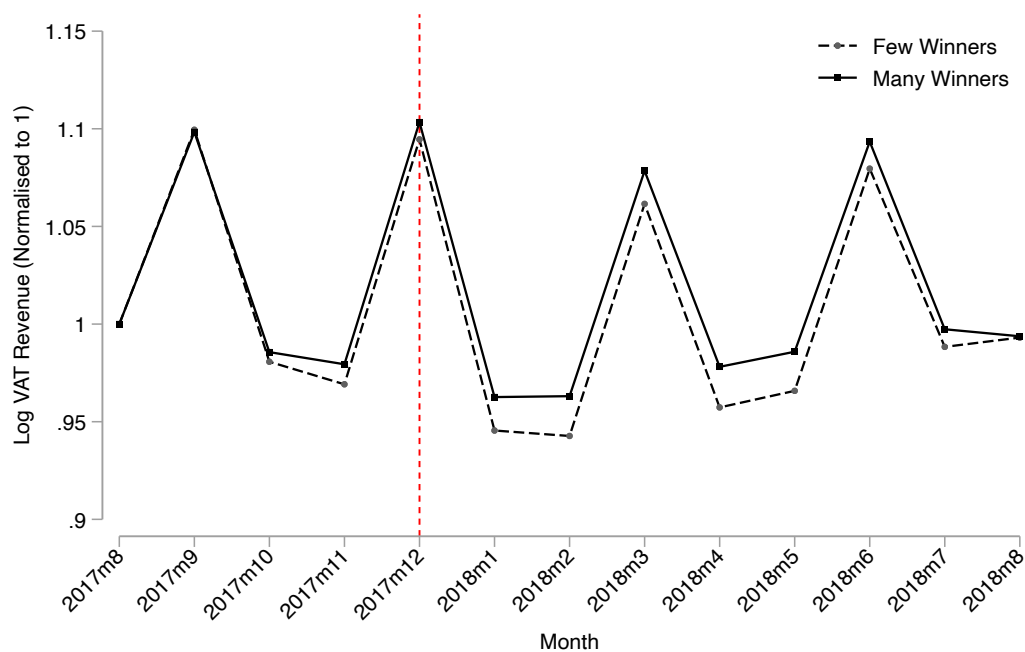
Notes: The figure presents the distribution of winners as a percentage of the population in each tax office. To construct this, winners in each tax office are divided by the tax office population and multiplied by 100. This gives the percentage of winners (x-axis). For example taking the most frequent 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 4.9.2. The distribution includes 96 tax offices.

Fig. 4.10 Variation of Winners in Postcode Population



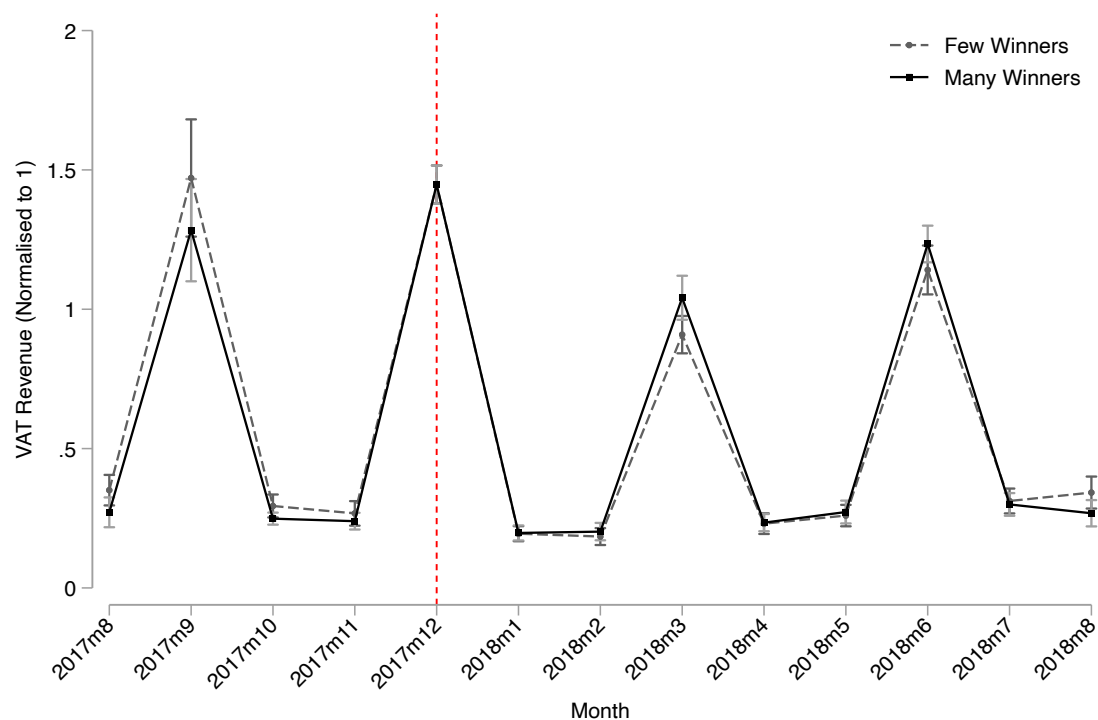
Notes: The figure presents the distribution of winners as a percentage of the population in each postcode. To construct this, winners in each tax office are divided by the postcode population and multiplied by 100. This gives the percentage of winners (x-axis). For example taking the highest observation, 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 4.9.2. The distribution includes 1,099 postcodes.

