ESSAYS ON FIRM
RESPONSES TO TAX POLICY

Annika Nivala

ESSAYS ON FIRM RESPONSES TO TAX POLICY

Annika Nivala

University of Turku

Turku School of Economics
Department of Economics
Economics
Doctoral Programme of Turku School of Economics

Supervised by

Professor Janne Tukiainen
University of Turku
Finland

Professor Kaisa Kotakorpi
Tampere University
Finland

Reviewed by

Professor Sebastian Siegloch
University of Mannheim
Germany

Cand. Polit. Thor Olav Thoresen
Statistics Norway
Norway

Dr Hanna Pesola
VATT Institute for Economic Research
Finland

Custos

Professor Janne Tukiainen
University of Turku
Finland

Opponent

Professor Sebastian Siegloch
University of Mannheim
Germany

ABSTRACT

This compilation thesis studies firm responses to tax policy. The essays use empirical methods and administrative firm-level data to study issues in public and labor economics: how do firms adjust their behavior to incentives caused by policy. The results can provide information to guide the design of public policy.

The first essay studies the effects of a first-employee subsidy in Finland. In 2007—2011 Finland provided a regional subsidy for hiring the first employee. The subsidy was supposed to encourage entrepreneurs to become employers, as over half of firms in Finland do not have employees. Using the geographic differences in the eligibility to the subsidy and full-population firm data, I study the effects on the probability of becoming an employer and several other firm outcomes. I find a zero effect on average. This seems to be explained by low take-up of the subsidy: only 2% of eligible firms that became employers used the subsidy. Descriptive evidence supports that restricting the subsidy to full-time employees and low awareness decreased the take-up and, hence, the effectiveness of the subsidy.

The second essay describes a plan for a government-run recruitment subsidy experiment. The randomized experiment is supposed to evaluate the effects of a recruitment subsidy on the probability of becoming an employer. Following the evidence from the first essay, the subsidy is not restricted to a type of employment contract and the subsidy covers a percentage of the wage costs in the first year up to a threshold. Therefore, expected take-up and effects are larger, and the results of the experiment can complement the knowledge from the first essay. The essay makes use of administrative firm-level data to plan the experiment and calculate the statistical power.

The third essay studies how firms respond to tax audits. We combine data from the Finnish tax administration on all tax audits conducted on firms in 2003—2016 to annual firm tax returns. This unique, novel data allows us to follow the tax reporting behavior of firms before and after being audited. By comparing the development of the audited firms to similar non-audited firms with similar prior development, we find that the audited firms significantly increase reported profits and revenue persistently after an audit. This means that auditing more firms could increase tax revenue in the long run, by increasing firms' tax compliance.

KEYWORDS: business taxation, firm behavior, public policy, employment subsidies, tax audits

TIIVISTELMÄ

Tämä kokoomaväitöskirja tutkii veropolitiikan vaikutuksia yrityksiin. Esseissä tutkitaan julkisen ja työn talouden kysymyksiä käyttämällä empiirisiä menetelmiä ja yritystason rekisteriaineistoja: miten yritykset reagoivat julkisen politiikan luomiin kannustimiin. Tuloksia voidaan käyttää apuna julkisen politiikan suunnittelussa.

Ensimmäinen essee tutkii ensimmäisen työntekijän palkkaustuen vaikutuksia Suomessa. Vuosina 2007—2011 Suomessa oli tarjolla alueellinen tuki ensimmäisen työntekijän palkkaamiseksi. Tuen oli tarkoitus rohkaista yrittäjiä ryhtymään työnantajaksi, kun yli puolet Suomen yrityksistä ovat yksinyrittäjiä. Tutkin tuen vaikutuksia työnantajaksi ryhtymisen todennäköisyyteen ja muihin yritysten tulemiin käyttämällä alueellisia eroja yritysten kelpoisuudessa tukeen ja koko yrityspopulaation aineistoa. Tulosten perusteella tuella ei ollut vaikutuksia keskimäärin. Tämän näyttäisi selittävän tuen alhainen käyttöaste: vain 2 % työnantajaksi ryhtyneistä yrityksistä käytti tukea. Kuvailevan analyysin perusteella rajoitus kokoaikaisiin työsuhteisiin ja alhainen tietoisuus tuesta laskivat tuen käyttöastetta ja siten tuen vaikuttavuutta.

Toisessa esseessä kuvaillaan suunnitelma valtion toimeenpanemasta rekrytointitukikokeilusta. Satunnaistetun kokeen on tarkoitus arvioida rekrytointituen vaikutuksia työnantajaksi ryhtymisen todennäköisyyteen. Ensimmäisen esseen näytön perusteella rekrytointitukea ei ole sidottu työsuhteen muotoon ja tuki korvaa osuuden palkkakustannuksista ensimmäisen vuoden aikana ylärajaan saakka. Tämän takia tuen odotettu käyttöaste ja vaikutukset ovat suuremmat ja kokeilun tulokset voivat täydentää ensimmäisen esseen tuottamaa tietoa. Esseessä käytetään yritystason rekisteriaineistoja kokeilun suunnittelussa ja kokeen tilastollisen voiman laskemisessa.

Kolmannessa esseessä tutkitaan verotarkastusten vaikutuksia yrityksiin. Yhdistämme ainutlaatuisen aineiston kaikista Suomen veroviraston tekemistä yritysten verotarkastuksista vuosilta 2003—2016 yritysten vuosittaisiin veroilmoituksiin. Näin voimme seurata yritysten veroraportointia ajassa ennen ja jälkeen verotarkastuksen. Tarkastettujen yritysten kehityksen vertaaminen samankaltaisiin samalla tavalla kehittyneisiin ei-tarkastettuihin yrityksiin paljastaa, että tarkastetut yritykset lisäävät merkittävästi ja pysyvästi raportoituja voittoja ja liikevaihtoa tarkastuksen jälkeen. Tämä tarkoittaa, että verotarkastusten lisääminen voisi tuottaa lisää verotuloja pitkällä aikavälillä lisäämällä yritysten verojen raportointia.

ASIASANAT: yritysverotus, yritysten käyttäytyminen, politiikkavaikutukset, työllistämistuet, palkkatuet, verotarkastukset

Acknowledgements

This thesis, like all scientific endeavors, was enabled by support and help from various people. My work owes to the many people that inspired and taught me along the way. I did not start with a passion for economics, nor would have imagined a career in it. However, little by little I discovered that this field is not only a powerful tool to understand and improve policy but also a rewarding field of study.

Firstly, I want to express my gratitude to my supervisors, as I had the opportunity to learn from excellent researchers. Kaisa Kotakorpi helped me through the whole PhD process, but she already supported me during the Master studies and provided me the opportunity to start the PhD with funding in her project. Her influence in my work is undeniable. Later, Janne Tukiainen started as my supervisor, providing plenty of excellent feedback, support and motivation. Previously, Heikki Kauppi was my supervisor and offered support in the academic work including teaching, and always made me believe in myself. Thank you all, for the encouragement, guidance, and faith in me during this process. Although Jarkko Harju and Tuomas Matikka are not included in the official list of supervisors, I consider them as equally important supervisor figures. They are co-authors in one of the essays, but even my first essay would not have happened without them. Thank you for providing me the opportunity for the internship at VATT that brought about the first essay, and inspired me to do research, and for the mentorship through these years.

I want to thank my pre-examiners Sebastian Siegloch, Hanna Pesola and Thor Olav Thoresen for reading my thesis and giving insightful comments. The comments helped me improve the thesis and will greatly help in advancing the papers for publication. Additional thank you for Sebastian Siegloch who agreed to act as my opponent.

