CHARLES UNIVERSITY
FACULTY OF SOCIAL SCIENCES
Institute of Economic studies

The impact of the electronic revenue registry on Czech firms

Bachelor's thesis

Author: Ondřej Varadi
Study program: Economics and Finance
Supervisor: Matěj Bajgar M.Sc., DPhil
Year of defense: 2023

Declaration of Authorship

The author hereby declares that he or she compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis in whole or in part and agrees with the thesis being used for study and scientific purposes.

Prague, August 1, 2023

Ondřej Varadi

Abstract

This thesis studies the effects of the electronic revenue registry introduced in the Czech Republic in December 2016 on reported firm output, purchases, and on firm entry/exit. The policy aimed to reduce tax evasion via improved reporting of cash transactions at the end of the supply chain. Firms in affected industries were required to use special electronic revenue registers that automatically send information to tax authorities. We conduct a difference-in-differences estimation on representative firm-level data provided by the Czech Statistical Office and construct control groups from industries that were unaffected by the policy but were otherwise similar to the affected industries. We construct multiple alternative control groups for each treated industry to test the sensitivity and robustness of the results to the choice of the control group. The thesis finds that the policy increased reported output in industries characterized by a high volume of small-ticket sales, which include *Food and beverage service activities*, *Accommodation*, and *Retail*. Additionally, in *Food and beverage service activities*, reported purchases increased by 16%, which could imply that the firms offset the increase in output by reporting greater purchases. Firm entry decreased and firm exit increased around and after the introduction of the policy, but there was a positive spike in firm entry in 2016. No clear effects were found on *Wholesale* and *Wholesale and retail trade and repair of motor vehicles and motorcycles* industries. The results of the thesis should, however, be interpreted in the light of the limitations of the data used and, in particular, the challenges of constructing a truly comparable control group.

JEL Classification	D22, D24, E26, K34,
Keywords	electronic revenue registry, tax regulation, difference in differences, fixed effects

Title	The impact of the electronic revenue registry on Czech firms

Author's e-mail	`42559143@fsv.cuni.cz`
Supervisor's e-mail	`matej.bajgar@fsv.cuni.cz`

Abstrakt

Tato bakalářská práce se zabývá efektem elektronické evidence tržeb na vykázané příjmy a náklady firem, a na firemní vstupy a výstupy z trhu. Elektronická evidence tržeb byla poprvé zavedena v prosinci roku 2016 a měla za cíl snížit množství nevykázaných firemních příjmu za pomoci elektronických pokladen, které automaticky zasílají informace finančnímu úřadu. K práci jsou využita firemní data poskytnutá Českým statistickým úřadem, která jsou analyzována metodou rozdílů v rozdílech. Pro každé ovlivněné odvětví je vytvořeno několik kontrolních skupin skládajících se z odvětví neovlivněných elektronickou evidencí tržeb, která se daným ovlivněným odvětvím co nejvíce podobají. Cílem této metodologie je otestování citlivosti výsledků na jednotlivé kontrolní skupiny. Výsledek analýzy naznačuje, že elektronická evidence tržeb způsobila zvýšení vykazovaných příjmů v odvětvích, která se vyznačují větším množstvím peněžních plateb. Tato odvětví zahrnují *Stravování a pohostinství*, *Ubytování*, a *Maloobchod*. Dalším zjištěním je, že u *Stravování a pohostinství* došlo ke zvýšení vykázaných nákladů o 16 %. Tento jev mohl být způsoben tím, že se firmy v reakci na zvýšení vykázaných příjmů rozhodly zvýšit i vykázané příjmy. Zároveň se v tomto odvětví snížil vstup a zvýšil výstup firem z trhu, a také došlo ke zvýšení vstupů firem na trh specificky v roce 2016. Nebyly zjištěné žádné jasné efekty elektronické evidence tržeb na odvětví *Velkoobchod* a *Velkoobchod, maloobchod a opravy motorových vozidel.* Výsledky analýzy by ovšem měly být brány s ohledem na omezení vyplývající z použitých dat, a zejména na náročnost sestavení kontrolních skupin, které by byly skutečně srovnatelné s ovlivněnými odvětvími.

Klasifikace JEL	D22, D24, E26, K34,
Klíčová slova	elektronická evidence tržeb, daňová regulace, rozdíly v rozdílech, fixní efekty

Název práce	Dopad elektronické evidence tržeb na české firmy

E-mail autora	`42559143@fsv.cuni.cz`
E-mail vedoucího práce	`matej.bajgar@fsv.cuni.cz`

Acknowledgments

The author is grateful especially to Matěj Bajgar M.Sc., DPhil, for his invaluable assistance and feedback during the analysis and writing of the thesis. Furthermore, the author would like to thank the Czech Statistical Office, which has provided access to the data necessary for conducting this research.

Typeset in FSV LaTeX template with great thanks to prof. Zuzana Havrankova and prof. Tomas Havranek of Institute of Economic Studies, Faculty of Social Sciences, Charles University.

Bibliographic Record

Varadi, Ondřej: *The impact of the electronic revenue registry on Czech firms.* Bachelor's thesis. Charles University, Faculty of Social Sciences, Institute of Economic studies, Prague. 2023, pages 89. Advisor: Matěj Bajgar M.Sc., DPhil

Contents

List of Tables

List of Figures

Acronyms

EET “Elektronická evidence tržeb”

ERR Electronic revenue registry

VAT Value-added tax

GDP Gross domestic product

IMF International monetary fund

CZK Czech koruna

EUR Euro

EU European Union

US United States

B2C Business-to-customer

B2B Business-to-business

IRS Internal Revenue Service

OECD Organisation for Economic Co-operation and Development

DiD Difference-in-differences

CZSO Czech Statistical Office

SBS Structural business statistics

NA Not Available

OLS Ordinary least squares

Chapter 1

Introduction

Improving tax compliance and reporting of businesses are common and important issues faced by governments and tax authorities in all countries. Firms have incentives to misreport earnings, and governments have to find an efficient way to de-incentivize firms from this behavior. This problem is more prevalent in developing countries, as developed countries often already have effective tax compliance policies and much more digital payments occur in their economies, which are easily monitored through payment providers and digital footprints. However, cash transactions, especially sales to final customers, are much harder to track and present far more opportunities for firms to misreport. This means that, even in a developed country, tax compliance could be improved by targeting industries characterized by being Business-to-customer (B2C) and having a high volume of cash transactions.

The Czech electronic revenue registry ("Elektronická evidence tržeb") system, introduced in December 2016, was designed to do exactly that. It aimed to improve tax compliance and increase reported firm revenues through electronic revenue registries, which are special registers that automatically send information to the tax authority each time a transaction is recorded, combined with a receipt lottery, which was supposed to incentivize customers to request a receipt upon purchase, submit it into an online system, and provide tax authorities with additional information. This combination was supposed to make discovering misreported output easier while making firms more fearful of their misreporting getting discovered, thus making them file more truthful reports. It was inspired by other electronic revenue registry systems in Europe (Croatia, Hungary, Slovakia) that have been adopted in the past and appear to have increased reported firm output (Lovics *et al.* 2019). The policy came in two waves. The first one (December 2016) included *Food and beverage service activities* and *Accommodation*, and the second wave (March 2017) added *Retail trade, except of motor vehicles and motorcycles*, *Wholesale trade, except of motor vehicles and motorcycles*, and *Wholesale and retail trade and repair of motor vehicles and motorcycles* industries. Additional waves were meant to include more industries, but they never came to be. Considering it started near the end of 2016, we expect that it would start affecting firms primarily in 2017.

This thesis will attempt to analyze and study the effects of "Elektronická evidence tržeb" (EET) on the affected industries. Specifically, the focus will be on reported output, reported purchases, and firm entry/exit. Because EET is a form of anti-tax evasion policy, we would expect the reported output of the affected industries to rise. This assumption is supported by academic papers studying tax regulation by Carrillo *et al.* (2017) and Slemrod *et al.* (2017). Naritomi (2019) also suggest that the effect of tax regulation is stronger for smaller firms with a high number of different consumers and a high volume of small ticket transactions, which supports our assumptions even more. However, these papers also highlight the fact that firms simply increased their reported purchases to mitigate the increase in reported output. Reported firm purchases will therefore also be a focus of this study, as we would expect them to increase along with output. The effect on firm entry/exit is also of interest, as Braunerhjelm *et al.* (2021), Klapper *et al.* (2006), and Scarpetta *et al.* (2002) suggest that increased tax regulation and administrative burden decrease firm entry, while findings by McGowan & Kneller (2012) and Da Rin *et al.* (2011) imply that exit rates are less affected by tax policies because already existing firms are less responsive to policy changes in this regard. We would therefore expect EET to lead to lower entry and, potentially, higher exit in the periods around its implementation. However, it is also possible that firm entry could have increased right before the start of EET. That is because firms could have feared being under suspicion from tax authorities if their reported output suddenly increased. They could have therefore chosen to re-enter the market as a new legal entity to avoid that. We will therefore look into this possible effect as well.

These expected relationships are investigated using Czech Structural business statistics (SBS) microdata from 2012-2020 provided by Czech Statistical Office (CZSO) and accessed through SafeCentre, which marks the first time this data is used for academic research. The effects will be estimated by creating difference-in-differences models combined with firm and year fixed effects for output and purchases, and industry and year fixed effects for entry/exit. The treated industries are compared to multiple control groups, which have different methods of selection, in order to get more robust results.

We find that EET is linked to increases in reported output in *Food and beverage service activities*, *Accommodation*, and *Retail trade, except of motor vehicles and motorcycles.* Furthermore, in *Food and beverage service activities*, reported purchases increased even more than reported output, which suggests that firms offset the increase in reported output by reporting more purchases. Also, the likelihood of firm entry decreased, while the likelihood of exit increased. Interestingly, 2016 was the only year where the likelihood of entry spiked and was higher. This could support the hypothesis about firms re-entering as different legal entities. We have not discovered a clear effect on *Wholesale trade, except of motor vehicles and motorcycles* and *Wholesale and retail trade and repair of motor vehicles and motorcycles* industries, which generally have a higher volume of Business-to-business (B2B) non-cash transactions.

It is important to note that the results could be influenced by multiple limitations. One of them is the imperfection of the control groups used for the Difference-in-differences (DiD) analysis, as there are no non-treated industries that are directly comparable to the treated ones in the data. Another could be the incompleteness of the data set, as some firms are not included in it, and in the case of *Wholesale and retail trade and repair of motor vehicles and motorcycles*, there is one sub-industry unaffected by EET that cannot be distinguished from the affected ones. Therefore, the results should be interpreted carefully and conservatively.

This thesis adds to the literature on tax regulation and anti-evasion policies, and their effects and effectiveness. It also adds to the literature on electronic revenue registries in developed countries and Europe, as there are not that many research papers on this topic. It also adds to the literature about Czech EET specifically, as it is the first research on the impact of the policy using firm SBS data. To the author's knowledge, this is only the second study of a European Electronic revenue registry (ERR) system using firm-level microdata, and the first study that applies difference-in-differences on such a system. Our research could also provide useful insight to policymakers designing anti-tax evasion policies.

Chapter 2

Theoretical background

This chapter introduces important concepts for the topic of the thesis. Specifically, it covers the characteristics and implementation of the Czech ERR policy, as well as the nature of the shadow economy and its connection to the policy. Throughout the thesis, ERR is used to address electronic revenue registries in general, while EET is used specifically to mean the Czech policy.

2.1 The electronic revenue registry in the Czech Republic

The Czech electronic revenue registry ("Elektronická evidence tržeb"), introduced by Act no. 112/2016 ("zákon č. 112/2016 Sb.") in December 2016, was an electronic revenue registry system with the aim of combating the shadow economy, making tax collection more effective, and eliminating unfair competition stemming from firms avoiding paying the Value-added tax (VAT). The system made it mandatory for businesses to send a data message containing information about a transaction to the tax administrator immediately after the transaction occurs, usually through an electronic revenue register. Businesses were also obligated to provide a receipt to the customer upon request, which also contained information about the transaction.[1] During its initial implementation, all types of transactions had to be reported in this way. The update to the policy in Act no. 256/2019 ("zákon č. 256/2019 Sb.") in October 2019 included a simplified offline method of reporting transactions for businesses with under 600 thousand CZK p.a. in revenues and less than 2 employees, and made reporting mandatory for cash transactions only. EET was inspired by

[1] Taking the receipt was not mandatory for the customer.

similar policies from other countries such as Hungary, Slovakia, and Austria. However, the most inspiration was taken from Croatia. Croatia introduced its electronic revenue registry system in 2012, with many similarities to the system the Czech Republic would later adopt. It was also intended to combat the shadow economy and ultimately increase the amount of taxes collected. Taxpayers also had to send information about a transaction immediately after it occurred, and customers could also verify the transaction with their receipts on a special website. Contrary to the Czech EET system, it was mandatory for the customer to take the receipt from the vendor (Leckéši 2020).

Table 2.1: Industries affected by EET

CZ-NACE	Stage	Description	Date
55	1	Accommodation	December 1, 2016
56	1	Food and beverage service activities	December 1, 2016
45.1	2	Sale of motor vehicles	March 1, 2017
45.2	-	Maintenance and repair of motor vehicles	No EET
45.3	2	Sale of motor vehicle parts and accessories	March 1, 2017
45.4	2	Sale, maintenance and repair of motorcycles and related parts and accessories	March 1, 2017
46	2	Wholesale trade, except of motor vehicles and motorcycles	March 1, 2017
47	2	Retail trade, except of motor vehicles and motorcycles	March 1, 2017

Table 2.1 shows the stages of the adoption of EET and affected industries. The system was to be rolled out in four phases. Each phase would affect different types of businesses differentiated by CZ-NACE, a classification number used to sort firms by the area of their economic activities. However, the rollout of the last two phases was postponed due to the COVID-19 pandemic. The third and fourth phases are therefore not included in the table. In the first phase, which was in effect from December 2016, EET applied to *Accommodation*, and *Food and beverage service activities*. The second phase started in March 2017 and included *Retail*, *Wholesale*, and *Wholesale and retail trade and repair of motor vehicles and motorcycles* (*Maintenance and repair of motor vehicles* was the only sub-industry unaffected by the policy). The third and fourth phases were supposed to include a large number of additional industries. However, after the COVID-19 pandemic, the policy was terminated by the newly elected government, and by the start of 2023, EET was shut down altogether.

The costs needed to initiate EET were predicted to be 370 million CZK for building the IT infrastructure of the system with an additional 170 million CZK yearly to maintain, and 130 million CZK for other expenses to support the policy, one of them being a receipt lottery, which, using monetary and other prizes, incentivized consumers to report their transactions by turning in receipts. There were also costs incurred on the affected firms, as they had to purchase special electronic revenue registers and maintain them. All in all, the system was expected to raise collected tax revenue by approximately 4,9 billion CZK from *Retail* and *Wholesale*, and 0,8 billion CZK from *Accommodation* and *Food and beverage service activities* (Ministry of interior of the Czech Republic 2016). After a year since its implementation, around 50 000 businesses from *Accommodation* and *Food and beverage service activities*, and around 105 000 from *Retail* and *Wholesale* were registered in the system (Ministry of interior of the Czech Republic 2018).

Considering the aim and the characteristics of EET, it is reasonable to expect that businesses from the affected industries would report more output. In response to having to report higher output and therefore pay higher taxes, some taxpayers may have also increased their reported purchases to reduce their tax liabilities. If that were to be true, we would observe an upward trend both in reported firm output and purchases after the implementation of the policy, especially in the first year.

The introduction of EET could have also affected firm entry and exit. Firm exit could have increased as some firms may have chosen to exit the market due to increased costs and oversight by public authorities. Fewer businesses could also have chosen to enter the market under the policy, which would lead to lower firm entry after its introduction. Firm entry and exit could have also been affected by EET even before its implementation. Some entrepreneurs may have closed their existing businesses and started as a new legal entity to avoid suspicion if their reported output would suddenly increase. In that case, we could observe an upward trend in firm entry just before but possibly also shortly after EET. This thesis will therefore analyze the effects of EET on reported firm output and purchases after its introduction, and on firm entry and exit both before and after its implementation.

2.2 Shadow economy

The shadow economy is a part of the economy that includes all economic activities hidden from the tax authorities.[2] There are several different reasons why an individual or a firm would choose to operate in the shadow economy. One of them is simply the allure of avoiding paying taxes on income, another could be regulatory-related (strict and/or poorly designed regulation). The economic activity could also be illegal under the jurisdiction, so it is rational to operate in the shadow economy (Medina & Schneider 2018). This thesis will be mainly focused on legal economic activities with monetary and regulatory incentives affecting the choice of being in the shadow economy.