Fig. 4.11 Regional Tax Offices with Few/Many Winners (Unconditional Means)



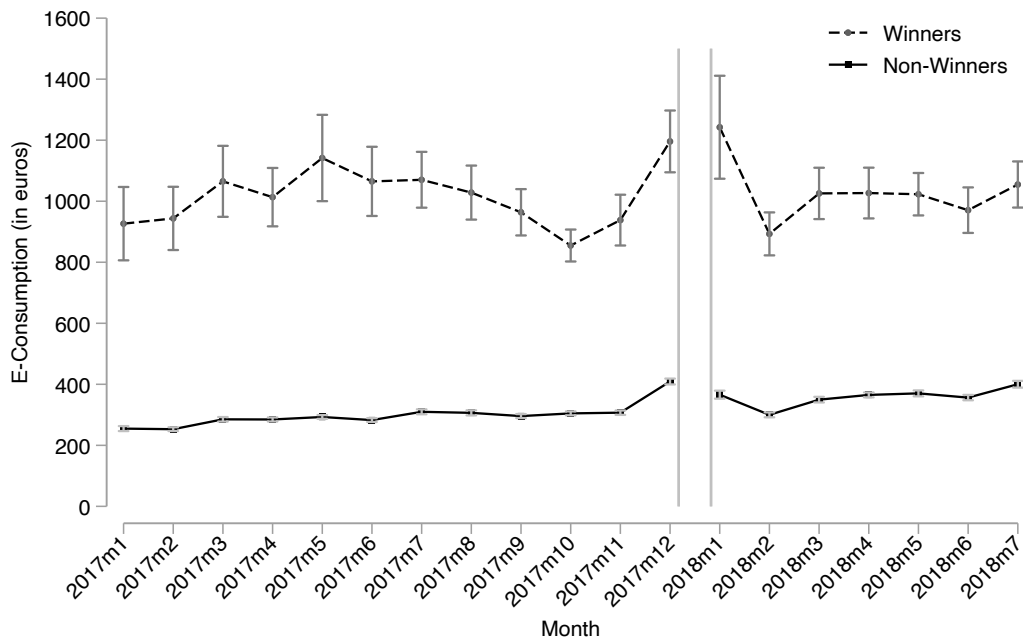
Notes: The figure presents unconditional monthly differences in VAT revenue (logarithmic form) between tax offices with many winners (solid line) and tax offices with few winners (dashed line). The red line is a reference point to December 2017, the month of the lottery and before winners received the 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. 4.12 Regional Tax Offices with Few/Many Winners (Poisson Regression)



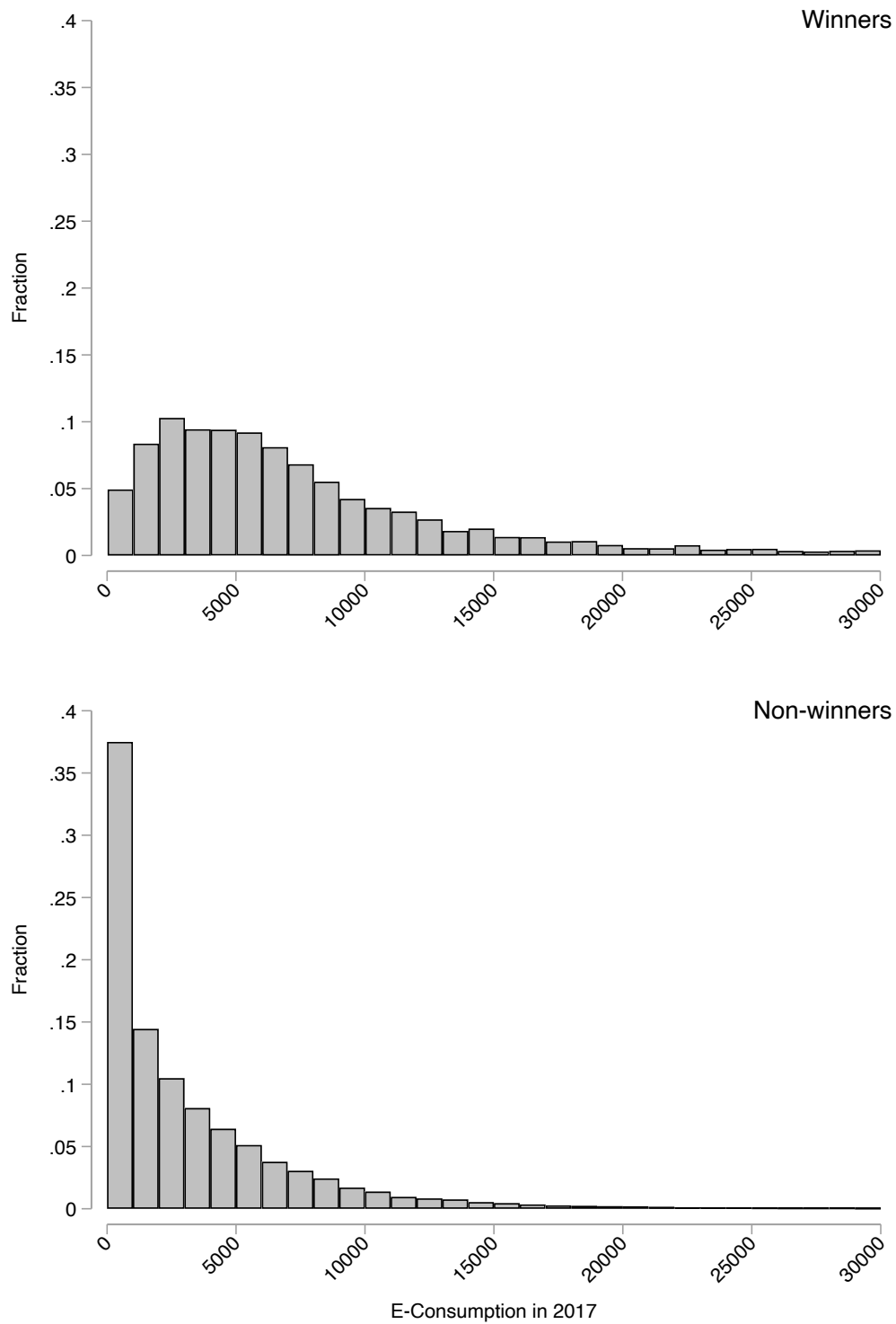
Notes: The figure presents monthly differences in VAT revenue between tax offices with many winners (solid line) and tax offices with few winners (dashed line). They are obtained from fitted values in a Poisson regression with interacted months (using absolute values) and after controlling for month and tax offices fixed effects. Regression estimates are shown in Table 4.4, 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. 4.13 Electronic Consumption (Unconditional Means)