My main motivation in economics is producing knowledge that is relevant for guiding policy. Hence, I am grateful for the concrete opportunities and co-operation that have made this possible while doing the PhD. A previous version of the first essay was included as a background report for the Economic Policy Council Report 2016, and the paper benefited from the discussions with and comments by Roope Uusitalo, thank you for the comments and opportunity. The second essay builds on a very exciting policy-oriented project for the Ministry of Economic Affairs and Employment, conducted with Elias Einiö and Oskari Nokso-Koivisto. The project was to plan a field experiment to study the effects of a recruitment subsidy for non-

in me, and to my brothers Joni and Teemu. Great thanks to my aunt Anne, for hosting me in my many visits to Helsinki and being interested in my work.

Last but not least, thank you Mikael. Words cannot express my gratitude for your endless support, kindness, and companionship. Thank you for your patience in listening to my troubleshooting and thinking out loud, for inspiring me with your different perspectives, and for the counterbalance to work. I am lucky and grateful to have you by my side.

September 2021
Annika Nivala

TABLE OF CONTENTS

LIST OF ORIGINAL RESEARCH PAPERS

(1) Nivala, Annika. (No) Effects of Subsidizing the First Employee: Empirical Evidence of a Low Take-up Puzzle Among Firms in Finland, *Published manuscript*.

(2) Nivala, Annika. Planning a Field Experiment: How Do Firms Respond to a Subsidy for Becoming an Employer?, *Unpublished manuscript*.

(3) Harju, Jarkko, Kotakorpi, Kaisa, Matikka, Tuomas and Nivala, Annika. Firms and Tax Audits: Evidence from Risk-Based Tax Audits in Finland, *Unpublished manuscript*.

Part I

INTRODUCTION

INTRODUCTION

This compilation thesis studies firm responses to tax policy in order to understand the consequences of policies on firms and society. Governments can use tax policy to collect revenue or to change behavior in order to, for example, correct market failures or improve equity. Regardless of the motivation, there is always a trade-off in policies: any policy has costs in addition to the benefits created. The trade-off can only be identified by knowing the causal effects of policies. Consequently, the effects of policies need to be known in order to understand whether the policies are desirable or not. In addition, understanding the mechanisms or reasons for the observed effect can provide valuable information in designing more effective policies.

This thesis consists of an introduction and three empirical essays on the effects of tax policy on firms. Tax policy is defined here rather broadly; in addition to taxes, tax policy includes subsidies, which can be considered negative taxes, and tax system features such as tax enforcement tools, which are used to ensure that the correct amount of taxes is paid. The theme of the first and second papers is a subsidy for hiring the first employee or becoming an employer. The subsidy is meant as an incentive for firms to become employers with the motivation that it could encourage job creation, and thus increase total employment. The third essay studies tax audits on firms as a tax enforcement tool. Consequently, the theme of the third paper is tax evasion and tax audit as a tool used to improve tax compliance.

The essays in this thesis aim to identify the causal effects of the tax policies studied, using extensive administrative panel data on the full population of Finnish firms and causal identification methods. The research is enabled by the existence of and access to register data on the firms. This kind of micro-level data enables the development of individual firms to be tracked over time, which is beneficial for the identification of causal effects. The second paper is an example of a methodological benchmark for causal identification, as it describes a plan for a randomized field experiment. The first and third essays make use of quasi-experimental methods. The idea is to find a group of firms that can provide a counterfactual for the firms that face a specific tax policy treatment.

The essays in this thesis contribute to the field of public economics, especially empirical public economics and firm behavior. Indeed, tax policy is at the core of public economics. The topic of tax compliance has received a lot of attention in the past 20 years and is still a topical research area. Nevertheless, there is little evidence on firm responses to business-as-usual tax audits. The topic of the first

and second essays also overlaps with labor economics. The effects of employment incentives are extensively studied. However, there is still a lack of evidence on the effects of employment incentives on firms and, especially, non-employer firms.

The rest of this chapter is organized as follows. Firstly, section 1 introduces the study of tax policy: how to evaluate the welfare effects of different tax policies using causal effects, a methodological approach to estimating the causal effects of policies, and different margins of response and the role of studying mechanisms behind the causal effects. Secondly, section 2 discusses literature that studies employment incentives and provides overviews of essays 1 and 2. Lastly, section 3 discusses literature on tax evasion and overviews essay 3. The first section sets the general framework and motivation for studying firm responses to tax policy. The second and third sections briefly describe the conceptual background and literature related to the essays in this thesis.

1 STUDYING TAX POLICY

1.1 Welfare evaluation of policies

How to improve economic welfare in society is the ultimate concern in public economics as well as many other fields in economics and, arguably, in policy design.[1] Governments can use different tools, including tax policy, in order to improve social welfare. For example, governments may want to support the creation of jobs to provide employment and economic prosperity to more people. Whatever the motive or the tool, there is always a trade-off: any tax policy can cause costs in addition to the potential benefits. Consequently, a welfare evaluation must consider the costs and benefits of the policy, in other words the efficiency of the policy.

In this dissertation I consider two tax policies: a subsidy for hiring the first employee and tax audits. First, a subsidy for hiring the first employee can increase job creation by firms that become employers because of the subsidy. Thus, it can increase the welfare of employees as well as entrepreneurs if the subsidy increases their income. However, the subsidy may cause deadweight spending, as some firms that would have become employers in the absence of the subsidy also use it. This spending must be financed either by reducing spending on other policies or by increasing government revenue through taxation, which can have further

---

[1] Although in reality the objectives of different agents in policy-making may differ from this general goal of improving economic welfare, as studied in the field of public choice.

costs through decreased incentives to work. Second, tax audits may increase tax compliance, increasing government revenue. However, increased tax compliance means higher *effective* tax rates, which, again, can have a negative effect on revenue through weaker incentives. How, then, can the ultimate welfare effects be evaluated?

A simple, unified framework for evaluating the welfare effects of policies is described in Hendren (2016), Hendren and Sprung-Keyser (2020) and Finkelstein and Hendren (2020). They argue that the causal effects of a policy are sufficient for welfare evaluation. The factor that determines the welfare impacts of any policy is called the marginal value of public funds (MVPF), which can be calculated using the causal, behavioral effects of the policy. The framework is closely related to the result in Feldstein (1999), which shows that (under some simplifying assumptions) the welfare effect of taxation can be determined by a sufficient statistic: the *elasticity of taxable income* with respect to marginal tax rates.

The MVPF is defined by the ratio of the benefits created by a policy to its effect on government spending:

$$MVPF = \frac{Beneficiaries' \ willingness \ to \ pay}{Net \ cost \ to \ government} \tag{1}$$

(Hendren and Sprung-Keyser, 2020). Consequently, the MVPF measures how much the policy increases welfare per unit of spending on it, or, as formulated in Finkelstein and Hendren (2020), the "bang for the buck".

The denominator in equation 1 is the net cost of the policy or, in other words, the causal effect of the policy on government revenue. Hendren (2016) calls this total effect on government revenue the *policy elasticity*. Consequently, it measures how much a change in the policy changes revenue – a close relative of the elasticity of taxable income, which is a sufficient statistic for the welfare effects of taxation. The total net effect of a policy on government revenue can be divided into a mechanical effect and fiscal externality (FE). This can be normalized by the spending on the policy. Thus, the normalized net cost to the government is $1 + FE$. The mechanical effect simply means the spending on the policy. In the case of the first-employee subsidy, it is the total amount of the subsidies paid to firms.

The fiscal externality arises through the behavioral responses to the policy that affect government revenue. This is determined by the causal effects of the policy. (Finkelstein and Hendren, 2020) For example, the first-employee subsidy can increase employment, and thus the wage costs of a firm, which increases payroll and income taxes related to the wage, increasing government tax revenue. In addition, the subsidy may help entrepreneurs to grow their businesses and profitability, increasing firm taxes paid on value added and profit. Consequently, the fiscal externality arises from the firm responses to the policy: employment, value added, and turnover that generate tax revenue for the government. In the case of the subsidy, the fiscal externality is negative and thus decreases the net costs of

the policy. To determine the fiscal externality, only the causal effect on the taxable income of the new employees and the firm is needed. In the case of the tax audits, the denominator consists of the costs of the tax audits (mechanical cost) and the subsequent effect on the reported taxes (behavioral effect).