Due to its nature of being more informal and having little to no official submitted reports, it is difficult to estimate the size of the shadow economy, and the impact it has on reported output, collected taxes, and the economy as a whole. Rais *et al.* (2015) estimated that the size shadow economy in the European Union (EU) was up to 2 billion EUR (approximately 14,3 % of the official economy) in 2014, while in the Czech Republic, its size was 15,1 % of the national Gross domestic product (GDP) in the same year. Firms in the industries affected by EET are probably not entirely in the shadow economy, as they are formally registered and pay some taxes and only part of their activity is in the shadow economy. However, it makes sense for the Czech government to try to reduce the size of this specific part of the shadow economy, as it could increase the country's tax revenue and GDP. For the analysis, it is crucial to understand which sectors of the economy are affected the most by the shadow economy and how EET was designed to combat it.

A substantial portion of the "legal" part of the shadow economy is made up of sectors characterized by having a higher amount of smaller cash transactions. This includes *Retail*, *Accommodation*, and *Food and beverage service activities*, which were targeted by EET. Conducting business via cash transactions makes it easier for firms to under-report their actual output and avoid paying taxes because monitoring cash transactions is much harder for the authorities than monitoring non-cash transactions, such as credit card transactions.

Awasthi & Engelschalk (2018) analyze the relationship between the level of the shadow economy and tax collection, policy complexity, and electronic payments by analyzing data from European countries. Their findings suggest that tax collection has a negative correlation with the size of the shadow economy,

[2] It includes the sale of goods and services, and also labor.

while more complexity in tax policies has a positive correlation. This implies that it is important to design policies aimed at tax collection so that they do not pose an excessive burden on the taxpayers with their complexity. Otherwise, their effectiveness could be diminished. The paper also found the level of electronic payments to be negatively correlated with the shadow economy. Higher levels of electronic payments (thus lower levels of cash payments) lead to shrinkage of the shadow economy. Focusing on cash transactions, which EET did to a large extent, could therefore be an effective way to combat and shrink the informal economy.

Considering the characteristics of the shadow economy in the Czech Republic, it is reasonable to expect that firms from industries affected by the policy could have at least partly operated in the shadow economy. *Accommodation*, *Retail*, and especially *Food and beverage service activities* industries can be characterized by having a higher amount of smaller sales to final consumers, of which a non-negligible part is made up of cash transactions. On the other hand, *Wholesale* is characterized by selling mostly to other businesses and having fewer cash transactions. In theory, there should therefore be fewer firms operating in the shadow economy in *Wholesale* compared to the other affected industries, which could make the effects of EET smaller in this industry.

It is possible that the shadow economy in Czechia shrunk as a result of the policy. However, Djankov *et al.* (2002) argues that countries with higher regulation of entry have larger informal economies. Analyzing the impact of EET on the shadow economy itself is beyond the scope of this thesis.

2.3 "Last mile" problem of VAT

VAT is a consumption tax imposed at each stage of the production chain. As a particular good or service moves up the chain, some value gets added to the product at every stage. This ensures that each firm that is a part of this chain gets taxed on the individual value added to the product and creates an efficient and more equitable tax system. Transactions along the supply chain are usually B2B and leave a paper trail that can be followed. Firms have incentives to report all of their purchases to reduce their tax abilities, which include transactions with firms below them in the supply chain. This also means that there is a higher risk of getting audited when misreporting B2B transactions. This effectively makes VAT have built-in self-enforcement properties and makes it easier for tax authorities to observe transactions in the economy and tax them, as Pomeranz (2015) points out.

However, there is a problem with VAT at the end of the supply chain, where a lot of sales are B2C. Slemrod (2007) mentions that there is great difficulty in monitoring these transactions and the tax compliance of firms at the end of the supply chain. Because customers are not obliged to report their transactions with a firm and receive no benefit from doing so, it is much easier for businesses to misreport B2C transactions and governments have to utilize other forms of tax compliance policies to target them. This weakness of VAT is referred to as the "last mile" problem of VAT. Credit card sales mitigate this problem somewhat because information about the transaction can be stored and provided to authorities by payment providers. However, information about B2C cash transactions is held only by the seller and customer, therefore there is a higher opportunity for firms to falsely report their sales and a lower chance of governments finding it out.(Naritomi 2019).

Understanding the nature of VAT is important to the analysis, as EET was designed to curb this "last mile" problem of VAT by trying to make customers report these transactions at the end of the supply chain, which would allow tax authorities access to useful information and make monitoring easier. Although EET was originally concerned with both credit card and cash transactions, it should have had a greater impact on cash transactions. The difference between the effects on the two types of transactions could be analyzed as well and bring interesting insight into how specific policies and regulations affect credit card sales and cash sales. However, the data used in this thesis do not allow making such distinctions between transactions.

Chapter 3

Literature review

This chapter aims to introduce and review literature that is relevant to the analysis. The first section concerns tax compliance of businesses, and how it is affected by different tax reporting policies. The second section covers the relationship between taxes, tax regulation, and firm entry and exit. The third section looks at literature analyzing the effects and effectiveness of EET, as well as ERRs in other countries. The final section discusses the contribution of this thesis to already existing literature.

3.1 Tax compliance and reporting policies

Understanding tax compliance of firms and how governments and tax authorities design and utilize tax reporting policies is important. The papers on this topic are mainly focused on third-party reporting policies, as they are quite common in developed countries and are being adopted by developing countries. Third-party reporting works on the basis of collecting information about transactions obtained from an unbiased third party (thus the name third-party reporting) that is not directly associated with the transactions and using these reports from third parties to monitor firm reporting compliance. This takes advantage of the fact that third parties have fewer incentives to create false reports. In most cases, the third party is the payment settlement provider (banks, credit card companies, PayPal) for the transactions themselves. Although EET was not a third-party reporting policy, the literature concerning this topic is still relevant, as the aim of such policies is the same - to improve tax compliance.

Carrillo *et al.* (2017) look at firm misreporting in the context of taxes, the positives and possible pitfalls of utilizing third-party information, and how companies respond to tax monitoring in Ecuador. The paper stresses the importance of verifying taxpayer self-reports against third-party reports, as firms that were informed by tax authorities about discrepancies in their tax reports increased their reported revenues. On the other hand, this went hand in hand with an increase in reported expenses, meaning that taxable income did not increase as much.[1] A large number of businesses also did not respond at all to the discrepancies notifications, which shows the importance of having functional and credible tax enforcement system, otherwise the positive effects of using third-party information will be diminished. This, however, applies primarily to developing countries, whereas the Czech Republic is a developed country with fairly solid enforcement capabilities.

Slemrod *et al.* (2017) investigate Form 1099-K, a tax form issued by the United States' Internal Revenue Service (IRS) and introduced in 2011 to curb tax evasion. Banks and third-party payment settlement entities such as PayPal send this form to the IRS and businesses if the gross payments received by the firm through the settlement entity exceed 600 USD in a given year. This provides the IRS with useful information about entities, as well as makes small businesses more likely to truthfully report their receipts and income. The authors state that the number of reported receipts increased by up to 24% among certain businesses after the form was implemented. However, reported expenses also increased by as much as 13%, which could suggest that the firms most affected by the policy chose to increase their reported expenses to offset the increase in reported revenue. The authors also note that there was not a significant effect on the aggregate level, which could be caused by large businesses. This is because large businesses, which have a great impact on aggregate numbers, already reported their gross receipts close to the amount reported on Form 1099-K, because they were already being monitored due to their size before the policy was implemented. Although this paper focuses on only credit card payments, it gives insight into using third-party information to combat tax evasion in a developed country, and how it affected reported revenues and expenses.

Naritomi (2019) analyzes an anti-tax evasion policy in Brazil that utilized a receipt verification system and financial incentives to induce customers to turn in their receipts to the tax authorities. The program was used in São

[1] The increase in expenses was by up to 96 cents for every dollar of increased revenue.

Paulo and included tax rebates and lottery prices to participating customers. Similarly to EET, customers could input codes from their receipts on a website to verify the transactions and also file a complaint against a firm. The purpose of this system was to increase the value of requesting and submitting a receipt, as normally there is little incentive to take a receipt, and many customers may choose not to do it. The paper investigated the impact of this policy on reported revenue and expenses in *Retail* by comparing it to *Wholesale*.[2] These sectors were chosen for comparison because they are similar in nature, but usually have different final customers. The author argues that the system should have had a greater impact on *Retail* when compared to *Wholesale* because *Retail* is characterized by having individuals as consumers. Reported revenue increased by on average 21% over four years in *Retail* as a result of the system, although reported expenses have increased as well. All in all, the increase in revenue was on average greater than the increase in expenses. The effect on revenue was also stronger for smaller firms with a high number of different consumers and a high volume of small ticket transactions. This makes sense, as misreporting firms with high volumes of transactions to customers should be at a higher risk of getting caught. Also, after a firm received a complaint, its number of reported receipts and reported revenue increased. These results suggest that implementing a third-party verification system with financial incentives for participants could increase tax revenue. However, it is also important to take into account the costs associated with such a system, and whether they are not higher than the expected increase in collected taxes. Although Brazil is quite different from the Czech Republic, their anti-tax evasion policy has similar characteristics to EET and provides valuable insight into the topic.

3.2 Tax, regulation, and firm entry/exit

Another important aspect to analyze is the relationship between taxes, regulation, and firm entry/exit. Increased regulatory tax policies lead to higher administrative costs incurred on firms, stemming from having to spend time and resources to adhere to stricter regulatory measures. This could in turn impact the decisions of businesses to enter or exit the market. There is not much academic literature about the effect of tax compliance regulatory policies on firm entry, and especially on firm exit. However, understanding the effects of

[2]Under the Brazilian classification of economic activity, *Retail* includes food services as well.

tax rates and other regulatory policies will help paint a picture of the possible effects of EET on firm entry/exit.

Braunerhjelm *et al.* (2021) analyze the effects of tax regulations and administrative burden on entrepreneurial behavior by studying data from developed Organisation for Economic Co-operation and Development (OECD) countries. Although the study is aimed more at the effect at each stage of a firm's life cycle, the authors find that the administrative burden has significant negative effects on the activity of entrepreneurs intending to enter the market, and argue that tax compliance represents a barrier of entry for starting firms. This barrier may cause potential new entrepreneurs to choose not to enter the market, thus negatively affecting firm entry. This paper does not focus on firm exit. However, the findings about the effects of administrative burden on firm entry are still valuable to the analysis.

Canare *et al.* (2019) support this theory in their analysis of the effects of cost and ease of entry on business creation. Their findings suggest that a lower cost of entry is positively associated with firm creation.

Another study of European firms by Klapper *et al.* (2006) finds that costly market entry regulations hamper the creation of new businesses in high-entry industries, while Scarpetta *et al.* (2002) identify burdensome regulations as having a negative effect on entry of small firms in multiple OECD countries.

These articles do not focus on firm exit. However, the findings about the effects of administrative burden on firm entry are still valuable to the analysis and can give us an idea as to what effect EET might have had on firm entry.

Using industry data from 19 OECD countries, McGowan & Kneller (2012) study the effects of corporate and personal income tax reforms on firm entry and exit rates. The results of the study suggest that already existing firms are less responsive to changes in tax policy than starting firms and that changes in corporate income tax rates affect entry rates, while exit rates are largely unaffected. Increases in corporate income tax rates were found to decrease entry rates. A similar conclusion about entry rates among European firms is reached by Da Rin *et al.* (2011) in their analysis of corporate income tax on European firms. Finally, Gurley-Calvez & Bruce (2008) find evidence that cutting marginal tax rates increase entrepreneurial longevity on United States (US) tax return panel data.

Although changes in tax rates are different from compliance regulatory policies, they still represent an effect on firm entry/exit stemming from higher tax-associated costs imposed on firms. It is therefore possible that the effect of EET could be similar because it incurred extra costs and administrative burdens on entrepreneurs.

3.3 Impact and effectiveness of electronic revenue registries

There are several papers analyzing the effects of ERRs in different economies, but these usually study non-European and mostly less developed countries. The author of this thesis was unable to find any literature conducting an econometric analysis of the Croatian ERR system, which EET was largely based on. However, the analyses of other ERRs are still relevant to this thesis.

In a master thesis from the Institute of Economic Studies, Besedová (2020) analyzes the effect of EET on GDP per capita, unemployment rate, and harmonized consumer price index by doing a synthetic control method analysis on OECD data. The results show that EET had a significant positive impact on GDP per capita in 2018, and statistically insignificant effects on the other two macroeconomic predictors. The author suggests that this increase in GDP per capita is caused by the shadow economy shrinking.

Another master thesis from Masaryk University examines the Czech EET system in great detail and also tries to assess the impact and effectiveness of the policy by analyzing multiple different economic indicators. This master's thesis does not conduct an econometric analysis, it simply looks at the changes and development of economic indicators over time on data obtained from various sources, including publicly available data obtained from CZSO. It also looks at and discusses statements made by politicians and web articles. The analysis concludes that collected VAT increased in most of the affected sectors, the number of newly registered firms also increased, and the number of firm exits increased after the implementation of EET. The increased amount of firm entries and exits could be partly explained by some taxpayers choosing to end their business while starting a new one to avoid possible investigations by the tax administration, due to the increase in reported revenue after EET was implemented. However, it is up to debate whether the increase in the number of firm exits was caused more by this possible strategy of some firms, or by the increased costs incurred on them because of the policy (Leckéši 2020).

It is also important to take into account the effect other ERR systems have had in other countries. Casey & Castro (2015) look at ERRs introduced in multiple developed and developing countries and analyze the effectiveness of the system using a survey of different tax administrations. The paper points out that ERRs can be truly effective only if they are part of a larger and comprehensive compliance involvement strategy, and they also require solid legislative support and the ability of the authorities to punish non-compliance. The paper also points out several flaws in ERRs stemming from their analysis. The implementation of ERRs in the surveyed countries was not associated with noticeable increases in VAT as a percentage of GDP, and the system could be circumvented by hacking the registries and altering stored information about transactions, or by simply not reporting cash sales in the registers.

Lovics *et al.* (2019) analyze the impact of ERR system in Hungary on firm turnover using a microeconomic data set that was created by linking data obtained from VAT returns, individual electronic cash registers, and individual corporate income tax returns. The authors conducted a firm and year fixed effects analysis with firm turnover as the dependent variable, but they do not utilize the DiD method. The results imply that turnover increased by a significant amount in *Retail* (23%), and *Accommodation and Food services* (35.1%) industries, especially among smaller businesses. The impact was more pronounced in the latter two sectors and was diminishing as the size of firms increased. Because EET was inspired by the Hungarian system, and Czechia and Hungary are fairly comparable countries, these results are relevant to anticipating the effects of EET. Also, this thesis will focus in part on firm output, which is similar to turnover. This is the only paper (besides this thesis) that studies ERRs in Europe using micro-data.

Bostan *et al.* (2017) look at an ERR system in Romania and the effects it had on VAT collection. They find that the implementation of ERRs had a positive effect on collected VAT, although it decreased the overall efficiency of fiscal collection. The authors argue that the diminished efficiency was caused by Romania using older types of electronic registers that store information about transactions in their memory and do not send them to the government electronically. It is also possible that some firms may alter the memory of those devices, in which case the system would be more inefficient. The Czech EET is slightly different from this, as the registers are connected to the tax authority.

It is important to consider the implications of these findings. Because the Czech Republic is a fairly developed country, the problem of having to have a solid legislative and authority framework does not apply that much for EET. The possibility of altering the stored information on the registers also does not apply in this case, because the transaction information had to be sent immediately to the tax authority, which stored this information by itself. However, the taxpayers simply not recording their transactions is a possible weakness of EET. The receipt lottery introduced to support the system was aimed at mitigating this weakness by incentivizing consumers to participate in the monitoring process. If the lottery was successful at complementing EET, revenues and VAT would be expected to increase as a result of the policy. On the other hand, if the lottery was ineffective, the effectiveness of EET and the legislative framework surrounding it could be disputed.

3.4 Contribution to existing literature

This thesis will contribute to the existing literature about tax compliance regulations and their effects on reported firm output and purchases, as well as firm entry/exit. As discussed above in this chapter, EET was a form of regulation aimed at improving tax compliance and is therefore connected to this literature. It will also contribute to literature specifically about ERRs and their effects on firms. Because papers about ERRs are often focusing on developing countries, there is a lack of literature about the effects of these systems in developed countries, and specifically in Europe. To the author's knowledge, this thesis is only the second study of a European ERR system using firm-level microdata, and the first study that applies difference-in-differences on such a policy.

Chapter 4

Data & Methodology

This chapter presents the author's hypotheses, the data used in the analysis, what econometric method will be used and why, and the limitations encountered in the analysis.

4.1 Hypotheses

Based on the discussion of EETs in Chapter 2 and on the effects of ERRs and third-party reporting policies suggested by relevant literature discussed in Chapter 3, several hypotheses can be made about the effects of EET.[1]

Hypothesis #1: Reported revenue of affected firms increased after the implementation of EET as it became harder for the firms to misreport their revenues.