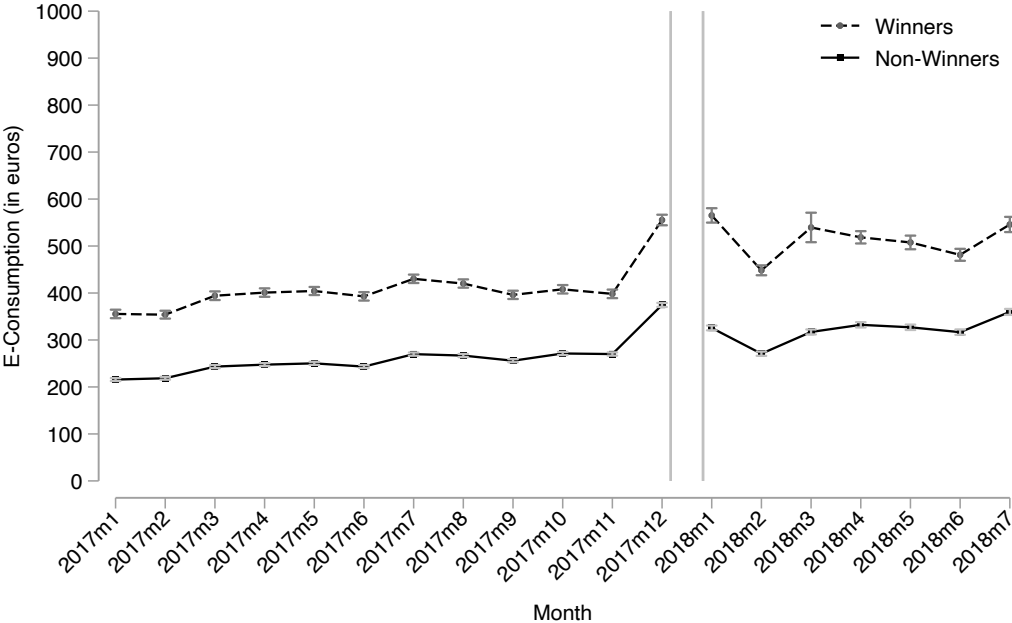
Notes: The figure shows unconditional mean electronic consumption for winners (dashed line) and non-winners (solid line). Winners include the 9,000 superdraw winners and non-winners the 50,000 randomly drawn sample. Electronic consumption is shown in the y-axis and months in the x-axis. The sample includes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

Fig. 4.14 Electronic Consumption in 2017



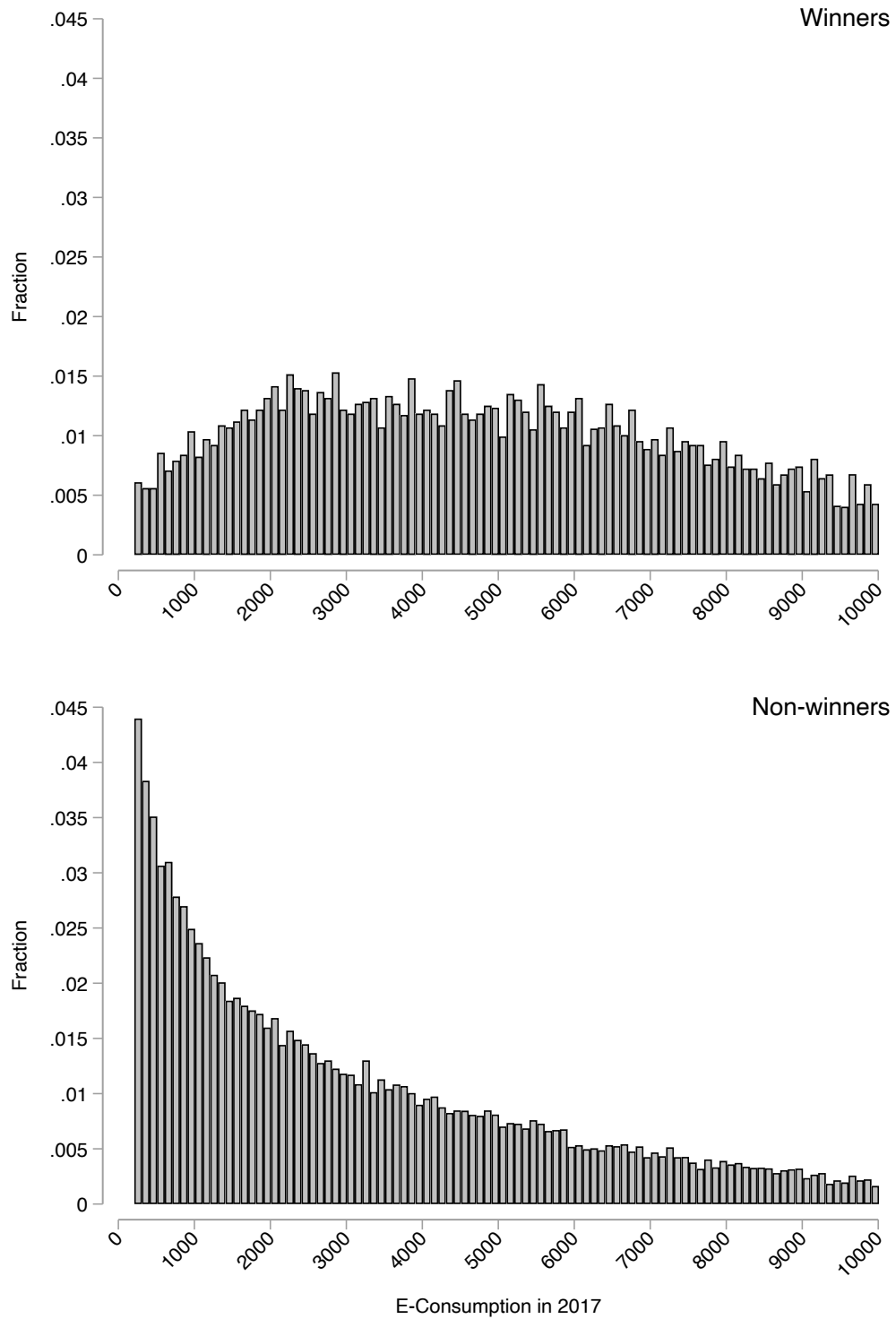
Notes: The figures plot the distributions of annual electronic consumption in 2017 for winners (top panel) and non-winners (bottom panel). The annual electronic consumption is shown in the x-axis. The sample includes 9,000 winners and 50,000 non-winners. The bins of the distributions are at €1,000 and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