The numerator in equation 1 measures the marginal benefits of the policy defined as the marginal willingness to pay (MWTP) of the beneficiaries of the policy.[2] The beneficiaries include those directly affected by the policy, but can also include those indirectly affected if the policy has externalities. The MWTP, however, does not need to depend on the causal effects of the policy. The simple case is when the policy is a cash transfer, for example a change in taxes: the marginal willingness to pay is simply equal to the amount of the cash transfer. This can, again, be normalized by the total spending on the policy. The normalized MWTP of a tax change, for example, is just one euro (or dollar). (Finkelstein and Hendren, 2020)

The above description of the marginal willingness to pay holds when i) the policy change is small, ii) the behavior of the agents targeted is privately optimal, and iii) the value of the policy equals its costs. Firstly, for large effects, the demand for the policy needs to be known to determine the MWTP. Secondly, additional assumptions about measuring welfare from behavioral welfare economics is needed if there are behavioral agents, as their welfare may differ from their willingness to pay. [3] Third, if the policy is an in-kind transfer (e.g. education, food stamps) it may not be equal to the spending on the policy, because the beneficiaries may not be willing to pay the cost of the in-kind transfer to get it. (Finkelstein and Hendren, 2020) These issues relate to measuring how much the policy increases the welfare of the recipients.

In addition, the MWTP, i.e. the marginal value created, may differ from the spending on the policy if there are market failures that the policy is able to address. For example, there may be externalities to others than the direct beneficiaries of the policy. In this case the total MWTP of all the beneficiaries needs to be measured. In addition, the direct beneficiaries may be willing to pay more than the spending on the policy if the policy can reduce the impact of market failures. Hendren (2016); Finkelstein and Hendren (2020) For instance, if the first-employee subsidy creates new jobs for people who would otherwise not be able to get jobs, the increased welfare of the new employees adds to the total MWTP. Furthermore, if the first-employee subsidy helps entrepreneurs overcome liquidity constraints and

---

[2] Here I abstract from social preferences to different groups of individuals for simplicity. However, conventional welfare weights for groups can be easily included to weight the marginal willingness to pay of different groups. This may be needed especially when a policy affects the welfare of different groups of people. Alternatively, one can calculate the MVPFs for each group separately and compare them.

[3] See e.g. Bernheim and Taubinsky (2018) and Farhi and Gabaix (2018) for a behavioral public economics welfare evaluation.

become employers, the subsidy may help them increase profits by more than the subsidy. In this case, the subsidy amount understates the MWTP for the subsidy. Consequently, the causal effects of the policy on the income of the new employees and the entrepreneur need to be known to calculate the MWTP.

To summarize, the causal effects of a policy can be used to calculate the MVPF of a particular policy that measures how effectively the policy increases welfare. Consequently, the framework allows for welfare evaluation of a policy and comparing the welfare effects of different policies. For example, the MVPFs of tax policy changes can be used evaluate the welfare effects. This allows us, for example, to measure the difference in the welfare effects of two policies or to evaluate the benefits of a policy compared to the costs of raising funds using a different policy.

Consider as an example increasing tax rates to finance a subsidy for hiring the first employee. First, the MVPF of increasing tax rates can be calculated using only the causal effect on taxable income.[4] This is because the MWTP of increasing taxes equals the tax increase, which can be normalized to one. Then, the MVPF is defined by the policy elasticity or the elasticity of taxable income. Let us assume a fiscal externality of 0.3[5] which means that one euro of mechanical tax increase decreases taxes paid by 0.3, thus the net increase in taxes is 0.7. Then, the MVPF of a tax rate change amounting to one euro is 1.43. Now, this can be compared to the MVPF of the first-employee subsidy. Unfortunately, the results I find in the study of the effects do not allow the fiscal externality to be calculated[6]. Consequently, the following calculation is not based on empirical estimates but is purely a demonstration. Let us assume that the subsidy does not decrease any market frictions, so that the benefits of the subsidy are equal to the mechanical cost of the subsidy. Then, the MWTP is equal to one. Assume that for 20% of the firms that used the subsidy the subsidy increases the taxes paid by 50%, in total by 10%. Thus, the fiscal externality is 0.1, leading to an MVPF of 1.11. With these values, the cost of increasing one more euro of tax revenue is 1.42, which can be compared to the benefit of providing a subsidy for hiring the first employee, which is 1.11. This may not necessarily mean that the subsidy is not desirable. The MVPFs only measure that the tax increase is "more expensive" than the value provided by the subsidy. However, if the welfare of the beneficiaries of the subsidy has a larger social welfare weight than the welfare of the taxed individuals, this can still be desirable.

Notably, the fiscal externality may be larger than the mechanical cost of a policy. Then, the net cost is negative. In this case, the policy is self-financing and

---

[4] With many simplifying assumptions such as no spillovers, see e.g. Saez, Slemrod and Giertz (2012) and Hendren (2016).

[5] Hendren (2016), suggests midpoints between 25 to 50%.

[6] Due to a low observed take-up of the subsidy, I cannot credibly identify the results for the firms that used the subsidy, which is needed for the MWTP.

Hendren and Sprung-Keyser (2020) define the MVPF as infinite. They also calculate the MVPFs of various policies and find that the MVPFs of many policies targeting education or children's health are, indeed, infinite. These kinds of policies are self-financing investments, and they are desirable irrespective of social values if the MWTP if positive. The MWTP can, however, also be negative. In fact, this is the case with the example above of a tax increase, in which case the cost is also negative (there is an increase in government revenue), so they cancel out (Finkelstein and Hendren, 2020). However, it is possible that the spending on the policy does not benefit anyone or even causes costs to people. This would make the MVPF negative if the costs are positive.

To conclude, the framework for calculating the MVPF of a policy is a simple way to evaluate and compare the welfare effects of different policies. The causal effect of the policy on government revenue is the key, and, in simple cases, the sufficient statistic for the evaluation. In the case of large changes, externalities, or in-kind transfers measuring the MWTP require more assumptions and, for instance, structural estimation or additional estimates of causal effects may be needed. All in all, the framework clearly points out what causal effects of policies need to be known to measure the welfare effects.

1.2   Causal identification of firm responses

This dissertation investigates the causal effects of tax policies on firms, which need to be understood to evaluate their welfare consequences. As Finkelstein and Hendren (2020) discuss, much of empirical work in economics in recent decades has concentrated on identifying causal effects. Consequently, the evaluation framework discussed above is useful and is undoubtedly motivated by the "credibility revolution" in economics, meaning to the focus on causal identification methods. The essays in this dissertation make use of the standard methodologies used for causal identification since the credibility revolution.

Angrist and Pischke (2010) discuss how the credibility revolution in empirical economics refers to a focus on a credible research design and utilizing more data. These can be complementary: often a credible design is only feasible in the presence of large amount of data, especially panel data. Data availability is key for economic research and determines what can be studied. For example, Almunia, Harju, Kotakorpi, Tukiainen and Verho (2019) discuss how research can benefit from access to administrative data. The essays in this dissertation make use of administrative panel data on the full population of Finnish firms. This data has many advantages: firms can be followed over time, the data is representative because it includes the full population, and the information is as reliable as possible because it is the official information gathered and used by government authorities. In particular, the tax return data that I use is the data that determines firms' tax

liability.