Hypothesis #2: Reported expenses of affected firms increased after the implementation of EET as the firms attempted to offset the increase in reported revenue.

Hypothesis #3: Firm exit increased after (or immediately before) the implementation of EET due to the higher (effective) taxation and extra administrative costs, or due to firms intending to reopen as a different legal entity.

Hypothesis #4a: Firm entry decreased after (or immediately before) the implementation of EET due to the higher (effective) taxation and extra administrative costs.

[1] Based on review of related literature, the hypotheses have been slightly modified from the ones presented in the Thesis proposal.

Hypothesis #4b: Firm entry increased after (or immediately before) the implementation of EET due to firms intending to reopen as a different legal entity.

4.2 Data

This thesis will use SBS panel data provided by the CZSO and accessed through SafeCentre. The provided data set contains anonymized yearly microeconomic data about incorporated firms in the Czech private sector between 2012 and 2020. This is the first time Czech SBS microdata are accessed through the CZSO SafeCentre for the purposes of academic research. The panel data set is unbalanced, as not all firms have observations for every year of the observed period, and has 2 408 982 observations in total. However, because the Covid-19 pandemic in 2020 would distort the results, this year is removed from the data set, which now has 2 097 867 observations over 8 years.

There are multiple variables of different firm attributes that will be useful to the analysis. In particular, these variables will be used:

- Activity

  This categorical variable has 4 possible values. The first and most common one is ACTIVE, which means that the unit was normally active in the period. ENTRY means that a firm was active in a given period but inactive in the previous 2 years, and signalizes that the firm entered the market. EXIT means that a firm was active in a given period but inactive in the following 2 years, and signalizes that the firm exited the market. The last one, INACTIVE, represents an anomaly in the data when a firm is considered active while being shown as inactive in the firm registry. However, this affects only 7 units in the data set, and will therefore not have a significant effect on the analysis. This variable will be useful for studying firm entry and exit.

- CZ-NACE

  Each firm has a 2-digit CZ-NACE code that specifies which industry it belongs to.[2] Section A of the NACE classification, which includes *Crop and animal production, hunting and related service activities* (CZ-NACE 01), *Forestry and logging* (CZ-NACE 02), and *Fishing and aquaculture*

[2] A more disaggregated version of CZ-NACE was not provided in the data set

(CZ-NACE 03), is not included in the data set. Also, *Activities of households as employers of domestic personnel* (CZ-NACE 97), *Undifferentiated goods- and services-producing activities of private households for own use* (CZ-NACE 98), and *Activities of extraterritorial organizations and bodies* (CZ-NACE 99) are not included in the data set. The absence of these industries should not pose a problem because they are arguably not very relevant to the analysis. However, the fact that we can only distinguish industries by the 2-digit form of CZ-NACE presents a much greater limitation. That is because the *Maintenance and repair of motor vehicles* sub-division (CZ-NACE 45.2), which was not affected by EET, cannot be distinguished from the other sub-divisions in the *Wholesale and retail trade and repair of motor vehicles and motorcycles* (CZ-NACE 45) industry.

- Value of Output

  Value of output (referred to simply as *output* in the remainder of the thesis) measures the value of the total output of a firm in thousands of Czech koruna (CZK). It is the sum of net turnover, change in the stock of goods, income from product-related subsidies, and capitalized output. The value of purchases of goods and services purchased for resale is then subtracted from this sum (Eurostat 2021). While it would be preferable to directly observe net turnover, output is highly correlated with turnover and can, thus, serve as its close proxy.

- Total purchases of goods and services

  Total purchases of goods and services (referred to simply as *purchases* in the remainder of the thesis) is the total amount of goods and services purchased in a given year in the form of expenses and current assets in accounting (Eurostat 2021). It is a suitable representation of a firm's operating expenses. Similarly to the Value of output, it is also reported in thousands of CZK.

- Value added

  Value added is the difference between the Value of output and Total purchases of goods and services and represents net operating income adjusted for depreciation, amortization, and employee benefits.

- Number of employees and self-employed persons

This variable (further called simply as the *number of employees*) is equal to the average number of employees and self-employed in a given period and represents the labor inputs of a firm. It also includes working owners, family workers, and outworkers. It will be used for choosing the control groups for the fixed effects DiD analysis because it can be used as an indicator of firm size.

- Tangible assets

  This attribute refers to tangible fixed assets and is also reported in thousands of CZK. Contrary to the other variables, this one is not an SBS indicator.

All of the variables measured in thousands of CZK were deflated using deflators from the OECD STAN database that are grouped by 2-digit CZ-NACE codes and have 2015 as the base year. Which type of deflators were used on each variable can be seen in Table 4.1. Doing this should avoid the spurious relationship problem caused by the tendency of economic variables to increase over time.

Table 4.1: Deflator STAN codes

Variable	Deflator
Output	PRDP
Purchases	INTP
Value added	VALP
Tangible assets	GFCP

By using the deflated variables already present in the data set and the number of individuals units, additional deflated variables were created that will be used to compare the characteristics of firms and help choose the control groups:

- Entry

  Entry is a dummy variable that signalizes whether a firm has entered the market in a particular year. Its value is set to 1 if the Activity variable says entry, or if there were no observations from the previous years and the current year was not 2012. This was done in case there were any wrongly assigned Activity values.

- Exit

  Exit is another dummy variable created in a similar way to the Entry dummy by using Activity. Firms were also described as having exited the market if there were no observations in the period following a particular year, except for 2020. This variable has a lead of 1 year. That is because if a firm exits the market in 2017, the exit will be shown in 2016 in the data. This is an important fact to consider during the analysis.

- Number of firms per industry and year.

  This value represents the number of firms in each industry in each year and is simply calculated by counting the number of unique units grouped by CZ-NACE classification and year. It can be further used to compute entry, exit, and churn rates for all industries in all years.

- Churn rate

  Churn rate represents the rate at which firms enter and exit the market. It is calculated by summing firm entry and exit rates together. Entry and exit rates at the industry level can be obtained by simply dividing the number of entries and exits by the total number of firms in a particular industry and year.

- Labor productivity

  We calculate Labor productivity by dividing value added by the number of employees. It tells us the value added produced by a single employee in a firm.

- Capital to worker ratio

  Capital to worker ratio is computed by dividing tangible assets by the number of employees and simply represents the capital intensity of firms.

- Average cost of production

  The average cost of production can be calculated by dividing purchases by output and stands for the average cost of producing a single unit of output.[3] It is another firm characteristic that will be useful for choosing control groups.

[3] In this case, one unit represents thousands of CZK.

Because some of the denominators used in formulas can have values of zero, the additional variables may not be computable for certain units. In that case, the value of the new variable is set to Not Available (NA) and will therefore be omitted from further equations.

Table 4.2: Summary statistics of industries affected by EET

Industry	**Year**	**Output**				**Purchases**				**Entries**	**Exits**
		Mean	*Median*	*Q1*	*Q3*	*Mean*	*Median*	*Q1*	*Q3*		
56	*2012*	5931	1361	392	4242	4937	1291	357	3800	630	511
	2013	5945	1278	390	4135	4917	1196	328	3723	1143	430
	2014	5741	1294	394	4175	4912	1262	358	3928	1061	489
	2015	5902	1367	372	4298	5238	1406	408	4076	962	534
	2016	5616	1224	364	4065	5062	1318	384	4085	1652	749
	2017	6381	1631	448	4962	5765	1691	462	4705	1497	733
	2018	6429	1642	444	5137	6060	1728	465	4981	1178	815
	2019	6508	1720	457	5282	6062	1813	464	5047	1131	1103
55	*2012*	11443	1435	288	6256	7829	1107	224	4714	46	146
	2013	12227	1490	311	6988	8112	1129	245	4992	170	140
	2014	12689	1454	321	6959	8374	1133	246	5072	146	130
	2015	13804	1929	424	8743	9036	1457	304	6193	158	120
	2016	13911	1779	408	8608	9165	1406	328	6240	253	155
	2017	14417	1780	429	8792	9380	1346	348	6339	350	147
	2018	13485	1630	428	8351	9082	1247	327	6093	340	172
	2019	12961	1357	421	7522	8759	1082	326	5544	401	243
45	*2012*	10811	1777	417	6615	47366	2818	603	12590	204	296
	2013	11017	1739	416	6570	50248	2830	602	12634	433	259
	2014	11507	1813	450	6715	56424	2725	596	12387	404	311
	2015	12728	1896	481	7001	65256	2958	696	13045	398	275
	2016	13187	1972	494	6995	68332	2870	664	12780	477	280
	2017	13293	1969	523	6984	69715	2902	616	12841	474	284
	2018	13146	1928	507	6767	65776	2820	623	11995	534	293
	2019	12616	1859	473	6521	63725	2774	613	12330	520	372
46	*2012*	10070	1076	184	4843	52001	2316	365	13631	3115	2237
	2013	9746	1024	187	4678	49175	2119	356	12437	6256	2681
	2014	9698	987	181	4513	47648	1960	328	11519	5721	3013
	2015	9919	1036	193	4589	47519	1987	338	11739	5229	2782
	2016	10477	1033	174	4755	49150	2113	321	12699	4410	3309
	2017	10681	1091	196	4816	49937	2064	323	12239	4817	2842
	2018	11221	1168	218	5145	51763	2120	326	12800	3077	2896
	2019	11551	1201	222	5225	52126	2130	322	12921	2823	3560
47	*2012*	10822	830	172	3043	38204	2693	631	10157	523	1207
	2013	10720	803	175	2917	37822	2590	621	9793	1715	1109
	2014	10992	813	171	2959	37780	2572	607	9621	1629	1362
	2015	11481	825	177	3034	39022	2625	615	9730	1838	1180
	2016	11404	864	193	3083	37511	2336	544	9065	2453	1390
	2017	11867	972	233	3233	37402	2538	614	9023	2968	1299
	2018	12466	978	243	3288	38281	2449	595	8874	2218	1325
	2019	12953	1001	234	3363	38564	2404	574	8733	2206	1711

Table 4.2 shows summary statistics grouped by industries and years for every industry affected by EET. Mean, median, first quartile (Q1), third quartile

(Q3) for firm output and purchases in thousands of CZK, and the number of firm entries and exits are shown.[4].

There is a noticeable jump in output and purchases in the *Food and beverages service activities* (CZ-NACE 56) industry in the year 2017, the first whole year that EET was active. There was also a large increase in the number of firm entries in 2016 with a slight decrease in 2017 and a decline back closer to the original levels in the subsequent years. Firm exit increased slightly in 2016 and the following years. This could suggest that the hypotheses about increases in firm revenue, expenses, and firm entry could be true for this particular industry.

In *Accommodation* (CZ-NACE 55), there is only a slight increase in output, purchases, and exit in 2017. It remains to be seen whether these increases are significant and caused by EET in further analysis. Firm entry increased much more in 2016 and especially in the following years. However, this could be caused by other factors causing a surge in firm entry such as the rise in popularity of AirBnb. If that were the case, the analysis of entry in this particular industry would be distorted. It is therefore crucial to be very conservative when interpreting the results.

For the remaining 3 industries (*Retail*, *Wholesale*, *Wholesale and retail trade and repair of motor vehicles and motorcycles*) affected by EET, there do not seem to be large jumps when the policy was implemented, the only exception being *Retail* (CZ-NACE 47), where firm entry increased in 2016 and 2017. Also, the output variable is much lower in these industries. That is because purchases of goods and services purchased for resale (which are relevant in these industries) are subtracted from output in SBS calculation. This greatly lowers the value of output, as the majority of sales in *Retail* and *Wholesale* consists of resold products. However, if EET had the intended effect on these industries, an increase in output should still be seen.

Firm exit increased in 2019 across all industries. However, because firm exit effectively has a 1 year lead (as discussed earlier in this section), this is most certainly caused by the start of the Covid-19 pandemic. This should not distort the results significantly, as the main focus is on the year around the start of EET

Next, the logs of output and purchases will be computed, as they will be useful for the analysis. If a variable for a particular firm has a negative or zero value, it will be instead set as NA. This is done to avoid having errors and infinite values in the data set. It will also cause firms that are inactive with

[4]Output and purchases are rounded to the nearest thousand in the table.

respect to their output and purchases to be omitted from estimations using log values. Only using active firms in this way is desirable for the analysis.

The created logs are then cleaned of outliers. This is done by calculating the difference between the maximum and minimum log value for each firm separately (does not have to be for adjacent years). The difference between these logs represents the highest percentage change a given firm has seen during the sample period. Subsequently, the highest 1% of values of the firm log differences are omitted from the analysis as outliers. This ensures that a firm that, for example, had a large log difference because of a period of inactivity paired with very active periods, will not distort the results of the analysis. The exact use of logs in the estimation will be explained in a later section of this chapter.

It is important to note that the author does not have permission to share the provided data set, and is limited on which statistics he was able to export from SafeCentre due to its rules and restrictions on data exporting.

4.3 Fixed effects difference-in-differences

This section will introduce the methodology of the econometric analysis used in this thesis to test the hypotheses on the data set provided by CZSO.

Originally, the synthetic control group method was chosen to be the primary method of analysis, as was indicated in the thesis proposal. It is a useful tool for estimating the effects of treatment on a treatment group by creating a synthetic control group that simulates the treatment group without the treatment and comparing post-treatment differences (Abadie 2021). However, this method is most effective when used on macroeconomic data with a high number of pre-treatment periods. Therefore, based on the characteristics of the firm data, the difference-in-differences method combined with fixed effects was deemed to be more appropriate.

The DiD method is used to examine the effects of treatment on a particular treatment group by comparing it to a control group unaffected by the treatment. In the case of EET being the treatment, the basic equation for the DiD model in panel data would look like this:

$$y_{it} = \beta_0 + \beta_1 treat + \beta_2 eet + \beta_3 treat * eet + \underbrace{a_i + u_{it}}_{\nu_{it}}$$

where y_{it} represents the dependent variable on which the effects of the treatment are estimated, *treat* is a dummy variable whose value is 1 if a unit is part of the treatment group, and *eet* is a time dummy equal to 1 if the observation is in the post-treatment period. For output and purchases, the *eet* period is selected to be the years 2017-2019. During this period, EET was active for all 5 treated industries. In the case of firm entry and exit, the *eet* period included 2016, as well. This is because exit has a lead of one year in the data set and because we are also interested in the effects of EET on firm entry in the periods around the implementation of the policy. β_0 is simply the intercept. The i and t indices represent the unit (firm) and time (years 2012-2019), respectively. The effect of the treatment on the dependent variable should be the estimate of β_3, the coefficient of the interaction between the two dummies. The composite error ν_{it} is composed of the time and unit specific idiosyncratic error u_{it}, and the unit specific and time-invariant unobserved effect a_i.

If the time-invariant unobserved effect a_i is correlated to our explanatory variables, it will cause our estimates to be inconsistent. Because the data is collected for the same sample each period, this correlation likely exists in the model. It is therefore crucial to control the unobserved time-invariant effects, and that can be done by using fixed effects.

Next, the equation will be combined with the fixed effects transformation. For output and purchases, the DiD equation will be combined with firm and year two-way fixed effects transformation. Controlling for firm fixed effects is important because of the unbalanced nature of a panel, and, in particular, the fact that if the smallest firms exit due to EET, this will lead to a higher output/purchases among the remaining firms, even if there are no within-firm increases. By adding the firm fixed effects, we can study purely within-firm variation in output and purchases. Although the data are deflated, controlling for year fixed effects is also important, as it will capture unobserved effects on the outcome variable that are different each year.[5] This can be achieved by including all unit (firm) dummies and all time (year) dummies except for 1 in each group, and removing $\beta_1 treat$ along with $\beta_2 eet$ from the equation to avoid multicollinearity with the added firm and year fixed effects.

In the case of firm entry and exit, industry fixed effects are used rather than firm fixed effects. Each firm has, in a sample period, either entered/exited or remained active, thus the firm fixed effect transformation would remove the

[5]This could, for example, be the overall state of the economy each year. Some years could see in total more reported output/purchases than others.

variation in entry and exit we are trying to estimate. Using industry fixed effects (adding a dummy variable for each CZ-NACE industry) is therefore more appropriate.