Fig. 4.15 Electronic Consumption - Winsorized Sample (Unconditional Means)



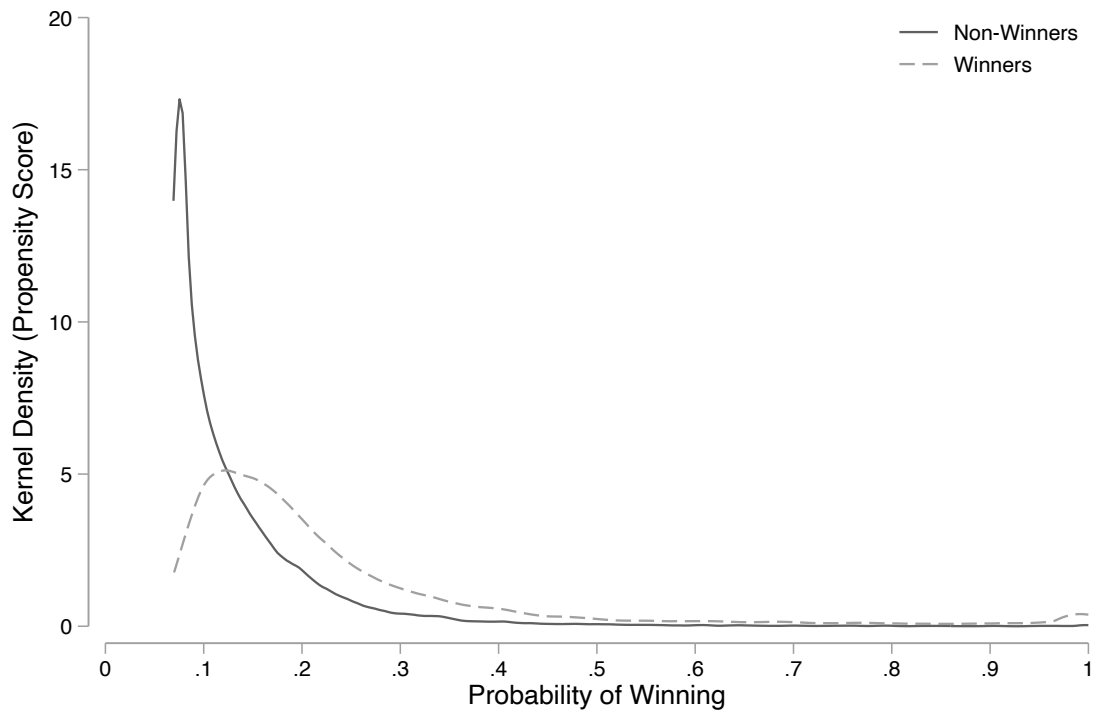
Notes: The figure shows unconditional mean electronic consumption for winners (dashed line) and non-winners (solid line). The sample is winsorized at top/bottom 10% based annual electronic consumption. Electronic consumption is shown in the y-axis and months in the x-axis. The sample includes 19 months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

Fig. 4.16 Electronic Consumption in 2017 - Winsorized Sample



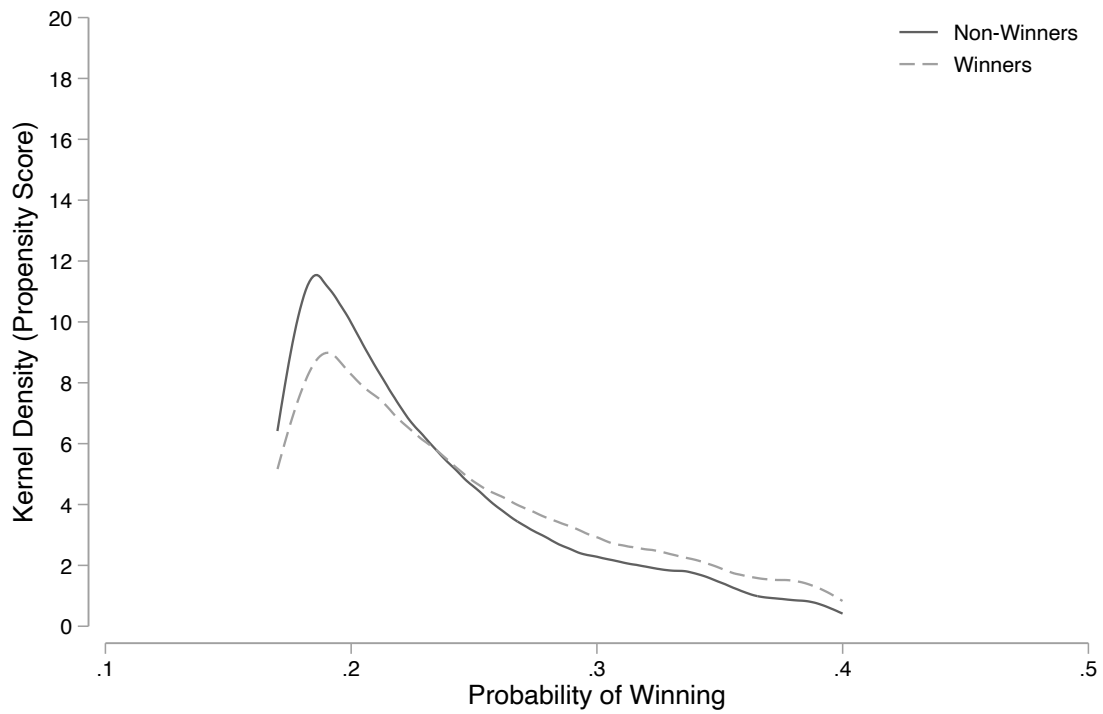
Notes: The figures plot the distributions of annual electronic consumption in 2017 for winners (top panel) and non-winners (bottom panel). The sample is winsorized at top/bottom 10% based annual electronic consumption. The annual electronic consumption is shown in the x-axis. The bins of the distributions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