The experimental ideal or benchmark for identifying the causal effect is a randomized trial (Angrist and Pischke, 2009, Angrist and Pischke, 2010, Kremer, 2020). Random assignment of a treatment (or policy) ensures that the treatment is independent of the outcome studied. Consequently, the difference between the average outcome of the treatment group versus the control group equals the average causal effect. In fact, randomized experiments both in laboratories and in the field have been widely used in economics (Harrison and List, 2004, Kremer, 2020) because of the clean identification. Moreover, a randomized experiment works as a benchmark for additional identification methods that exploit variation that is in some sense "as good as random".

The second essay in this dissertation relates to the methodological ideal of identifying a causal effect: the essay describes a randomized field experiment studying the effects of subsidizing becoming an employer. The essay describes a plan for an experiment to be conducted by the Finnish government. At the time of writing this, the implementation and timing of the experiment are not yet certain. However, the essay provides a pre-trial plan, which is needed for transparency of the analysis of experiments to be transparent, and describes the various aspects in planning a field experiment. While the causal identification is simple, there are many aspects to consider, from the choice of the experiment's target group to the sample size calculations and designing the treatment to ensure that the experiment provides credible, relevant information on the policy.

In addition to randomized experiments, the identification of causal effects often uses so-called quasi-experiments or natural experiments (Angrist and Pischke, 2009, Angrist and Pischke, 2010). The idea is to find natural variation in a policy, for example a policy change that simulates random assignment in a specified setting or for a limited group. Essentially, the idea is to find similar, comparable groups that face a different policy. If there is no reason, other than the policy, for the similar groups to have different outcomes, the difference between the groups may identify the causal effect of the policy.

A canonical, and widely applied, example of an identification strategy in a natural experiment is the difference-in-differences method (Angrist and Pischke, 2009, 2010; Abadie and Cattaneo, 2018). This design exploits panel data and a situation where there is a policy change for a group (treatment group) and no change for a similar group (control group). The assumption behind the difference-in-differences identification is common trends; the groups would have developed similarly without the policy change.[7] In this case, the development of the con-

---

[7] Often additional assumptions of no spillovers and composition change are mentioned. Assumption of no spillovers means that the policy does not affect the control group. Assumption of no composition change means that the policy does not affect which individuals are in the groups or, in other words, individuals cannot manipulate their treatment status. These are needed for the control group to provide a counterfactual development, i.e., the development if there was no policy change.

trol group is a *counterfactual* for the development of the treatment group if there was no policy change, and the difference in the development or, in other words, change in the difference (difference-in-differences) measures the effect of the policy. There can be a difference between the groups prior to the policy change – the pre-policy difference between the groups accounts for it. All that is needed is data before and after a policy change and a group comparable to the group that experiences a policy change. Additional popular research designs utilizing natural experiments include regression discontinuity and instrumental variables, which I do not discuss here.

The first essay in this dissertation uses a standard difference-in-differences identification strategy. The essay studies the effects of a regional first-employee subsidy that was in force in Finland in 2007—2011. Consequently, there is a change in a policy, i.e. the introduction of a new subsidy for firms in the subsidy area, while the firms in the area without the subsidy are not eligible for a similar subsidy. Thus, the effect can be identified by comparing the change in employment between firms in the treatment (subsidy) area to firms in the control (no subsidy) area. The assumption is that the firms in the areas would have developed similarly without the subsidy, which I directly examine in the essay.

Extensive micro-level data can provide for a credible research design even without a policy change or when a counterfactual is not readily available. Common research designs exploit e.g. i) variation in treatment timing (event-study) or ii) matching techniques to make the treatment and control groups more comparable. First, different individuals (or units) may be exposed to a policy at different times. In these cases, a researcher may use an event study approach or so-called "staggered" difference-in-differences, described in Borusyak and Jaravel (2021). The event study approach relies on the randomness of the timing of the treatment. Then, the effect of the treatment can be identified as the common change in outcomes after the treatment for the individuals treated at different times. This is enabled by the different treatment timing, because then the common time trend can be accounted for. Staggered difference-in-differences means a difference-in-differences strategy where the treated individuals are exposed to the treatment, like a policy change, at different times. The identifying assumptions and challenges of these research designs are discussed in e.g. Borusyak and Jaravel (2021) and Goodman-Bacon (2020). Notably, the treatment does not have to be a policy change – it can be any event that affects the individual outcome and that satisfies the identifying assumptions.

It should be noted that there are important differences between the standard difference-in-differences and the event study methods. Firstly, in canonical difference-in-differences the treatment assignment or the policy is at the group level rather than the individual level. This means that the standard errors should be corrected for the correlation within the groups (Bertrand, Duflo and Mullainathan, 2004). In the event study, the assignment to treatment is at the individual level.

Secondly, the identifying assumptions are somewhat different when there is variation in treatment timing. In the standard difference-in-differences, the parallel trends assumption implies similar development over time without the treatment. In an event study without a control group, the identifying assumption is that the timing of the event is random, i.e. there is no change related to the timing of the event. Staggered difference-in-differences or an event study with a control group rely on similar development of the treatment and control groups relative to the timing of the treatment, not just over time. Many recent papers (e.g., Borusyak and Jaravel, 2021 and Goodman-Bacon, 2020) discuss the assumptions and the appropriate methodologies when using variation in treatment timing.

Secondly, matching is a common method for improving comparability between treated and non-treated groups based on observable variables (Angrist and Pischke, 2009; Abadie and Cattaneo, 2018). There are many different matching techniques, but the idea is to pre-process the data so that the treatment and control groups are similar to each other, except for the treatment status. This is supposed to account for the selection into treatment, i.e. that some individuals end up in the treatment group because of their characteristics. Now, if the difference in characteristics before the treatment is accounted for by the matching, the difference between the treatment and control groups may identify the effect of the treatment. The availability of extensive data is key for successful matching: differences in characteristics can only be reduced to the extent of the observable characteristics, thus data on many variables is needed.

The third essay in this dissertation makes use of matching methods combined with staggered difference-in-differences (or event study with a control group) to identify the effect of tax audits, which are specific firm-level events. Consequently, the treatment is not a policy change but the event of being audited as a part of regular auditing policy. The research design is enabled by unique data that combines the firm population tax returns with information on firm tax audits to identify the tax audit events of firms and the development of these over time relative to the audit year. The development of the audited firms around the audit is compared to the development of non-audited firms around a randomly assigned fixed point in time. In other words, we have assigned a "pseudo audit" for the non-audited firms to analyze their development relative to a fixed year. Using a Coarsened Exact Matching (CEM) method (Iacus, King and Porro, 2012), the estimation is restricted to firms similar in their observable variables prior to the audit (or the pseudo audit for the non-audited firms). The matching is crucial for improving the comparability of the audited and non-audited firms because they differ in their characteristics and development prior to being audited. The identification method, then, is the difference-in-differences method, which identifies the change in development between the audited and non-audited firms after the (pseudo) audit. The assumption behind the identification method is the common trends relative to the audit year, which is examined in the essay.

1.3 Mechanisms of the response and compliance

Firms can respond to tax policy in different ways, e.g. changes in (real) revenue, substituting labor for capital, or changing their compliance with a policy. In addition, tax policy has multiple dimensions including the tax base, tax rate, salience, and compliance costs, which firms may respond to (Gillitzer and Slemrod, 2014). Consequently, this dissertation, as well as most empirical studies, is interested in the *mechanisms* of the observed responses – the "why" and "how" behind the response. However, the causal effect on government revenue is the only response needed for welfare evaluation, as described above in section 1.1. Is there, then, any other reason to study the mechanisms than academic curiosity in understanding the response?