Below is the year and firm fixed effects DiD equation that will be used for output and purchases:

$$y_{it} = \beta_0 + \beta_1 treat * eet + \sum_{i=2}^{k} \delta_i \boldsymbol{D}_i + \sum_{t=2}^{T} \gamma_t \boldsymbol{T}_t + u_{it} \tag{4.1}$$

And here is the year and industry fixed effects DiD equation that will be used for firm entry and exit:

$$y_{it} = \beta_0 + \beta_1 treat * eet + \sum_{i=2}^{f} \delta_i \boldsymbol{F}_i + \sum_{t=2}^{T} \gamma_t \boldsymbol{T}_t + a_i + u_{it} \tag{4.2}$$

In the output and purchases version of the equation, the unobserved effect a_i got removed from the equation with this transformation and is contained in the firm fixed effects $\sum_{i=2}^{k} \delta_i \boldsymbol{D}_i$, where $\boldsymbol{D}_i$ is the firm dummy, δ_i is the coefficient, and k is the number of individual units. In the case of entry and exit, the unobserved effect a_i was not removed because there are no firm dummies that would contain in. The industry fixed effects are represented by $\sum_{i=2}^{f} \delta_i \boldsymbol{F}_i$, where $\boldsymbol{F}_i$ is the industry dummy, δ_i the coefficient, and f in the total number of industries in the treatment and control groups. $\sum_{t=2}^{T} \gamma_t \boldsymbol{T}_t$ represents the individual year dummies for 2012-2019, and are the same in both equations. γ_t is the coefficient and $\boldsymbol{T}_t$ is a year dummy in year t out of T total years. The year 2015 was chosen as the base year and was not included among the dummies. Now, the effect of the treatment in DiD is the estimate of β_1.

For calculating the estimators, most statistical software uses the within transformation, which de-means all variables by subtracting their averages across all time periods from them. This eliminates all time-constant variables from the equation, including the time-invariant unobserved effect a_i and the intercept. The estimates are then obtained by pooled Ordinary least squares (OLS).

There will be multiple different versions of the final fixed effects DiD equation because the effect of EET has to be estimated for studied dependent variables separately, and because firm fixed effects are used for output and purchases, while industry fixed effects are used for entry and exit. Below are the 4 final equations that will be used to analyze the effect of EET on chosen firm characteristics:

$$log(out) = \beta_0 + \beta_1 treat * eet + \sum_{i=2}^{k} \delta_i \boldsymbol{D}_i + \sum_{t=2}^{T} \gamma_t \boldsymbol{T}_t + u_{it} \tag{4.3}$$

$$log(prchs) = \beta_0 + \beta_1 treat * eet + \sum_{i=2}^{k} \delta_i \boldsymbol{D}_i + \sum_{t=2}^{T} \gamma_t \boldsymbol{T}_t + u_{it} \tag{4.4}$$

$$entry = \beta_0 + \beta_1 treat * eet + \sum_{i=2}^{f} \delta_i \boldsymbol{F}_i + \sum_{t=2}^{T} \gamma_t \boldsymbol{T}_t + a_i + u_{it} \tag{4.5}$$

$$exit = \beta_0 + \beta_1 treat * eet + \sum_{i=2}^{f} \delta_i \boldsymbol{F}_i + \sum_{t=2}^{T} \gamma_t \boldsymbol{T}_t + a_i + u_{it} \tag{4.6}$$

Each equation focuses on different firm characteristics of interest. $log(out)$ and $log(prchs)$ are the created logs of the output and purchases variables, and will serve to represent the effect of EET on firm revenue and expenses, respectively. Using log values was deemed more appropriate than normal level values, as exceptionally large firms could drive the results in normal values were to be used. The log values have also been cleaned of outliers. *entry* and *exit* are the created dummy variables related to firm entry/exit. The β_1 estimator is of interest here, as it will represent the effect of EET on revenue, expenses, and firm entry and exit of the treated industries.

However, it is possible that the effects of EET could have varying levels of impact in different periods. For example, it is reasonable to expect that firm output and purchases would see an increase in 2017, the first full year of the treatment, and then remain increased in the subsequent periods. On the other hand, we expect only a temporary spike or dip in firm entry and exit in the periods around the introduction of EET, and then a return to standard values. The effect could be possibly seen in the periods before EET because firms were aware of the policy in advance. It is therefore desirable to also create output and purchases models with included individual post-treatment year dummies (2017-2019) interactions with the treatment group, and entry and exit models with a 2016 dummy interaction along with the post-treatment dummies (2016-2019):

$$\begin{aligned} log(out) = \beta_0 &+ \beta_1 treat * y17 + \beta_2 treat * y18 + \beta_3 treat * y19 \\ &+ \sum_{i=2}^{k} \delta_i \boldsymbol{D}_i + \sum_{t=2}^{T} \gamma_t \boldsymbol{T}_t + u_{it} \end{aligned} \tag{4.7}$$

$$\begin{aligned} log(prchs) = \beta_0 + \beta_1 treat * y17 + \beta_2 treat * y18 + \beta_3 treat * y19 \\ + \sum_{i=2}^{k} \delta_i \boldsymbol{D}_i + \sum_{t=2}^{T} \gamma_t \boldsymbol{T}_t + u_{it} \end{aligned} \tag{4.8}$$

$$\begin{aligned} entry = \beta_0 + \beta_1 treat * y16 + \beta_2 treat * y17 + \beta_3 treat * y18 + \beta_4 treat * y19 \\ + \sum_{i=2}^{k} \delta_i \boldsymbol{F}_i + \sum_{t=2}^{T} \gamma_t \boldsymbol{T}_t + a_i + u_{it} \end{aligned} \tag{4.9}$$

$$\begin{aligned} exit = \beta_0 + \beta_1 treat * y16 + \beta_2 treat * y17 + \beta_3 treat * y18 + \beta_4 treat * y19 \\ + \sum_{i=2}^{k} \delta_i \boldsymbol{F}_i + \sum_{t=2}^{T} \gamma_t \boldsymbol{T}_t + a_i + u_{it} \end{aligned} \tag{4.10}$$

In these equations, $y16$, $y17$, $y18$, and $y19$ are dummy variables of specific years. Therefore, the β estimators of the interactions between year and treated dummies will indicate the effects of EET in each year separately. This should allow the results to be interpreted more in-depth.

Finally, to be able to test for pre-trends, a set of similar equations with year and treatment interactions for all periods will be created as well:

$$\begin{aligned} log(out) = \beta_0 + \beta_1 treat * y12 + \beta_2 treat * y13 + \beta_3 treat * y14 + \beta_4 treat * y16 \\ + \beta_5 treat * y17 + \beta_6 treat * y18 + \beta_7 treat * y19 \\ + \sum_{i=2}^{k} \delta_i \boldsymbol{D}_i + \sum_{t=2}^{T} \gamma_t \boldsymbol{T}_t + u_{it} \end{aligned} \tag{4.11}$$

$$\begin{aligned} log(prchs) = \beta_0 + \beta_1 treat * y12 + \beta_2 treat * y13 + \beta_3 treat * y14 + \beta_4 treat * y16 \\ + \beta_5 treat * y17 + \beta_6 treat * y18 + \beta_7 treat * y19 \\ + \sum_{i=2}^{k} \delta_i \boldsymbol{D}_i + \sum_{t=2}^{T} \gamma_t \boldsymbol{T}_t + u_{it} \end{aligned} \tag{4.12}$$

$$\begin{aligned} entry = \beta_0 &+ \beta_1 treat * y12 + \beta_2 treat * y13 + \beta_3 treat * y14 + \beta_4 treat * y16 \\ &+ \beta_5 treat * y17 + \beta_6 treat * y18 + \beta_7 treat * y19 \\ &+ \sum_{i=2}^{k} \delta_i \boldsymbol{F}_i + \sum_{t=2}^{T} \gamma_t \boldsymbol{T}_t + a_i + u_{it} \end{aligned} \tag{4.13}$$

$$\begin{aligned} exit = \beta_0 &+ \beta_1 treat * y12 + \beta_2 treat * y13 + \beta_3 treat * y14 + \beta_4 treat * y16 \\ &+ \beta_5 treat * y17 + \beta_6 treat * y18 + \beta_7 treat * y19 \\ &+ \sum_{i=2}^{k} \delta_i \boldsymbol{F}_i + \sum_{t=2}^{T} \gamma_t \boldsymbol{T}_t + a_i + u_{it} \end{aligned} \tag{4.14}$$

The year 2015 is chosen as the base year and its interaction with *treat* is removed from the equations. These equations will serve as a form of robustness checks to check the DiD parallel trends assumption. Ideally, the interactions in the pre-treatment years should be close to 0, and then diverge after the treatment was implemented. This would indicate that the treatment and control groups behaved similarly before EET, and started to diverge after the policy was in effect. Also, for additional robustness, clustered standard errors by industry will be used. Standard errors clustered by firm would probably be much smaller, and, if used in the interpretation, would make many results significant, even though reality may be different. If we cluster standard errors by industry, they should be generally larger and lead to more robust results.

To sum up, 12 different models will be used to analyze the effect of EET. Equation 4.3, Equation 4.7, and Equation 4.11 will be used for firm output, and Equation 4.4, Equation 4.8, and Equation 4.12 will be used for firm purchases. The impact on firm entry will be analyzed in Equation 4.5, Equation 4.9, and Equation 4.13, and on firm exit in Equation 4.6, Equation 4.10, and Equation 4.14. Also, each model will be estimated for the 5 treated industries separately. That is because each industry exhibits different characteristics, and EET could have therefore had different levels of effect on each of them.

The last important step of the fixed effects DiD analysis is choosing a suitable control group that should be as similar to the treatment group in the pre-treatment period as possible but not be affected by the treatment itself.

4.4 Control groups

Choosing an appropriate control group is no easy task due to the nature of the data set and the CZ-NACE classification. The policy affected entire industries, so it is not feasible to use untreated firms in the same industry as the treated group. Therefore, firms from industries with different CZ-NACE codes will have to be used for the control groups. However, no industries in the CZ-NACE classification are truly directly comparable to the treated industries.

To try and mitigate these issues, each observed treatment group will have three different control groups, that being a Broad control group, a Narrow control group, and an Algorithmic control group. Except for the Broad group, which will be the same for every treated group, all the chosen control groups are different for each treatment group. Below are the summaries of each control group type, and how they were created.

4.4.1 Broad control group

This control group contains the majority of CZ-NACE industries included in the data set. To create it, all industries contained in the CZSO data set are taken. Then, all industries with less than 50 active firms in the data set in any period are removed. This is to adhere to SafeCentre rules concerning data exports, as well as to remove industries with a low number of firms from the control groups. During this step, *Mining of coal and lignite* (CZ-NACE 5), *Extraction of crude petroleum and natural gas* (CZ-NACE 6), *Mining of metal ores* (CZ-NACE 7), *Mining support service activities* (CZ-NACE 9), *Manufacture of tobacco products* (CZ-NACE 12), *Manufacture of coke and refined petroleum products* (CZ-NACE 19), *Remediation activities and other waste management services* (CZ-NACE 39), *Water transport* (CZ-NACE 50), *Air transport* (CZ-NACE 51), *Public administration and defence; compulsory social security* (CZ-NACE 84), *Residential care activities* (CZ-NACE 87), *Libraries, archives, museums and other cultural activities* (CZ-NACE 91), and *Activities of membership organisations* (CZ-NACE 94) industries are removed. Additional industries that could prove problematic in the analysis are removed as well. *Electricity, gas, steam and air conditioning supply* (CZ-NACE 35), *Water collection, treatment and supply* (CZ-NACE 36), *Sewerage* (CZ-NACE 37), and *Waste collection, treatment and disposal activities; materials recovery* (CZ-NACE 37) industries are removed because they are highly regulated and often in the public sphere. *Land transport and transport via pipelines* (CZ-NACE 49), *Postal and courier activities* (CZ-NACE 53), *Education* (CZ-NACE 85), *Human health activities* (CZ-NACE 86), and *Social work activities without accommodation* (CZ-NACE 88) are also removed because they are largely in the public sphere. Finally, finance-related industries *Financial service activities, except insurance and pension funding* (CZ-NACE 64), *Insurance, reinsurance and pension funding, except compulsory social security* (CZ-NACE 65), and *Activities auxiliary to financial services and insurance activities* (CZ-NACE 66) are omitted, as they are usually heavily regulated by the government. The final Broad control group contains 52 different industries and is the same for all treatment groups.

4.4.2 Narrow control group

This control group is created from the Broad control group and is different for every treated CZ-NACE industry. Industries were selected for this group based on their characteristics and similarities with the treated groups.

For *Wholesale and retail trade and repair of motor vehicles and motorcycles* (CZ-NACE 45), *Manufacture of motor vehicles, trailers and semi-trailers* (CZ-NACE 29) is included in the control group because it is related to motor vehicles. *Other manufacturing* (CZ-NACE 32) is included because it concerns the manufacturing of products, that are partly sold to customers at the end of the supply chain. Finally, *Repair and installation of machinery and equipment* (CZ-NACE 33) and *Repair of computers and personal and household goods* (CZ-NACE 95) industries are included because they are concerned with repairs of products. Overall, this Narrow control group includes 4 industries.

For *Wholesale trade, except of motor vehicles and motorcycles* (CZ-NACE 46), industries that sell products and services primarily to other businesses are selected. These include *Other manufacturing* (CZ-NACE 32), *Warehousing and support activities for transportation* (CZ-NACE 52), *Computer programming, consultancy and related activities* (CZ-NACE 62), *Other professional, scientific and technical activities* ((CZ-NACE 74), *Employment activities* (CZ-NACE 78), and *Office administrative, office support and other business support activities* (CZ-NACE 82). Overall, this Narrow control group includes 6 industries.

For *Retail trade, except of motor vehicles and motorcycles* (CZ-NACE 47), industries that are related to the sale of products and services to final customers are selected. These include *Other manufacturing* (CZ-NACE 32), *Publishing activities* (CZ-NACE 58), *Creative, arts and entertainment activities*

(CZ-NACE 90), *Repair of computers and personal and household goods* (CZ-NACE 95), and *Other personal service activities* (CZ-NACE 96). Overall, this Narrow control group contains 5 industries.

For *Accommodation* (CZ-NACE 55), industries related to accommodation activities and housing are selected. These include *Real estate activities* (CZ-NACE 68),*Rental and leasing activities* (CZ-NACE 77), *Travel agency, tour operator and other reservation service and related activities* (CZ-NACE 79), and *Services to buildings and landscape activities* (CZ-NACE 81). In addition, *Sports activities and amusement and recreation activities* (CZ-NACE 93), *Repair of computers and personal and household goods* (CZ-NACE 95), and *Other personal service activities* (CZ-NACE 96) are included because they provide service activities. Overall, the Narrow control group for *Accommodation* consists of 7 industries.

For *Food and beverage service activities* (CZ-NACE 56), industries that provide service activities are selected for the control group. The chosen industries are *Telecommunications* (CZ-NACE 61), *Travel agency, tour operator and other reservation service and related activities* (CZ-NACE 79), *Creative, arts and entertainment activities* (CZ-NACE 90), *Sports activities and amusement and recreation activities* (CZ-NACE 93), *Repair of computers and personal and household goods* (CZ-NACE 95), and *Other personal service activities* (CZ-NACE 96). Overall, the Narrow control group for *Food and beverage service activities* consists of 6 industries.

4.4.3 Algorithmic control group

Similarly to the Narrow control group, this control group is created from the Broad control group and is different for each affected CZ-NACE industry. It is created via Mahalanobis distance (Mahalanobis 1936), which is characterized by the following equation:

$$D_{ij} = \sqrt{x_i' V^{-1} y_j}$$

where x_i and y_j are observations with k variables, V is a $k * k$ covariate matrix describing the variance and covariance of all variables, and D_{ij} is the Mahalanobis distance. It measures the multivariate distance between two observations. The advantage of the Mahalanobis distance as a measure of multivariate distance is that it does not depend on used units in the variables, and it takes into account the covariance between each variable. Essentially, Mahalanobis distance tells us how far are two values of a certain variable in standard deviations from each other, and generalizes it for cases with multiple variables.

For these reasons, it is used to create an Algorithmic control group of industries based on pre-treatment values of multiple firm characteristics variables. The variables used for distance matching are Labor productivity, Capital to worker ratio, Average cost of production, Employees, and Churn rate. First, we compute the median of these variables for each period and calculate the mean of the median values in 2012-2015 (the pre-treatment period) for each CZ-NACE industry separately. Next, we calculate the Mahalanobis distance between these pre-treatment means for each treated group (separately) and industry in the Broad control group. For each treated group, 5 industries with the smallest Mahalanobis distance are chosen for the control group. This should, in theory, help us create a control group where the industries have somewhat similar characteristics in the pre-treatment period as the treated industry. However, this approach is not perfect, as industries are selected based on median values from each period and some industries could have been selected for the control group even though they may not be truly similar to the treatment group.

For *Wholesale and retail trade and repair of motor vehicles and motorcycles* (CZ-NACE 45), the selected industries are *Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials* (CZ-NACE 16), *Manufacture of furniture* (CZ-NACE 31), *Publishing activities* (CZ-NACE 58), *Travel agency, tour operator and other reservation service and related activities* (CZ-NACE 79), and *Sports activities and amusement and recreation activities* (CZ-NACE 93).

For *Wholesale trade, except of motor vehicles and motorcycles* (CZ-NACE 46), the selected industries are *Manufacture of food products* (CZ-NACE 10), *Construction of buildings* (CZ-NACE 41), *Motion picture, video and television programme production, sound recording and music publishing activities* (CZ-NACE 59), *Creative, arts and entertainment activities* (CZ-NACE 90), and *Repair of computers and personal and household goods* (CZ-NACE 95).

For *Retail trade, except of motor vehicles and motorcycles* (CZ-NACE 47), the selected industries are *Construction of buildings* (CZ-NACE 41), *Publishing activities* (CZ-NACE 58), *Travel agency, tour operator and other reservation service and related activities* (CZ-NACE 79), *Creative, arts and entertainment activities* (CZ-NACE 90), and *Sports activities and amusement and recreation activities* (CZ-NACE 93).

For *Accommodation* (CZ-NACE 55), the selected industries are *Manufacture* *of beverages* (CZ-NACE 11), *Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials* (CZ-NACE 16), *Manufacture of furniture* (CZ-NACE 31), *Rental and leasing activities* (CZ-NACE 77), and *Sports activities and amusement and recreation activities* (CZ-NACE 93).

For *Food and beverage service activities* (CZ-NACE 56), the selected industries are *Manufacture of food products* (CZ-NACE 10), *Manufacture of wearing apparel* (CZ-NACE 14), *Specialised construction activities* (CZ-NACE 43), *Creative, arts and entertainment activities* (CZ-NACE 90), and *Other personal service activities* (CZ-NACE 96).

4.5 Limitations

This section describes the limitations faced during the analysis.

Firstly, the fact that only the 2-digit version of CZ-NACE is contained in the SBS microdata poses a limitation. If a more disaggregated version of CZ-NACE was available, *Maintenance and repair of motor vehicles* sub-division (CZ-NACE 45.2) could be removed from the treated group and could potentially serve as a control group. The effect of EET could be also examined more thoroughly on different sub-divisions. This could be a possible extension of the analysis in the future.

Secondly, creating a control group that closely matches the pre-treatment trends of the treated groups is a difficult task. That is because the CZ-NACE classification separates industries into quite distinct groups that often do not share many similarities. This limitation is attempted to be mitigated by using multiple control groups, and by estimating models that allow us to conduct analyses of pre-treatment trends. However, it is crucial to closely check the parallel trends assumption of DiD and to make conservative conclusions about the results.

Thirdly, the number of pre-treatment periods in the data set was not particularly high. If data from additional years before 2012 was available, it could be used to create more effective control groups and to study the pre-treatment trends in more detail. This could also be a possible extension of the analysis.

Lastly, because only incorporated firms are included in the data set, self-employed people ("živnostníci" in Czech) are missing from it. This can distort our results somewhat, as there are many self-employed businesses in, for example, the *Food and beverage service activities* (CZ-NACE 56) industry. Also, EET was in large part aimed at these businesses.

Chapter 5

Results

This chapter presents the results of the analysis. The Narrow control group is used for the interpretation of the results, while The Algorithmic and Broad control groups serve to check the similarity in trends and results across different control groups (results for the latter two control groups are shown in the Appendix). The Narrow control group is our methodologically preferred control group because we manually selected it based on similar characteristics to the treated industries. The Broad control group contains the majority of CZ-NACE industries and many of them are characteristically different from the treated industries. The industries in the Algorithmic group should be more similar to the treated industries compared to the Broad group, and represent an important robustness check for the results obtained from the Narrow control group.

This approach should help achieve more robust results considering the limitations we face. As each of the 5 industries affected by the EET is rather different and may have been differently affected by EET, we discuss the results for each of these industries in a separate subsection.

5.1 *Food and beverage service activities* (CZ-NACE 56)

The estimates of the interactions of the treated group with year dummies (Equation 4.11, Equation 4.12, Equation 4.13, Equation 4.14) are plotted in Figure 5.1 to allow for easier interpretation of the results and pre-treatment trends. The values on the y-axis represent the percentage change relative to 2015 when comparing the treated and control groups. The whiskers coming

Figure 5.1: *Food and beverage service activities*: Narrow

from each point represent the 95% confidence intervals. If the value 0 is inside a whisker, it essentially means that the estimate is not significantly different from 0 at the 5% significance level.

Relative to 2015, Output was declining in the pre-treatment period, then sharply increased in 2017 and remained elevated. Purchases started increasing slightly in 2016, but saw the biggest jump in 2017 and subsequently stayed at this elevated level. The results for entry and exit, which can be interpreted as the change in the likelihood of a firm entering or exiting the market, are less clear because they varied quite a bit in the pre-treatment period. However, entry still has a temporary spike in 2016, and then a decline in subsequent years. Exit slightly increased at the beginning of and during the treatment period, although the increase does not appear to be particularly large at first glance. Now, let us compare this to the trends observed in the other control groups (see Figure A.1 and Figure A.2). Output has higher pre-treatment values relative to 2015, which could make the interpretation of the actual trends more difficult and is most likely caused by the control groups being imperfect. However, we still see a noticeable jump from 2017 onward. Purchases appear to have similar trends in the other control group, as they increased drastically in 2017 and remained elevated. Entry was also declining and had a positive spike in 2016. Exit also saw a slight increase in the post-treatment period.

The results of Table 5.1 can be examined to analyze the effects only in the treatment period (i.e. relative to all pre-treatment years). While the overall increase in Output of 7% after the start of EET is not statistically significant, the increase of 9% in 2017 is. This suggests that our Hypothesis #1 could be true because reported firm output appears to have increased after EET, although only slightly. This could mean that EET was not as effective as expected at increasing reported firm output. However, the results could be influenced by only incorporated firms being in the data set or by possible imperfections of the control group.

Purchases increased by over 16% and all the results for individual years and the treatment period itself are statistically very significant. These findings suggest that EET is linked to increases in reported firm purchases. This would go in line with Hypothesis #2 of firms wanting to offset the increase in reported output by increasing reported purchases as well.

The likelihood of a firm entering the market decreased by 2 percentage points after EET. The decrease is statistically significant along with the decreases in 2017-2019. This could support Hypothesis #4a that EET had a negative effect on firm entry, as firms may have chosen not to enter the market because of the increased regulation. 2016 is the only year where the change in the likelihood of firm entry was positive, with an over 1 percentage point increase. This could, on the other hand, support Hypothesis #4b that firm entry increased around the implementation of EET, as firms entered the market as a new legal entity because of fear of causing suspicion by having sudden spikes in reported output.

The likelihood of firm exit increased by close to 1 percentage point, and this value is statistically significant. All of the years saw statistically significant increases except 2019. The largest statistically significant increase was almost 1 percentage point in 2016. This could support Hypothesis #3 of EET causing firms to exit the market because of increased regulation and administrative costs.

If we compare the results to the results from the other control groups (see Table A.1 and Table A.2), we can see that they suggest similar effects of the treatment. Output seems to have increased slightly overall, although the result is statistically significant only in the Broad control group. But the increase in 2017 is still statistically significant relative to all control groups. Purchases increased quite a bit relative to all control groups, and all the results are statistically significant. Although the absolute change in entry is overall statistically insignificant in the two groups, there are statistically significant positive spikes in 2016 and 2017, and then a sharp decline. This could suggest that the likelihood of firm entry was indeed greater around the start of EET. The results for exit are similar, as the likelihood of firms exiting increased overall, and the results are statistically significant.

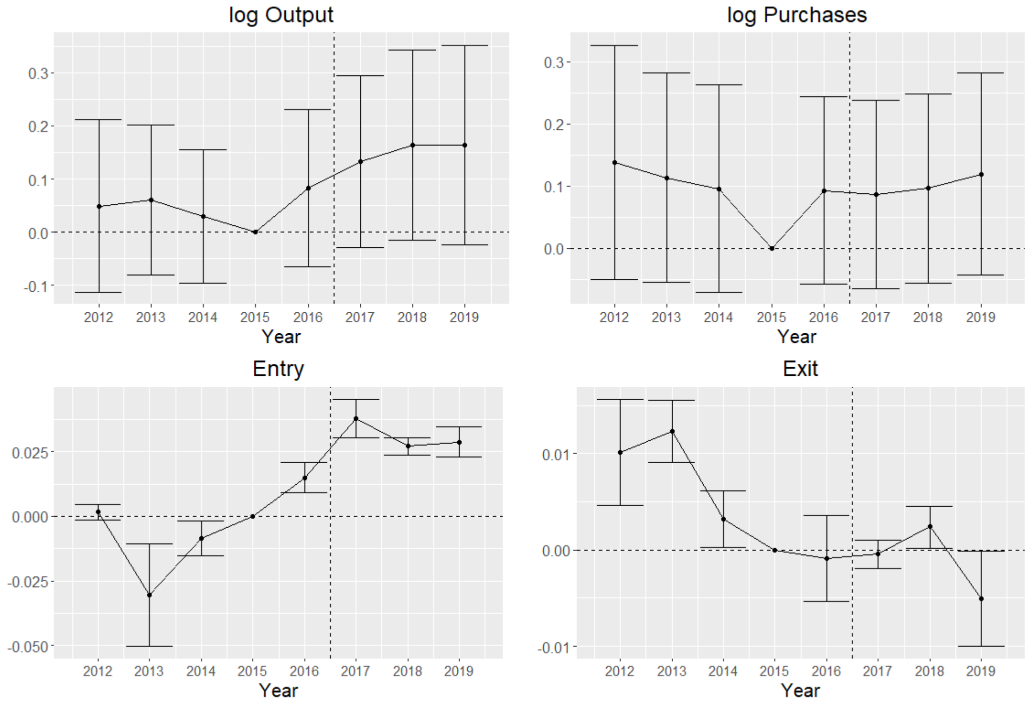
Table 5.1: *Food and beverage service activities*

	Narrow control group							
	log Output		log Purchases		Entry		Exit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treat × eet	0.0749		0.1678**		-0.0199**		0.0084**	
	(0.0454)		(0.0340)		(0.0051)		(0.0022)	
treat × y2016						0.0157		0.0095*
						(0.0081)		(0.0034)
treat × y2017		0.0935*		0.1699***		-0.0117.		0.0048***
		(0.0330)		(0.0259)		(0.0056)		(0.0006)
treat × y2018		0.0658		0.1728**		-0.0350***		0.0091**
		(0.0476)		(0.0401)		(0.0044)		(0.0023)
treat × y2019		0.0625		0.1601**		-0.0428***		0.0099
		(0.0628)		(0.0422)		(0.0058)		(0.0053)
Fixed-effects								
firm	Yes	Yes	Yes	Yes				
year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
industry					Yes	Yes	Yes	Yes
Observations	129 239	129 239	128 957	128 957	133 781	133 781	133 781	133 781
R^2	0.82120	0.82121	0.83390	0.83390	0.01321	0.01397	0.00444	0.00445
Within R^2	0.00615	0.00618	0.00531	0.00532	0.00482	0.00558	0.00157	0.00158

Clustered (industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

5.2 *Accommodation* (CZ-NACE 55)

Figure 5.2: *Accommodation*: Narrow



By looking at Figure 5.2, we can see that Output increased in post-treatment years and remained elevated. However, the increase started gradually in 2016, meaning that the true effect could be less clear to identify. Similarly to Output, Purchases saw an increase relative to the base year. However, the increase was very slight, the pre-treatment period has larger values relative to 2015 compared to the post-treatment trend, and 2015 was lower compared to all other periods. This makes interpretation of the results more difficult and the causal effects of EET less clear. Entry had large increases from 2016 onward. However, this could be caused by factors unrelated to EET, such as the AirBnb boom, and the pre-treatment periods having a drop compared to the base year. The results should therefore be interpreted very conservatively, and it is likely that they do not represent the true effect of EET. There does not seem to be a clear effect on exit, which was higher in the pre-treatment period and did not seem to increase after EET. Exit could be influenced by external factors as well. When we compare these trends to the ones in the other control groups (Figure A.3 and Figure A.4), we see that the trends in output and purchases look similar, with a very slight increase in Output. Entry and exit also have somewhat similar trends, as entry sharply increased from 2016 onward, and the effect on exit is less clear.

Next, let us analyze the results in Table 5.2. Output increased significantly by almost 11%, and all years are significant too (the largest increase is 12% in 2019). There does not seem to be a significant difference in Purchases. On the other hand, entry increased by over 3 percentage points overall, and each year is statistically significant. Exit decreased by less than 1 percentage point overall, and each period is statistically significant. However, as discussed above, we have to be very careful while interpreting these values. If the rise of AirBnb caused more firms to enter the market and made existing firms less likely to exit, it could drive the results of the analysis, thus making the interpretation of effects of EET on firm entry and exit in *Accommodation* problematic. These external factors could also influence output and purchases. A possible imperfection of the control group could also make the interpretation of the results more difficult. Relative to the other control groups (Table A.3 and Table A.4), there are statistically significant increases in output in 2017, although the overall effect seems to be insignificant. The overall effect on Purchases was also statistically insignificant, however, 2017 had statistically significant increases in both control groups. This could suggest that EET had some effect on reported purchases in this industry. However, this effect is unclear and difficult to estimate because of the statistical insignificance of the results in the Narrow control group and observed trends in graphs. Entry increased significantly relative to all control groups, however, as mentioned before, the results cannot be properly interpreted because of possible external factors, and the true effect of EET on entry is thus unclear. The effects on exit are also unclear and hard to interpret, as there is an increase relative to the Algorithmic control group, while there is no significant effect compared to the Broad control group.

Table 5.2: *Accommodation*

	Narrow control group							
	log Output		log Purchases		Entry		Exit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treat × eet	0.1086*		0.0220		0.0366***		-0.0075**	
	(0.0329)		(0.0265)		(0.0031)		(0.0019)	
treat × y2016						0.0242**		-0.0073*
						(0.0046)		(0.0024)
treat × y2017		0.0887*		0.0078		0.0470***		-0.0068***
		(0.0263)		(0.0250)		(0.0055)		(0.0010)
treat × y2018		0.1197*		0.0187		0.0363***		-0.0040*
		(0.0351)		(0.0266)		(0.0016)		(0.0017)
treat × y2019		0.1200*		0.0420		0.0379***		-0.0115**
		(0.0388)		(0.0290)		(0.0023)		(0.0027)
Fixed-effects								
firm	Yes	Yes	Yes	Yes				
year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
industry					Yes	Yes	Yes	Yes
Observations	322 328	322 328	333 793	333 793	351 388	351 388	351 388	351 388
R^2	0.83844	0.83844	0.83347	0.83347	0.00773	0.00775	0.00374	0.00374
Within R^2	0.00140	0.00140	0.00125	0.00126	0.00571	0.00573	0.00148	0.00149

Clustered (industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

5.3 *Retail trade, except of motor vehicles and motorcycles* (CZ-NACE 47)

Figure 5.3: *Retail trade, except of motor vehicles and motorcycles*: Narrow

In Figure 5.3, there seems to be an increase in Output in post-treatment, while the values in pre-treatment periods seem to be close to the base year. The effect on Purchases is less clear, as the base year has lower values compared to almost every other period. There is a spike in 2017 for firm entry, while other periods are closer to the base year. This spike could support the hypothesis of increased entry around EET. Firm exit was decreasing from 2017 onward. However, the pre-treatment values seem to vary quite a bit from the base year. These results should be, therefore, interpreted conservatively. These trends are less clear when we compare them to the other control groups (Figure A.5 and Figure A.6). Although Output seemed to increase relative to the Algorithmic control group, the trend in the Broad case is a bit strange, as all periods are quite lower than the base year. The values in post-treatment years were still higher compared to periods before 2015, but we will have to interpret the results carefully. The effect on Purchases is unclear, as the base year is quite lower compared to other periods. There seems to be a very slight increase in 2017, but the effect of EET is still unclear. There seems to be a slight increase in entry and a slight decrease in exit in 2017 compared to all control groups.

In Table 5.3, we see that the overall effect of 10% on Output is narrowly statistically insignificant. However, the effects in 2017 (10%) and 2018 (9%) have a p-value of less than 0.1, they are therefore significant at the 10% significance level. The effects on Purchases, entry, and exit are all insignificant except for exit in 2019. Entry in 2017 is quite higher than in the other periods, however, it is still insignificant. When looking at results from other control groups (Table A.5 and Table A.6), we can see that there was a significant increase in output overall and in all periods compared to the Algorithmic control group. On the other hand, when compared to the Broad group the results are statistically insignificant. This could still imply that EET caused an increase in output in this industry, as the Broad group is the least ideal control group. However, we have to interpret these results carefully, and the true magnitude of the effect is unclear. There is a statistically significant increase in purchases relative to the Algorithmic group, but the results are insignificant for the Broad case (although the estimates are positive). Based on this, it is unclear whether EET caused an increase in reported purchases in this industry. Entry had statistically significant increases in 2017 and 2018 compared to the Algorithmic and Broad groups. This could imply that firm entry increased around the start of EET. However, the effect is larger in the Broad case (which is the less ideal control group) and insignificant in the Narrow case. Therefore, the interpretation of the results is not clear. A similar case is firm exit, where there was a statistically significant decrease in 2017 compared to the two control groups, but insignificant compared to the Narrow group, therefore the effects of EET on exit are also unclear. The slight decrease in the likelihood of exit could be caused by imperfections in the control groups, or by possible external factors.

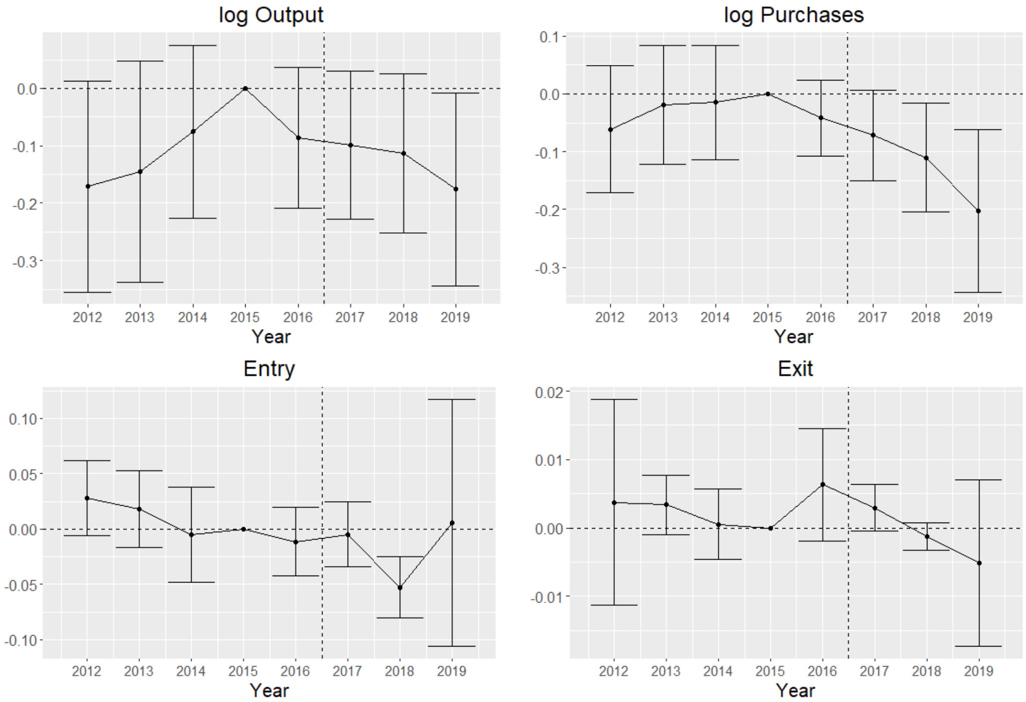
Table 5.3: *Retail trade, except of motor vehicles and motorcycles*

	Narrow control group							
	log Output		log Purchases		Entry		Exit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treat × eet	0.1034		0.0214		0.0102		-0.0039	
	(0.0532)		(0.0325)		(0.0092)		(0.0025)	
treat × y2016						0.0114		0.0010
						(0.0131)		(0.0044)
treat × y2017		0.1037.		0.0355		0.0239		-0.0038
		(0.0486)		(0.0422)		(0.0132)		(0.0033)
treat × y2018		0.0889.		0.0143		0.0022		-0.0048
		(0.0378)		(0.0279)		(0.0053)		(0.0026)
treat × y2019		0.1184		0.0120		0.0044		-0.0073.
		(0.0801)		(0.0368)		(0.0081)		(0.0034)
Fixed-effects								
firm	Yes	Yes	Yes	Yes				
year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
industry					Yes	Yes	Yes	Yes
Observations	204 191	204 191	209 647	209 647	214 814	214 814	214 814	214 814
R^2	0.86946	0.86946	0.88182	0.88182	0.00936	0.00945	0.00146	0.00148
Within R^2	0.00904	0.00906	0.00509	0.00511	0.00851	0.00860	0.00053	0.00055

Clustered (industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

5.4 *Wholesale trade, except of motor vehicles and motorcycles* (CZ-NACE 46)

Figure 5.4: *Wholesale trade, except of motor vehicles and motorcycles*: Narrow



The trends in this industry appear to be mostly unclear (Figure 5.4). Output and purchases are lower relative to the base year in all periods, and decrease as they get further from that year. This could be caused by non-ideal control groups, and the results most likely cannot be interpreted effectively. There also does not appear to be a noticeable on entry and exit, except for a decrease in entry in 2018. When combined with the trends compared with the other control group (Figure A.7 and Figure A.8), the effects are still unclear. The trends seem somewhat different with each control group, which makes the interpretation of them hard and unclear.

Looking at Table 5.4, some of the effects in post-treatment years became significant, but the overall effects still appear insignificant. Although the overall decrease in Purchases (-10%) is significant, this result is difficult to properly interpret due to the reasons discussed before. The 6 percentage points decrease in entry in 2018 is significant, but it could be also caused by other unobserved factors. The effects on output are statistically insignificant in the other control group cases (Table A.7 and Table A.8), and mostly in purchases as well. Purchases had a statistically significant increase in 2017 in the Algorithmic case. However, the estimate is quite different from the one obtained in the Narrow case, making it difficult to interpret. Entry had a statistically significant decrease in both the Broad and Algorithmic cases. This could imply that EET negatively affected entry in this industry. However, because the trends in the graphs were somewhat different, the true effect is unclear. It could have been caused by having imperfect control groups, or by other external factors. The case is similar with exit, as it had a statistically significant decrease compared to the two control groups. However, the results in the table and trends in the graphs differ from the ones in the Narrow case, which also makes the effect of EET unclear and difficult to interpret.

Table 5.4: *Wholesale trade, except of motor vehicles and motorcycles*

	Narrow control group							
	log Output		log Purchases		Entry		Exit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treat × eet	-0.0499		-0.1000*		-0.0252		-0.0013	
	(0.0511)		(0.0392)		(0.0261)		(0.0032)	
treat × y2016						-0.0207		0.0046
						(0.0180)		(0.0044)
treat × y2017		-0.0219		-0.0474		-0.0139		0.0012
		(0.0348)		(0.0248)		(0.0176)		(0.0021)
treat × y2018		-0.0357		-0.0860.		-0.0618**		-0.0030.
		(0.0494)		(0.0372)		(0.0165)		(0.0015)
treat × y2019		-0.0987		-0.1781*		-0.0036		-0.0070
		(0.0776)		(0.0639)		(0.0509)		(0.0062)
Fixed-effects								
firm	Yes	Yes	Yes	Yes				
year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
industry					Yes	Yes	Yes	Yes
Observations	497 651	497 651	508 769	508 769	529 065	529 065	529 065	529 065
R^2	0.84607	0.84609	0.87141	0.87145	0.02120	0.02183	0.00419	0.00423
Within R^2	0.00842	0.00853	0.00154	0.00184	0.00676	0.00739	0.00134	0.00137

Clustered (industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

5.5 *Wholesale and retail trade and repair of motor vehicles and motorcycles* (CZ-NACE 45)

Figure 5.5: *Wholesale and retail trade and repair of motor vehicles and motorcycles*: Narrow

Looking at Figure 5.5, there do not appear to be clear trends in this industry except for exit. Output jumped in 2016. However, the second phase of EET, which includes this industry, has not been rolled out yet this year. Also, the values in pre-treatment years relative to 2015 are of a similar magnitude. The results should, therefore, be interpreted carefully. Purchases and entry seem to have no major effects in post-treatment periods. On the other hand, some of the pre-treatment periods have a larger difference relative to the base year. There is a slight decrease in firm exit in the post-treatment period. However, the values in the pre-treatment period vary somewhat. When looking at the trends compared to the other control groups (Figure A.9 and Figure A.10), we can see that the trends in each control group are not exactly similar, which could be caused by having imperfect control groups. This makes interpreting the trends difficult.

Table 5.5 pretty much confirms the unclearness of effects of EET on Output and Purchases, as none of their estimates are statistically significant. Overall, entry increased significantly by 1 percentage point, while 2017 (1 percentage point increase) and 2018 (1 percentage point increase) were statistically significant years. However, because the pre-treatment years had lower values compared to the base year in the graphs, this makes the true effects somewhat unclear. Exit is a similar case, as it decreased overall by around a quarter of a percentage point, but it also cannot be interpreted clearly. The fact that some of the firms in this industry were not affected by EET but are still included in the regression could distort the results. There could also be issues stemming from the imperfection of control groups. When looking at the cases with other control groups (Table A.9 and Table A.10), we confirm that the effects of EET are unclear in this industry. There are statistically significant increases in output compared to both control groups (in all periods compared to the Algorithmic group). However, the true effects cannot be clearly interpreted, as the results are quite different from the Narrow case. There are also no clear and statistically significant effects on purchases except for 2017 compared to the Algorithmic control group. The effects on entry are even less clear because the effects seem to be vastly different for each control group and thus are difficult to interpret.

Table 5.5: *Wholesale and retail trade and repair of motor vehicles and motorcycles*

	Narrow control group							
	log Output		log Purchases		Entry		Exit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treat × eet	-0.0093		-0.0748		0.0111*		-0.0025*	
	(0.0458)		(0.0604)		(0.0026)		(0.0007)	
treat × y2016						0.0100.		0.0021
						(0.0041)		(0.0028)
treat × y2017		0.0340		-0.0431		0.0095		-0.0028
		(0.0348)		(0.0252)		(0.0087)		(0.0047)
treat × y2018		-0.0170		-0.0726		0.0123*		-0.0038*
		(0.0440)		(0.0642)		(0.0033)		(0.0011)
treat × y2019		-0.0503		-0.1131		0.0124***		-0.0050
		(0.0722)		(0.1020)		(0.0010)		(0.0026)
Fixed-effects								
firm	Yes	Yes	Yes	Yes				
year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
industry					Yes	Yes	Yes	Yes
Observations	94 282	94 282	95 279	95 279	97 296	97 296	97 296	97 296
R^2	0.89706	0.89708	0.90507	0.90508	0.00625	0.00626	0.00094	0.00096
Within R^2	0.00661	0.00679	0.00307	0.00319	0.00181	0.00181	0.00032	0.00034

Clustered (industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

5.6 Discussion

To summarize the results, EET seems to have had noticeable effects on some industries, especially on industries with a higher volume of small ticket transactions. In the case of *Food and beverage service activities* (CZ-NACE 56), output jumped in 2017, the first year when EET was in full effect, and remained elevated. This is expected, as this industry has a large number of cash transactions, which were the primary target of EET. However, importantly, Purchases increased even more than Output. It seems as if the firms increased their reported purchases as a response to the policy. This fact could put in question the effectiveness of EET. That is because although taxable output increased, it got offset by the increase in purchases. This ultimately means that tax revenue, which the Czech government wanted to increase with EET, could have increased by far less than the value of the extra reported output. Firm entry decreased overall, which suggests that the increased regulation caused potential new firms to not enter the market. However, in 2016 firm entry had a positive spike. Although the estimate was narrowly insignificant when compared to the Narrow control group, it still slightly supports our hypothesis that firms entered the market before the start of EET as a new legal entity to avoid possible suspicions. This could be a potential area of research for further literature. Firm exit increased slightly in this industry, which could support our hypothesis that the extra administrative burden and costs made some firms choose to exit the market.

In *Accommodation*, output increased slightly after EET, but there was no clear effect on purchases. Entry and exit had somewhat strange results, with entry having a sharp increase and exit being higher in the pre-treatment period. This is most probably influenced by the AirBnb boom, in which case the effects of EET are hard to interpret.

Output increased in *Retail trade, except of motor vehicles and motorcycles* after EET as well. The effects on other variables are, however, somewhat unclear.

There were no clear effects of EET in *Wholesale trade, except of motor vehicles and motorcycles* and *Wholesale and retail trade and repair of motor vehicles and motorcycles* industries. We expected EET to have a lesser effect on them because these industries have far fewer cash transactions when compared to the other ones. *Wholesale* is made up of mostly B2B transactions, which EET did not affect after its revision. In the case of *Sale and repair of motor* *vehicles*, there could be issues caused by only the 2-digit version of CZ-NACE being available in the data, thus firms unaffected by EET were included in the treatment group. Also, there were differences in trends when different control groups were used, which suggests that the control groups are not ideal. This means effects on these industries cannot be clearly interpreted from this analysis.

Chapter 6

Conclusion

In conclusion, this thesis studied the effects of EET on reported firm output and purchases, and firm entry and exit in multiple affected industries by using the difference-in-differences method combined with firm and year fixed effects for output and purchases, and industry and year fixed effects for entry and exit.

We found that EET had an impact on industries characterized by having a higher volume of small ticket transactions, which includes *Food and beverage service activities*, *Accommodation*, and *Retail trade, except of motor vehicles and motorcycles* industries. In these industries, we saw an overall increase in reported output after EET was implemented. Furthermore, in the *Food and beverage service activities* industry, we saw a sizeable overall increase of over 16 % in reported purchases. The likelihood of firm entry in this industry was lower by 2 percentage points overall, while the likelihood of exit was nearly 1 percentage point higher. Interestingly, the likelihood of entry in 2016 was actually higher by over 1 percentage point compared to the pre-treatment period. The effects on these variables in the other two industries are less clear. In *Accommodation*, for example, the results about firm entry and exit could be influenced by the rise of AirBnb.

We did not find clear effects of EET on *Wholesale trade, except of motor vehicles and motorcycles* and *Wholesale and retail trade and repair of motor vehicles and motorcycles*. Even though some results were statistically significant, they could not be clearly attributed to EET. This could be caused by imperfections in the control groups, and in the case of *Wholesale and retail trade and repair of motor vehicles and motorcycles* by the treated group containing a sub-industry not affected by EET.

The effects observed in the *Food and beverage service activities* industry support the findings in related literature. Carrillo *et al.* (2017) and Slemrod *et al.* (2017) also suggest that increased tax regulation increases output, but firms also increase their reported purchases as well. In the case of entry, Braunerhjelm *et al.* (2021), Klapper *et al.* (2006), and Scarpetta *et al.* (2002) linked higher tax regulation and administrative burden decrease firm entry, a relationship that could be compared to the effects found in this industry.

The effects on output in *Accommodation* and *Retail* (also *Food and beverage service activities*) support the findings of Lovics *et al.* (2019) about the effectiveness of a ERR policy in Hungary, where output increased in *Accommodation and food services*, as well as the findings of Naritomi (2019), who found that a tax-evasion policy in Brazil with certain similarities to EET caused output to increase in the *Retail* industry.

This thesis adds to the literature on tax regulation and anti-evasion policies and how they impact reported output and purchases, as well as firm entry and exit. In particular, it adds to the literature on ERRs, especially in Europe, where there is little literature doing micro-data analyses.

There are several potential extensions to this work. Firstly, an analysis of the effects on individual sub-industries using more disaggregated industry classification could be conducted, which would allow for a more detailed analysis. Secondly, the analysis could be conducted again with small non-incorporated firms being included in the data set. This would likely lead to more robust and precise results. Lastly, the analysis could be extended to other European countries with similar ERR systems (if suitable data would be available), which would allow for a more broad and robust interpretation of the effects of such policies.

Bibliography

Abadie, A. (2021): "Using synthetic controls: Feasibility, data requirements, and methodological aspects." *Journal of Economic Literature* **59(2)**: pp. 391–425.

Awasthi, R. & M. Engelschalk (2018): "Taxation and the shadow economy: how the tax system can stimulate and enforce the formalization of business activities." *World Bank Policy Research Working Paper* **(8391)**.

Besedová, M. (2020): *The Effects of Electronic Records of Sales: A Synthetic Control Method Analysis.* Univerzita Karlova, Fakulta sociálních věd.

Bostan, I., C. Popescu, C. Istrate, I.-B. Robu, & I. Hurjui (2017): "The impact of taxation of the domestic economic transactions on the vat collection through electronic fiscal devices." *Amfiteatru Economic* **19(45)**: p. 581.

Braunerhjelm, P., J. E. Eklund, & P. Thulin (2021): "Taxes, the tax administrative burden and the entrepreneurial life cycle." *Small Business Economics* **56**: pp. 681–694.

Canare, T., J. P. Francisco, & J. F. Morales (2019): "Long-and short-run relationship between firm creation and the ease and cost of doing business." *International Journal of the Economics of Business* **26(2)**: pp. 249–275.

Carrillo, P., D. Pomeranz, & M. Singhal (2017): "Dodging the taxman: Firm misreporting and limits to tax enforcement." *American Economic Journal: Applied Economics* **9(2)**: pp. 144–164.