Fig. 4.17 Propensity Score - Whole sample



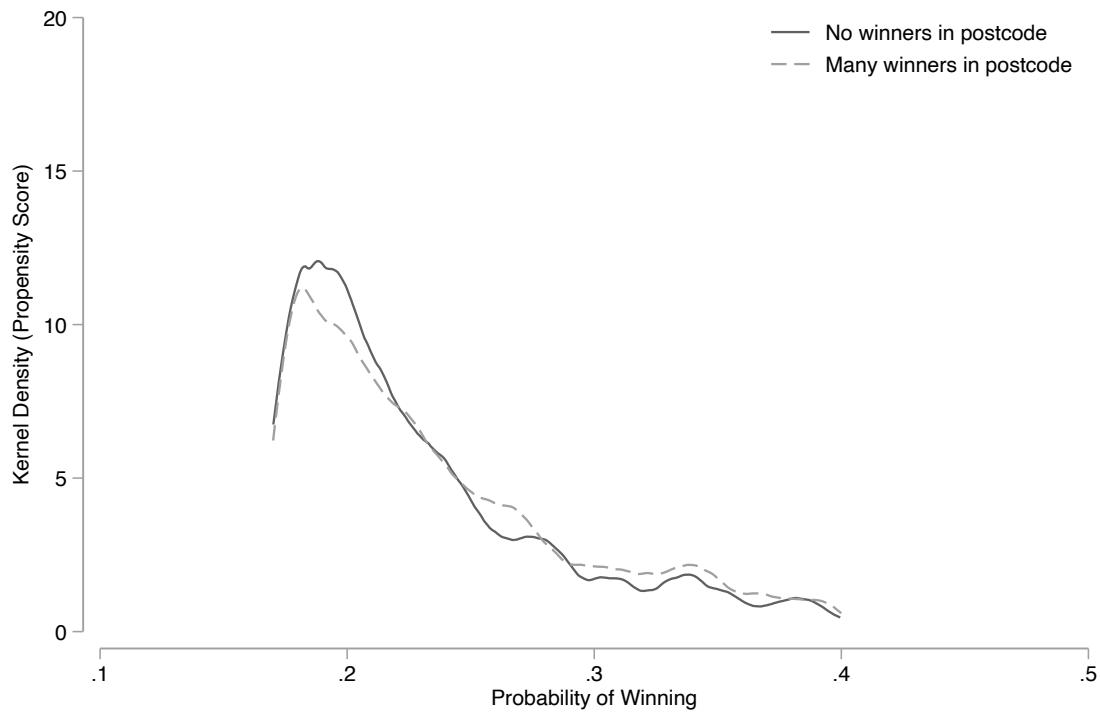
Notes: The graph plots kernel density functions of the propensity scores generated by Equation 4.2 for winners (dashed line) and non-winners (solid line). The propensity scores indicate the probability of individuals in winning in the superdraw based on the tickets generated in the months of January to September 2017 (which corresponded to the lottery draws).

Fig. 4.18 Propensity Score - Matching e-consumption



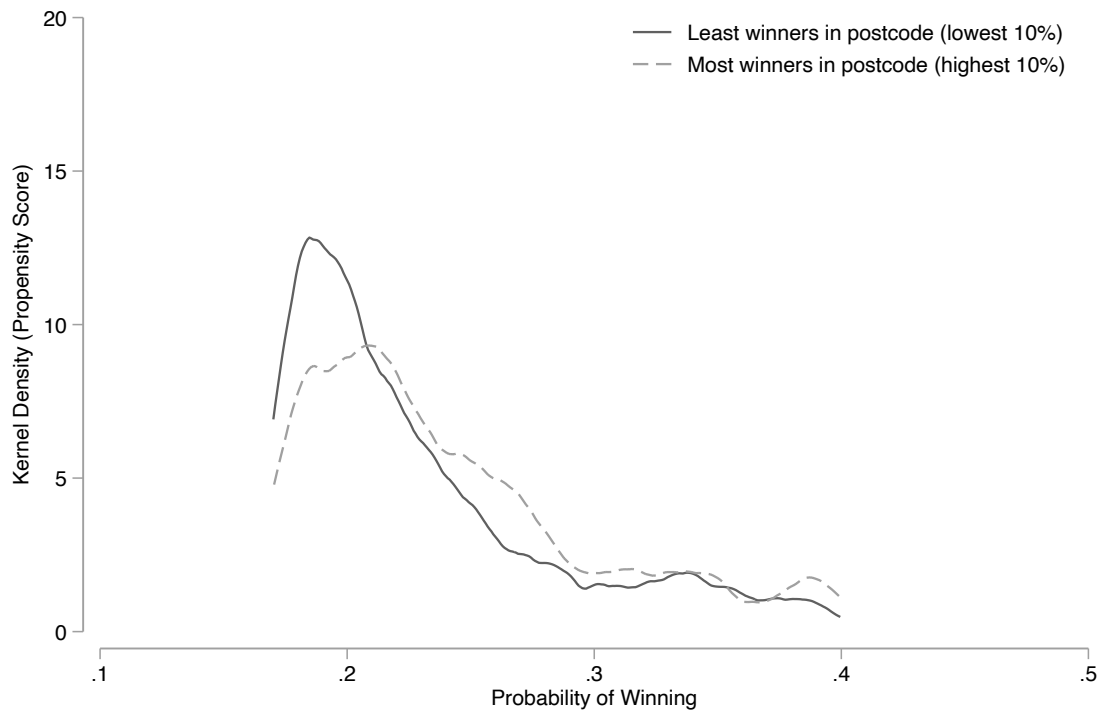
Notes: The graph plots kernel density functions of the propensity scores generated by Equation 4.2 for winners (dashed line) and non-winners (solid line). The propensity scores indicate the probability of individuals in winning in the superdraw based on the tickets generated in the months of January 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. 4.19 Propensity Score - Non-winners with no winners in postcode



Notes: The graph plots kernel density functions of the propensity scores generated by Equation 4.2 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. 4.20 Propensity Score - Non-winners with 10% lowest winners in postcode



Notes: The graph plots kernel density functions of the propensity scores generated by Equation 4.2 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.