In effect, understanding the mechanisms can make the empirical results useful in improving or evaluating future policies and not only in revealing the welfare effects of existing or past policies. This is due to an embedded limitation of causal identification: only the effect of a *specific* treatment can be estimated. This is referred to as the external validity of the results – results in any empirical study only show the effect in the context of the study, which may not be the same as the effect in another context. Consequently, the causal effect is sufficient for *ex post* policy evaluation, i.e. evaluation after the adoption of the policy. Understanding the mechanisms is necessary for *ex ante* policy evaluation. For this purpose, more structural estimation strategies may be helpful to construct the counterfactual before the adoption of a policy (Abadie and Cattaneo, 2018). In addition, the mechanisms may help to identify the MWTP, as discussed above in section 1.1.

Related to understanding the mechanisms of the response is understanding the treatment; in other words, the effect of *what* is measured. As mentioned briefly above, tax policy includes many aspects. Gillitzer and Slemrod (2014) describe a tax system including, for example, compliance costs (costs of reporting taxes), salience (how transparent the tax system is, or how well the tax rates are known by individuals), remittance and withholding (the agent responsible for collecting the tax), and enforcement (punishments for non-compliance, tax monitoring by authorities) in addition to the more canonical aspects of the tax base (what is taxed) and the tax rate. Consequently, a tax policy treatment can include many different "treatments" if multiple aspects change simultaneously. This can be controlled for in a randomized trial but not in a natural experiment. Even when the effect of one change can be isolated, the other aspects of the policy can affect interpretation of the results. For example, a tax rate change may not be equal to price change if there are compliance effects or imperfect salience. For example, Chetty, Looney and Kroft (2009) show that people react differently to tax changes depending on the salience of the tax. Subsidy programs in particular often include many aspects: a firm may be eligible for a price decrease but there may be compliance costs or salience issues. These issues complicate interpretation of the results.

Firm compliance with tax policy is a specific theme addressed in this dissertation. Compliance, in essence, refers to whether the firm behaves in accordance with the tax policy or not. For example, tax compliance refers to reporting taxes correctly and, accordingly, paying the correct amount of taxes. In this case, non-compliance is illegal and subject to punishment. The third essay in this paper directly studies tax non-compliance and tax audits as a tool for improving compliance. However, compliance can also be voluntary. For example, using subsidies is not compulsory but, in a sense, not using an available policy is non-compliance with the policy. The first essay addresses this type of non-compliance or under-use of the first-employee subsidy as a potential reason for the observed effects. In addition, randomized experiments need to account for the potential non-compliance with the treatment, which means that not all individuals in the treatment group comply with the treatment.

For example, the first-employee subsidy, studied in the first essay, creates an incentive for firms to become employers by reducing the net labor costs of the first employee. However, this price decrease is not the only component of the policy, because i) the subsidy had to be applied for, and ii) the eligible employment contracts for receiving the subsidy were restricted. First, the subsidy had to be applied for, meaning that a firm has to know about the subsidy to experience the price decrease, and the firm may face administrative costs in applying for and using the subsidy. Consequently, the salience and compliance costs of the policy may affect firms' responses. Second, the subsidy restricted the type of employment contract, which increases the compliance costs. Importantly, these are issues that are directly affected by the policy design. Hence, the compliance response is important for interpretation of the observed firm responses and for guiding future policy.

In fact, the second essay uses the results on compliance provided in the first essay in designing the treatment in the experiment. The subsidy for becoming an employer in the experiment is supposed to have lower administrative costs, be more salient, and have no restrictions on the type of employment contract. This can make the subsidy more effective in decreasing the costs of becoming an employer and, consequently, be more effective. The plan for the experiment also examines how compliance affects the credibility of the experiment as measured by the statistical power.

In contrast, the third essay directly studies tax compliance by firms and the effectiveness of tax audits in increasing tax compliance. The essay studies the mechanisms behind the response, which may shed light on the source of tax non-compliance. This, in turn, can help in targeting the tax administration's resources. In addition, studying the anatomy of the response may be directly relevant for fiscal externality if there are effects on the reporting of multiple tax bases.

2 EMPLOYMENT INCENTIVES

There is a wide literature in public and labor economics that studies the effects of employment incentives since Kaldor (1936) suggested wage subsidies as a means to reduce unemployment. Employment incentives refer to public policies meant to increase the demand for or supply of labor, usually by subsidies or tax reductions. Accordingly, there are various types of policies with different goals.[8] For example, there are subsidies (or tax reductions) targeted at specific groups of employees, such as young, old or low-wage individuals (studied in e.g. Huttunen, Pirttilä and Uusitalo, 2013; Kaiser and Kuhn, 2016; Groh, Krishnan, Mckenzie and Vishwanath, 2016; and Saez, Schoefer and Seim, 2019b,a). In addition, active labor market policies, which are policies meant to support the unemployed to find jobs or those at risk of losing their jobs to retain them, often include wage subsidies targeted at the individuals (see e.g. Card, Kluve and Weber, 2010, and Brown and Koettl, 2015 for reviews of active labor market policies).

Additionally, employment incentives are sometimes targeted at firms to encourage job creation or decrease the destruction of jobs. These can be general job-creation subsidies (Betcherman, Daysal and Pagés, 2010) or, for example, regionally targeted tax subsidies (Korkeamäki and Uusitalo, 2009; Bennmarker, Mellander and Ockert, 2009). Employment incentives targeted at firms are of specific interest during economic depressions as they can help to reduce labor market separations during the depression and thus accelerate the recovery (Neumark, 2013; Cahuc, Carcillo and Le Barbanchon, 2019; Bruhn, 2020; Benzarti and Harju, 2021). In fact, during the COVID-19 crisis, employment subsidies were widely used to encourage firms to retain employment (e.g. Hamilton 2020 and Neilson, Humphries and Ulyssea 2020 study subsidies during COVID-19). This can be helpful if labor market separations are decreased because losing a job can result in long-term unemployment if finding a new job is difficult, as it is during depressions. Long-term unemployment, in turn, can result in loss of human capital or reduced employer perception of labor productivity. For example, Saez et al. (2019a) find permanent employment effects from a temporary employer payroll tax reduction on youth workers.

In a simple theoretical framework of a partial labor market equilibrium model, a wage subsidy (or a similar decrease in labor taxes) decreases the cost of labor increasing labor, demand (or in the case of a subsidy to the individual increases the net wage, increasing labor supply). Then, the subsidy *increases employment* and

---

[8] See e.g. Brown, Merkl and Snower (2011) for an evaluation of different policies.

wages, with the sizes of these effects depending on the labor demand and supply elasticities. In addition, employers can *substitute* using more employment instead of capital as the relative price of labor decreases. If the wage subsidy is targeted at a specific group of employees, it also reduces the relative labor costs of this group compared to other groups of employees. This can cause employers to *substitute* subsidized employees instead of the non-subsidized group. Often wage subsidies affect only a small number of jobs in the labor market. In this case the effects are concentrated in the subsidized firms and employees, and increases in wages are less likely (because the labor supply for the specific subsidized jobs is elastic due to the large amount of non-subsidized jobs).

However, the simple model ignores many potential margins of response. The effects can differ for two main reasons i) imperfections in the labor market and ii) the features of the subsidy design make it different from a pure price decrease. Firstly, there are many imperfections in the labor markets that can affect the effectiveness of employment incentives. For example, liquidity-constrained firms may use the extra cash provided by the incentives to grow their business. In addition, pay equity concerns or labor market regulations may prevent firms from paying a wage according to the labor productivity of an employee. Then, a wage subsidy may reduce the gap between the wage and labor productivity. For example, the results in Saez et al. (2019b) support these mechanisms. Secondly, many wage subsidies include administrative costs or may not be fully salient, which may decrease their effectiveness. Also, some employee-targeted wage subsidies or vouchers may include signals of lower labor productivity or stigma so that employees are less willing to use them, or employers may not be willing to hire employees with these subsidies. For example, Brown (2015) mentions employer take-up as important.