Casey, P. & P. Castro (2015): *Electronic Fiscal Devices (EFDs) An Empirical Study of their Impact on Taxpayer Compliance and Administrative Efficiency.* International Monetary Fund.

Da Rin, M., M. Di Giacomo, & A. Sembenelli (2011): "Entrepreneurship, firm entry, and the taxation of corporate income: Evidence from europe." *Journal of public economics* **95(9-10)**: pp. 1048–1066.

Djankov, S., R. La Porta, F. Lopez-de Silanes, & A. Shleifer (2002): "The regulation of entry." *The quarterly Journal of economics* **117(1)**: pp. 1–37.

Eurostat (2021): "Methodological manual on european structural business statistics. 2021 edition." *Europa*, url: https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-21-016. [accessed 21.7.2023].

Gurley-Calvez, T. & D. Bruce (2008): "Do tax cuts promote entrepreneurial longevity?" *National Tax Journal* **61(2)**: pp. 225–250.

Klapper, L., L. Laeven, & R. Rajan (2006): "Entry regulation as a barrier to entrepreneurship." *Journal of financial economics* **82(3)**: pp. 591–629.

Leckéši, D. (2020): *Dopad zavedení Elektronické evidence tržeb v ČR [online].* Diplomová práce, Masarykova univerzita, Ekonomicko-spárvní fakulta Brno. SUPERVISOR: Ing. Bc. Jana Hvozdenská, Ph.D.

Lovics, G., K. Szîoke, C. G. Tóth, & B. Ván (2019): "The effect of the introduction of online cash registers on reported turnover in hungary." *Technical report*, MNB Occasional Papers.

Mahalanobis, P. (1936): "Mahalanobis distance." In "Proceedings National Institute of Science of India," volume 49, pp. 234–256.

McGowan, D. & R. Kneller (2012): "Tax policy and firm entry and exit dynamics: Evidence from oecd countries." *Bangor Business School Research Paper* **(12/006)**.

Medina, L. & M. F. Schneider (2018): *Shadow economies around the world: what did we learn over the last 20 years?* International Monetary Fund.

Ministry of interior of the Czech Republic (2016): "Závěrečná zpráva hodnocení dopadů regulace ria." *etržby*, url: https://www.etrzby.cz/assets/cs/prilohy/Zaverecna-zprava-hodnoceni-dopadu-regulace-RIA-k-zakonu-o-evidenci-trzeb-2016.pdf. [accessed 5.7.2023].

Ministry of interior of the Czech Republic (2018): "Závěrečná zpráva hodnocení dopadů regulace ria." *Europa*, url: `https://www.etrzby.cz/assets/cs/prilohy/Zaverecna-zprava-hodnoceni-dopadu-regulace-RIA-k-zakonu-c-256-2019-Sb-kterym-pdf`. [accessed 5.7.2023].

Naritomi, J. (2019): "Consumers as tax auditors." *American Economic Review* **109(9)**: pp. 3031–3072.

Pomeranz, D. (2015): "No taxation without information: Deterrence and self-enforcement in the value added tax." *American Economic Review* **105(8)**: pp. 2539–2569.

Rais, J., O. Klička, & A. Rod (2015): "Stínová ekonomika v České republice 01/2015." *Centrum ekonomických a tržních analýz, z. ú.* **1**.

Scarpetta, S., P. Hemmings, T. Tressel, & J. Woo (2002): "The role of policy and institutions for productivity and firm dynamics: Evidence from micro and industry data." .

Slemrod, J. (2007): "Cheating ourselves: The economics of tax evasion." *Journal of Economic perspectives* **21(1)**: pp. 25–48.

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

Appendix A

Regressions results of Algorithmic and Broad control groups

Figure A.1: *Food and beverage service activities*: Algorithmic

Figure A.2: *Food and beverage service activities*: Broad

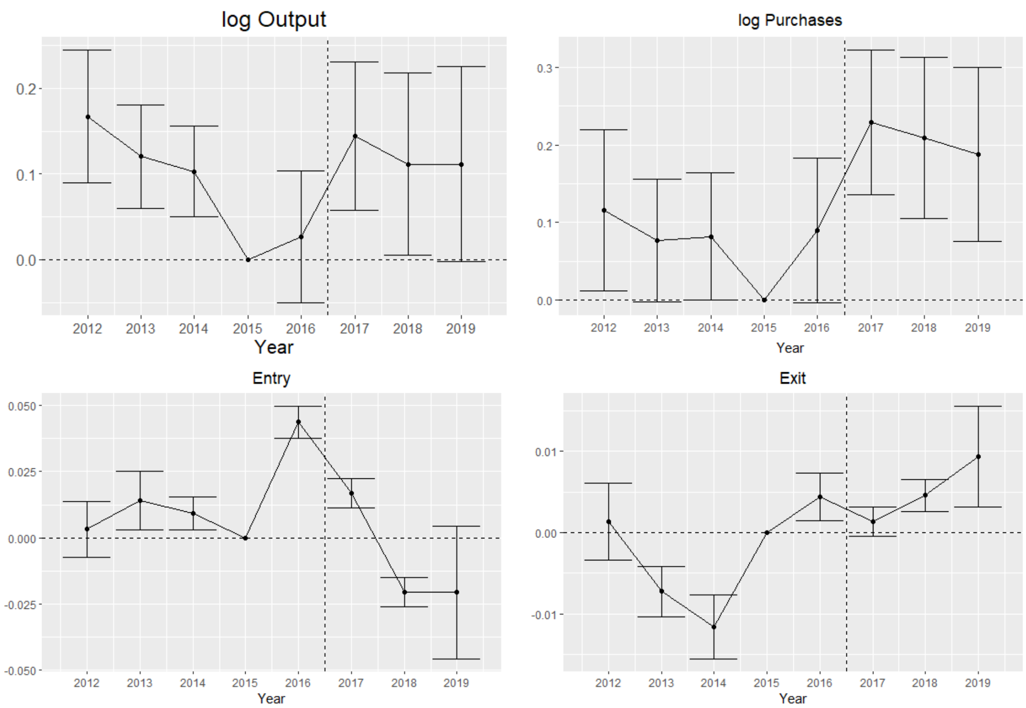


Table A.1: *Food and beverage service activities*

	Algorithmic control group							
	log Output		log Purchases		Entry		Exit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treat × eet	0.0372		0.1232***		-0.0045		0.0167***	
	(0.0244)		(0.0067)		(0.0024)		(0.0021)	
treat × y2016						0.0377***		0.0146***
						(0.0033)		(0.0019)
treat × y2017		0.0634*		0.1390***		0.0138*		0.0095**
		(0.0163)		(0.0065)		(0.0053)		(0.0017)
treat × y2018		0.0235		0.1308***		-0.0236***		0.0162***
		(0.0227)		(0.0055)		(0.0026)		(0.0018)
treat × y2019		0.0199		0.0960***		-0.0380***		0.0255***
		(0.0367)		(0.0104)		(0.0020)		(0.0034)
Fixed-effects								
firm	Yes	Yes	Yes	Yes				
year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
industry					Yes	Yes	Yes	Yes
Observations	256 565	256 565	256 969	256 969	265 627	265 627	265 627	265 627
R^2	0.84164	0.84164	0.84727	0.84728	0.00628	0.00744	0.00191	0.00198
Within R^2	0.00532	0.00536	0.00400	0.00404	0.00274	0.00390	0.00081	0.00088

Clustered (industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

Table A.2: *Food and beverage service activities*

	Broad control group							
	log Output		log Purchases		Entry		Exit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treat × eet	0.0629.		0.1444***		-0.0034		0.0094***	
	(0.0322)		(0.0257)		(0.0051)		(0.0025)	
treat × y2016						0.0372***		0.0088***
						(0.0034)		(0.0024)
treat × y2017		0.0836***		0.1633***		0.0103**		0.0058***
		(0.0238)		(0.0222)		(0.0034)		(0.0014)
treat × y2018		0.0509		0.1434***		-0.0270***		0.0090***
		(0.0345)		(0.0265)		(0.0031)		(0.0020)
treat × y2019		0.0509		0.1231***		-0.0272*		0.0138**
		(0.0412)		(0.0322)		(0.0122)		(0.0042)
Fixed-effects								
firm	Yes	Yes	Yes	Yes				
year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
industry					Yes	Yes	Yes	Yes
Observations	1 279 920	1 279 920	1 299 841	1 299 841	1 346 123	1 346 123	1 346 123	1 346 123
R^2	0.86268	0.86268	0.86663	0.86663	0.01894	0.01921	0.00647	0.00648
Within R^2	0.00450	0.00451	0.00279	0.00280	0.00271	0.00298	0.00079	0.00080

Clustered (industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

Figure A.3: *Accommodation*: Algorithmic

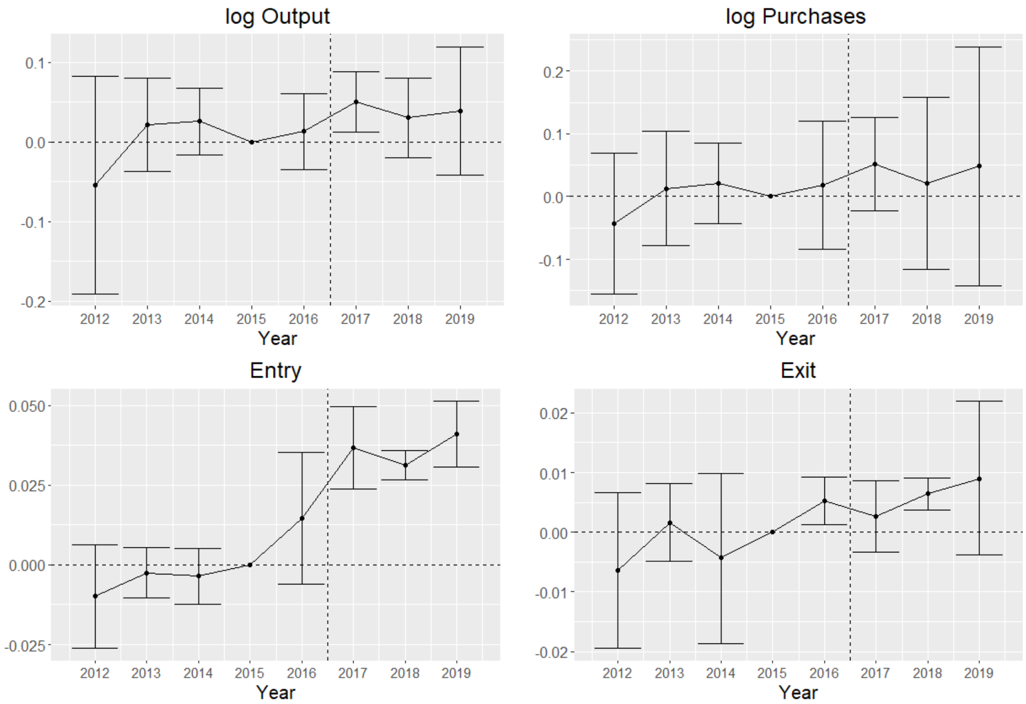


Figure A.4: *Accommodation*: Broad

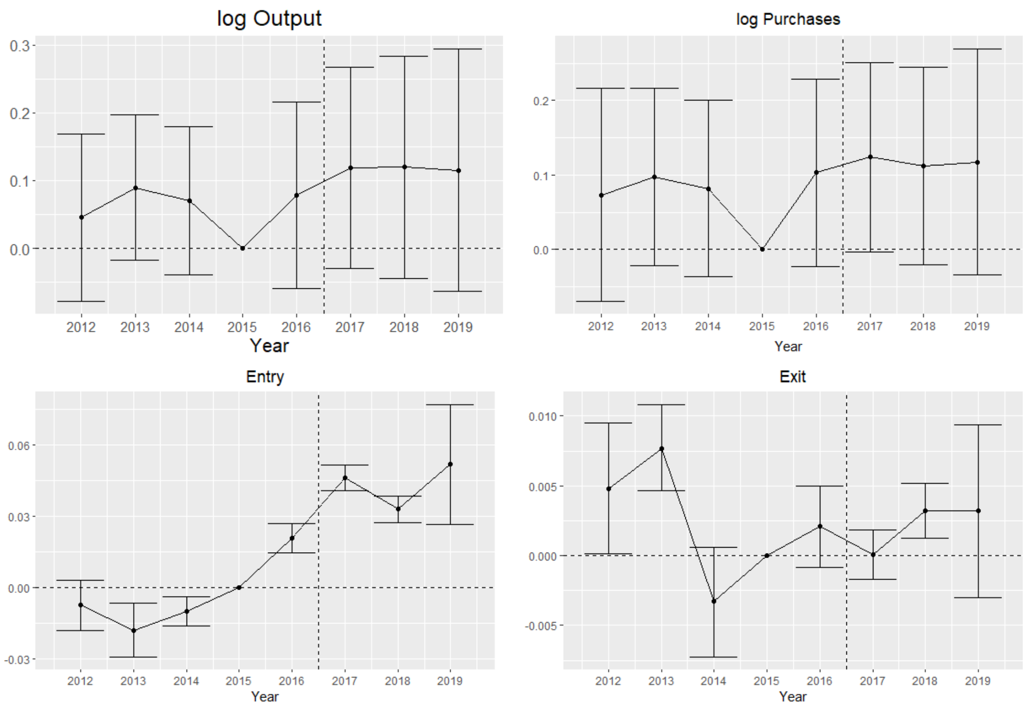


Table A.3: *Accommodation*

	Algorithmic control group							
	log Output		log Purchases		Entry		Exit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treat × eet	0.0363		0.0354		0.0355**		0.0083.	
	(0.0222)		(0.0356)		(0.0056)		(0.0034)	
treat × y2016						0.0186		0.0076
						(0.0095)		(0.0040)
treat × y2017		0.0464*		0.0464*		0.0407**		0.0050
		(0.0127)		(0.0123)		(0.0068)		(0.0030)
treat × y2018		0.0265		0.0158		0.0353***		0.0088*
		(0.0272)		(0.0410)		(0.0036)		(0.0028)
treat × y2019		0.0349		0.0434		0.0450***		0.0114.
		(0.0337)		(0.0595)		(0.0055)		(0.0055)
Fixed-effects								
firm	Yes	Yes	Yes	Yes				
year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
industry					Yes	Yes	Yes	Yes
Observations	85 062	85 062	86 325	86 325	88 258	88 258	88 258	88 258
R^2	0.88834	0.88834	0.88675	0.88675	0.01155	0.01172	0.00105	0.00106
Within R^2	0.00536	0.00537	0.00305	0.00307	0.00645	0.00662	0.00069	0.00070

Clustered (industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

Table A.4: *Accommodation*

	Broad control group							
	log Output		log Purchases		Entry		Exit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treat × eet	0.0672		0.0539		0.0473***		-0.0001	
	(0.0448)		(0.0332)		(0.0051)		(0.0025)	
treat × y2016						0.0294***		-0.0002
						(0.0033)		(0.0024)
treat × y2017		0.0676.		0.0588*		0.0548***		-0.0022
		(0.0364)		(0.0289)		(0.0034)		(0.0015)
treat × y2018		0.0692		0.0485		0.0416***		0.0009
		(0.0452)		(0.0316)		(0.0031)		(0.0020)
treat × y2019		0.0647		0.0539		0.0604***		0.0009
		(0.0549)		(0.0416)		(0.0121)		(0.0042)
Fixed-effects								
firm	Yes	Yes	Yes	Yes				
year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
industry					Yes	Yes	Yes	Yes
Observations	1 231 406	1 231 406	1 252 259	1 252 259	1 295 560	1 295 560	1 295 560	1 295 560
R^2	0.86699	0.86699	0.87016	0.87016	0.01924	0.01926	0.00648	0.00648
Within R^2	0.00438	0.00438	0.00242	0.00242	0.00286	0.00287	0.00074	0.00074

Clustered (industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

Figure A.5: *Retail trade, except of motor vehicles and motorcycles*: Algorithmic

Figure A.6: *Retail trade, except of motor vehicles and motorcycles*: Broad

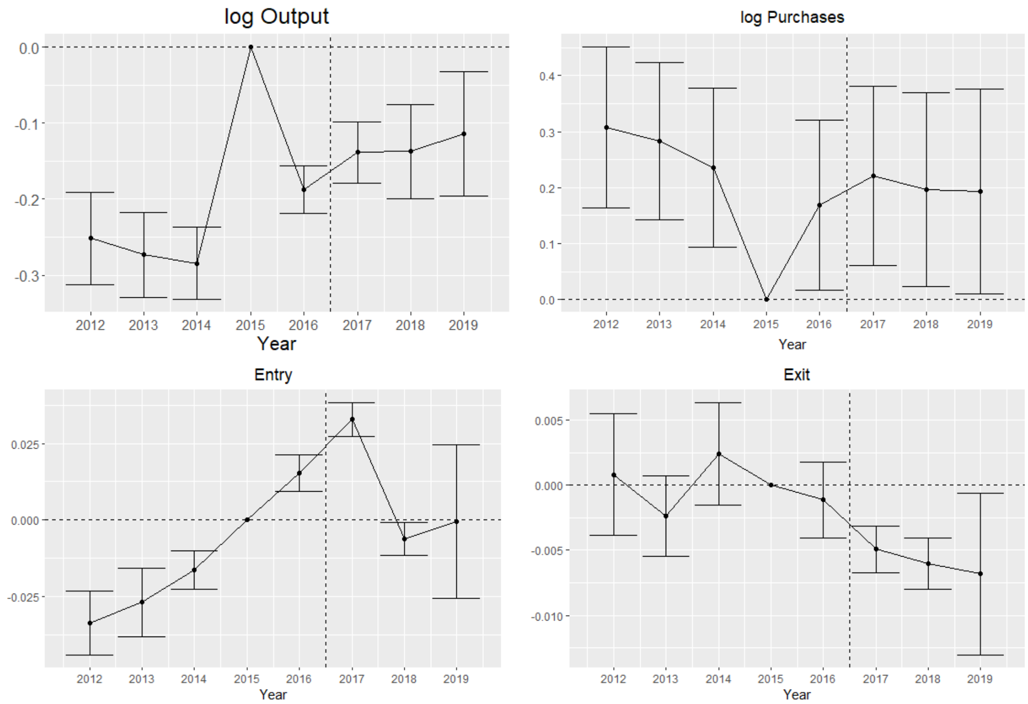


Table A.5: *Retail trade, except of motor vehicles and motorcycles*

	Algorithmic control group							
	log Output		log Purchases		Entry		Exit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treat × eet	0.1276**		0.0664**		0.0305		-0.0025	
	(0.0260)		(0.0108)		(0.0153)		(0.0031)	
treat × y2016						0.0342		-0.0008
						(0.0201)		(0.0009)
treat × y2017		0.1227**		0.0921***		0.0500*		-0.0061***
		(0.0216)		(0.0106)		(0.0194)		(0.0005)
treat × y2018		0.1035*		0.0415		0.0232.		-0.0023
		(0.0284)		(0.0206)		(0.0112)		(0.0036)
treat × y2019		0.1595**		0.0616**		0.0162		-0.0008
		(0.0344)		(0.0139)		(0.0110)		(0.0074)
Fixed-effects								
firm	Yes	Yes	Yes	Yes				
year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
industry					Yes	Yes	Yes	Yes
Observations	275 051	275 051	281 708	281 708	288 998	288 998	288 998	288 998
R^2	0.86263	0.86264	0.86742	0.86743	0.00869	0.00896	0.00119	0.00120
Within R^2	0.00681	0.00687	0.00621	0.00627	0.00602	0.00628	0.00048	0.00049

Clustered (industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

Table A.6: *Retail trade, except of motor vehicles and motorcycles*

	Broad control group							
	log Output		log Purchases		Entry		Exit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treat × eet	0.0069		0.0703		0.0290***		-0.0050*	
	(0.0253)		(0.0603)		(0.0050)		(0.0025)	
treat × y2016						0.0344***		-0.0014
						(0.0034)		(0.0024)
treat × y2017		0.0009		0.0852		0.0519***		-0.0051***
		(0.0174)		(0.0544)		(0.0034)		(0.0015)
treat × y2018		0.0003		0.0619		0.0128***		-0.0063**
		(0.0263)		(0.0619)		(0.0031)		(0.0020)
treat × y2019		0.0212		0.0614		0.0185		-0.0071
		(0.0361)		(0.0676)		(0.0121)		(0.0042)
Fixed-effects								
firm	Yes	Yes	Yes	Yes				
year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
industry					Yes	Yes	Yes	Yes
Observations	1 361 302	1 361 302	1 386 921	1 386 921	1 433 625	1 433 625	1 433 625	1 433 625
R^2	0.86156	0.86156	0.86657	0.86657	0.01798	0.01813	0.00595	0.00595
Within R^2	0.00530	0.00531	0.00257	0.00257	0.00329	0.00345	0.00070	0.00071

Clustered (industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

Figure A.7: *Wholesale trade, except of motor vehicles and motorcycles*: Algorithmic

Figure A.8: *Wholesale trade, except of motor vehicles and motorcycles*: Broad

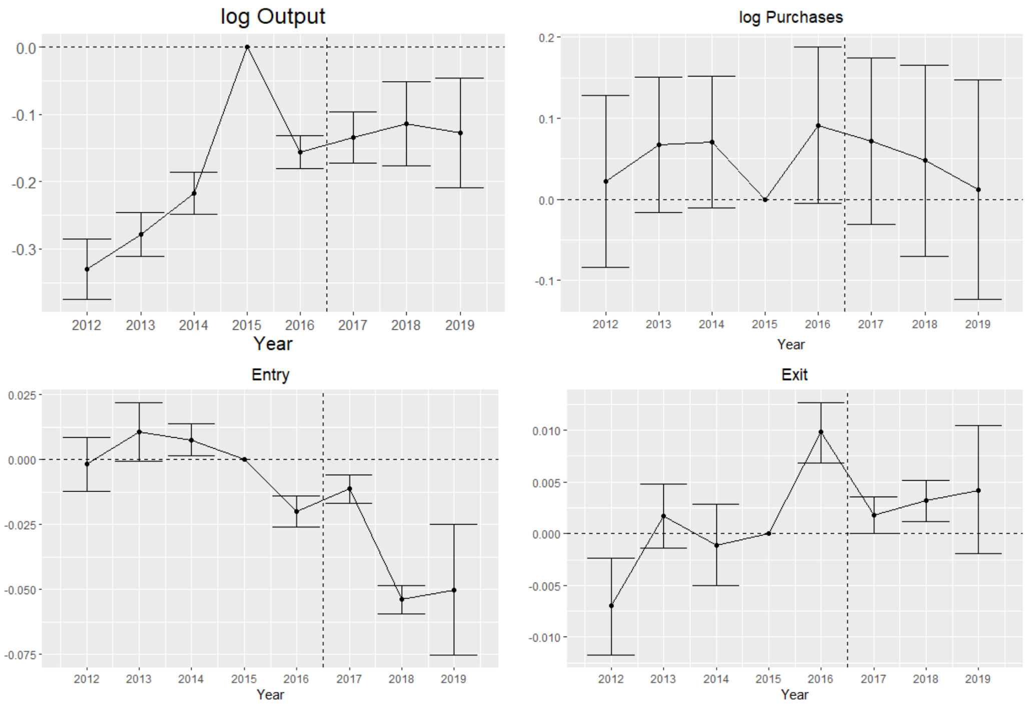


Table A.7: *Wholesale trade, except of motor vehicles and motorcycles*

	Algorithmic control group							
	log Output		log Purchases		Entry		Exit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treat × eet	0.0758		0.0179		-0.0281*		0.0107**	
	(0.0537)		(0.0265)		(0.0096)		(0.0018)	
treat × y2016						-0.0104		0.0138***
						(0.0110)		(0.0014)
treat × y2017		0.0822		0.0644.		-0.0070		0.0026*
		(0.0448)		(0.0254)		(0.0124)		(0.0010)
treat × y2018		0.0722		0.0027		-0.0390***		0.0106**
		(0.0459)		(0.0159)		(0.0056)		(0.0024)
treat × y2019		0.0722		-0.0212		-0.0540**		0.0156*
		(0.0742)		(0.0406)		(0.0112)		(0.0054)
Fixed-effects								
firm	Yes	Yes	Yes	Yes				
year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
industry					Yes	Yes	Yes	Yes
Observations	419 566	419 566	429 749	429 749	445 285	445 285	445 285	445 285
R^2	0.85185	0.85185	0.87420	0.87422	0.00791	0.00832	0.00188	0.00192
Within R^2	0.00523	0.00523	0.00077	0.00087	0.00592	0.00633	0.00092	0.00096

Clustered (industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

Table A.8: *Wholesale trade, except of motor vehicles and motorcycles*

	Broad control group							
	log Output		log Purchases		Entry		Exit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treat × eet	0.0195		0.0015		-0.0381***		0.0062*	
	(0.0265)		(0.0359)		(0.0050)		(0.0024)	
treat × y2016						-0.0243***		0.0112***
						(0.0034)		(0.0024)
treat × y2017		0.0114		0.0276		-0.0155***		0.0032*
		(0.0156)		(0.0294)		(0.0034)		(0.0014)
treat × y2018		0.0312		0.0037		-0.0581***		0.0046*
		(0.0279)		(0.0360)		(0.0031)		(0.0020)
treat × y2019		0.0165		-0.0316		-0.0543***		0.0057
		(0.0390)		(0.0452)		(0.0122)		(0.0042)
Fixed-effects								
firm	Yes	Yes	Yes	Yes				
year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
industry					Yes	Yes	Yes	Yes
Observations	1 517 697	1 517 697	1 546 964	1 546 964	1 602 260	1 602 260	1 602 260	1 602 260
R^2	0.85556	0.85556	0.86637	0.86638	0.01694	0.01728	0.00583	0.00585
Within R^2	0.00549	0.00549	0.00189	0.00193	0.00354	0.00388	0.00080	0.00081

Clustered (industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

Figure A.9: *Wholesale and retail trade and repair of motor vehicles and motorcycles*: Algorithmic

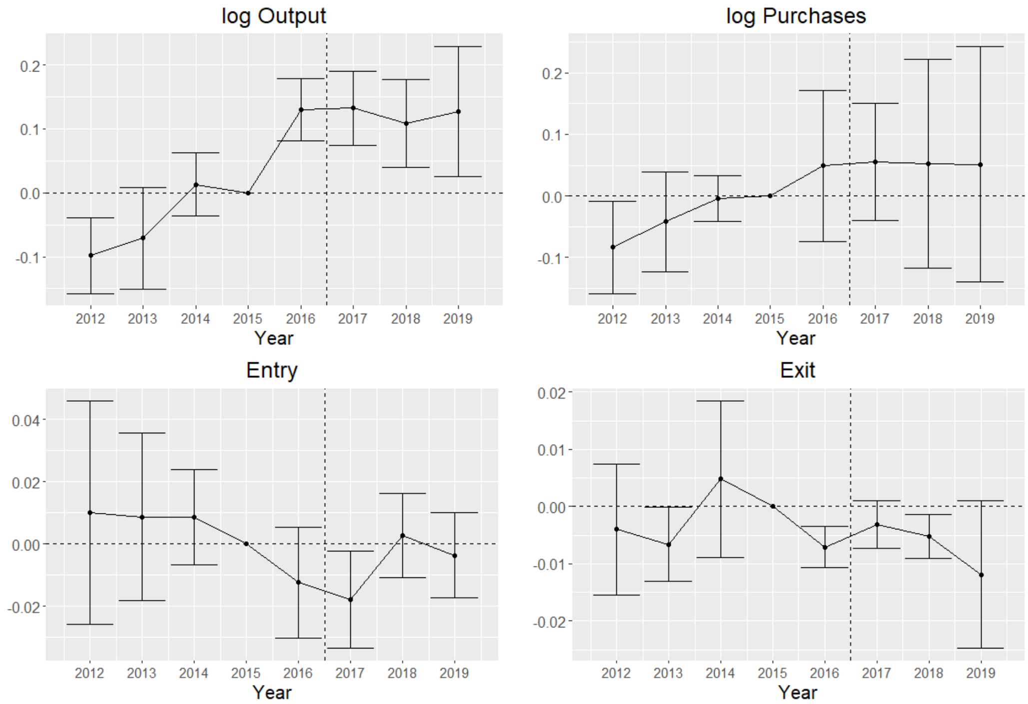


Figure A.10: *Wholesale and retail trade and repair of motor vehicles and motorcycles*: Broad

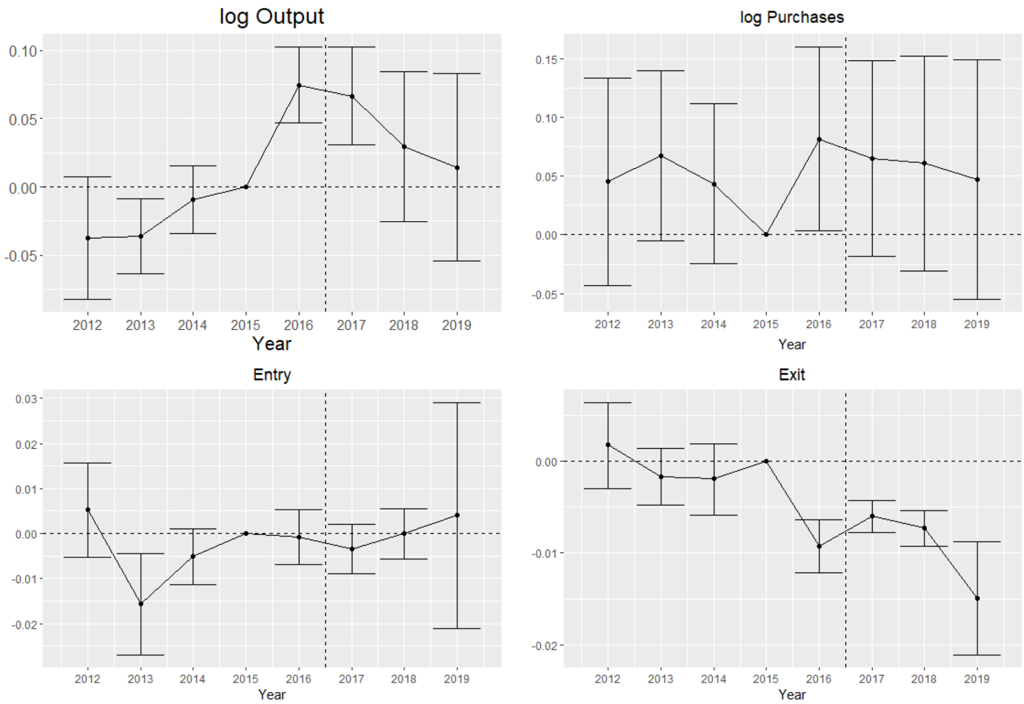


Table A.9: *Wholesale and retail trade and repair of motor vehicles and motorcycles*

	Algorithmic control group							
	log Output		log Purchases		Entry		Exit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treat × eet	0.1102*		0.0600		-0.0144.		-0.0055	
	(0.0294)		(0.0445)		(0.0071)		(0.0043)	
treat × y2016						-0.0192		-0.0057
						(0.0112)		(0.0037)
treat × y2017		0.1191**		0.0619.		-0.0246.		-0.0017
		(0.0244)		(0.0278)		(0.0111)		(0.0037)
treat × y2018		0.0960*		0.0594		-0.0042		-0.0039
		(0.0302)		(0.0514)		(0.0039)		(0.0037)
treat × y2019		0.1149*		0.0583		-0.0104.		-0.0105
		(0.0411)		(0.0610)		(0.0049)		(0.0069)
Fixed-effects								
firm	Yes	Yes	Yes	Yes				
year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
industry					Yes	Yes	Yes	Yes
Observations	125 516	125 516	127 071	127 071	129 959	129 959	129 959	129 959
R^2	0.87449	0.87449	0.88970	0.88970	0.00965	0.00978	0.00097	0.00100
Within R^2	0.00449	0.00451	0.00341	0.00341	0.00379	0.00393	0.00043	0.00046

Clustered (industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

Table A.10: *Wholesale and retail trade and repair of motor vehicles and motorcycles*

	Broad control group							
	log Output		log Purchases		Entry		Exit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treat × eet	0.0309		0.0121		0.0039		-0.0090***	
	(0.0247)		(0.0251)		(0.0051)		(0.0025)	
treat × y2016						0.0031		-0.0088***
						(0.0034)		(0.0024)
treat × y2017		0.0595***		0.0189		0.0005		-0.0055***
		(0.0164)		(0.0226)		(0.0034)		(0.0015)
treat × y2018		0.0230		0.0151		0.0038		-0.0068**
		(0.0254)		(0.0249)		(0.0031)		(0.0020)
treat × y2019		0.0078		0.0015		0.0079		-0.0145**
		(0.0343)		(0.0312)		(0.0121)		(0.0042)
Fixed-effects								
firm	Yes	Yes	Yes	Yes				
year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
industry					Yes	Yes	Yes	Yes
Observations	1 260 967	1 260 967	1 282 244	1 282 244	1 326 185	1 326 185	1 326 185	1 326 185
R^2	0.86637	0.86638	0.87149	0.87149	0.01909	0.01909	0.00648	0.00649
Within R^2	0.00439	0.00440	0.00239	0.00239	0.00266	0.00266	0.00072	0.00072

Clustered (industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*