4.8.2 Tables

Table 4.2 Summary Statistics

	Winners			Non-winners		
	(1) N	(2) Mean	(3) St.Dev.	(4) N	(5) Mean	(6) 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 4.13.

Table 4.3 Effect of Winning on VAT Revenue - Detailed

	Monthly (1) Log Revenue	(2) Log Revenue	(3) Log Revenue	Quarterly (4) Log Revenue	(5) Log Revenue	(6) Log Revenue
Superdraw Winners	0.0017*** (0.0004)	0.0012*** (0.0004)	0.0010** (0.0004)	0.0023*** (0.0005)	0.0012*** (0.0004)	0.0012*** (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*** (0.0031)	14.3048*** (0.0045)	-10.0551 (8.6603)	15.8568*** (0.0092)	15.7121*** (0.0143)	22.0163*** (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 4.1. 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 4.1 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 4.4 Few versus Many Winners (Bottom/Top 20% of Tax Offices)

	Linear (1) Log VAT Revenue	Poisson (2) VAT Revenue
Many Winners Interaction with:		
August 2017	-0.0088 (0.0145)	-0.2571* (0.1442)
September 2017	-0.0100 (0.0096)	-0.1361 (0.1146)
October 2017	-0.0039 (0.0077)	-0.1682* (0.0893)
November 2017	0.0014 (0.0092)	-0.1103 (0.1116)
January 2017	0.0083 (0.0105)	0.0141 (0.1047)
February 2018	0.0115 (0.0114)	0.0919 (0.1250)
March 2018	0.0081* (0.0045)	0.1360** (0.0561)
April 2018	0.0119 (0.0106)	0.0157 (0.1167)
May 2018	0.0112 (0.0104)	0.0481 (0.1166)
June 2018	0.0050 (0.0037)	0.0786 (0.0544)
July 2018	0.0002 (0.0103)	-0.0406 (0.1153)
August 2018	-0.0082 (0.0133)	-0.2445* (0.1418)
Constant	1.0988*** (0.0028)	1.6200*** (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 4.12 in Column (2). These are generated from Regression 4.1 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 4.5 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 4.2. 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 4.17.

Table 4.6 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 4.2 and plotted in Figure 4.17 for the top "Whole Sample" panel and in Figure 4.18 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 4.7 Winners E-Consumption

	Linear (1)	Poisson (2)
	Log E-Consumption	E-Consumption
Winners Interaction with:		
January 2017	-0.0201 (0.0323)	0.00285 (0.0233)
February 2017	-0.0128 (0.0291)	-0.00975 (0.0220)
March 2017	-0.0369 (0.0284)	-0.00468 (0.0235)
April 2017	-0.0297 (0.0266)	-0.0145 (0.0211)
May 2017	-0.0741*** (0.0272)	-0.0506** (0.0217)
June 2017	-0.0302 (0.0257)	-0.0210 (0.0206)
July 2017	-0.0402 (0.0262)	-0.0307 (0.0208)
August 2017	-0.0345 (0.0262)	-0.0260 (0.0206)
September 2017	-0.0533** (0.0263)	-0.0431** (0.0208)
October 2017	-0.0255 (0.0261)	-0.0223 (0.0220)
November 2017	-0.0265 (0.0229)	-0.0416** (0.0211)
January 2018	0.129*** (0.0225)	0.101*** (0.0224)
February 2018	0.115*** (0.0238)	0.0862*** (0.0255)
March 2018	0.0635** (0.0265)	0.0562 (0.0368)
April 2018	0.0519** (0.0250)	-0.00235 (0.0220)
May 2018	0.0591** (0.0262)	0.0242 (0.0282)
June 2018	0.0420 (0.0275)	0.0142 (0.0229)
July 2018	-0.00115 (0.0287)	-0.00386 (0.0291)
Constant	6.543*** (0.00938)	6.927*** (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 4.3. 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 4.2. 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 4.18. Robust standard errors are clustered at the individual level.