Consequently, the effects of wage subsidies are largely dependent on the many potential responses by firms. Recently there have been many studies on the effects on firms of employment subsidies or payroll tax reductions. The evidence is quite mixed, which may support the importance of policy design as well as the complexity in firm responses. For example, Korkeamäki and Uusitalo (2009) Bennmarker et al. (2009) Betcherman et al. (2010) study the effects of employment incentives targeted at firms. The estimated effects in Korkeamäki and Uusitalo (2009) and Bennmarker et al. (2009) are moderate or insignificant. On the other hand, Betcherman et al. (2010) find a significant increase in employment, which may, however, be largely due to job formalization rather than increased real employment. Cahuc et al. (2019) study the effects of a hiring credit during a depression, finding significant employment effects. Benzarti and Harju (2021) find that a payroll tax reduction had limited effect prior to an economic recession, but increased employment and sales in the firms affected during the recession. Kangasharju (2007) and Lombardi, Nordström Skans and Vikström (2018) find positive effects on the firm performance of wage subsidies for the unemployed,

but Lechner, Wunsch and Scioch (2013) find they may even harm firms. Saez et al. (2019b) find that a payroll tax reduction for young people increased youth employment, and increased the growth of firms through larger windfall gains, but Huttunen et al. (2013) find moderate employment effects from a wage subsidy for older low-wage workers, mainly increasing the hours of those already employed. Additionally, many papers study the effects of employment incentives in developed countries. For example, Mckenzie, Woodruff and De Mel (2019) study the effects of a wage subsidy on micro firms. Grimm and Paffhausen (2015) survey the evidence of different policies targeted at small and medium-sized firms.

2.1 Overview of essay 1

The first essay studies the effects on firm outcomes of subsidizing the first employee, using a regional subsidy program in Finland in 2007-2011 as a natural experiment. Firms in the subsidy area were eligible for a subsidy that amounted to 30% of the wage costs of the first employee in the first year and 15% in the second year if they hired the employee on a permanent contract with at least 25 hours of work per week. To be eligible, the firm had to have had no employees for at least 12 months. The subsidy targeted a large group of firms: over half of Finnish firms do not have employees other than the entrepreneur.

The essay makes use of register data on the full population of firm tax returns in 2000-2013 and the subsidy grant decisions. By combining the data on firm outcomes with the subsidy decisions, I am able to study the take-up of the subsidy among the firms that became employers while eligible for the subsidy.

By comparing the firms in the eligible area to firms in the similar neighboring but ineligible area, I find that the subsidy had zero effect on the probability of becoming an employer and additional firm outcomes. The estimates are precisely estimated and there are no large differences between the types of firms studied.

The zero estimated effect seems to be due to a low take-up of only 2% among the eligible firms that became employers. This raises a puzzle as to why firms do not use the subsidy. Calculating the ex post monetary gains of using the subsidy, I find that firms with larger gains are more likely to use the subsidy, but the take-up rate increases to at most 12% for firms with an ex post calculated subsidy between €8,000 and €12,000. By comparing the subsidized firms to firms that did not use the subsidy, I find descriptive evidence to support low awareness as a reason for the low take-up rather than administrative costs of using the subsidy. Furthermore, the restriction to full-time employment seems to be an important reason for the low take-up. The full-time restriction increases the opportunity costs of using the subsidy for firms that are only willing to hire their first employee on a temporary or part-time employment contract.

The essay contributes to the literature that studies how firms respond to em-

ployment incentives. Firstly, it is the first paper, to my knowledge, to study the effects of employment incentives specifically on the probability of becoming an employer among non-employer firms. This is important, as non-employer firms form the majority of all firms in Finland and large shares in other countries as well. Previously, Lechmann and Wunder (2017) and Fairlie and Miranda (2016) have studied the probability and dynamics of becoming an employer. Secondly, the paper raises firm take-up of subsidies and employment incentives as an important issue, and this has received little attention in previous literature. Relatedly, Neilson et al. (2020) show that small firms are less likely and slower to take up paycheck protection program loans, and incomplete take-up especially among smaller firms is mentioned in Korkeamäki and Uusitalo (2009) and Huttunen et al. (2013). In fact, policy design is an important factor in the effectiveness of employment incentives. The low take-up puzzle is also related to recent studies providing evidence that firms seem to make mistakes (e.g. Zwick and Almunia, Hjort, Knebelmann and Tian 2020).

2.2   Overview of essay 2

The second essay sets out a plan for a large-scale field experiment to study how firms respond to a subsidy for becoming an employer. The experiment is designed to be conducted by the Finnish government as close as possible to a policy that could be implemented as a universal policy. Hence, the experiment could provide evidence for evaluating the policy prior to its full implementation.

In planning the experiment, the paper uses full population tax return data to describe the population of non-employer firms. This can be used to assess how to target the subsidy and provide estimates for calculating the statistical power of the experiment. The same data can be used to identify the target population when running the experiment and the statistical analysis of the experiment.

In the experiment, firms picked from tax registers that had no employees for at least 12 months are randomized to receive eligibility for i) a recruitment subsidy, ii) training or iii) no treatment in order to serve as a control group. The recruitment subsidy equals 50% of the actualized labor costs in the first year up to a maximum of €10,000. The subsidy is paid automatically according to information provided by the firm for payroll taxation, after the firm has claimed the subsidy. The design of the subsidy is supposed to have minimal administrative costs for the firm, and consequently make it more attractive to firms than the previous first-employee subsidy described above. The training is aimed at helping the firm to become an employer (such as information on the legal responsibilities of employers) and can be customized for firms.

The preliminary budget and assumed take-up rates enable the recruitment subsidy to be offered to 13,087 firms and training to 14,700. Based on these numbers

and descriptive statistics from the data, the calculated minimum detectable effect (MDE) on the probability of becoming an employer is 11% for the subsidy and 10.5% for the training treatment. The paper confirms the validity of the statistical power calculations by simulating the power using bootstrap samples of the empirical data. The paper also assesses the multiple hypothesis correction using simulations.

The main contribution of the second essay is a public plan for a field experiment that serves the transparency of conducting randomized experiments. The paper proposes an approach of dividing the experiment into stages, to manage the uncertainty of subsidy take-up. It also presents descriptive evidence on the population of non-employer firms and the probability of becoming an employer. If the experiment is run, it could shed light on how firms respond to a subsidy for becoming an employer. This complements the results from the first essay, by showing how the program design affects the effectiveness of the subsidy. The experiment could be used to evaluate the effects of a potential policy program that aiming to reduce the costs of becoming an employer so as to increase job creation.

3 TAX EVASION

Recently, there has been increasing interest in public economics in studying tax compliance or tax evasion. Gillitzer and Slemrod (2014) describe tax compliance as an important feature of tax systems, and Slemrod (2019) states that the topic of tax evasion is related to the "bread and butter concerns of public economics", as tax compliance affects the efficiency and equity of the a system. Tax evasion refers to illegal forms of reducing tax liability, usually by misreporting economic activity. Tax compliance is the opposite: the true reporting of activity that determines correct tax liability. Tax evasion is related but needs to be distinguished from tax avoidance or tax planning, which refers to legal activities for reducing tax liability, e.g. by shifting income between different tax bases. In addition, tax non-compliance can refer to intentional and unintentional misreporting. Empirical study of tax evasion is challenging because tax evasion is, by definition, not observed in administrative records. However, as Slemrod (2019) describes, access to administrative data, new methods and cooperation with tax authorities have led to an increase in credible empirical studies on tax evasion. Especially, random audits and "traces" of income have been used to study the extent of tax evasion (e.g. Kleven, Knudsen, Kreiner, Pedersen and Saez, 2011; Artavani, Morse and Tsoutsoura, 2016), and studies using randomized field experiments or random audits

(Kleven et al., 2011; Gemmell and Ratto, 2012; Carrillo, Pomeranz and Singhal, 2017; Bérgolo, Ceni, Cruces, Giaccobasso and Perez-truglia, 2017) to study the effects of tax enforcement.