Table 4.8 Non-winners E-Consumption

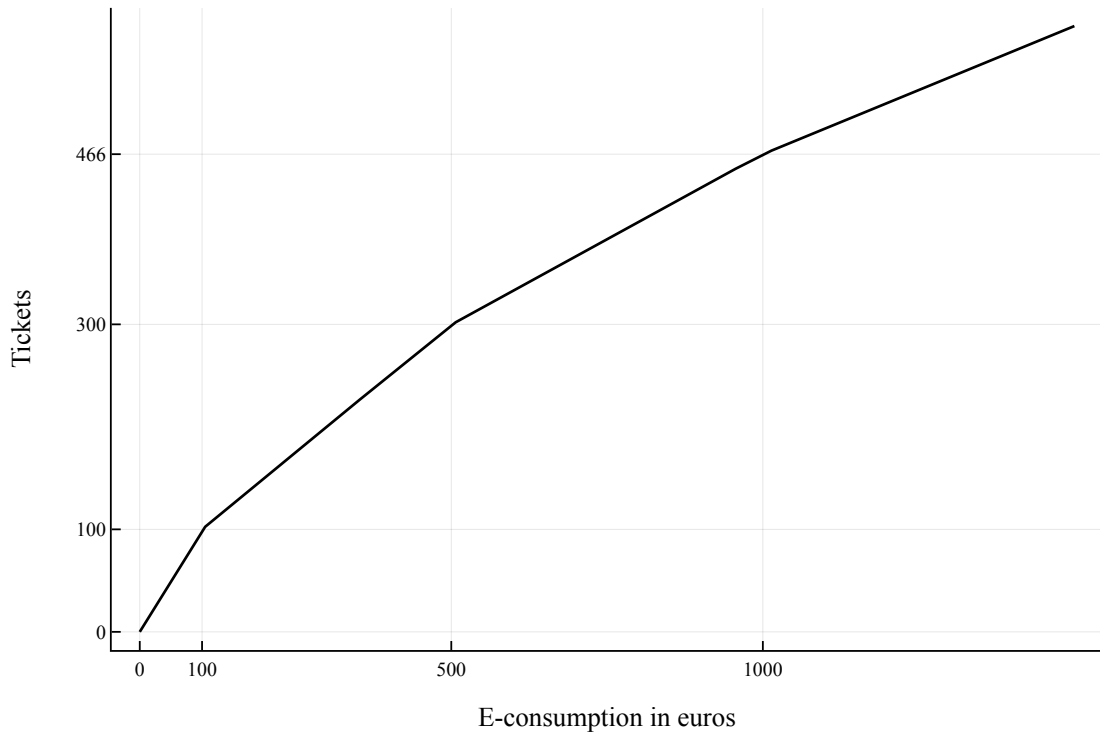
	Lowest/Highest 10% (1) Log E-Consumption	No/Many Winners (2) Log E-Consumption
Non-winners with Many Winners in Postcode Interaction with:		
January 2017	0.00972 (0.0796)	0.0771 (0.0792)
February 2017	0.0322 (0.0729)	0.135* (0.0692)
March 2017	0.134* (0.0761)	0.116* (0.0695)
April 2017	0.0706 (0.0663)	0.137** (0.0586)
May 2017	0.0858 (0.0717)	0.0771 (0.0658)
June 2017	0.0216 (0.0682)	0.0701 (0.0620)
July 2017	0.0808 (0.0738)	0.0927 (0.0648)
August 2017	0.0814 (0.0750)	0.178*** (0.0670)
September 2017	0.0581 (0.0827)	0.0131 (0.0689)
October 2017	0.116 (0.0703)	0.107* (0.0627)
November 2017	0.0251 (0.0715)	0.0483 (0.0596)
January 2018	0.00910 (0.0729)	0.0106 (0.0614)
February 2018	0.136* (0.0761)	0.0914 (0.0664)
March 2018	0.0819 (0.0877)	-0.0335 (0.0708)
April 2018	0.0692 (0.0784)	0.0796 (0.0657)
May 2017	0.135* (0.0716)	0.125* (0.0670)
June 2017	0.187** (0.0791)	0.195*** (0.0704)
July 2017	0.123 (0.0805)	0.180** (0.0713)
Constant	6.554*** (0.0238)	6.504*** (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 4.4. Column (1) includes a sample with the lowest/highest percentage of winners in their population as illustrated in Figure 4.4. 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 4.6. 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 4.20. For Column (2) the propensity scores correspond to Figure 4.19. The depended variable is electronic consumption of individuals in logarithmic form. Robust standard errors are clustered at the postcode level.

4.9 Appendix: Tax Lottery Information

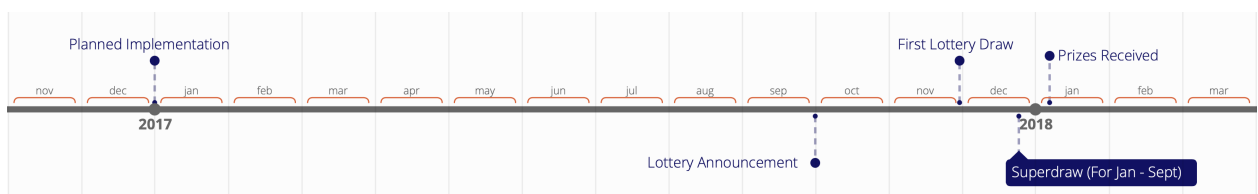
4.9.1 Lottery

Fig. 4.21 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 4.2.

Fig. 4.22 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 2018.

Fig. 4.23 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 transferred 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 4.21. 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.

4.9.2 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:

$$\sum_{m=1}^{12} \bar{T}_m = \sum_{m=1}^{12} \sum_{i=1}^{N_1} T_{i,m,1} + \sum_{m=1}^{12} \sum_{i=1}^{N_1} \hat{T}_{i,m,1} + \omega \sum_{m=1}^{12} \sum_{i=1}^{N_2} T_{i,m,2} \quad (4.5)$$

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 ω , 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 ω to be 129. Expanding the random sample gives a total taxpayer population of 6.45 million ($50,000 \times 129$), 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

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

Bibliography

- ADAM, S., BROWNE, J., PHILLIPS, D. and ROANTREE, B. (2021). Frictions and taxpayer responses: evidence from bunching at personal tax thresholds. *International Tax and Public Finance*, **28**, 612–653.
- 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.
- ALLINGHAM, M. G. and SANDMO, A. (1972). Income tax evasion: a theoretical analysis. *Journal of Public Economics*, **1** (3), 323–338.
- ALMUNIA, M. and RODRIGUEZ, D. L. (2014). Heterogeneous Responses to Effective Tax Enforcement: Evidence from Spanish Firms, Banco de Espana Working Paper No. 1419.
- ARTAVANIS, N., MORSE, A. and TSOUTSOURA, M. (2016). Measuring income tax evasion using bank credit: Evidence from Greece. *The Quarterly Journal of Economics*, **131** (2), 739–798.
- BASTANI, S. and SELIN, H. (2014). Bunching and non-bunching at kink points of the Swedish tax schedule. *Journal of Public Economics*, **109**, 36–49.
- 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 SÁENZ SOMARRIBA, M. (2022). Electronic Payment Technology and Tax Compliance: Evidence from Uruguay’s Financial Inclusion Reform, CEPR Discussion Paper No. DP17097.

- BROWNING, M., BOURGUIGNON, F., CHIAPPORI, P.-A. and LECHENE, V. (1994). Income and Outcomes: A Structural Model of Intrahousehold Allocation. *Journal of Political Economy*, **102** (6), 1067–1096.
- 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.
- CARROLL, C. D., SLACALEK, J. and TOKUOKA, K. (2014). The Distribution of Wealth and the MPC: Implications of New European Data. *The American Economic Review*, **104** (5), 107–111.
- CHENG, G. *et al.* (2020). *The 2012 private sector involvement in Greece*. Publications Office of the European Union.
- CHETTY, R., FRIEDMAN, J. N., OLSEN, T. and PISTAFERRI, L. (2011). Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records. *The Quarterly Journal of Economics*, **126** (2), 749–804.
- CHIAPPORI, P.-A., GIMÉNEZ-NADAL, J. I., MOLINA, J. A. and VELILLA, J. (2020). *Household Labor Supply: Collective Evidence in Developed Countries*, Cham: Springer International Publishing, pp. 1–28.
- 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*, **66** (3), 198–220.
- 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.
- 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).
- ENGSTRÖM, P. and HOLMLUND, B. (2009). Tax evasion and self-employment in a high-tax country: evidence from Sweden. *Applied Economics*, **41** (19), 2419–2430.
- FELDSTEIN, M. (1999). Tax Avoidance and the Deadweight Loss of the Income Tax. *The Review of Economics and Statistics*, **81** (4), 674–680.
- 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.

- GELBER, A. M., JONES, D. and SACKS, D. W. (2020). Estimating Adjustment Frictions Using Nonlinear Budget Sets: Method and Evidence from the Earnings Test. *American Economic Journal: Applied Economics*, **12** (1), 1–31.
- GORODNICHENKO, Y., MARTINEZ-VAZQUEZ, J. and SABIRIANOVA-PETER, K. (2009). Myth and Reality of Flat Tax Reform: Micro Estimates of Tax Evasion Response and Welfare Effects in Russia. *Journal of Political Economy*, **117** (3), 504–554.
- GUPTA, S., KEEN, M., SHAH, A. and VERDIER, G. (2017). *Digital Revolutions in Public Finance*. International Monetary Fund.
- HAICHAO FAN, N. Q., YU LIU and WEN, J. (2018). The Dynamic Effects of Computerized VAT Invoices on Chinese Manufacturing Firms, NBER Working Paper 24414.
- HEINZE, G. and SCHEMPER, M. (2002). A solution to the problem of separation in logistic regression. *Statistics in Medicine*, **21** (16), 2409–2419.
- HONDROYIANNIS, G. and PAPAIOKONOMOU, D. (2017). The effect of card payments on VAT revenue: New evidence from Greece. *Economics Letters*, **157**, 17 – 20.
- HURST, E., LI, G. and PUGSLEY, B. (2014). Are Household Surveys Like Tax Forms? Evidence from Income Underreporting of the Self-Employed. *The Review of Economics and Statistics*, **96** (1), 19–33.
- IMMORDINO, G. and RUSSO, F. F. (2018). Cashless payments and tax evasion. *European Journal of Political Economy*, **55**, 36–43.
- KAPLANOGLU, G. and RAPANOS, V. T. (2013). Tax and trust: The fiscal crisis in Greece. *South European Society and Politics*, **18** (3), 283–304.
- and — (2015). Why do people evade taxes? New experimental evidence from Greece. *Journal of Behavioral and Experimental Economics*, **56**, 21–32.
- KLEVEN, H. J. (2016). Bunching. *Annual Review of Economics*, **8** (1), 435–464.
- , 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.
- and WASEEM, M. (2013). Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan. *The Quarterly Journal of Economics*, **128** (2), 669–723.
- KOPCZUK, W. and SLEMROD, J. (2006). Putting Firms into Optimal Tax Theory. *The American Economic Review*, **96** (2), 130–134.
- KOURDOUMPALOU, S. (2016). Do corporate governance best practices restrain tax evasion? Evidence from Greece. *Journal of Accounting and Taxation*, **8** (1), 1–10.
- LOVICS, G., SZÍÓKE, 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.

- LUNDBERG, S., POLLAK, R. and WALES, T. (1997). Do Husbands and Wives Pool Their Resources? Evidence from the United Kingdom Child Benefit. *Journal of Human Resources*, **32** (3), 463–480.
- MADZHAROVA, B. (2020). Traceable Payments and VAT Design: Effects on VAT Performance. *CEifo Economic Studies*, **66** (3), 221–247.
- NARITOMI, J. (2019). Consumers as Tax Auditors. *American Economic Review*, **109** (9), 3031–3072.
- 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.
- PIKETTY, T., SAEZ, E. and STANTCHEVA, S. (2014). Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities. *American Economic Journal: Economic Policy*, **6** (1), 230–71.
- 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 ŠMIETANKA, A. (2021). VAT gap in the EU : report 2021, Directorate General Taxation and Customs Union, European Commission.
- PROVOPOULOS, G. A. (2014). The Greek economy and banking system: recent developments and the way forward. *Journal of Macroeconomics*, **39**, 240–249.
- QUISUMBING, A. R. and MALUCCIO, J. A. (2003). Resources at Marriage and Intrahousehold Allocation: Evidence from Bangladesh, Ethiopia, Indonesia, and South Africa. *Oxford Bulletin of Economics and Statistics*, **65** (3), 283–327.
- SAEZ, E. (2010). Do Taxpayers Bunch at Kink Points? *American Economic Journal: Economic Policy*, **2** (3), 180–212.
- , SLEMROD, J. and GIERTZ, S. H. (2012). The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review. *Journal of Economic Literature*, **50** (1), 3–50.
- SLEMROD, J. (2007). Cheating Ourselves: The Economics of Tax Evasion. *Journal of Economic Perspectives*, **21** (1), 25–48.
- , 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.
- TAUBINSKY, D. and REES-JONES, A. (2017). Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment. *The Review of Economic Studies*, **85** (4), 2462–2496.
- 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. (2022). The role of withholding in the self-enforcement of a value-added tax: Evidence from Pakistan. *The Review of Economics and Statistics*, **104** (2), 336–354.