The canonical model of tax evasion is set out in Allingham and Sandmo (1972). In the model, individuals have income, pay taxes on their reported income, face the possibility of an audit and a punishment if they are caught evading taxes. Based on these, individuals choose the optimal reported income. The probability of an audit and the punishment are tax enforcement tools that crucially determine the extent of tax evasion for a given tax rate. The model alone is not able to explain high levels of tax compliance and the low tax audit rates observed in many developed countries. Consequently, there are theoretical suggestions that aim to explain the lack of tax evasion, e.g. in terms of moral considerations.

Recent theoretical and empirical evidence emphasizes the role of third-party reporting in tax compliance. In particular, Kleven et al. (2011) use randomized tax audits in Denmark to document very high levels of tax compliance for third-party reported income but lower levels of compliance for self-reported income such as income from self-employment. Third-party reporting of income, for example wage income reporting by an employer, allows the tax administration to easily detect tax evasion when an individual's reported income does not match with an employer's reports. For this reason, tax enforcement of self-employed persons, who report their own income, is especially challenging. Consequently, reporting and remittance rules are important tax enforcement tools in the tax system. For example, VAT is often argued to be self-enforcing due to the feature that a firm's sales are expenses for another firm. Consequently, firms report their expenses to reduce tax liability. This can provide an incentive for the seller to report sales in order not to get caught for tax evasion. This feature, however, has some weak links, especially when the firm sells to final consumers rather than other firms (see e.g. Naritomi, 2019).

Because of the role of third-party reporting, firms have a central role in overall tax compliance. Kleven, Kreiner and Saez (2016) even argue that tax compliance and firm size in an economy are connected, because collusive tax evasion is more difficult to maintain in large firms. Consequently, there is an increasing interest in studying the tax compliance of firms and their interaction with employees and consumers. For example, Bérgolo et al. (2017) and Carrillo et al. (2017) use field experiments to study how firm tax reporting responds to communications sent by the tax authority. Almunia and Lopez-Rodriguez (2018) study the effects of monitoring intensity on firm tax compliance in Spain. In addition, Naritomi (2019) studies tax enforcement of VAT by incentivizing consumers to report purchases in exchange for lottery tickets in Brazil. Bjørneby, Alstadsæter and Telle (2018) use randomized tax audits to study collusive tax evasion between firms and employees in Norway.

3.1 Overview of essay 3

The third essay studies how firms respond to real-life tax audits using data on operational tax audits combined with firm tax returns in Finland. Based on theory, it is not clear how firms respond. A tax audit may cause a firm to update its perceived probability of an audit either upwards or downwards, which would increase or decrease future tax compliance, respectively. If a tax audit increases tax compliance, it increases the effective tax rate faced by a firm, which may have a negative real effect on the firm's business activity. In addition, tax audits are costly. Consequently, the effectiveness of tax audits is an empirical question but there is limited evidence in the literature on how firms respond to tax audits.

The essay makes use of novel data on all operational tax audits in Finland in 2003-2016. This data is combined with the full population of firm tax returns in Finland in 2000-2016. Importantly, this makes it possible to follow the audited and non-audited firms over time. In the interest of introducing new data, we describe the audited firms and audit outcomes in some detail. In the paper, we construct a control group for the audited firms using CEM matching. Then, we use a difference-in-differences estimation strategy to estimate how the audited firms respond to being audited relative to a similar group of firms, with similar tax return development that do not get audited.

We find a statistically significant, immediate and persistent increase of 13.5% in firm profits for five years following an audit (including the audit year). The response in profits seems to be rather due to an increase in reported revenue than a decrease in costs, which implies under-reporting of revenue as a channel of tax evasion. We find that the audited firms with a positive detected tax deficit have a larger increase in profits, but there is a small decrease in their probability of reporting positive profits and revenue. This may be a sign of a negative real effect due to the higher effective tax rates. Consistent with evasion of payroll taxes or collusive evasion of income taxes with employees, we find an increase in reported labor costs and the number of employees following an audit.

The essay contributes to the literature on the effects of tax enforcement tools on firm tax compliance. The paper adds to the limited literature on firm responses to tax audits. Importantly, the paper studies how firms respond to operational tax audits, which are the usual tax enforcement tool. The effects of random audits are studied more, but the effects may differ because i) the response may be different if the firm knows it is being audited by random and ii) the target group may be different. Previously, the effects have been studied in DeBacker, Heim, Tran and Yuskavage (2015) and D'Agosto, Manzo, Pisani and D'Arcangelo (2018). The advantage of the essay compared to previous studies is the availability of the full population of audited and non-audited firms and an extensive analysis of tax return outcomes. The evidence on firm responses to operational tax audits can complement the literature using random audits or field experiments (Bjørneby et al., 2018;

Carrillo et al., 2017; Bérgolo et al., 2017). The empirical evidence is directly relevant to assessing the extent to which auditing one more firm affects government tax revenue.

References

Abadie, A. – Cattaneo, M. D. (2018) Econometric Methods for Program Evaluation. *Annual Review of Economics*, Vol. 10, 465–503.

Allingham, M. – Sandmo, A. (1972) Income Tax Evasion: A Theoretical Analysis. *Journal of Public Economics*, Vol. 1 (3-4), 323–338.

Almunia, M. – Harju, J. – Kotakorpi, K. – Tukiainen, J. – Verho, J. (2019) Expanding access to administrative data: the case of tax authorities in Finland and the UK. *International Tax and Public Finance*, Vol. 26 (3), 661–676.

Almunia, M. – Hjort, J. – Knebelmann, J. – Tian, L. (2020) Strategic or Confused Firms? Evidence from " Missing " Transactions in Uganda.

Almunia, M. – Lopez-Rodriguez, D. (2018) Under the Radar : The Effects of Monitoring Firms on Tax Compliance. *American Economic Journal: Economic Policy*, Vol. 10 (1), 1–38.

Angrist, J. D. – Pischke, J.-S. (2009) *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Angrist, J. D. – Pischke, J. S. (2010) The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives*, Vol. 24 (2), 3–30.

Artavani, N. – Morse, A. – Tsoutsoura, M. (2016) Measuring Income Tax Evasion Using Bank Credit: Evidence From Greece. *Quarterly Journal of Economic*, Vol. 131 (2), 739–798.

Bennmarker, H. – Mellander, E. – Ockert, B. (2009) Do Regional Payroll Tax Reductions Boost Employment? *Labour Economics*, Vol. 16 (5), 480–489.

Benzarti, Y. – Harju, J. (2021) Can payroll tax cuts help firms during recessions? *Journal of Public Economics*, Accepted.

Bérgolo, M. L. – Ceni, R. – Cruces, G. – Giaccobasso, M. – Perez-truglia, R. (2017) Tax Audits As Scarecrows: Evidence from a Large-Scale Field Experiment. Working paper, National Bureau of Economic Research.

Bernheim, B. D. – Taubinsky, D. (2018) Behavioral Public Economics. Working Paper 24828, National Bureau of Economic Research.

Bertrand, M. – Duflo, E. – Mullainathan, S. (2004) How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, Vol. 119 (1), 249–257.

Betcherman, G. – Daysal, N. M. – Pagés, C. (2010) Do employment subsidies work? Evidence from regionally targeted subsidies in Turkey. *Labour Economics*, Vol. 17 (4), 710–722.

Bjørneby, M. – Alstadsæter, A. – Telle, K. (2018) Collusive Tax Evasion by Employers and Employees: Evidence from a Randomized Field Experiment in Norway. Working Paper 7381, CESifo.

Borusyak, K. – Jaravel, X. (2021) Revisiting Event Study Designs.

Brown, A. (2015) Can hiring subsidies benefit the unemployed? *IZA World of Labor*, Vol. 163, 1–10.

Brown, A. J. – Koettl, J. (2015) Active labor market programs - employment gain or fiscal drain? *IZA Journal of Labor Economics*, Vol. 4 (12).

Brown, A. J. G. – Merkl, C. – Snower, D. J. (2011) Comparing the effectiveness of employment subsidies. *Labour Economics*, Vol. 18 (2), 168–179.

Bruhn, M. (2020) Can Wage Subsidies Boost Employment in the Wake of an Economic Crisis? Evidence from Mexico. *Journal of Development Studies*, Vol. 56 (8), 1558–1577.

Cahuc, P. – Carcillo, S. – Le Barbanchon, T. (2019) The Effectiveness of Hiring Credits. *Review of Economic Studies*, Vol. 86 (2), 593–626.

Card, D. – Kluve, J. – Weber, A. (2010) Active labour market policy evaluations: A meta-analysis. *Economic Journal*, Vol. 120 (548), 452–477.

Carrillo, P. – Pomeranz, D. – Singhal, M. (2017) Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement. *American Economic Journal: Applied Economics*, Vol. 9 (2), 144–164.

Chetty, R. – Looney, A. – Kroft, K. (2009) American Economic Association Salience and Taxation: Theory and Evidence. *The American Economic Review*, Vol. 99 (4), 1145–1177.

D'Agosto, E. – Manzo, M. – Pisani, S. – D'Arcangelo, F. M. (2018) The Effect of Audit Activity on Tax Declaration: Evidence on Small Businesses in Italy. *Public Finance Review*, Vol. 46 (1), 29–57.

DeBacker, J. – Heim, B. T. – Tran, A. – Yuskavage, A. (2015) Legal Enforcement and Corporate Behavior: An Analysis of Tax Aggressiveness after an Audit. *The Journal of Law and Economics*, Vol. 58 (2), 291–324.

Fairlie, R. W. – Miranda, J. (2016) Taking the Leap: The Determinants of Entrepreneurs Hiring their First Employee.

Farhi, E. – Gabaix, X. (2018) Optimal Taxation with Behavioral Agents. Working paper 21524, National Bureau of Economic Research.

Feldstein, M. S. (1999) Tax Avoidance and the Deadweigth Loss of the Income Tax. *The Review of Economics and Statistics*, Vol. 81 (4), 674–680.

Finkelstein, A. – Hendren, N. (2020) Welfare Analysis Meets Causal Inference.

Journal of Economic Perspectives, Vol. 34 (4), 146–167.

Gemmell, N. – Ratto, M. (2012) Behavioral responses to taxpayer audits: Evidence from random taxpayer inquiries. *National Tax Journal*, Vol. 65 (1), 33–58.

Gillitzer, C. – Slemrod, J. (2014) *Tax Systems*. The MIT Press, Cambridge.

Goodman-Bacon, A. (2020) Difference-in-differences with Variation in Treatment Timing, URL `http://goodman-bacon.com/pdfs/ddtiming.pdf`.

Grimm, M. – Paffhausen, A. L. (2015) Do interventions targeted at microentrepreneurs and small and medium-sized firms create jobs? A systematic review of the evidence for low and middle income countries. *Labour Economics*, Vol. 32, 67–85.

Groh, M. – Krishnan, N. – Mckenzie, D. – Vishwanath, T. (2016) Do Wage Subsidies Provide a Stepping-Stone to Employment for Recent College Graduates? Evidence From a Randomized Experiment in Jordan. *The Review of Economics and Statistics*, Vol. 98 (3), 488–502.

Hamilton, S. (2020) A Tale of Two Wage Subsidies: The American and Australian Fiscal Responses to COVID-19. *National Tax Journal*, Vol. 73 (3), 829–846.

Harrison, G. W. – List, J. a. (2004) Field Experiments. *Journal of Economic Literature*, Vol. 42 (4), 1009–1055.

Hendren, N. (2016) The policy elasticity. *Tax Policy and the Economy*, Vol. 30 (1), 51–89.

Hendren, N. – Sprung-Keyser, B. (2020) A Unified Welfare Analysis of Government Policies. *Quarterly Journal of Economics*, Vol. 135 (3), 1209–1318.

Huttunen, K. – Pirttilä, J. – Uusitalo, R. (2013) The employment effects of low-wage subsidies. *Journal of Public Economics*, Vol. 97, 49–60.

Iacus, S. M. – King, G. – Porro, G. (2012) Causal inference without balance checking: Coarsened Exact Matching. *Political Analysis*, Vol. 20 (1), 1–24.

Kaiser, U. – Kuhn, J. M. (2016) Worker-level and firm-level effects of a wage subsidy program for highly educated labor: Evidence from Denmark. *Research Policy*, Vol. 45 (9), 1939–1943.

Kaldor, N. (1936) Wage Subsidies as a Remedy for Unemployment. *Journal of Political Economy*, Vol. 44 (6), 721–742.

Kangasharju, A. (2007) Do Wage Subsidies Increase Employment in Subsidized Firms? *Economica*, Vol. 74, 51–67.

Kleven, H. – Kreiner, C. T. – Saez, E. (2016) Why Can Modern Governments Tax So Much? An Agency Model of Firms as Fiscal Intermediaries. *Economica*, Vol. 83, 219–246.

Kleven, H. J. – Knudsen, M. B. – Kreiner, C. T. – Pedersen, S. – Saez, E. (2011)

Unwilling Or Unable To Cheat? Evidence From A Tax Audit Experiment In Denmark. *Econometrica*, Vol. 79 (3), 651–692.

Korkeamäki, O. – Uusitalo, R. (2009) Employment and wage effects of a payroll-tax cut – evidence from a regional experiment. *International Tax and Public Finance*, Vol. 16 (6), 753–772.

Kremer, M. (2020) Experimentation, Innovation, and Economics. *American Economic Review*, Vol. 110 (7), 1974–1994.

Lechmann, D. S. – Wunder, C. (2017) The dynamics of solo self-employment: Persistence and transition to employership. *Labour Economics*, Vol. 49, 95–105.

Lechner, M. – Wunsch, C. – Scioch, P. (2013) Do firms benefit from active labour market policies? Discussion Paper 7614, IZA.

Lombardi, S. – Nordström Skans, O. – Vikström, J. (2018) Targeted wage subsidies and firm performance. *Labour Economics*, Vol. 53, 33–45.

Mckenzie, D. – Woodruff, C. – De Mel, S. (2019) Labour Drops: Experimental Evidence on the Return to Additional Labour. *American Economic Journal: Applied Economics*, Vol. 11 (1), 202–235.

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

Neilson, C. – Humphries, J. E. – Ulyssea, G. (2020) Information Frictions And Access To The Paycheck Protection Program. Working Paper 27624, National Bureau of Economic Research.

Neumark, D. (2013) Spurring Job Creation in Response To Severe Recessions: Reconsidering Hiring Credits. *Journal of Policy Analysis and Management*, Vol. 32 (1), 142–171.

Saez, E. – Schoefer, B. – Seim, D. (2019a) Hysteresis from Employer Subsidies. Working paper, National Bureau of Economic Research.

Saez, E. – Schoefer, B. – Seim, D. (2019b) Payroll Taxes, Firm Behavior, and Rent Sharing: Evidence from a Young Workers' Tax Cut in Sweden. *American Economic Review*, Vol. 109 (5), 1717–1763.

Saez, E. – Slemrod, J. – Giertz, S. H. (2012) The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review. *Journal of Economic Literature*, Vol. 50 (1), 3–50.

Slemrod, J. (2019) Tax compliance and enforcement. *Journal of Economic Literature*, Vol. 57 (4), 904–954.

Zwick, E. () The Costs of Corporate Tax Complexity. *American Economic Journal: Economic Policy*, Forthcoming.

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU