Three Essays on the Economics of Taxation

by

William C. Boning

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in the University of Michigan
2019

Doctoral Committee:

Professor Joel Slemrod, Chair
Professor Charles Brown
Professor James R. Hines Jr.
Professor Matthew Shapiro

William C. Boning

wcboning@umich.edu

ORCID iD: 0000-0003-0426-3943

DEDICATION

To Claire Roberts-Thomson and to my parents, with thanks for your love and support.

ACKNOWLEDGEMENTS

I thank my advisers, Joel Slemrod, Jim Hines, Charlie Brown, and Matthew Shapiro, for guidance, wisdom, and support.

I am grateful for the dedication and knowledge of my co-authors on chapters of this dissertation, John Guyton, Ron Hodge, Joel Slemrod, Ugo Troiano, and Alex Turk.

This work has benefited from comments from people including, but not limited to: Brian Erard, John Friedman, Jacob Goldin, Henrik Kleven, Jeffrey Liebman, Zachary Liscow, Michael Lovenheim, Joana Naritomi, Alari Paulus, Juan Carlos Suarez Serrato, Dmitry Taubinsky, Christian Traxler, Mazhar Waseem, Danny Yagan, Eric Zwick and seminar participants at the University of Michigan, University of North Carolina, National Tax Association, International Institute of Public Finance, the IRS-Tax Policy Center conference, and various potential employers.

I also greatly appreciate those who work to make the university a supportive environment in which to do research. These people include, among others, Mary Ceccanese, Laura Flak, Julie Heintz, Olga Mustata, Mary Mangum, Michael Lee, and Cassandra Kelly. The infrastructure and environment they have built has made my time at the university and my work better.

I thank my academic mentors, James Poterba, Greg Mankiw, and Ed Glaeser, for helping me to reach this point in my career.

The data for this dissertation were accessed while I was a Student Volunteer Employee through the Joint Statistical Research Program of the Internal Revenue Service, and I thank the great many people who have worked to make these projects and that program a success. Thanks to, among others, John Guyton, Alex Turk, Alicia Miller, Ron Hodge, Ben Herndon, Pat Langetieg, Stacy Orlett, Victoria Bryant, Alex Contos, Anne Herlache, Juan Mendez, Brian Best, Jeff Butler, Barry Johnson, Michael Weber, Davy Leighton, Joanne Gburek, and Andrew Dettling. Their many efforts included reviewing this work to ensure that confidential information about taxpayers and about IRS methods was not disclosed.

This research was in part supported through an NSF Graduate Research Fellowship Program grant, no. DGE 1256260, and through a Haber Fellowship. The opinions and conclusions expressed herein are those of the author or authors, and do not necessarily reflect the positions of the IRS or NSF.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

<u>**Table**</u>

LIST OF APPENDICES

Appendix

ABSTRACT

This dissertation applies natural and randomized experiments to understand how tax systems shape the behavior of people and firms. Chapter one studies how marginal changes in employer withholding change whether and when tax liabilities are paid. It uses evidence from a policy change in 2009 to show that more tax is paid on time when employers, rather than employees, are responsible for turning more of it in. The patterns of behavior across taxpayers suggest frictions, rather than liquidity constraints, explain this behavior. Chapter two presents evidence from a randomized experiment to show that letters and visits from the IRS increase the employment taxes firms turn in, and that visits also have network effects, increasing tax remitted by firms sharing a tax preparer with the visited firms. Chapter three also uses a randomized experiment to study whether contacting firms that are likely to be financially constrained affects their ability to operate. It finds that tax enforcement substantially and persistently reduces employment at visited firms, demonstrating that tax enforcement and compliance can alter the structure of production.

CHAPTER I

Paying Taxes Automatically: Behavioral Effects of Withholding Income Tax

1.1 Introduction

Employers in the U.S. and in nearly all countries surveyed by OECD (2015) must withhold income taxes from their employees' paychecks. Despite large gross tax collections through withholding, which were more than $1.3 trillion in FY 2017, late payments are extensive among taxpayers who owe more tax than the amount withheld—8.2 million taxpayers filed returns with balances not paid by the due date in FY 2017 (Internal Revenue Service (2018), Tables 1 and 16). On the margin, how much does a reduction in withholding lead to late payment, and does withholding help or hinder taxpayers by preventing late payment?

This paper assesses how and why changes in withholding affect late tax payment, using IRS administrative data to examine a 2009 policy change that reduced the amount withheld for some taxpayers. Affected taxpayers and a control group received the same change in total income from the policy, but at different times: affected taxpayers had larger paychecks but smaller refunds or larger balances due at filing. I show that on the margin additional withholding leads many taxpayers to pay on time rather than late, reducing the costs of collecting late payments. Patterns in who responds reveal why people respond. If withholding prevents late payments by fully rational but liquidity-constrained taxpayers, then taxpayers with liquid assets should be less responsive than those without liquid assets. This explanation is not supported by the evidence: I find that taxpayers earning interest income, who are more likely to have liquid assets, are no less responsive than others to a withholding change. If taxpayers pay late because of frictions that affect several decisions at once, like a tendency to procrastinate, late payers should be more likely to make other errors. I find that late payers are more likely to file returns with errors that both overstate and understate their tax liability, consistent with the explanation that greater withholding reduces late payments by removing frictions from the payment decision.

In the episode I study, policy changes effectively cut 2009 withholding by $250 for households that both earned wages and received certain benefits. Regardless of benefit receipt, a combination of payroll tax credits and one-time direct payments provided households earning wages with $400 ($800 if married filing jointly) of additional after-tax income, but benefit receipt changed how and when the additional income arrived. Households receiving Social Security retirement, Supplemental Security Income, or veterans' disability benefits received an Economic Recovery Payment of $250 in April or May 2009. The payment reduced these households' payroll tax credits by $250, but withholding tables included the full payroll tax credit for all households. While households without benefits had no change in their refunds or balances due at filing, households with benefits had refunds $250 smaller or balances due $250 larger in April 2010. The difference in payment timing between households with and without benefits, holding after-tax income fixed, amounts to a $250 withholding cut for households with benefits.

I analyze the effects of this policy change using IRS administrative panel data. The panel combines information from tax returns with information from other IRS data sources, including records of IRS-initiated error corrections and notices sent to taxpayers, the timing and amount of direct tax payments, and how long taxes go unpaid, none of which are available in public-use samples. Records of stimulus payments and demographic information come from the Social Security Administration. I select a 2000-2013 sample panel for analysis from data on the universe of taxpayers.

To identify the effects of changing withholding separately from other policy changes and economic shocks, I compare households earning wages with and without a recovery payment–and thus a withholding cut–in years before and after the change. Controlling for pre-treatment demographics including age, marital status, spouse's age, and number of dependents addresses the concern that recovery payment recipients, who earned Social Security retirement, Supplemental Security Income, or veterans' disability benefits, respond differentially to other changes over time. Trends in tax payments are parallel before the policy change during the 2009 tax year. Data from prior years rule out the Great Recession as an explanation for the effect when taxpayers file their tax year 2009 returns in early 2010: neither the prior recession in 2001 nor the early part of the Great Recession affects the parallel trends.

I find that cutting withholding substantially increases late tax payment. Cutting withholding by $250 leads an additional 1.4 percent of taxpayers to pay late. For each dollar withholding decreases, taxpayers fail to pay five cents of their balances by the due date. Most taxpayers receive a refund and cannot pay late, so the impact is larger among taxpayers with a balance due, for whom decreasing withholding by $250 leads to a 5.7 percentage point increase in late payment and whose overdue balances rise by 20 cents per dollar withholding decreases. Taxpayers whose withholding is cut are more likely to owe money to IRS over the twelve months after filing, at which point the following tax year's refunds provide the IRS with an opportunity to collect some taxpayers' remaining balances due. In tax years following the one-year cut to withholding in 2009, the effect reverses sign: taxpayers subject to the withholding cut are less likely to make late payments in future years, which may reflect taxpayer adjustments, for example opting to have more tax withheld. These results imply that cutting withholding raises the cost of collecting late taxes, increasing the overall burden of taxation.

I obtain predictions about which taxpayers respond to changes in withholding from separate models in which people pay late because of binding liquidity constraints or because of frictions in submitting payments. The liquidity constraint model predicts that late payers would prefer higher present consumption and have exhausted their ability to borrow at interest rates below the interest rate on late payments. Given prevailing interest rates, this model predicts late payers do not also hold savings.[1] Cutting withholding relaxes the liquidity constraint facing those people who would like to pay late. Adding to the model the ability to adjust withholding subject to a minimum preserves this result for a policy change that cuts the minimum level of withholding.

If, instead, frictions such as hassle costs or procrastination lead taxpayers to pay late, those who pay late when withholding is cut may also be more affected by frictions on other margins of behavior. For example, taxpayers whose time or attention is especially valuable when returns and payments are due face higher hassle costs both of making payments on time and of completing accurate tax returns.

Testing the models' predictions using heterogeneous responses to the 2009 withholding cut, I find that the evidence is inconsistent with late payment due to liquidity constraints and instead supports late payment due to frictions. If liquidity constraints lead to late payments, taxpayers holding liquid assets should respond less to a withholding cut. I use interest income as a measure of liquid assets, and find that cutting withholding increases late payments by just as much among taxpayers earning substantial interest income as among other taxpayers. Given lower interest rates on saving than late payment, late payers leave money on the table. This result is robust to capturing liquidity with various measures of interest and dividend income. Late payers are also more likely to file returns with inaccuracies that would both

[1]The effective interest rate on late tax payments in 2009 was 9.1 percent annually, compounded on a monthly basis, exceeding the return to most forms of saving but lower than the interest rates then charged on credit card balances or on unsecured personal loans.

raise and lower their tax bills, failing to claim the MWP payroll tax credit altogether or to report credit-reducing Economic Recovery Payments.

In sum, I find that on the margin withholding reduces the total burden of taxing income through two channels without imposing costs on liquidity-constrained taxpayers. Withholding lowers the costs of collecting late tax payments, and insulates taxpayers from the effects of frictions in making tax payments.

1.2 Related Literature

While this paper provides the first evidence of the effect of income tax withholding on late tax payment, existing work finds that withholding has effects on other aspects of taxpayer behavior. Chang and Schultz (1990) show a correlation between the amount withheld and income reported, and Rees-Jones (2017) and Engström, Nordblom, Ohlsson, and Persson (2015) find that many taxpayers report liability just below the tax already withheld, in the region where no additional payment is due, a result they attribute to loss aversion. Engström, Nordblom, Ohlsson, and Persson (2015) also use identifying variation in withholding from tax rate rounding to show that even a small balance due leads taxpayers to claim a dubious deduction. While Feenberg and Skinner (1989) find that taxpayers who owe a balance due at filing save more in tax-preferred retirement accounts, Feldman (2010) shows that the 1992 withholding cut instead *reduced* tax-preferred saving in a manner consistent with mental accounts. Although Brockmeyer and Hernandez (2016) study withholding from businesses under a sales tax rather than from individuals under an income tax, they find that a reform that increased withholding led to greater revenue because in the context they study there are high hassle costs of claiming refunds.

Why people pay taxes late remains an open question, but this paper shows that some late payments result from frictions rather than from liquidity needs. Although the model in Andreoni (1992) uses liquidity needs to explain why taxpayers underreport income rather than why they pay late, the model is a natural starting point for a liquidity-driven theory of late payment. Existing empirical work on late tax payments focuses on how to collect payments once they are late, not on why they are late to begin with. Perez-Truglia and Troiano (2018) find that letters emphasizing both financial penalties and shaming work, while Hallsworth, List, Metcalfe, and Vlaev (2017) find that letters with descriptive social norm appeals, financial information, and messages regarding public services speed collection. Taken as a whole, the empirical work on late payments emphasizes that when people pay taxes depends on both financial penalties and other factors.

The connection between frictions and timely tax payment parallels the role of frictions in take-up of a range of tax and non-tax benefits. The evidence shows that institutional and informational details matter. Simplified mailings increase EITC take-up (Bhargava and Manoli (2015)), as do reminders targeted at inattention (Guyton, Langetieg, Manoli, Payne, Schafer, and Sebastiani (2017)). The earnings response to the EITC depends on information about the tax code (Chetty and Saez (2013)). How information is presented affects tax-advantaged retirement saving (Saez (2009)). Behavioral factors may reduce take-up of unemployment insurance (Blank and Card (1991)). Simplicity and user-friendliness are key design principles not only for distributing benefits, but also for collecting information and revenue.

Withholding allows tax authorities to collect revenue from wages and salaries at lower cost. In comparison to income from sources subject only to information reporting, taxpayers report more of their income from wages, which are subject to both information reporting and withholding (Internal Revenue Service (2016), Kleven, Knudsen, Kreiner, Pedersen, and Saez (2010)). Dusek and Bagchi (2017) find that revenues increase 28 percent when U.S. states adopt both information reporting and employer withholding.

Perhaps because it raises revenue at low cost, withholding is at the heart of modern tax systems. The U.S. government raises revenue mostly through withheld income taxes and payroll taxes, both remitted by employers. The U.S. is not alone. Fifty of the fifty-five countries surveyed by OECD (2015) require employers to withhold income tax from wages and salaries. Jensen (2019) shows that, in U.S. history and across countries, the income tax is most often applied only to income groups that are at least 80 percent employees.

This paper's findings contribute to a recent empirical literature that challenges the traditional view in public finance that the burden of taxation does not depend on who remits the tax. Kopczuk, Marion, Muehlegger, and Slemrod (2016), for example, find that diesel tax evasion changes when remittance responsibility is shifted to a different level of the production chain. Placing the remittance responsibility with employers reduces the burden of taxation by removing the additional friction-driven costs taxpayers bear when making payments and the additional administrative cost of collecting late taxes.

1.3 Employer Withholding in the U.S.

U.S. policy aims to match the income tax each household pays during the tax year to the total tax liability on their tax return, in keeping with the legal requirement that tax is due when income is earned. During the tax year, tax is paid in two ways: people directly pay estimated tax, and employers withhold tax from their employees' paychecks. The amount withheld depends on marital status and on wages, from which allowances for deductions and credits are subtracted. The default amount withheld assumes the employee is single and has no allowances, and so exceeds most employees' liability. Employees can file a form with their employers to adjust the amount withheld, updating marital status and allowances or claiming an exemption from withholding. There is no upper limit on the tax employees can request be withheld–they can write in an arbitrary additional amount to withhold. There is, however, a lower limit on withholding, because employees cannot claim allowances for more credits and deductions than they receive, and can only claim an exemption from withholding if they owe no tax. The IRS sends letters to employers locking in higher levels of withholding for employees who set withholding too low.

After the tax year, taxpayers file returns and settle up with the government. Returns are due in the middle of the following April. If the amount withheld or submitted directly as estimated taxes exceeds tax liability, the taxpayer receives a refund of the excess payment, without interest. If payments during the tax year are less than tax liability, the remaining tax is due as a direct payment by the mid-April filing deadline. In cases where payments during the tax year are much lower than tax liability, the IRS charges interest.[2]

During the 2009-2010 period, the penalties and interest charged on late taxes totaled 9.1 percent over the first year, compounded monthly[3]. The interest rate charged on late tax payments is therefore less than the rate charged by credit cards, which averaged 13 to 15 percent in 2009-2010 (Board of Governors of the Federal Reserve System (2018)), but much higher than the interest rate on bank deposits or other liquid savings.

IRS sends people who owe late tax a bill, then a second bill, before turning to other means to collect. After the second bill, the IRS will subtract the unpaid amount from any future refunds. An IRS employee may also visit the taxpayer in person, and can attempt to collect through levies against sources of income or assets and through liens against or seizures of property.

The withholding system relies on employers' cooperation, which is strictly enforced. The executives responsible can be personally liable for any withheld tax the employer does not remit. In this paper, I set employer behavior aside. In practice most but not all employers cooperate. In Boning, Guyton, Hodge, Slemrod, and Troiano (2018), my coauthors and I find that enforcement contact has large effects on the withheld tax remitted by the few potentially uncooperative employers, and smaller spillover effects on the tax remitted by

[2]Interest, called the estimated tax penalty, is typically due if payments during the tax year are less than ninety percent of the year's tax, with some exceptions.

[3]The penalty for late payment is 0.5 percent of the unpaid tax per month (up to a maximum of 25 percent), plus monthly interest at the federal short-term rate plus three percentage points. This calculation also assumes the taxpayer files on time, avoiding the substantial penalty for failure to file of five percent of the tax due per month up to a maximum of 25 percent.

other employers sharing the same tax preparer.

1.4 Policy Variation and Data

Interactions between two stimulus policies in the 2009 American Recovery and Reinvestment Act effectively cut withholding by $250 for some employees, who then had either smaller refunds or larger balances due than a control group. The act otherwise had similar consequences for the two groups. I follow both groups in a panel of administrative tax data. A difference-in-difference approach identifies the effect of collecting more tax directly from taxpayers instead of through employer withholding.

1.4.1 Policy Changes

Households with income from both wages and certain government benefits received an additional $250 from stimulus policies during the 2009 tax year but also had refunds that were $250 smaller or balances due that were $250 larger when filing their 2009 tax returns. These changes are relative to households earning wages but not receiving the relevant benefits. The two groups had their taxes collected at different times, through different methods, but eventually the policy changes gave both groups the same amount of additional income. The difference was due to an interaction between the Making Work Pay payroll tax credit and one-time Economic Recovery Payments to benefit recipients.

The Making Work Pay (MWP) payroll tax credit cut fully eligible employees' 2009 and 2010 tax bills by up to $400 per year, or $800 per year if married and filing jointly. The credit was refundable. The credit amount was a function of earned income, which included wages, certain kinds of self-employment income, and non-taxable combat pay. The credit was phased in at a rate of 6.2 percent up to the full amount of credit at earned income of $6,451 ($12,903 if married filing jointly), remained at the maximum at earned income up to $75,000 ($150,000), and then phased out until households with earned income of $95,000 ($190,000) received no credit.

Each person eligible for Social Security or railroad retirement benefits, Supplemental Security Income benefits[4], or veterans disability or survivorship payments received an Economic Recovery Payment (ERP) of $250 in April or May 2009. Each person received at most one Economic Recovery Payment, though a couple filing jointly could receive two. The government agencies providing benefits sent payments to 55.2 million taxpayers, mostly by direct deposit, with the rest sent by mail. It is unlikely that taxpayers selected into receiving a payment. To receive a payment, one had to be eligible for benefits during November 2008,

[4]because of blindness, disability, or low income while over 65

December 2008, or January 2009, the final three months before the Recovery Act passed, but the payments were added to the act in conference committee in February 2009.

The Economic Recovery Payment counted against the MWP payroll tax credit. Households with both wages and an Economic Recovery Payment were eligible for at most $150 ($550) of MWP credit. MWP credit amounts less than $250 were reduced to zero.

Tax year 2009 withholding fell by the full MWP credit amount for all employees earning wages, regardless of eligibility. A single employee whose only income was wages of $60,000, for example, received an additional $400 in her paychecks. As she was eligible for the full $400 MWP credit, her withholding and tax liability changed by the same amount. Her refund or balance due for tax year 2009 was similar to her refund or balance due in prior tax years. The amount withheld did not depend on the amount of the MWP credit for which the employee was eligible. As a result, taxpayers who were not fully eligible for the MWP credit, including those receiving Economic Recovery Payments, had either larger balances due or smaller refunds with their 2009 returns.

Consider a single taxpayer whose wage income was $60,000 and who received an Economic Recovery Payment. His MWP credit was $150 and the Economic Recovery Payment gave him another $250, so he kept final benefits totaling $400. If he had not received an Economic Recovery Payment, he would have kept the same amount, distributed only through the full MWP credit. His withholding was incorrectly cut by the full MWP credit of $400, but the MWP credit he could claim on his 2009 tax return was only $150. As a result, his 2009 tax return would show a refund $250 smaller or balance due $250 larger than if he had not received an Economic Recovery Payment.

Employees could have had more tax withheld to offset the policy change, but such changes are uncommon. Withholding policy sets a default subject to taxpayer adjustment, and Jones (2012) finds substantial inertia around the default–after a policy change cut withholding in 1992, employees adjusted withholding slowly or not at all. The form employees use to adjust withholding is not filed with IRS, but with employers, so I do not observe withholding adjustments. Employees' withholding adjustments would cause the effect of withholding estimated in this paper to be a lower bound relative to the effect if employees could not adjust withholding. Given that taxpayers can adjust, the effect estimates in this paper are the relevant estimates for withholding policy.

1.4.2 Data and Sample Construction

I bring together information from several IRS administrative databases to construct a panel that follows households in the treatment and control groups over tax years 2000-2013. To obtain comparable numbers of control and treated households, I randomly sample 1 percent of the overall population and 10 percent of the population of Economic Recovery Payment recipients. I restrict both groups to those with enough wages to have $250 of tax credit for the payment to offset, producing a reduction in withholding. Pre-treatment differences between treated and control taxpayers are consistent with payment receipt depending on retirement or disability benefit receipt. These differences are addressed by the difference-in-difference empirical strategy I employ.

Data come from taxpayers' returns as well as from the results of IRS automatic error-detection checks, IRS records of payments received and unpaid taxes, and Social Security Administration records of demographics and Economic Recovery Payment receipt[5]. Data are at tax-year frequency, with the exception of monthly data on amounts owed to IRS. Data are available from 2000-2013 for most variables, though payment amount data are only comprehensive beginning in 2003.

By definition, households in the treatment group received exactly one Economic Recovery Payment, and households in the control group received no Economic Recovery Payment. I restrict to households by keeping only the primary filer in couples filing jointly. I limit panel imbalances across tax years by restricting the sample to taxpayers between the ages of 30 and 70 in 2009. I restrict both groups to households who earned wages between $4033 and $82,500 ($177,500 if married filing jointly). This restriction is necessary for the the policy interaction to have the full $250 effect on how and when tax was collected. Instead of wages from tax year 2009, which may respond to the MWP payroll tax credit, the restriction uses wages from tax year 2008. Powell (2015) and Mortenson and Whitten (2016) find that taxpayers report self-employment income or incorrect 2009 wage amounts to obtain the full MWP payroll tax credit. Wages from the two years are highly correlated, and the restriction is broad, so results do not substantially depend on whether the restriction uses 2008 or 2009 wages.

The outcomes I study are an indicator for whether the taxpayer paid the full amount due with the return, the dollar amount of payment by the deadline net of the balance due with the return, which is winsorized at the 1st and 99th percentiles, and the percentage of the remaining balance due with the return that is not paid by the deadline. Taxpayers owed a refund are counted as fully paid and counted as having zero payment net of balance due.

1.4.3 Characteristics of Treatment and Control Group Households

Households in the treatment and control groups differ before treatment, although these differences are mitigated by the difference-in-differences strategy I adopt. One would expect that households eligible for retirement or other benefits would differ from other households,

[5]The payment receipt records are comprehensive regardless of the agency distributing the payment

and in fact two thirds of households in the the treatment group receive Social Security retirement benefits in 2008, as shown in Table 1.1. The treated group are older, averaging 58 years of age versus 41. Treated households are ten percentage points more likely to be married filing jointly, and on average claim 0.5 fewer dependents. They average $41,000 of wage income, less than the control group's $59,000. There is a smaller difference in adjusted gross income, which measures taxable income from a wider range of sources, from $61,000 in the treated group to $68,000 in the control group. Treated households earn more interest income from both taxable and tax-exempt sources, on average about $1,400 versus $800. Differences between the treatment and control groups are addressed by my empirical strategy, in which I use a difference-in-difference estimation approach and controls for time-varying effects of pre-treatment demographics. When considering whether the empirical results for the treated group might apply to taxpayers more generally, it is worth noting that the treatment group are on average younger than retirement age and earn substantial wage income.

A large majority of both groups pay on time. The treated group on average receive smaller refunds, but are equally or slightly more likely to pay and file on time before treatment. In both groups in 2008, about 95.5 percent of taxpayers fully pay the tax due with their returns. The control group on average pay $129 less than the balance due, failing to turn in 4.05 percent of the balance of tax due with the return, while the treated group on average pay $113 less than the balance due, paying all but 3.99 percent of the balance due. Among the treated group, 96.4 percent file on time, slightly more than the 95.3 percent of control group timely filers.

Conditional on owing a balance due, households in both groups are much less likely to pay on time. Only 80 percent of treated households owing a balance due pay it in full in 2008, and 72 percent of control group households. On average, treated households owing a balance due leave $500 unpaid, compared to $800 in the control group.

The treatment affects treated households' returns as expected, reducing the treated group's Making Work Pay payroll tax credit in 2009 relative to the control group. Table 1.2 shows that the treated group on average receive $196 less ($351 rather than $545) Making Work Pay tax credit than the control group. This $196 is less than the full $250 reduction the policy would otherwise induce because I predict a credit that depends on 2009 earned income using 2008 wages, so that changes in wages between the two years or certain other income items could attenuate the extent of treatment. However, for the majority of the treatment group the credit is reduced by exactly the full $250 from the maximum possible credit, with 16 percent of treated households receiving a credit of exactly $150 and another 47 percent receiving exactly $550. Most treated taxpayers would receive the full credit but for the interaction with the ERP check, which would result in a credit of $150 if single and $550 if married filing jointly[6].

1.5 Empirics: Main Effects

1.5.1 Difference-in-Difference Specification

I use a difference-in-difference approach to compare the changes in the tax compliance behavior of the treated and control groups over time. The identifying assumption is that trends in behavior between the treatment and control groups would remain parallel absent the policy change, controlling for time-varying effects of pre-treatment characteristics.

The stimulus act also increased the earned income tax credit available to certain households with three or more dependents beginning in 2009, so I conduct a robustness check in which I exclude households with three or more dependents in 2008, reported in Table A.2. To my knowledge, no other legislative or IRS policy changes in 2009 could confound the results by differentially affecting the treatment and control groups.

Restricting to the treatment and control groups defined above, I estimate the average treatment effect with the following event-study version of the difference-in-difference specification, which interacts baseline demographics with year fixed effects.

$$y_{it} = \alpha_i + \sum_t \beta_t ERP_i + \sum_t \gamma_t X_i + \varepsilon_{it}, \tag{1.1}$$

where y_{it} is an outcome measure for household i in time period t, α_i is a household fixed effect, β_t is the effect of a $250 reduction in 2009 withholding on the outcome variable y in time period t, ERP_i is an indicator equal to one for if the household received exactly one Economic Recovery Payment in 2009 and zero otherwise, $\gamma_t X_i$ is a vector of time-varying effects of pre-determined covariates that are not perfectly co-linear with treatment, and ε_{it} is an error term.

The treatment and control groups differ in background characteristics such as age that affect the time path of taxpayer behavior over several years time, so I control for time-varying effects of characteristics determined before the policy was announced. Specifically, I control for interactions between year fixed effects and: age fixed effects, age of spouse in 2008, if any, filing status in 2008, and dependents in 2008. In the monthly data on debts to IRS, I address seasonality in the difference between the treated and control groups by removing

[6] Households could also have exactly this much credit due to earned income on the phase-in or phase-out regions of the credit schedule, but such situations are rare, as illustrated by the fact that less than one half of one percent of control group households earn exactly this much credit.

month-of-year-by-treatment fixed effects estimated using the data from the pre-treatment period.

The coefficients of interest are the effect of the reduction in withholding beginning in 2009, when the reduction took place, captured by β_t where $t > 2008$. Rather than imposing a single coefficient for 2009 and later years, these coefficients capture the effect for each tax year separately. The specification also directly provides a placebo test: the trends in the outcomes between the treated and control groups should be parallel prior to treatment in the 2009 tax year.

1.5.2 Main Effects Results

Estimating specification 1.1 reveals four facts. First, taxpayers whose withholding is reduced by $250 in 2009 are more likely to pay late, and pay larger amounts late, with larger effects among taxpayers owing a balance due. Second, the resulting tax debts last up to one year. Third, taxpayers whose withholding is reduced are less likely to pay late in future years. Finally, the evidence from pre-treatment years supports the identifying assumption that absent the policy change the difference in late payment between the treatment and control groups would be similar across years.

When taxpayers are required to make more tax payments directly, they are less likely to pay in full, leave larger balances due unpaid, and pay a smaller fraction of the balance due. The difference-in-difference estimates are reported in Table 1.3 and depicted in Figure 1.1, Figure 1.2, and Figure 1.3. Treated taxpayers are 1.43 percentage points less likely to remit the full amount due with their 2009 returns. The treated group on average pay $13 less of their 2009 taxes by the due date, which is five percent of the $250 reduction in withholding. Averaging across taxpayers, the treated group pay 0.7 percentage points less of the dollar amount due with the return. All of these results are highly statistically significant.

The magnitudes of the late payment effects are about four times larger after excluding the majority of taxpayers due a refund, who cannot pay late. If one restricts each year's sample to taxpayers with a balance due, as in Table 1.4, the withholding change leads taxpayers with a balance due to be 5.6 percentage points less likely to pay in full. This is a large change, but is reasonable when compared to the 72 percent of control group taxpayers with a balance due who pay in full in 2008. Among taxpayers with a balance due, the effect of the $250 reduction in 2009 withholding is a $50 decline in average net payments relative to the balance due, a large effect but reasonable given that the control group taxpayers owing a balance due left an average of $822 unpaid in 2008. Late payment is highly responsive to the balance of tax due at filing, and therefore to withholding.

Taxpayers whose withholding was reduced by $250 are significantly less likely to be free of debts to IRS for a year after the deadline. Figure 1.4 shows that the two groups' likelihood of being free of tax debt is roughly comparable prior to a dramatic spike downward for the control group following treatment, in which many treated taxpayers owe tax debts. The period of the spike, depicted in detail in Figure 1.5, begins during the 2009 tax year filing season in March and April 2010. The spike in tax debts peaks in May and June 2010, which trails the filing and payment deadline in mid-April because late payments are not added to the month-by-month debt data until after an initial notice period. The effect then diminishes over time and finally becomes no longer significant in either economic or statistical terms in May 2011. The unwinding of tax debt among the treatment group relative to the control group during the 2010 filing season suggests that some of the additional late payments may have been collected by reducing the refund amounts for taxpayers with a 2010 refund and remaining tax debt from tax year 2009. Excess withholding, and the resulting refunds, enables IRS to more cheaply collect late taxes from prior years. The magnitudes of the point estimates for no debt to IRS in Table 1.5 are smaller than for the full payment indicator because some taxpayers who do not pay in full by the deadline pay in full during the notice period that precedes addition to the debt data.

I assume that tax is paid eventually, and do not explicitly address the few taxpayers who may set withholding as low as possible with the goal of never paying the tax they owe. I find that treated taxpayers are not more likely to owe tax debts beyond one year after the deadline, which suggests that few taxpayers may attempt to never pay, or that the change I study may not affect them on the margin.

In the years after withholding was cut, treated taxpayers become *more* likely to pay on time relative to the control group, a reversal of the results for 2009. The magnitudes of the reversed effects are much smaller than the 2009 effects. The treated group are 0.2 percentage points more likely to pay in full in 2010, after having been 1.4 percentage points less likely to pay in full in 2009. This reversal is consistent with taxpayers requesting that their employers make adjustments to withholding or changing other behaviors to avoid the negative consequences of the withholding cut or of paying late. It is also consistent with a taxpayer preference for higher withholding, perhaps as a safeguard against the costs or frictions of making a direct payment.

For all three measures of late payments, the parallel pre-trends assumption is supported by data on the several years before treatment. While in some pre-treatment periods individual coefficients are statistically significant at the $p < 0.05$ level, no coefficient is statistically significant at the $p < 0.01$ level, and this may reflect the power the large sample size provides to detect even economically insignificant deviations from perfectly parallel pre-trends. The magnitudes of the pre-treatment coefficients are much smaller than the 2009 effects: for the fully paid indicator no pre-treatment coefficient is more than one tenth the magnitude of the 2009 coefficient, while for the dollar value and percentage of balance due underpaid the pre-treatment coefficients are no larger than approximately one fifth the size of the 2009 coefficient.

The estimated effects do not appear to be driven by differential exposure to the Great Recession between treated and control groups. The evidence from years before 2009 is not consistent with differences driven by recessions. During the previous recession in 2001, the treated group is if anything more likely to pay in full, not less, relative to the 2008 difference. The Great Recession took place over multiple years, including parts of 2007, 2008, and 2009. If the treatment group are especially likely to pay late when economic circumstances worsen, then one would expect that the treatment group's compliance would be substantially higher in relative terms in good years, with lower values in the 2007 and 2008 tax years. This is not what the figures show: the 2009 effect is isolated to a single year, with no evidence that late payments by treated and control taxpayers diverged in the earlier years of the Great Recession.

These results are robust to restricting the analysis to taxpayers who are married and file jointly or to households with fewer than three dependents in 2008. Restricting to married taxpayers filing joint returns, as in Table A.1, removes the concern that the estimated effects could be due to filing status differences between the treated and control groups. Including only households with two or fewer dependents in 2008, as in Table A.2, addresses the concern that the 2009 expansion of the EITC for those with three or more dependents could affect the results.

1.6 Withholding as Forced Saving and Automatic Payment

Withholding, or "pay as you earn", has two meanings. In one sense, paying as you earn means paying *when* you earn. Withholding forces payment on time, and can have similar consumption effects to forced saving. Forcing fully rational employees whose ability to borrow is limited by liquidity constraints to pay on time harms them because it overrides their preference. In another sense, paying as you earn means that you pay simply *by earning*. Payment happens automatically as employees earn income and the payment decision is out of their hands. Automatic payment may benefit taxpayers who would otherwise pay late because of frictions or costs, including hassle costs, time costs of making accurate payments, costly attention, or procrastination.

1.6.1 Liquidity Constraints and Withholding as Forced Saving

Additional withholding harms taxpayers subject to binding liquidity constraints, who are unable to borrow as much as they would like to finance current consumption. In a model of liquidity constraints, changing withholding only affects those taxpayers who cannot borrow at interest rates lower than the interest rate charged on late tax payments, and does not affect taxpayers who choose to save. This is still the case if taxpayers are allowed to can adjust the amount withheld subject to a lower limit.

Consider a simple model in which there are two periods. The first period includes the tax year up through the deadline for tax payment, while the second period is a later date at which late payments are made with interest. The taxpayer's utility is $U(C_1, C_2) = u(C_1) + \delta u(C_2)$. The taxpayer earns taxable income Y_1 during the tax year, which is subject to income tax $\tau(Y_1)$, and receives an untaxed bequest B_2 in the second period, which may lead the taxpayer to wish to borrow to smooth consumption.

Financial market frictions cap the taxpayer's ability to borrow at a fixed sum, which I set to zero without loss of generality. The taxpayer chooses savings $S \geq 0$, which earn a gross interest rate R_S.

An amount set by government policy H is withheld from the taxpayer's income in the first period, and credited against the taxpayer's tax liability. The taxpayer chooses an additional amount of tax to pay in the first period, $\Pi \geq 0$. Excess payments are refunded withhout interest, and tax that is not paid in the first period incurs interest at rate R_D. I assume $R_D > R_S$, as the interest charged on late tax payments exceeds interest on savings in practice. This immediately implies that the taxpayer will either save or pay taxes late, but not both, as such points are strictly inside the budget constraint. A taxpayer who both saves and pays late could engage in arbitrage by instead devoting the saved income to timely payment, earning $R_D - R_S > 0$ and increasing consumption in one or both periods.

If there were no financial frictions to limit borrowing, withholding an amount less than the tax liability due would have no effect in this model, and withholding in excess of the tax due would have only a negative income effect due to forgone interest. Private borrowing at rate $R_S < R_D$ would dominate late tax payment. The taxpayer's first order condition would only depend on withholding in excess of tax liability,

$$u'(C_1) = \delta R_s u'(R_S[Y_1 - \tau(Y_1) - C_1] + B_2 - [R_S - 1] \max\{0, H - \tau(Y_1)\}), \tag{1.2}$$

and only because of an income effect from forgone interest on the refund amount.

Returning to a model with liquidity constraints, the taxpayer cannot pay late if withholding exceeds liability, $H > \tau(Y_1)$, so to study late payment I consider the case where $H < \tau(Y_1)$. The intertemporal budget constraint consists of three segments, with two kink points. In the first period, the taxpayer's budget constraint is $C_1 \leq Y - S - H - \Pi$. Given the chosen levels of saving S and direct tax payments Π, the taxpayer's second-period budget constraint is $S \geq 0$ and:

$$C_2 \leq R_S S + B_2 - R_D(\max\{\tau(Y_1) - H - \Pi, 0\}) + \max\{H + \Pi - \tau(Y_1), 0\}. \tag{1.3}$$

The maximum possible second-period consumption occurs when the taxpayer pays the full tax bill on time and saves the remaining income, at the point where $C_1 = 0$ and $C_2 = R_S(Y_1 - \tau(Y_1)) + B_2$. At this point, the taxpayer can obtain an additional dollar of consumption in the first period by saving less, forgoing R_S in the second period. The budget constraint has constant slope $-R_S$ until the point at which consumption in the first period is $Y_1 - \tau(Y_1)$ and the remaining income is devoted to paying the full tax bill on time, with saving $S = 0$ and consumption in the second period $C_2 = B_2$. At this point there is a kink in the budget constraint, as further first-period consumption requires late tax payment and thus forgoing second-period consumption R_D. From this point the budget constraint is linear until first-period consumption is maximized by setting $S = 0$, $\Pi = 0$, and $C_1 = Y - H$, with second-period consumption $B_2 - R_D(\tau(Y) - H)$. At this point, further borrowing is prevented by financial frictions, and the budget constraint kinks again, becoming vertical until it connects to the horizontal axis.

In this model, withholding changes only affect taxpayers for whom the liquidity constraint is binding - it does not enter the utility function, and appears in the realized budget constraint only when preferences lead the taxpayer to make no additional non-withheld tax payment. The possible solutions occur at the two kink points and on the segments with slope $-R_S$ and $-R_D$, and depending on preferences, the interest rate charged on late tax payments, and the size of the second-period bequest relative to first period net of tax income. On the segment where the taxpayer saves, solutions occur with the standard tangency between the indifference curve and budget constraint, and satisfy

$$R_S = \frac{u'(C_1)}{\delta u'(B_2 + R_S[Y_1 - \tau(Y_1) - C_1])}. \tag{1.4}$$

At the kink point where the taxpayer neither saves nor pays late, the following inequality holds:

$$R_S \leq \frac{u'(Y_1 - \tau(Y_1))}{\delta u'(B_2)} \leq R_D. \tag{1.5}$$

On the segment where the taxpayer does not save and pays late, the solution satisfies:

$$R_D = \frac{u'(C_1)}{\delta u'(B_2 - R_D[\tau(Y_1) - Y_1 + C_1])}. \tag{1.6}$$

Finally, withholding does enter the solution condition when the liquidity constraint binds and the taxpayer would like to pay even the withheld tax late, where

$$R_D \leq \frac{u'(Y_1 - H)}{\delta u'(B_2 - R_D[\tau(Y_1) - H])}. \tag{1.7}$$

The comparative static with respect to an increase in withholding given that the taxpayer's full tax liability is not already withheld is as expected: withholding does not affect taxpayers for whom it is not a binding constraint, while it reduces the welfare of taxpayers for whom the liquidity constraint binds by restricting their ability to pay late, forcing first period consumption lower relative to second period consumption in contrast to their preferences. For taxpayers making any additional non-withheld tax payment, the additional payment falls one-for-one with the increase in withholding, with no further changes. For taxpayers choosing to make no non-withheld tax payment, $\frac{\partial C_1}{\partial H} = -1$, $\frac{\partial C_2}{\partial H} = R_D$, and the effect on utility is obtained by substituting into 1.7:

$$\frac{\partial U}{\partial H} = -u'(C_1) + \delta R_D u'(C_2) \leq 0. \tag{1.8}$$

1.6.1.1 When Taxpayers Can Adjust Withholding

Allowing taxpayers to adjust withholding subject to a lower limit (enforced, for example, by rules requiring employers to withhold at at least this level) does not substantially change the results of this model, except to recast it in terms of the lower limit $\bar{H}$ rather than the previous policy-fixed level H. The potential budget constraint given flexibility to choose $H \geq \bar{H}$ mirrors the above, except with the liquidity constraint binding at the lowest possible level of withholding $\bar{H}$.

Taxpayers who would pay non-withheld tax given withholding of $\bar{H}$ are indifferent to the level of withholding so long as it is below their preferred total first-period tax payments. Taxpayers who would like to pay less tax on time than the minimum level of withholding $\bar{H}$ are liquidity constrained, with the result that increasing the minimum level of withholding harms these taxpayers.

While this model omits the estimated tax penalty employees owe if withholding (and estimated tax payments made during the tax year) is too low relative to tax liability, adding the penalty does not substantially change the model. For example, if the penalty applied to the difference between withholding and tax liability, taxpayers would choose to set withholding equal to their total desired first-period tax payments to minimize the penalty, making no additional non-withheld tax payments. The estimated tax penalty would then be absorbed into a higher R_D charged on late payments.

The liquidity-constraint model does not explain why many taxpayers are overwithheld, receiving refunds following tax filing. Adding uncertain non-labor income not subject to withholding in combination with the estimated tax penalty for insufficient withholding could generate this result. Alternatively, taxpayers might choose to be overwithheld as a commitment device that forces saving, or due to inertia around a high default level of withholding, as in Jones (2012).

1.6.2 Frictions and Withholding as Automatic Payment

A wide variety of costs or decision-making frictions could explain why withholding reduces late payments. These costs or frictions should explain why taxpayers pay, but not on time, given the penalties and interest late payment incurs. One example is hassle costs that are convex in the time spent preparing and filing taxes, and procrastination is another example, though many other non-monetary costs or frictions would fit within this framework.

1.6.2.1 A General Model of Payment Frictions

Beginning with the model developed above, suppose instead that the taxpayer is not liquidity-constrained and can save and borrow freely at rate $R_S < R_D$. The taxpayer behaves as though making a direct tax payment Π costs not only the amount paid, but also an additional time-dependent amount $f_t(\Pi)$, with $f_t(0) = 0$ and $f_t'(\cdot) \geq 0$. This payment may include a fixed cost of making any positive payment, a variable cost component increasing in the amount of payment, or both. The taxpayer's per-period utility is linear in income, and late payments are repaid in the second period, $C_1 - f_1(\Pi) + \delta[C_2 - f_2(\tau(Y_1) - H - \Pi)]$.

The timely payment amount Π is chosen to maximize

$$-H - \Pi - f_1(\Pi) - \delta R_D \max\{\tau(Y_1) - H - \Pi, 0\} - \delta f_2(\tau(Y_1) - H - \Pi). \tag{1.9}$$

At an interior solution where the taxpayer makes an additional tax payment that is less than the full balance due,

$$-1 - f_1'(\Pi) + \delta R_D + \delta f_2'(\tau(Y_1) - H - \Pi) = 0, \tag{1.10}$$

or equivalently $\delta R_D - 1 = f_1'(\Pi) - f_2'(\tau(Y_1) - H - \Pi)$. If the cost is independent of timing,

$f_1(\cdot) = f_2(\cdot)$, but has an intercept of zero is convex and increasing in the amount paid, with $f'(\cdot) > 0$ and $f''(\cdot) > 0$, then the result will be an interior solution.

Alternatively, if the friction is a fixed cost of making any payment regardless of the amount paid, then taxpayers will either pay the full amount on time or the full amount late. The corner solution where the taxpayer pays the full remaining balance due on time has value $-\tau(Y_1) - f_1(\tau(Y_1) - H)$, while the corner solution where the taxpayer makes no additional payment on time has value $-H - R_D(\tau(Y_1) - H) - f_2(\tau(Y_1) - H)$. Taking the difference gives $[R_D - 1](\tau(Y_1) - H) + f_2(\tau(Y_1) - H) - f_1(\tau(Y_1) - H)$. Delaying payment is worth the interest paid if making the payment on time results in a friction cost sufficiently high relative to making the payment late. If $f(\cdot)$ has both a fixed and variable component, then the optimum may lie at either the interior solution or at one of the corners, depending on parameter values.

Late payment occurs because payment costs or frictions are higher at filing time than afterward. There are many possible interpretations of $f(\cdot)$; it could, for example, be a fixed time or mental effort cost of making the payment, which combined with being busier during filing season than afterwards makes waiting until after the deadline to pay more attractive. Procrastination, another potential explanation, makes the present-period costs relevant for the decision of when to pay appear larger. The monetary of making the payment e.g. by check, online transfer, or wire transfer is likely small and unlikely to vary over time, absent liquidity constraints.

1.6.2.2 Hassle Costs of Tax Filing and Payment

Withholding can generate welfare gains by, for example, eliminating the hassle costs of making payments. Making payments on time involves much higher hassle costs than paying late because paying on time requires additional time and effort on top of the time and hassle of filing a return, while paying late allows the taxpayer to make a payment when time and effort are less costly. Taxpayers subject to high hassle costs at filing time spend less time not only on payment and verifying the payment amount, but also on verifying other items on the tax return. Higher hassle costs could be higher for higher-income taxpayers, whose opportunity cost of time is higher, and who thus are more likely to respond to a change in withholding.

Suppose that the hassle of spending time and effort to file and pay taxes is convex in the fraction of a time period spent on tax filing and payment. The hassle cost function is $\psi(e)$, with $\psi(0) = 0$, $\psi'(e) > 0$, and $\psi''(e) > 0$. The marginal cost of the additional effort spend to figure out the right payment amount and make the payment is higher before the tax filing and payment deadline in April because of the effort already exerted around this time to file a tax return. Delaying payment until after the deadline costs interest and penalties, but reduces the hassle cost of the effort payment requires. Specifically, suppose that filing requires effort $e_f > 0$, and payment requires effort $e_p > 0$. Then, the convexity of $\psi(e)$ implies $\psi(e_f + e_p) > \psi(e_f) + \psi(e_p)$.

If less than the full tax liability is withheld, the taxpayer bears a hassle cost of making the payment. The taxpayer chooses whether to exert the effort e_p to make a payment when filing, avoiding penalties and interest, or after filing, with penalties and interest on the tax paid late. The taxpayer chooses to pay late when

$$\psi(e_f) + \psi(e_p) + [R_D - 1](\tau(Y_1) - H) < \psi(e_f + e_p). \tag{1.11}$$

If the hassle cost of making a payment does not depend on the payment amount, increased withholding reduces the interest cost of paying late, conditional on needing to make a payment. This effect of withholding can make late payment more common.

Hassle costs of this form imply that the same taxpayers will pay late and make mistakes in filing. If some taxpayers have especially high costs of time or effort exerted when filing, for example, then similar logic applies both to their decision to pay late and to their decision to cease expending additional effort to verify that the tax return is correct. Let the expected cost of inaccuracies on the tax return given filing effort e_f be $i(e_f)$, with $i'(\cdot) < 0$ and $i''(\cdot) > 0$, so that effort has decreasing marginal returns. Then the first order condition for filing effort sets $\psi'(e_f) = i'(e_f)$. If taxpayer j has an attribute, for example a complicated tax situation, that shifts the effort cost of filing upward by a_j, then $\psi'(e_f + a_j) = i'(e_f)$ yields a lower e_f and higher mistakes, captured by a lower value of $i(\cdot)$. The same attribute would also make late payment more likely due to the convexity of ψ, which increases the savings from paying late:

$$\psi(e_f + a_j + e_p) - \psi(e_f + a_j) - \psi(e_p) > \psi(e_f + e_p) - \psi(e_f) - \psi(e_p). \tag{1.12}$$

Hassle costs thus generate the result that the same people will make mistakes on their returns and pay late.

Hassle costs are likely higher for higher-income taxpayers, who tend to have more complex tax situations, increasing both the effort required to file e_f and the effort cost of determining and making the correct payment e_p. Higher wage rates imply a higher opportunity cost of time, so that even a similar time investment in accurate payment imposes a larger burden on higher-wage and therefore higher-income taxpayers. Higher-income taxpayers may also have larger balances due, even though withholding increases with the tax liability due on wages, because of higher income from sources not subject to withholding, although this can be offset by adjusting withholding upward or by making larger estimated tax payments.

Higher hassle costs for higher-income taxpayers imply that for a given balance due they are more likely to pay late, and withholding that eliminates the need to make a payment will have a larger effect for higher-income taxpayers. Withholding can have substantial benefits if it removes hassle costs of payment altogether, as such costs indicate resources otherwise devoted to making payments.

1.6.2.3 Procrastination

Procrastination can lead taxpayers to pay late. Present bias or simple impatience overvalues the costs of paying today relative to the costs and benefits of paying tomorrow. The costs of making a timely payment include penalties and interest, but also include the time and effort involved in determining the payment due and to make the payment and psychological costs due to loss aversion. If enough is withheld that no additional payment is due, then withholding completely removes the opportunity for procrastination to interfere with making tax payments.

Consider a naive present-biased taxpayer deciding when to pay. Let R_S be the gross interest rate on private saving or borrowing, H the amount withheld, and $\tau(Y)$ the tax due. For simplicity, assume that payment is a binary choice in each period between paying the full balance due $\tau(Y) - H$ and making no payment. To make late payment possible, assume that $H < \tau(Y)$. In addition to the monetary cost, payment also costs time, effort, and the psychological cost due to loss aversion, which total $\psi(\tau(Y) - H)$. After the deadline at time $t = 0$, failure to pay incurs expected sanctions $R_D^{max\{0,t\}}[\tau(Y) - H)] + S(t)$ from the tax authority, which include not only interest and penalties, but also utility costs of anxiety induced by both ever-sterner admonitions to pay and the possibility that assets will be subject to a lien, levy, or seizure. The taxpayer discounts the future quasi-hyperbolically with discount factor δ and additional discount factor β applied to periods after the present.

Taxpayers deciding when to pay trade off the additional costs of paying today rather than tomorrow against the additional benefits of paying today. The taxpayer pays at the first t such that

$$- \psi(R_D^{max\{0,t\}}[\tau(Y) - H)]) - R_D^{max\{0,t\}}[\tau(Y) - H)] \geq \tag{1.13}$$

$$- \beta\delta\Big[\psi(R_D^{max\{0,t+1\}}[\tau(Y) - H)]) + (1/R_S)R_D^{max\{0,t+1\}}[\tau(Y) - H)] + S(t)\Big]. \tag{1.14}$$

Absent procrastination and assuming that the return to saving is lower than R_D, taxpayer behavior is governed by intertemporal substitution given discount rate δ, leading many taxpayers to pay at the deadline, maximizing the time value of delay, and some taxpayers with high values of δ to pay late. Taxpayers are especially likely to pay at times when the expected sanctions for continued failure to pay increase.

Procrastination leads more taxpayers to pay after the deadline, and by driving a wedge between the utility function they use when making the payment decision, which discounts at rate $\beta\delta$, and the utility they ultimately experience, which discounts at rate δ, procrastination makes taxpayers whose behavior it changes worse off.

Withholding changes the relative costs of payment today and tomorrow. Differentiating the net benefit of payment today with respect to H gives

$$\psi'(R_D^{max\{0,t\}}[\tau(Y)-H]) - \beta\delta\psi'(R_D^{max\{0,t+1\}}[\tau(Y)-H]) + R_D^{max\{0,t\}} - \beta\delta(1/R_S)R_D^{max\{0,t+1\}}. \tag{1.15}$$

Assume $\psi'(\cdot)$ is constant. Then, before the deadline, withholding can make earlier payment more likely because it reduces the non-monetary costs of making payments - the effect of withholding on the net benefit of paying today is $\psi'(\cdot)[1-\beta\delta]+1-\beta\delta(1/R_S)$. Withholding also reduces the incentive to pay at the deadline to maximize interest. At the deadline, withholding's effect on the net benefit is $\psi'(\cdot)[1-\beta\delta]+1-\beta\delta(R_D/R_S)$, which again can induce payment on time through non-monetary costs of payment. Withholding could also induce later payment by reducing the interest cost of paying late. The effect of increasing withholding is potentially ambiguous for taxpayers who still owe payments, but the effect if withholding increases enough that no additional payment is due is unambiguous, preventing late payment altogether and removing the non-monetary cost ψ of making payments.

These models make different predictions about who pays late, which I test empirically. The liquidity constraint model implies that a change in withholding only affects taxpayers who do not save in liquid assets. The frictions model implies that withholding disproportionately affects taxpayers with high actual or perceived cost of time and effort before the deadline relative to after the deadline, who will be especially likely to make errors before the deadline along several dimensions of choice. Some taxpayers may be affected by both liquidity constraints and costs of making payments directly. These taxpayers would attenuate the predicted patterns of behavior across taxpayers, blurring the sharpness of the predictions the models make.

1.7 Empirics: Who Pays Late and Why

To determine *why* people pay late, I test conflicting predictions about *who* pays late. I test the liquidity-constraint explanation for late payment using heterogeneity in the effect of the 2009 policy change by liquidity, and the payment-friction explanation for late payment by examining the relationship between late payment and errors on tax returns. The estimated patterns of response support a friction-driven rather than liquidity-constraint explanation for late payments. The estimated effects are no smaller, and if anything are larger, among taxpayers earning interest income, a proxy for liquidity. Late payers are especially likely to make errors on their returns that leave money on the table or underreport verifiable income.

1.7.1 Heterogeneity Specification

I study which taxpayers are more responsive to a reduction in withholding by interacting the coefficient on the time period-by-treatment interaction with a variable that captures heterogeneity before treatment, for example by whether the taxpayer earned at least \$100 of interest income in tax year 2008, and also interact the per-period effects of the baseline coefficients with the measure of heterogeneity. The resulting specification is

$$y_{it} = \alpha_i + \sum_t \eta_t ERP_i * int_{i,2008} + \sum_t \gamma_t X_i * int_{i,2008} + \varepsilon_{it}. \tag{1.16}$$

The notation is the same as the main specification, but the coefficients of interest are now the difference-in-difference-in-difference coefficients on the interaction between Economic Recovery Payment receipt and interest income (or another measure of heterogeneity), η_t. These coefficients quantify the difference in treatment effects between those taxpayers for whom the measure of heterogeneity takes on a value of one and a value of zero. These interaction coefficients capture the difference between causal effects of treatment for sub-populations defined by e.g. interest income, not the causal effect of changes in interest income.

1.7.2 Reducing Withholding Causes Taxpayers Earning Interest Income to Pay Late

I test the theory that late payment is due to liquidity constraints by examining heterogeneous treatment responses across access to liquid assets. In a model in which taxpayers are fully rational, taxpayers choose to pay taxes late because the opportunity costs of paying taxes on time exceed the interest and penalties due on late tax payments, and only taxpayers who do not hold low-interest assets pay late in response to withholding changes. This prediction is not consistent with the results of interacting measures of available funds with treatment, in which taxpayers with available funds are no less likely to be affected by the withholding cut as taxpayers without available funds. Instead, the late payment response among taxpayers with interest income is slightly larger, which could be the result of higher hassle costs for higher-income taxpayers with more complex tax situations and higher opportunity costs of time.

Although the tax data do not include a direct measure of low-interest asset holdings at filing time, I construct a variety of proxies for assets from data on the tax return. I combine taxable and tax-exempt interest income, then create an indicator variable equal to one if total interest income is greater than \$100 in 2008, which is my preferred proxy for low-interest asset holdings. Given the low-interest environment at this time, earning \$100 of 2008 interest income implies holding principal many times larger than the \$250 withholding reduction. In alternate specifications reported in Tables A.4-A.7 in the appendix, I use alternative measures of interest income, including total interest income greater than \$500, total interest income greater than zero, total interest income greater than zero in all tax years from 2005-2008, and taxable dividend income greater than \$100. In no case is the effect significantly smaller among those with higher liquidity than among those without.

Taxpayers with interest income respond *more* strongly to treatment, which is inconsistent with the liquidity-constraint explanation for late payment and could instead be due to higher hassle costs for higher-income taxpayers with more complex tax situations or to higher costs of the time spent checking payment accuracy. A reduction in withholding makes taxpayers with interest income greater than \$100 in 2008 even less likely to pay in full than those without by one percentage point, as reported in Table 1.6 and shown in Figure 1.6. Taxpayers with interest income leave an additional \$20 of their balances due not paid in 2009, shown in Figure 1.7, while there is no significant difference in the percentage of the balance due underpaid by interest income, as shown in Figure 1.8. The parallel pre-trends assumption for this triple-difference specification is supported by the fact that there is no statistically significant effect in any pre-treatment year for the dollar value underpaid, but there are pre-treatment effects in some years for the fully paid indicator and the percentage of the balance due underpaid. The interaction coefficients in 2004 and 2007 for the fully paid indicator are positive and strongly statistically significant, although with magnitudes less than half the one-percentage-point decline in 2009.

1.7.3 Late Payers Make More Errors on Their Returns

The same taxpayers who pay late are more likely to make various errors on their tax returns, lending support to the notion that the mechanism behind late payment is a friction that leads taxpayers to err along multiple dimensions due, for example, to limited time, lack of information, inattention, or procrastination. I focus on relatively common errors related to the Making Work Pay payroll tax credit. Millions of taxpayers made such errors in 2009.

In both the treated and control groups, late payers were more likely to make some error related to the Making Work Pay credit, as Table 1.7 shows. In the control group, the difference is 11 percent of late payers with some error versus 6 percent of those who paid in full, while in the treated group the difference is 37 percent among those who paid late versus 10 percent of those who paid in full.

Many taxpayers did not claim the 2009 Making Work Pay credit despite the fact that the credit of up to $800 could be claimed using a single page form and with information mostly reported elsewhere on the return[7]. Among the control group, 9 percent of those who did not pay on time did not claim the credit, while 5 percent of those who did pay on time did not claim the credit. Among the treated group, seven percent of those who paid late did not claim the credit, while five percent of those who paid on time did not claim.

Those taxpayers receiving an Economic Recovery Payment often do not indicate payment receipt when claiming the credit, a $250 error in their own favor. When IRS detected this error, it sent taxpayers either a smaller refund or a request for additional payment, which may lead to late payment. Among those who did not pay on time in the treated group, 19 percent did not correctly report their Economic Recovery Payment on their return, while two percent of those who paid on time in the treated group did not report the Economic Recovery Payment.

The higher error rates among taxpayers failing to pay on time are consistent with the notion that frictions such as lack of time, information, or attention make taxpayers more likely to make errors in general, and that late payment may be one result of these frictions.

1.8 Policy Implications

The results in this paper inform tax withholding policy, which is currently being reshaped to reflect changes in the 2017 Tax Cuts and Jobs Act. Withholding policy involves a trade-off between withholding too much tax from some taxpayers and too little tax from others. The current withholding tables do not have sufficient flexibility to withhold exactly for all taxpayers because they depend only on wages, a limited set of allowances and marital status. This paper suggests that on the margin there are more benefits than costs of withholding slightly more. The Tax Cuts and Jobs Act of 2017 will reshuffle taxpayers' tax liability relative to past levels of withholding, which may lead more taxpayers to owe balances due and therefore to pay late. Consistent with this concern, the IRS is currently engaged in an outreach campaign encouraging taxpayers to adjust withholding, and has created a withholding

[7]As the returns came in, in light of several million taxpayers not claiming the credit and the desire to distribute the credit as economic stimulus, the IRS decided to credit these taxpayers with the credit amount they would have received based on information elsewhere on their returns and administrative data on payment receipt.

calculator on its website.

Cutting withholding, holding tax liability constant, might appear to provide economic stimulus when it is needed without a cost in tax revenue. A 1992 withholding cut was meant to provide economic stimulus at no change in tax liability. Shapiro and Slemrod (1995) find that forty-three percent of consumers surveyed said that they would spend most of the income this policy shifted forward from their refunds or balances due to their paychecks. The 1992 policy increased balances due, which this paper suggests led to both additional administrative costs and burdens on taxpayers. The 2009 stimulus, in contrast, aimed to reduce taxpayers' tax liability and withholding by the same amount, and increases in balances due like those I study were unintentional. Providing stimulus by reducing tax liability alongside withholding does not change the balances taxpayers owe at filing, avoiding administrative and taxpayer burdens at a cost in tax revenue.

My results are likely to generalize to other marginal changes in withholding. Sahm, Shapiro, and Slemrod (2012) find that spending responses to stimulus policies are larger to one-time payments than increased paychecks, so one might expect that the withholding reduction via one-time payment could produce more liquidity-driven late payments than a withholding reduction implemented through increased paychecks. The Great Recession could also increase withholding's impact on late payments by tightening liquidity constraints. However, I do not find that liquidity constraints drove late payments after a 2009 policy change. Treated taxpayers who are on average older and mostly receive Social Security retirement benefits could be unusually vulnerable to frictions, but the impact is larger for treated taxpayers who instead received Supplemental Security Income or veterans' disability/survivorship benefits and are likely younger[8], perhaps because retirees have more time to devote to ensuring accurate payment.

More caution is due when generalizing these results to larger changes in withholding. While I find that the effects of a marginal change in withholding are due to frictions rather than liquidity, larger increases in withholding would be more likely to result in binding liquidity constraints, so larger changes in withholding could work through a different mechanism. Frictions are large relative to the consequences of failing to optimize in response to a small change in withholding, but would be smaller relative to the effects of failing to adjust after a large change in withholding.

[8]See Table A.3.

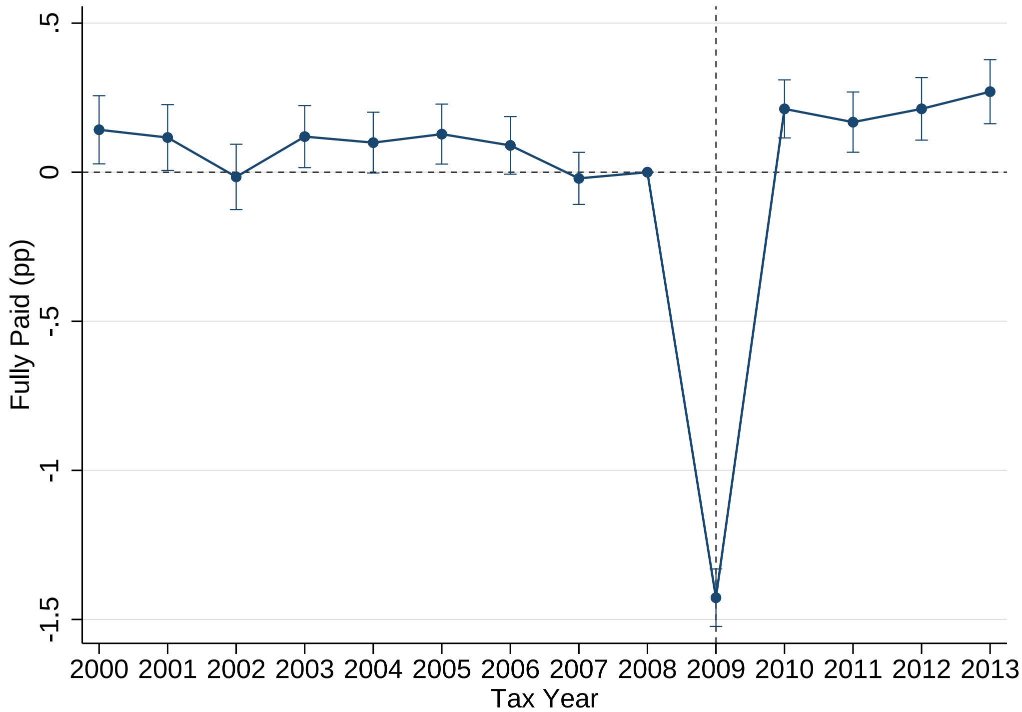
1.9 Conclusion

I find that if employers withhold less tax, and people are therefore responsible for paying more tax directly, late payments rise. An interaction between two 2009 stimulus policies, a payroll tax credit and direct payments to retirement and disability beneficiaries, provides identifying variation in withholding. Using a difference-in-differences strategy, I show that late payments increase by an average of five percent of the amount of a withholding cut. In a sample that excludes people due a refund, who cannot pay late, the estimated effects are four times larger. By cutting withholding for more than 5.7 million households, the policy I study caused an additional 80,000 households to pay taxes late. Taxpayers whose withholding is cut by $250 are more likely to owe late taxes until one year later, when the following year's refunds provide the IRS with an opportunity to collect.

While in theory liquidity constraints could explain late payment, I find instead that people pay late because making tax payments directly creates hassle costs or frictions. If people pay late because of liquidity constraints, then changing withholding should have less of an effect on taxpayers not subject to liquidity constraints, yet I find that taxpayers earning interest income respond at least as much to a change in withholding, contrary to the liquidity constraint explanation. Frictions including procrastination and limited time or attention would suggest that people who pay late also make more concurrent errors. I find that people who pay late are more likely to file incorrect tax returns, either leaving a tax credit on the table or not reporting government-provided income.

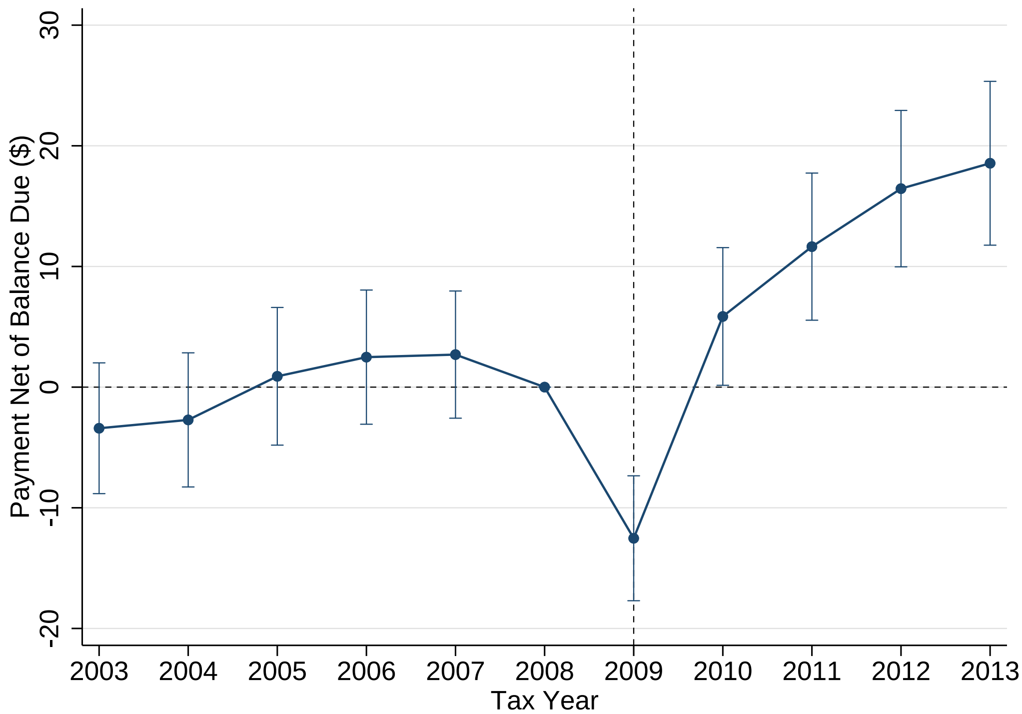
On the margin, increased withholding has two substantial benefits. Greater withholding reduces the number of taxpayers who pay late, saving the administrative costs incurred to collect the taxes they owe. Withholding also reduces the costs taxpayers bear when responsible for making tax payments, whether due to hassle costs or frictions.

Figure 1.1: Difference-in-Difference Estimate of Effect of Reducing Withholding by $250 on Probability Fully Paid (Percentage Points)



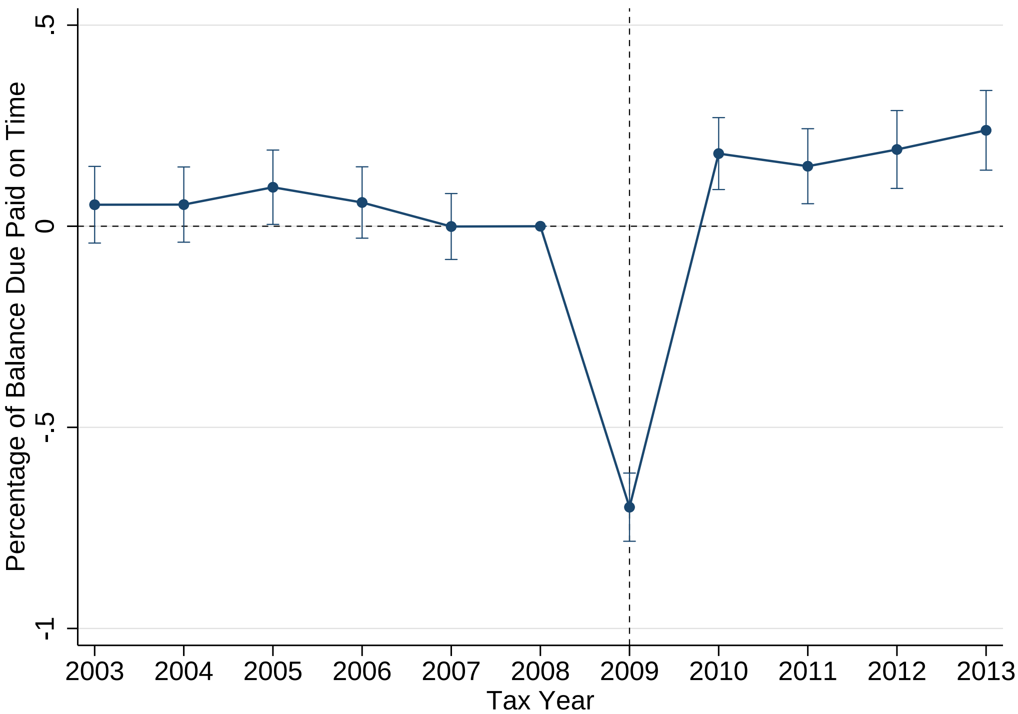
Notes: Outcome is an indicator for paying the full balance due with a return. Plots coefficients and 95% confidence intervals (where standard errors are clustered at the taxpayer level) from a regression with taxpayer fixed effects and interactions between year fixed effects and 2008 age, marital status, spouse's age, and dependents.

Figure 1.2: Difference-in-Difference Estimate of Effect of Reducing Withholding by $250 on Payment Net of Balance Due ($)



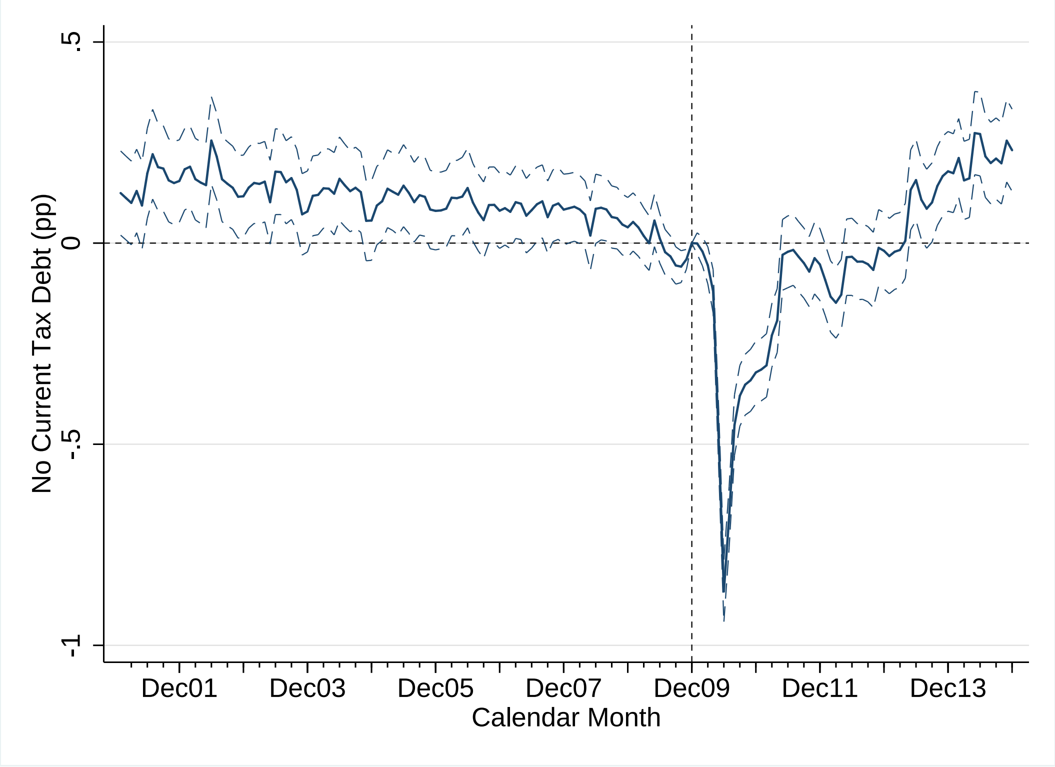
Notes: Outcome is the dollar amount paid with a return minus the balance due. Plots coefficients and 95% confidence intervals (where standard errors are clustered at the taxpayer level) from a regression with taxpayer fixed effects and interactions between year fixed effects and 2008 age, marital status, spouse's age, and dependents.

Figure 1.3: Difference-in-Difference Estimate of Effect of Reducing Withholding by $250 on Percentage of Balance Due Paid



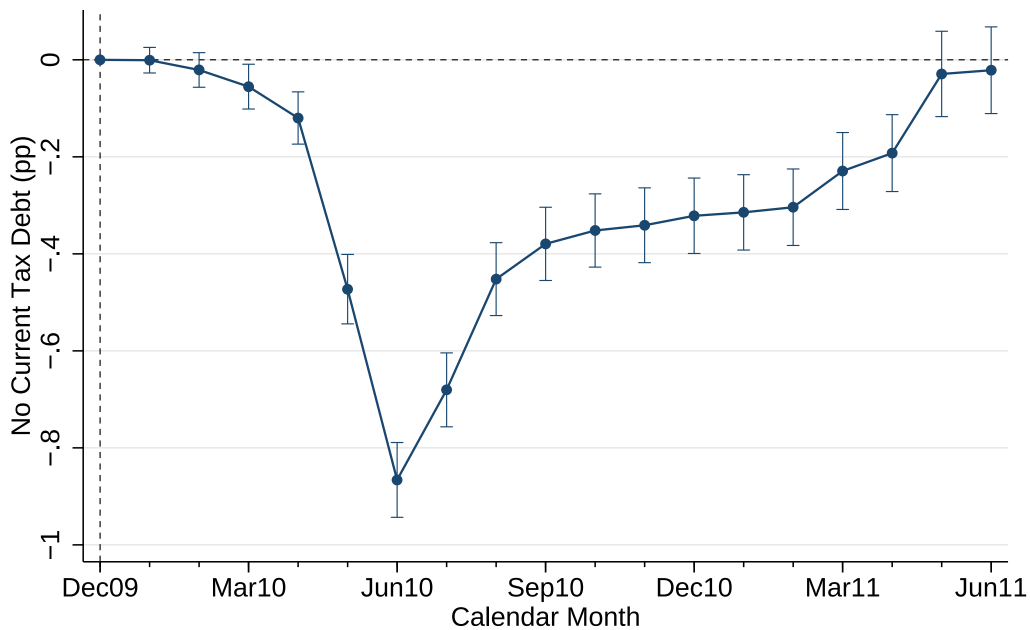
Notes: Outcome is the percentage of the balance due paid with the return. Plots coefficients and 95% confidence intervals (where standard errors are clustered at the taxpayer level) from a regression with taxpayer fixed effects and interactions between year fixed effects and 2008 age, marital status, spouse's age, and dependents.

Figure 1.4: Difference-in-Difference Estimate of Effect of Reducing Withholding by $250 on Probability of No Current Tax Debt (Percentage Points)



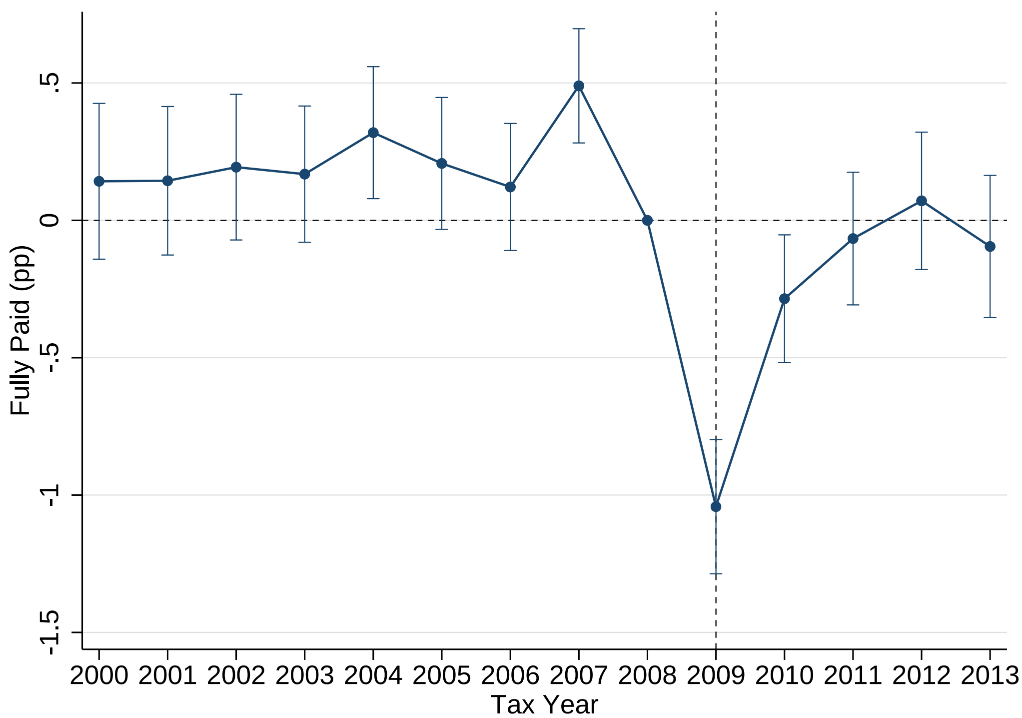
Notes: Outcome is an indicator for owing no current unpaid assessed tax. Plots coefficients and 95% confidence intervals (where standard errors are clustered at the taxpayer level) from a regression with taxpayer fixed effects, interactions between time fixed effects and 2008 age, marital status, spouse's age, and dependents, and treatment-by-month-of-year fixed effects estimated from the pre-treatment period.

Figure 1.5: Difference-in-Difference Estimate of Effect of Reducing Withholding by $250 on Probability of No Current Tax Debt (Percentage Points), Detail



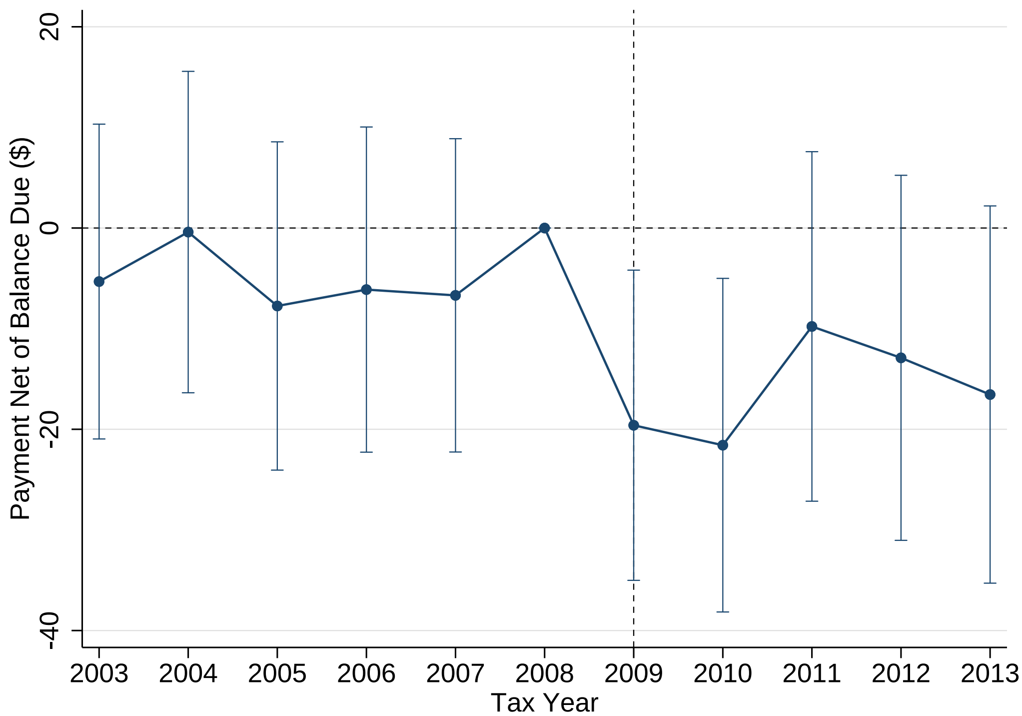
Notes: Outcome is an indicator for owing no current unpaid assessed tax. Plots coefficients and 95% confidence intervals (where standard errors are clustered at the taxpayer level) from a regression with taxpayer fixed effects, interactions between time fixed effects and 2008 age, marital status, spouse's age, and dependents, and treatment-by-month-of-year fixed effects estimated from the pre-treatment period.

Figure 1.6: Difference-in-Difference Estimate of Effect of Reducing Withholding by $250 on Probability Fully Paid (Percentage Points): Coefficient on Interaction with Interest Income Greater Than $100 in 2008



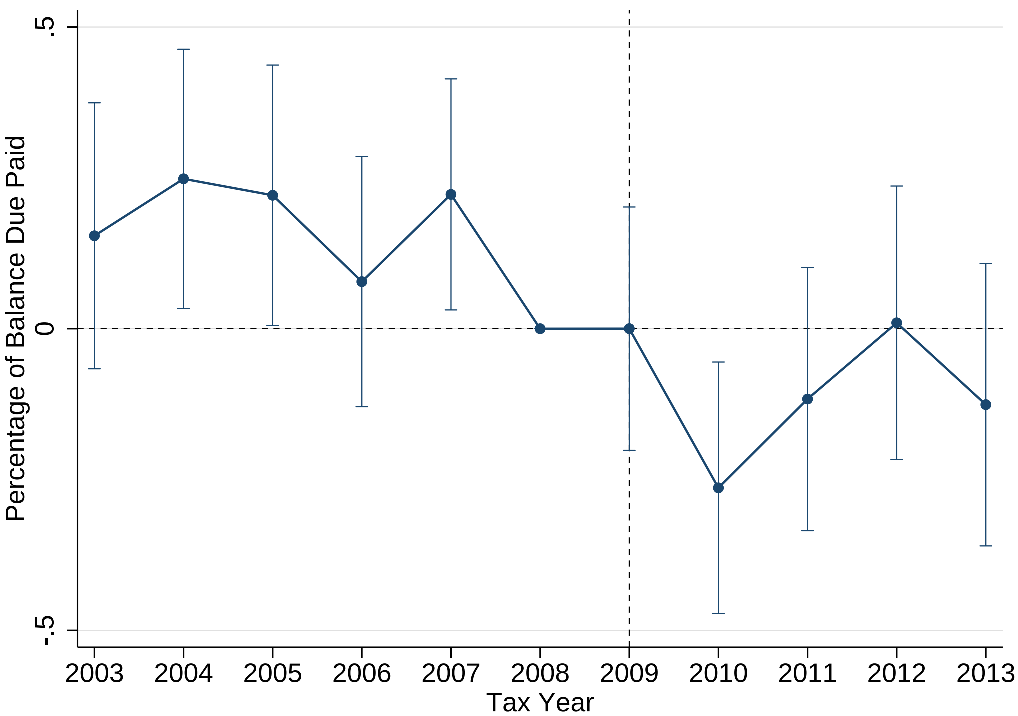
Notes: Outcome is an indicator for paying the full amount due with a return. Plots the interaction coefficients between treatment, the interest income indicator, and tax year, and 95% confidence intervals (where standard errors are clustered at the taxpayer level) from a regression with taxpayer fixed effects, interactions between time fixed effects, the interest income indicator and 2008 age, marital status, spouse's age, and dependents, and treatment-by-month-of-year fixed effects estimated from the pre-treatment period.

Figure 1.7: Difference-in-Difference Estimate of Effect of Reducing Withholding by $250 on Payment Net of Balance Due ($): Coefficient on Interaction with Interest Income Greater Than $100 in 2008



Notes: Outcome is the dollar amount paid with a return minus the balance due. Plots the interaction coefficients between treatment, the interest income indicator, and tax year, and 95% confidence intervals (where standard errors are clustered at the taxpayer level) from a regression with taxpayer fixed effects, interactions between time fixed effects, the interest income indicator and 2008 age, marital status, spouse's age, and dependents.

Figure 1.8: Difference-in-Difference Estimate of Effect of Reducing Withholding by $250 on Percentage of Balance Due Paid: Coefficient on Interaction with Interest Income Greater Than $100 in 2008



Notes: Outcome is the percentage of the balance due paid with the return. Plots the interaction coefficients between treatment, the interest income indicator, and tax year, and 95% confidence intervals (where standard errors are clustered at the taxpayer level) from a regression with taxpayer fixed effects, interactions between time fixed effects, the interest income indicator and 2008 age, marital status, spouse's age, and dependents.

Table 1.1: Summary Statistics before Treatment by Treatment Group

	Control	Treated
Single Filer (pp)	37.0	29.0
Married Filing Jointly (pp)	59.1	68.8
Age	46.20	58.04
Dependents Claimed	0.833	0.361
Wages ($)	59,047	41,265
Adjusted Gross Income ($)	67,802	60,763
Social Security (pp)	0.249	67.1
Total Interest Income ($)	832.7	1437.9
Positive Total Interest Income (pp)	50.5	58.3
Refund (pp)	84.0	77.0
Balance Due (Refund if Negative) ($)	-1,953.0	-1,187.6
Fully Paid on Time (pp)	95.52	95.41
Filed on Time (pp)	95.20	96.37
Payment Net of Balance Due ($)	-132.9	-119.0
Percentage of Balance Due Paid	95.95	96.01
Balance Due if Positive ($)	2,562.1	2,306.3
Refund if Positive ($)	2,874.5	2,304.5
Full Paid (pp) if Balance Due	72.28	80.12
Payment Net of Balance Due ($) if Balance Due	-822.8	-515.6
Random Sample Size	1%	10%
N	478,923	567,778

Means reported. (pp) indicates percentage points. Tax year 2008 variables unless otherwise indicated. Treated and control groups have 2008 wages between $4033 and $82,500 ($177,500 if married filing jointly) that would, if earned in 2009, qualify for a Making Work Pay payroll tax credit of at least $250. Treated units received exactly one Economic Recovery Payment of $250 in April or May 2009; control units did not. This payment reduced the treated group's credit by up to $250, close to the observed 2009 difference of $194. The difference is due to 2008-2009 wage changes and because the credit's definition of earned income differs somewhat from wages. Treatment groups are determined on the basis of 2008 wages because 2009 wages may respond to the policy change. Most treated units receive either exactly the maximum possible single credit of $150 or joint credit of $550, reduced by $250 from $400 ($800). As both groups received a withholding reduction in line with the full credit, this implies that the policies resulted in larger 2009 balances due (or smaller refunds) for the treated group than otherwise. I restrict to households in which tax filer(s) are between 30 and 70 years of age in 2008 and the primary filer is not deceased. Single and married filing jointly indicators do not add to one because some taxpayers are either married filing separately or filing as head of household. Dependents claimed are non-spouse dependents under age 25. Dollar amounts are winsorized at the 1st and 99th percentiles.

Table 1.2: Making Work Pay Tax Credits by Treated Group

	Control	Treated
2009 Making Work Pay Credit ($)	545.9	351.2
Exactly $150 of 2009 Making Work Pay Credit (pp)	0.282	16.0
Exactly $550 of 2009 Making Work Pay Credit (pp)	0.164	47.3
Random Sample Size	1%	10%
Households	483,291	577,901

Means reported. Treated and control groups have 2008 wages between $4033 and $82,500 ($177,500 if married filing jointly) that would, if earned in 2009, qualify for a Making Work Pay payroll tax credit of at least $250. Treated units received exactly one Economic Recovery Payment of $250 in April or May 2009; control units did not. This payment reduced the treated group's credit by up to $250, close to the observed 2009 difference of $194. The difference is due to 2008-2009 wage changes and because the credit's definition of earned income differs slightly from wages. Treatment groups are determined on the basis of 2008 wages because 2009 wages may respond to the policy change. Most treated units receive either exactly the maximum possible single credit of $150 or joint credit of $550, reduced by $250 from $400 ($800). As both groups received a withholding reduction in line with the full credit, this implies that the policies resulted in larger 2009 balances due (or smaller refunds) for the treated group than otherwise. I restrict to households in which tax filer(s) are between 30 and 70 years of age in 2008 and the primary filer is not deceased.

Table 1.3: Difference-in-Difference Estimates of Effects of $250 Withholding Reduction

	(1) Fully Paid (pp)	(2) Payment Net of Balance Due ($)	(3) Percentage of Balance Due Paid
2000 * Treated	0.142**		
	(0.0583)		
2001 * Treated	0.116**		
	(0.0563)		
2002 * Treated	-0.0158		
	(0.0559)		
2003 * Treated	0.119**	-3.407	0.0535
	(0.0531)	(2.765)	(0.0486)
2004 * Treated	0.0991*	-2.716	0.0538
	(0.0520)	(2.836)	(0.0477)
2005 * Treated	0.128**	0.897	0.0970**
	(0.0513)	(2.911)	(0.0472)
2006 * Treated	0.0899*	2.486	0.0591
	(0.0493)	(2.837)	(0.0453)
2007 * Treated	-0.0208	2.697	-0.000642
	(0.0446)	(2.689)	(0.0418)
2008 * Treated	(omitted)	(omitted)	(omitted)
2009 * Treated	-1.427***	-12.53***	-0.698***
	(0.0492)	(2.641)	(0.0432)
2010 * Treated	0.212***	5.856**	0.181***
	(0.0497)	(2.911)	(0.0456)
2011 * Treated	0.168***	11.64***	0.149***
	(0.0516)	(3.111)	(0.0476)
2012 * Treated	0.212***	16.45***	0.191***
	(0.0535)	(3.306)	(0.0494)
2013 * Treated	0.270***	18.55***	0.238***
	(0.0548)	(3.463)	(0.0505)
F-stat Pre-2008	3.278	1.649	1.202
p-value Pre-2008	<0.0001	0.143	0.305
F-stat Post-2008	301.6	23.68	112.6
p-value Post-2008	<0.0001	<0.0001	<0.0001
Household FE	Yes	Yes	Yes
Year FE *Demographics	Yes	Yes	Yes
Years	14	11	11
Household-Years	13,782,245	10,904,992	10,904,992
R-Squared	0.000142	0.0000186	0.0000712

Notes: standard errors (in parentheses) clustered by household. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Includes interactions between year fixed effects and 2008 age fixed effects, marital status, spouse's age, and dependents. Payment amount data begin in 2003. Payment net of balance due winsorized at the 1st and 99th percentiles. F-statistics report joint statistical significance of all pre-2008 or post-2008 coefficients.

Table 1.4: Difference-in-Difference Estimates of Effects of $250 Withholding Reduction Conditional on Owing a Balance Due

	(1) Full Paid (pp) if Balance Due	(2) Payment Net of Balance Due ($) if Balance Due
2000 * Treated	-0.395**	
	(0.193)	
2001 * Treated	-0.633***	
	(0.198)	
2002 * Treated	-0.855***	
	(0.195)	
2003 * Treated	-0.433**	-14.88
	(0.199)	(14.37)
2004 * Treated	-0.605***	-23.80*
	(0.185)	(13.61)
2005 * Treated	-0.247	-14.90
	(0.184)	(13.79)
2006 * Treated	-0.407**	-14.18
	(0.180)	(13.59)
2007 * Treated	0.0225	-8.327
	(0.172)	(12.94)
2008 * Treated	(omitted)	(omitted)
2009 * Treated	-5.647***	-49.76***
	(0.188)	(13.44)
2010 * Treated	-0.216	-17.30
	(0.181)	(13.88)
2011 * Treated	-0.105	1.752
	(0.180)	(14.25)
2012 * Treated	-0.157	2.849
	(0.179)	(14.28)
2013 * Treated	-0.00653	4.505
	(0.180)	(14.68)
F-stat Pre-2008	4.874	0.651
p-value Pre-2008	<0.001	0.661
F-stat Post-2008	276.9	4.482
p-value Post-2008	<0.001	<0.001
Tax Unit Fixed Effects	Yes	Yes
Year FE * Demographics	Yes	Yes
Years	14	11
Tax Unit-Years	2729742	2118187
R-Squared	0.000647	0.0000153

Notes: standard errors (in parentheses) clustered by household. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Includes interactions between year fixed effects and 2008 age fixed effects, marital status, spouse's age, and dependents. Payment amount data begin in 2003. Payment net of balance due winsorized at the 1st and 99th percentiles. F-statistics report joint statistical significance of all pre-2008 or post-2008 coefficients.

Table 1.5: Difference-in-Difference Estimates of Effect of $250 Withholding Reduction on Monthly Tax Debt Status

Calendar Month	No Current Tax Debt (pp)	
December 2009 * Treated	(omitted)	(omitted)
January 2010 * Treated	-0.000782	(0.0135)
February 2010 * Treated	-0.0209	(0.0182)
March 2010 * Treated	-0.0553**	(0.0236)
April 2010 * Treated	-0.120***	(0.0275)
May 2010 * Treated	-0.473***	(0.0365)
June 2010 * Treated	-0.866***	(0.0393)
July 2010 * Treated	-0.680***	(0.0389)
August 2010 * Treated	-0.452***	(0.0383)
September 2010 * Treated	-0.379***	(0.0385)
October 2010 * Treated	-0.352***	(0.0385)
November 2010 * Treated	-0.341***	(0.0393)
December 2010 * Treated	-0.322***	(0.0397)
January 2011 * Treated	-0.314***	(0.0396)
February 2011 * Treated	-0.304***	(0.0402)
March 2011 * Treated	-0.229***	(0.0405)
April 2011 * Treated	-0.192***	(0.0404)
May 2011 * Treated	-0.0291	(0.0449)
Household Fixed Effects	Yes	
Month Fixed Effects * Demographics	Yes	
Month-of-Year Fixed Effects * Treated Indicator	Yes	
Months	168	
Household-Months	232,753,860	
R-Squared	0.0000262	

Notes: standard errors (in parentheses) clustered by household. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Includes treatment * calendar month fixed effects estimated from the pre-2010 period and month FE * 2008 age fixed effects, marital status, spouse's age, and dependents.

Table 1.6: Difference-in-Difference Estimates of Effects of $250 Withholding Reduction: Coefficients on Interaction with 2008 Interest Income Greater than $100 Indicator

	(1) Fully Paid (pp)	(2) Payment Net of Balance Due ($)	(3) Percentage of Balance Due Paid
2000 * Treated * Interest Income	0.142		
	(0.145)		
2001 * Treated * Interest Income	0.144		
	(0.138)		
2002 * Treated * Interest Income	0.194		
	(0.135)		
2003 * Treated * Interest Income	0.168	-5.319	0.154
	(0.127)	(7.981)	(0.112)
2004 * Treated * Interest Income	0.319***	-0.401	0.248**
	(0.123)	(8.149)	(0.110)
2005 * Treated * Interest Income	0.207*	-7.745	0.221**
	(0.122)	(8.322)	(0.110)
2006 * Treated * Interest Income	0.121	-6.116	0.0780
	(0.118)	(8.243)	(0.106)
2007 * Treated * Interest Income	0.490***	-6.690	0.223**
	(0.106)	(7.943)	(0.0977)
2008 * Treated * Interest Income	(omitted)	(omitted)	(omitted)
2009 * Treated * Interest Income	-1.04***	-19.60**	0.0000775
	(0.125)	(7.864)	(0.103)
2010 * Treated * Interest Income	-0.285**	-21.58**	-0.264**
	(0.119)	(8.457)	(0.106)
2011 * Treated * Interest Income	-0.0663	-9.779	-0.117
	(0.123)	(8.861)	(0.111)
2012 * Treated * Interest Income	0.0711	-12.90	0.00971
	(0.128)	(9.254)	(0.116)
2013 * Treated * Interest Income	-0.0951	-16.55*	-0.126
	(0.132)	(9.564)	(0.119)
F-stat Pre-2008	3.544	0.414	1.813
p-value Pre-2008	0.000412	0.839	0.106
F-stat Post-2008	18.13	1.827	2.095
p-value Post-2008	<0.0001	0.104	0.0629
Household Fixed Effects	Yes	Yes	Yes
Year FE * Demographics	Yes	Yes	Yes
Years	14	14	14
Household-Years	13,327,835	10,703,607	10,703,607
R-Squared	0.00283	0.00109	0.00189

Notes: standard errors (in parentheses) clustered by tax unit. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Interest income includes both taxable and tax-exempt interest. Includes interactions between year fixed effects, interest indicator, and 2008 age fixed effects, marital status, spouse's age, and dependents. Payment amount data begin in 2003. Payment net of balance due winsorized at the 1st and 99th percentiles. F-statistics report joint statistical significance of all pre-2008 or post-2008 coefficients.

Table 1.7: Errors in Claiming and Calculating Making Work Pay Credit by Treatment Group and Whether Fully Paid in 2009

	Control, Not Fully Paid	Control, Fully Paid	Treated, Not Fully Paid	Treated, Fully Paid
Made any error related to credit (pp)	10.9	6.29	37.3	9.69
Did not claim credit(pp) (IRS filed ex post)	9.19	5.44	6.97	5.37
Claimed credit, but did not report Economic Recovery Payment (pp)	N/A	N/A	18.6	1.89
Random Sample Size	1%	1%	10%	10%
N	22,705	460,586	39,008	538,893

CHAPTER II

Heard it Through the Grapevine: The Direct and Network Effects of a Tax Enforcement Field Experiment on Firms

with John Guyton, Ronald Hodge II, Joel Slemrod, and Ugo Troiano

2.1 Introduction

The effects of tax enforcement directed at one taxpayer are not limited to that taxpayers behavior. Increased enforcement can deter evasion in the canonical Allingham and Sandmo (1972) model of tax evasion by changing other taxpayers perceptions of the probability that evasion will be detected and punished. Deterrence may be general to all taxpayers or limited to those who receive information about the level of enforcement from the treated taxpayer through a shared network connection. Beginning with the audit threat letters discussed in Blumenthal, Christian, and Slemrod (2001) and Slemrod, Blumenthal, and Christian (2001) and continuing with a large recent literature surveyed Hallsworth (2014) and Slemrod (forthcoming), field experiments in cooperation with tax authorities have provided substantial insights into the effects of feasible tax enforcement initiatives. The literature on tax enforcement still contains understudied issues and important gaps, which this paper begins to fill using results from a large-scale field experiment done in partnership with the Internal Revenue Service (henceforth the IRS).

The first gap this paper addresses is a lack of attention to compliance and enforcement for collecting what we call employment taxes, which include payroll taxes and employee income taxes withheld and remitted by employers. The lack of attention to collecting employment taxes is surprising given that they bring in a large amount of revenue. In fiscal year 2017, FICA[1] payroll tax revenue was $1.05 trillion and individual income tax withheld by employers

[1] FICA is the Federal Insurance Contributions Act, which covers contributions toward Social Security and Medicare. This number does not include other employment taxes: SECA, unemployment insurance, or railroad retirement.

was \$1.33 trillion, comprising 31.6 percent and 38.9 percent of total collections, respectively, or together over 70 percent of taxes collected by the IRS[2].

This paper also provides evidence about how the effects of tax enforcement spill over to firms connected to treated taxpayers through networks. We examine the response of firms connected to treated taxpayers through several distinct, sometimes-overlapping networks. The network effects capture responses driven by financial ties or by information about enforcement spread by word-of-mouth. The spillover effects contribute (or detract from) the total revenue impact of the enforcement initiatives. Understanding network effects could improve the cost-effectiveness of enforcement policy; for example, treating the most-connected taxpayers increases voluntary compliance in the agent-based model of Andrei, Comer, and Koehler (2014), and degl'Innocenti and Rablen (2019) find large simulated revenue gains from using some network information to target enforcement. Taking a broader perspective, the network effects are a crucial link between the specific deterrence effects (i.e., on the treated taxpayers) of an enforcement initiative and the general deterrent effect of changing all actual and prospective evaders perceptions of the likelihood that evasion will be detected and punished. While we are not the first to note the potential network effects of tax enforcement, existing field experiments focus primarily on geographic connections between households. One example is the study of the spillover effects on nearby households of in-person visits by Austrian TV tax inspectors in Rincke and Traxler (2011) and Drago, Mengel, and Traxler (2015), as opposed to the inter-firm links studied in this paper; neither the professional preparer nor parent-subsidiary links for which we find evidence have previously been examined. In contrast, Meiselman (2018) finds no evidence that sending letters to Detroit city income tax non-filers leads their neighbors to file. Pomeranz (2015) is a notable exception to the focus on household ties, in which an experiment shows that an audit threat increases the VAT declarations of the treated firms suppliers, but not treated firms clients. This pattern is consistent with the incentives greater VAT enforcement provides for treated firms to insist that transactions with suppliers are reported, and for treated firms suppliers to match reports with the treated firm, and is not informative about word-of-mouth diffusion of information in a payroll tax or income tax setting like the one we study[3].

Finally, this study contributes to the literature that examines to what extent the delivery mechanism of an enforcement intervention matters. Ortega and Scartascini (2018) show that in the context of Colombian taxes visits are more effective than emails, which are more effective than letters, and Ortega and Scartascini (2015) show that phone calls are more

[2]From Table 1 of the IRS Data Book 2017.

[3]Alstadster, Kopczuk, and Telle (2018) study how information about a legal tax avoidance scheme diffuses. Perez-Truglia and Troiano (2018) study how the visibility of shaming affects the rate of payment of tax delinquencies.

effective than letters, but this pattern has not been demonstrated for taxes in advanced economies.

We study both direct and network effects in a large-scale field experiment conducted in partnership with the IRS, in which 12,172 firms suspected of failure to remit all of the tax they owe, but not subject to any compliance intervention by the IRS, were assigned either to one of two treatment arms or to a control group. One treatment was an informational letter, while the other was a much more dramatic intervention, an in-person visit to the place of business by an IRS Revenue Officer.

We find that in-person visits have large, persistent direct effects on tax payments, while letters have small, fleeting direct effects. A visit from a Revenue Officer causes firms to remit an average of $3,686 in additional tax one quarter after the visit. This effect slowly diminishes to $1,652 four quarters after the visit. The visit also raises the probability of remitting any tax by 12.9 percentage points and log (tax remitted) by 13.2 log points one quarter after treatment. Receiving a letter does not cause firms to remit more tax on average, but it does increase the probability they remit any tax by three percentage points one quarter after treatment.

We also find evidence of network effects. Firms whose tax preparers' other clients receive an in-person visit eventually remit more tax, a network effect that lags the direct effect by three quarters. On average, firms in the experimental group share a tax preparer with 23 other firms. These 23 other firms each remit an average of an additional $243 four quarters after the visit, an effect that is highly statistically significant. This effect takes time to develop. Point effects on tax remitted in the first two quarters after treatment are $86 and $52 and statistically insignificant, while the point effect three quarters after treatment is $156 and is statistically significant at only the ten percent level. This delay is consistent with an informational story, in which tax preparers pass information to their clients only during infrequent contacts. Taking into account the large number of linked firms, the aggregate tax preparer network effect summing over the four quarters following the visit is 1.2 times the direct effect.

The paper proceeds as follows. In Section 2.2 we describe the experimental setting and treatments. In Section 2.3 we present the direct effects of our two tax enforcement interventions, the in-person visit and the letter. In Section 2.4 we describe the network effects. In Section 2.5 we discuss the economic significance of the estimates. Section 2.6 presents a conceptual framework to think about the welfare effects of the interventions and the consequences for policy design, and Section 2.7 concludes.

2.2 Setting and Treatments

More than 6.5 million U.S. firms deposited federal income tax withheld from wages and salaries, federal unemployment insurance taxes, and FICA taxes between the fourth quarter of 2013 and the fourth quarter of 2014. Firms report these tax remittances using Form 941, Employers Quarterly Federal Tax Return. Most employers are required to make semi-weekly or monthly Federal Tax Deposits (FTDs) of these employment taxes.

The IRS uses an algorithm to identify and prioritize firms at risk of falling behind on their required deposits in each quarter. The IRS assigns at-risk firms into categories called FTD Alerts. For firms with high priority alerts (Alert A or B status), the IRS assigns a Revenue Officer to contact the firm within fifteen days of the alerts issuance. The experiment we study was carried out on a third group of firms, designated as having Alert C status. These are firms for which the algorithm indicates a higher risk of falling behind on their deposits than the general population, but not as high a risk as firms designated Alert A or B. In some quarters prior to the experiment, Alert C firms may have received a letter about their deposits. Some, but by no means all, firms receive the same FTD Alert designation for more than one consecutive quarter[4]. It is especially relevant from a tax enforcement policy standpoint to understand the behavior of Alert C firms, because these firms are at the margin of enforcement action from the IRS, and are therefore the most relevant population when considering whether to expand or contract the set of firms the IRS contacts.

This paper uses a randomized experiment to study the effects of sending letters to and visiting at-risk firms at the margin of enforcement action. There were 12,172 such firms assigned Alert C status by algorithm based on payments before and during the fourth quarter of 2014. These firms were randomly assigned to one of three groups. A control group received no FTD Alert-related contact. A second group received an informational letter[5] early in the first quarter of 2015. The letter notes that the firms deposits have decreased, discusses the firms deposit responsibility and potential penalties, and provides information and resources about federal tax deposits and their payment. The third group of firms received an initial in-person contact at the place of business from an IRS Revenue Officer[6]. Initial contact procedures emphasize providing the taxpayer with information about the collection process,

[4]Due to high turnover from quarter to quarter (e.g., only 28 percent of control group firms continue to have the Alert C designation after one quarter), we expect that a few of the firms randomly assigned in the experiment we study would have received an enforcement action prior to the experiment because of an earlier Alert status. Random assignment makes this fact unlikely to bias our results, although it is relevant when considering how our results generalize to other contexts.

[5]A copy of the letter is included in the online appendix. If a taxpayer has filed a form giving a representative power of attorney, the representative also receives a copy of any written correspondence.

[6]IRS records indicate that Revenue Officers dedicated time to contacting nearly all assigned firms.

discussing the taxpayers deposit compliance status, and gathering basic information. In some cases, a Revenue Officer may use information from an initial contact to determine that further investigation or contact is warranted, following collection procedures.

Alert C firms show signs of noncompliance before treatment. As Table 2.1 shows, compared to the average firm filing a quarterly employment tax return, firms with Alert C status as of the fourth quarter of 2014 had more employees but remitted less tax and were more likely to have not remitted any tax at all. As expected due to randomization, the treatment groups are similar before treatment. All three groups' tax remittances, probability of remitting any tax, and log (tax remitted), depicted in Figure 2.1 fall sharply over the four quarters prior to treatment. Control firms' remittances also rebound to an extent one quarter after treatment, a pattern analogous to the Ashenfelter dip discussed by Heckman and Smith (1999) in the context of labor market interventions, wherein those who qualify for job training often have temporarily depressed earnings that tend to revert upward toward their longer-term mean even absent treatment. Without an experimental control group, it would be difficult to construct a control group from observational data that would not underestimate the control group's rebound in compliance and thus tend to overestimate the effect of treatment.

2.2.1 Follow-up Treatment

Recent work by Bhargava and Manoli (2015) and Guyton, Langetieg, Manoli, Payne, Schafer, and Sebastiani (2017) has shown that many enforcement initiatives have short-lived effects on taxpayer behavior and that reminders, essentially follow-up rounds of treatment, can boost the persistence of the policys effect. This inspired a novel (in the context of tax administration research) feature of the design of this experiment, drawing on practice in medicinewhere patients who are initially unresponsive to treatment may receive continued treatment[7].

At the end of the quarter during which treatment took place, the algorithm that determines whether firms are designated high risk (Alert C) ran again, and some of the 12,172 firms in the experiment were again designated high risk. Firms that were again designated high risk received a second dose of their assigned treatment in the following quarter. Thus, each firm assigned, for example, to the visit group received one visit early in Q1 2015 and, if the firm remained at high risk based on its payments through week twelve of Q1 2015, it received a second visit in the second quarter of 2015. The same procedure was followed

[7]See, for example, Zonder, Pemberton, Brandt, Mohamed, and Schiffer (2003) on leukemia and Diehl, Stein, Hummel, Zollinger, and Connors (2003) on treatment of refractory Hodgkins lymphoma with a second course of high-dose chemotherapy.

with the letter treatment. After the second quarter, no firm received further experimental treatment, although some businesses in the experiment might have been assigned to very high risk (Alert A or B) status and thereby been subject to routine enforcement action. Table 2.2 presents a treatment timeline.

Turnover in high risk status, detailed in Figure 2.3, is large. Only 28 percent of control group firms remained in this category one quarter after random assignment. Among firms assigned to receive a letter, 28 percent continued to have high risk status in the following quarter and received a second letter. Among firms assigned to receive a visit, just 19 percent-about one-third less–continued to have high risk status in the next quarter and therefore received a second visit. The lower fraction of firms assigned to receive a visit continuing in high risk status is consistent with the result, detailed below, that the visit increased remittances.

This follow-up treatment allows us to assess the effects of a realistic treatment protocol in which recalcitrant cases receive a follow-up intervention. If the treatment interventions we study were to become standard practice, follow-up treatment of unresponsive firms might well become tax administration procedure. We include firms regardless of follow-up treatment status, but the proper interpretation of our results includes the follow-up treatment administered to firms whose remittance behavior continued to indicate high risk. Beginning two quarters after treatment, the estimated impacts capture both the persistent component of the initial treatment administered to all firms in the treatment group and the effect of the follow-up treatment administered one quarter later to a subset of treatment group firms.

2.3 Direct Effects

2.3.1 Event Study Regression Design

Our preferred specification uses an event-study regression design that reduces residual variance and allows for a flexible time path of the treatment response[8]. This design rests on the assumptions that there are no contemporaneous changes that affect the treatment and control groups differentially, and that absent treatment the time paths of the outcome variables in the treated and control groups would evolve in a parallel fashion. In fact, there were no contemporaneous IRS policy changes that might affect the treatment groups differentially. Figure 2.1 illustrates that the trends in the outcome variables we study are similar across treatment groups for several quarters prior to treatment, which supports the assumption that these trends would continue to be parallel absent the experiment. We

[8]Results using a cross-sectional direct effects specification instead of our preferred difference-in-difference specification are similar and are reported in an appendix.

estimate models of the form

$$Y_{it} = \sum_j \sum_q \beta_{jq} 1(T_i = j) 1(t = q) + \eta_t + e_{it}, \tag{2.1}$$

where Y_{it} denotes the outcome of interest, e.g. the log amount of employment tax that firm i remitted with Form 941 in quarter t, β_{jq} is the coefficient that indicates the direct effect of treatment j on the outcome q quarters after treatment, $1(T_i = j)$ is an indicator variable equal to one if firm i received treatment j, $1(t = q)$ is an indicator equal to one if t is q quarters after treatment, η_t is a fixed effect for quarter t, and e_{it} is the regression error term. Standard errors are clustered at the firm level to account for possible serial correlation in the error term. We study how letters and visits affect tax remitted (in dollars, winsorized at the 98th percentile), the probability of remitting any employment tax, for which we use a linear probability model, and log (tax remitted), which omits firms that do not remit any tax.

2.3.2 Direct Effects Results

We find that in-person visits have large, lasting direct effects on tax payments. Figure 2.2 illustrates that the effect on tax remitted overall and on the probability of remitting any tax last four quarters after treatment. One quarter after a visit firms remit an additional \$3,686. Visited firms are 12.9 percentage points more likely to remit any tax one quarter after treatment; this effect is large relative to the 58 percent of control group firms that remitted any tax one quarter after treatment. This effect shrinks to 6.9 percentage points by four quarters after treatment. The effect on log (tax remitted) lasts only a single quarter. Although control firms' compliance does improve after treatment, which is consistent with mean reversion and the Ashenfelter dip, visited firms' compliance rebounds much more. Control firms rebounding suggests that observational studies comparing firms receiving a visit or letter to firms selected from the general population would likely overstate the effects of the compliance treatments, and further indicates the value of conducting randomized experiments.

Letters have much smaller, and fleeting, direct effects. Letters do not lead to substantially higher average tax payments or increases in log (tax remitted), as shown in Figure 2.2. Letters have an effect only on the probability that firms remit any tax one quarter after treatment, which rises by three percentage points. This effect is highly statistically significant, but does not persist beyond one quarter after initial treatment, suggesting that follow-up letters have little or no effect.

The causal effect of the initial visit beyond one quarter cannot be separated from the combined effect of the follow-up procedure in which continually non-compliant firms receive a second visit, but effects are largest one quarter after treatment, and a second letter appears to have no effect. Section 2.5 compares the estimated impact of these treatments to their cost to evaluate their impact on net revenue and assess them from a welfare economics perspective.

2.3.3 Direct Effects and Firm Size

We next explore whether larger firms respond more to treatment using a triple-difference regression specification that compares the direct effect for the largest ten percent of firms to the direct effect for the smallest ten percent of firms. We define size to be the number of employees in the calendar year before treatment, as measured by Forms W-2 filed with the IRS. The largest ten percent of firms have at least 67 employees, while the smallest ten percent of firms have at most two employees.

In dollar terms, the direct effect of a visit is much larger for firms in the top ten percent by pre-treatment size than for firms in the bottom ten percent. The direct effect of the visit on tax remitted one quarter after treatment is $6,595 dollars larger for the largest firms than for the smallest firms, as depicted in Figure 2.3, a difference that is highly statistically significant. The responses of the largest and smallest firms, summarized in Table 2.5 are otherwise similar, including for the letter. Holding the cost of contacting a firm constant, visiting larger firms uses the same resources to collect more additional revenue than visiting smaller firms.

2.4 Network Effects

The administrative data we use enable us to study the network deterrent effects of enforcement interventions, which operate through connections between untreated firms and firms directly receiving the enforcement intervention[9]. As discussed earlier, this analysis could provide insight about how information regarding enforcement actions diffuses to alter the generally perceived probability that tax evasion will be detected. Even if the per-linked-firm network effect is small, many linked firms per treated firm can still result in a substantial aggregate effect of network connections on total remittance behavior. Understanding the information network structure could also inform the design of information campaigns, as

[9]Network effects through connections between treated and untreated firms in the randomly assigned group violate the usual assumption in a randomized experiment that the untreated firms receive no treatment. This violation would tend to bias our estimates of the direct effects towards zero, but our direct effects estimates are unchanged when we control for network links.

models show higher voluntary compliance results from providing information to or about taxpayers with the most links (Andrei, Comer, and Koehler (2014), degl'Innocenti and Rablen (2019)). We investigate several types of network, some of which have been examined before (although usually with respect to households rather than firms), and some that have not been heretofore studied. Our rich data set allows us to examine certain links between firms that have not been rigorously studied before.

One connection between the firms in our sample and others is geographic. Geography ties together firms with addresses in the same ZIP Code or, at a more fine-grained level, a shared ZIP+4. The 42,000 five-digit ZIP Codes in the United States indicate a shared postal facility and are assigned to either geographic areas or post office boxes, while a ZIP+4 is a nine-digit designation for a small group of blocks or segment of a postal route (USPS, 2016). Firms in our experimental sample share a ZIP Code with an average of 659[10] other employers filing quarterly employment tax returns, and share a ZIP+4 with an average of just 3 other employers.

Firms also share tax preparers or tax preparation firms. Each individual tax preparer has a unique Preparer Tax Identification Number (PTIN), which that preparer includes on each return he or she prepares. If the preparer is part of a tax preparation firm, the firms unique Employer Identification Number (EIN) is also included on each prepared return. These identifiers allow us to identify when two firms returns are prepared by the same individual preparer or by preparers working at the same tax preparation firm. We consider two firms linked to a tax preparer or tax preparation firm if that tax preparer or tax preparation firm prepared at least one Form 941 for that firm in the four quarters prior to treatment; it is plausible that firms might have contact with a tax preparer or tax preparation firm they have used in the past year even if they are no longer using that preparer, especially if they are concerned about IRS enforcement action related to past filings. Each firm in our experimental sample shares a tax preparer with an average of 23 other employers and a tax preparation firm with an average of 98 other employers.

Network effects through shared tax preparers are of interest for two reasons beyond their implications for correctly estimating the revenue impact of enforcement initiatives. First, preparers may be an effective target for expanded information reporting or other enforcement treatments. Second, and related, the fact that the treatment spills over to other firms with the same tax preparer suggest that preparers play a role in firms decision-making, an issue addressed by Klepper, Mazur, and Nagin (1991) and recently by Klassen, Lisowsky, and

[10] As some firms are linked to more than one Alert C firm, the sample of firms linked by ZIP code to Alert C firms is somewhat smaller than the number of links per firm times the size of the Alert C sample (536 linked firms instead of 659), and similarly for the other network channels we study.

Mescall (2015), who analyze confidential data from the IRS to examine whether the party primarily responsible for a firms tax compliance function-an external auditor or the internal tax department-is related to the firms tax aggressiveness.

Finally, we investigate links between parent corporations and their subsidiaries. Parent/subsidiary relationships meet one of two sets of criteria in the year prior to treatment assignment. In the first case, the parent corporation files IRS Form 851, Affiliations Schedule, with a consolidated group annual tax return indicating that the parent owns stock with 80 percent or more of both the total value and voting power of the subsidiary directly or indirectly through other corporations in the consolidated group. In the second case, the parent corporation is a subchapter S corporation and has filed Form 8869, electing to treat a domestic corporation whose stock it wholly owns as a qualified subchapter S subsidiary which is deemed liquidated. This definition implies that firms have at most one parent and that parent firms cannot themselves have a parent, as parents in our sample are either the ultimate parent of a consolidated group or S corporations whose owners are required by law to be individual people. The business operations of the parent and subsidiary are presumably tightly linked, given the degree of ownership and filing of a consolidated annual tax return.

These three sets of networks capture a diverse range of relationships between firms. For example, the network effect per link to a firm visited by a Revenue Officer may be large for one channel but not others, and the network effect per link to a letter firm need not be large for that channel. One might expect that letters have network effects through ZIP and ZIP+4, as these links capture both geographic proximity and shared postal delivery, while visits might have especially strong effects through shared preparers or tax preparation firms, as the preparer or firm may interact directly with the Revenue Officer. Additionally, links to visited firms through a given channel, for example a shared preparer, may affect tax payments overall, only on the extensive margin captured by the indicator for remitting any tax, or only on the intensive margin captured by log (tax remitted).

2.4.1 Identifying Network Effects with Non-Random Selection into Network Linkages

We aim to identify the causal network deterrence effects of the letter and visit treatments. This causal effect captures the difference between a firms compliance behavior if its network "neighbors" happen to receive a letter or visit and that firms behavior if its network neighbors happen to receive no treatment. When estimating these effects, it is important to keep in mind that simply comparing the post-treatment behavior of firms with network neighbors that received a letter or visit to the post-treatment behavior of all firms without treated network neighbors would provide a biased estimate of the network effect. This is because having treated network neighbors requires having network neighbors with high-risk (Alert C) status, so that network links may very well not be random.

Firms with Alert C status are less likely than other employers to have remitted any Form 941-related tax, as Table 2.1 shows, and so it is natural to suppose that the network neighbors of firms with Alert C status might have systematically different remittance behavior compared to other firms network neighbors. For example, if adverse local economic shocks make firms in a neighborhood less likely to remit tax payments, firms in that neighborhood are both more likely to have Alert C status themselves and more likely to be linked to firms with Alert C status. The resulting correlation between connections to treated firms and lower tax payments would bias network effects estimates downward. The same concern arises for links through preparer networks; some preparers may be more experienced, or more sympathetic, or condoning, towards at-risk businesses and thus develop clienteles of such businesses. Parents and their subsidiaries are also likely to share similar compliance behavior.

To address the selection bias concern, we compare firms with the same number of Alert C neighbors. Consider the example of two firms, each sharing its own unique ZIP Code with exactly one Alert C firm in the experimental sample. Prior to random assignment, the likelihood of each firm sharing its ZIP Code with a firm that receives a visit is 1/3. Conditional on the number of links to Alert C firms, network treatment is randomly assigned and thus independent of firms characteristics and potential compliance outcomes. Comparing firms with the same number of links to Alert C firms allows us to identify an unbiased causal effect of being linked to a treated firm, because before treatment the network treated and control groups are equally likely to have low tax payments on the basis of their similar connections to Alert C firms. The regression approach we implement is a generalized version of the event-study approach used above to study direct effects, where we pool firms with different numbers of links to Alert C firms to produce a single treatment estimate, but control for differential patterns of compliance over time between firms based on their total links to Alert C firms. This approach relies on the assumption that, conditional on the number of total links to Alert C firms, the trends in compliance would be parallel across firms linked to different treatment groups absent treatment. Specifically, separately for each network channel c we run regressions of the form:

$$Y_{it} = \sum_j \sum_q \rho_{cjq} L_{cij} 1(t = q) + \sum_l \theta_{clt} + e_{it}, \tag{2.2}$$

where Y_{it} is the outcome for firm i in quarter t, ρ_{cjq} is the network effect through channel c of treatment j, q quarters after treatment, L_{cij} is the number of links through network channel c that firm i has to firms that received treatment j, $1(t = q)$ is an indicator equal to one if t is q quarters after treatment, θ_{clt} is a fixed effect common to all firms connected through network channel c to a total of l treated and control firms in quarter t, and e_{it} is the regression error term. Note that, conditioning on a fixed value of the total number of links to Alert C firms, this specification is a standard event-study specification with quarter fixed effects. The specification pools the event-study specifications across different numbers of total links to Alert C firms and constrains the estimated network effect to be linear in the number of links to treated firms. We do this in separate specifications for firms sharing a preparer, preparer firm, ZIP Code, or ZIP+4 with an Alert C firm and for the subsidiaries and parents of Alert C firms. We cluster the standard errors at the level of the channel used in that specification, e.g. ZIP Code, preparer, or parent, which addresses correlation in the error term between firms sharing, e.g., a preparer or parent as well as serial correlation in the error term.

2.4.2 Tax Preparer Network Effects Results

We find evidence that tax enforcement interventions have network effects transmitted through a shared tax preparer several quarters after treatment, but not immediately. In person visits increase the tax remitted by visited firms' tax preparers' other clients by an average of $156 three quarters after treatment and by $243 four quarters after treatment, as shown in Figure 2.4 and Table 2.6. The effect three quarters after treatment is statistically significant at the ten percent level, while the effect four quarters after treatment is statistically significant at the one percent level. We find that letters increase log (tax remitted) by letter recipients' tax preparers' other clients four quarters after treatment by 1.09 log points. This effect is statistically significant at the five percent level. The time delay between treatment and these tax preparer network effects is consistent with the low frequency with which most firms exchange information with their tax preparers.

2.4.3 Tax Preparation Firm Network Effects Results

In contrast to the network effects of shared individual tax preparers, we do not find evidence of network effects through shared tax preparation firms. The results shown in Figure 2.5 and in Table 2.7, are precisely estimated, ruling out large per-firm spillovers, and are not statistically significant at the five percent level. While Alert C firms on average share a tax preparer with 23 other firms, they share a tax preparation firm with an average of 659 other businesses, limiting the influence of a single contact with one of the preparation firms' clients on the firm's other clients.

2.4.4 Narrow Geographic (ZIP+4) Network Effects Results

We find mixed evidence of narrow geographic network effects between firms on the same postal route, presented in Figure 2.6 and Table 2.8, and no evidence of such network effects on tax remitted overall. One quarter after treatment, log (tax remitted), which excludes firms remitting nothing, rises by 3.26 log points for firms sharing a postal route with a visited firm, and three quarters after treatment the probability of remitting any tax falls by 1.39 percentage points for firms sharing a postal route with a letter recipient. These effects are statistically significant at the five percent level, and do not accompany substantial changes in tax remitted overall.

2.4.5 Geographic (ZIP Code) Network Effects Results

We do not find evidence of spillovers at the ZIP code level on tax remitted, although visits do have small, positive spillovers conditional on remitting any tax. As Figure 2.7 and Table 2.9 show, two quarters after treatment, firms in the same ZIP code as a visited firm have log(tax remitted), which is conditional on remitting any tax, that is 0.412 log points higher (with a p-value less than 0.05), and four quarters after treatment the effect on this outcome is 0.666 log points (with a p-value less than 0.01). These effects do not translate to higher tax payments overall.

2.4.6 Parent and Subsidiary Network Effects Results

There is limited evidence that letters and visits have effects on the parents of contacted firms. Few treated firms have parents, and therefore the estimates, presented in Figure 2.8 and Table 2.10 are imprecise. Three quarters after treatment, parents of visited firms remit an additional $4.15 million, but this effect is statistically significant only at the ten percent level, and as such is weak evidence. Across other quarters, outcomes, and both treatments, there is little evidence of an effect on treated firms' parents.

Contacting a parent firm has similarly ambiguous effects on its subsidiaries. Figure 2.9 and Table 2.11 show that the in-person visit raises tax remitted by the visited firm's subsidiaries in the quarter after treatment by $915,000, an effect that is highly statistically significant, yet at the same time decreases the probability the subsidiary remits any tax by 0.917 percentage points. The letter has no effects on subsidiaries that are statistically significant at the five percent level. Only 49 treated firms are parents, and there is evidence, shown in Table B.14, of a pre-treatment trend in subsidiaries' tax payments, so these results should be interpreted with caution.

To summarize the network effects results, focusing on overall tax remitted, visits have delayed spillover effects on tax remitted through shared individual tax preparers, while there is not strong evidence of effects through the other networks we study.

2.5 Comparison of Aggregate Network Effects to Direct Effects

Even if the network effect is small *per linked firm*, a large number of network links between firms can imply substantial network effects in the aggregate. To compare the aggregate network effects to the direct effects, we define the network multiplier, equal to the ratio of the aggregate network effect of a treatment to the treatment's direct effect. We multiply the coefficient on tax remitted (or sum of coefficients) by the average number of links per firm to obtain a per-letter or per-visit effect, which we divide by the direct effect to obtain a network multiplier. Limiting our focus to estimates with 95 percent confidence intervals that exclude zero, we find that the tax preparer network multiplier over the year following the visit is 1.2, and the subsidiary network multiplier one quarter after the visit is 8.1 (although violations of the parallel trends assumption before treatment for subsidiaries make us skeptical of this estimate). Multipliers for the quarter after treatment are reported in Table 2.12 and for the four quarters after treatment in Table 2.13. Over the following year, each visit leads the visited firm to remit an additional \$10,233, and also generates an additional \$12,258 from firms sharing a tax preparer with the visited firm.

These estimates depend on several simplifying assumptions. Although we multiply the mean effect per link by the mean number of links, both the effect and the number of links are unlikely to be distributed evenly across the population. These calculations also do not account for heterogeneous direct effects by firm size or for non-linear dose response to multiple links to treated firms. The networks we discuss above intersect, as firms for example may share both a neighborhood and a tax preparer, though in unreported results we find that including all of the networks in a single specification does not substantially change the tax preparer network effect of the visit. Despite these assumptions, the network multiplier calculations demonstrate that network effects may be economically substantial.

2.6 Implications for Policy

How do these findings inform resource allocation decisions? Should each treatment be expanded or cut back? To answer these questions, we need to consider all the costs and benefits of each treatment. Before proceeding, we note that all the estimated effects pertain to revenue remittance and not, as is true in most similar studies, reported tax liability (that may not be remitted in a timely way, or at all).

2.6.1 Would Net Revenue Rise?

A treatment boosts net revenue if the marginal revenue it raises exceeds its marginal administrative costs. There are three components to the revenue raised: the direct effect on the treated group, the network effect, and the general deterrent effect in the population at large, denoted as r_{Dt}, r_{Nt}, and r_{Gt}, respectively, where subscript t indicates the treatment, either V for visit or L for letter. In this paper, we have estimated the direct effect and the network effect, but not the general deterrent effect. The revenue raised should be compared to the marginal administrative cost, denoted a_t. The calculation for each treatment is simply whether $r_{Dt} + r_{Nt} + r_{Gt} \equiv r_t > a_t$.

To address these questions, we begin by referring to the dollar values for the year following treatment calculated in Table 2.13, where we show that $r_{DV} = \$10,233$ and $r_{DL} = \$322$. Based on IRS data, $a_V = \$220$ and $a_L = \$4$. Both treatments clearly increase net revenue without taking network or general deterrent effects into account. Assuming the general deterrent effect, which we cannot observe, were negligible, incorporating the statistically significant tax preparer network effect of the visit yields $r_{NV} = \$12,258$. Then $r_V = \$10,233 + \$12,258 = \$22,491 \gg \$220$. Similarly, we can calculate that $r_L = \$322 > \$4$. Even absent general deterrent effects, both treatments easily pass this simple net-revenue-increasing test.

There are, though, other issues to consider. These calculations ignore compliance costs incurred by treated taxpayers, which are likely higher for the visit. In addition, we have ignored any difference between the average effect we have estimated and the marginal effect, although this difference may not be large given that the population of firms we study are not the highest-risk firms routinely subject to treatment, but instead a group of firms that typically are not treated. These calculations should be done on a discounted present-value basis. Given that current interest rates are near zero, discounting itself is not a substantively important issue over the course of a single year. What is not known is whether the estimated effects would reverse sign if carried out past the year we examine. In other words, we will be overstating the net revenue gain to the extent that the treatments cause payments to accelerate but not increase in total; we see no sign of this over the course of a year but cannot be sure it is not an issue in the longer term.

2.6.2 Would Re-Allocating Resources Raise More Revenue?

Given a fixed resource budget, would more re-allocating resources between visits and letters raise more revenue? If the objective of the tax authority is to maximize revenue net of cost, then the answer depends on whether the following inequality holds: if it does, resources should be shifted from letters to visits[11]:

$$r_{DV} + r_{NV} + r_{GV} > \frac{a_V}{a_L}(r_{DL} + r_{NL} + r_{GL}). \tag{2.3}$$

In Expression 2.3, (a_V/a_L) represents the trade-off in the extent of alternative treatments: visiting one fewer firm enables the tax authority to send (a_V/a_L) more letters while staying within the given budget. Now the relative general deterrence effects of the two treatments can matter. If we are willing to assume that the general deterrence effects are proportional to the sum of the direct and network effects, $\frac{r_{GV}}{r_{DV}+r_{NV}} = \frac{r_{GL}}{r_{DL}+r_{NL}}$ then Expression 2.3 simplifies to:

$$r_{DV} + r_{NV} > \frac{a_V}{a_L}(r_{DL} + r_{NL}). \tag{2.4}$$

Using our values from above, the left-hand side of Expression 2.4 is $\$10,233 + \$12,258 = \$22,491$, while the right-hand side is $(220/4) * \$322 = \$17,710$. Because letters deliver about 1/70 of the visits return for 1/55 of the cost, the average per-dollar-spent return is slightly higher for the visit and thus a fiscally-constrained tax agency would increase revenue by shifting resources from letters to visits at the margin[12]. Given the degree of uncertainty surrounding our estimates of both the direct and network effects, however, we cannot confidently rule out that the per-dollar returns to the two interventions are the same.

2.6.3 Would Policy Changes Increase Welfare?

The evaluation of whether welfare would rise when a given policy changes is more complicated. For one thing, such an evaluation should account for marginal compliance costs (resource costs borne directly by private citizens in the form of time and expenditure), which are social costs that do not show up in government budgets. Second, the appropriate criterion is not whether revenue net of cost increases, because that ignores the fact that any additional tax remittance is a transfer from private hands, which has social value, to the government that provides services that are of value to the population. As shown in Keen and Slemrod (2017), which draws on Slemrod and Yitzhaki (1987) and Mayshar (1991), the

[11] All the point estimates have associated confidence bands, and thus the cost-benefit analyses are themselves subject to error.

[12] If the average return in our sample equals the marginal return, and in equilibrium the deterrent effects of the two treatments are not related, as they would be if for example firms that do not respond to a letter are later visited as a result. This possibility is not addressed by the experiments we conduct, because we do not vary the operational procedure in which populations judged to be higher-risk than our sample receive visits.

welfare impact of the intervention can be approximated by

$$\Delta W \equiv (v^{'} - 1)\Delta R - v^{'}\Delta a - \Delta c. \tag{2.5}$$

In Expression 2.5, $v^{'}$ is the marginal social value of an additional dollar of revenue. If the question is whether to increase administrative effort, ceteris paribus, then $v^{'}$ represents the marginal social value of raising a dollar of net revenue for public spending. If the question is whether to increase administrative effort while reducing, say, the tax rate in a revenue-neutral way, then $v^{'}$ represents the social cost saved by reducing the tax collected via the tax rate by one dollar, sometimes referred to as the marginal efficiency cost of raising funds. In either case, the first term on the right-hand-side of Expression 2.5 is the marginal social value of the additional net revenue collected when an administrative policy instrument increases by one unit. Because raising revenue is costly, the value of $v^{'}$ will exceed one. The other two terms on the right-hand-side of Expression 2.4 are the marginal social cost of increasing government spending and the marginal compliance cost; the former is multiplied by $v^{'}$ to reflect that government spending must be funded by raising distortionary, and therefore socially costly, taxation. To see the implications of Expression 2.5, following Mayshar (1991) we set $v^{'} = 1.17$ and assume that the marginal compliance cost is twice the marginal administrative cost. In addition, we assume that the general deterrent effect is zero. Then Expression 2.5 becomes the following for letters and visits respectively:

$$\Delta W_L = (1.17 - 1) * 322 - 1.17 * 4 - 8 = 42.1 \gg 0, \tag{2.6}$$

$$\Delta W_V = (1.17 - 1) * 22,491 - 1.17 * 220 - 440 = 3,126 \gg 0. \tag{2.7}$$

In these calculations, additional letters and visits each enhance welfare. To be sure, these illustrative calculations depend on arbitrary assumptions about the social value of marginal revenue, the marginal compliance cost, and the general deterrent effect of expanding enforcement instruments. The calculations do, though, illustrate the difference between subjecting enforcement initiatives to a net-revenue-maximizing criterion and subjecting enforcement initiatives to a welfare-maximizing criterion.

2.7 Conclusion

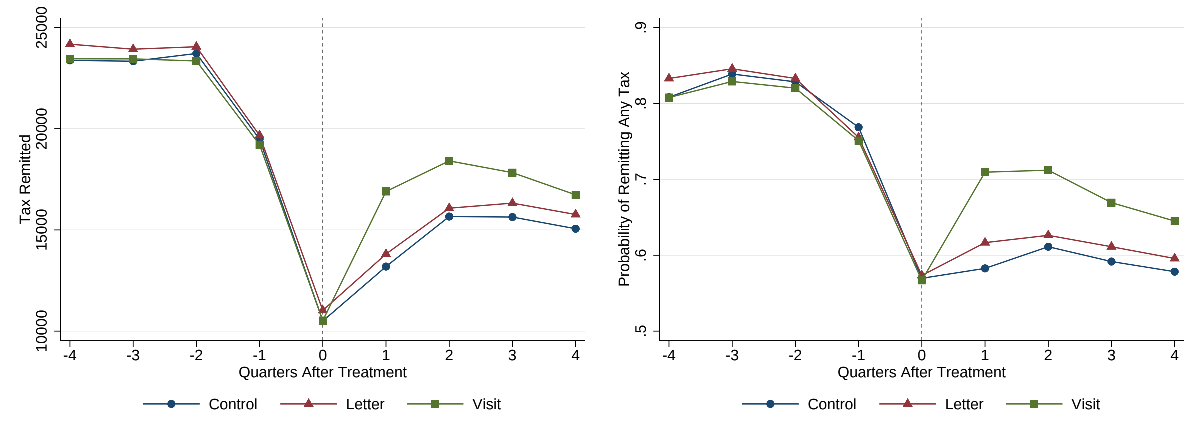
This paper uses a randomized experiment conducted in partnership with the IRS to estimate both the change in employment payroll and withholding taxes remitted caused by receiving a letter noting that the firms deposits have decreased, discussing the firms deposit responsibility and potential penalties, and providing general information, or caused by an in-person visit from an IRS Revenue Officer. In addition, we estimate the network, or spillover, effects on taxes remitted by firms linked to letter and visit recipients by geography, tax preparers, and parent-subsidiary relationships. To our knowledge, no previous research has investigated the effects of tax enforcement on firms sharing a tax preparer with the treated firm or on the treated firms parent or subsidiaries.

We find large, immediate effects of in-person visits on tax remitted that persist for at least four quarters and are transmitted through tax-preparer networks. Although the per-firm-link tax-preparer network effects of the visit are much smaller than the direct effects, their aggregate effect is 1.2 times the size of the direct effect. We find that letters increase the likelihood that firms remit any tax by three percentage points, but this effect lasts only one quarter, and the effect of the letter on tax remitted overall is not statistically significant. There is no evidence of network effects of the letter. Given the empirical results, both visits and letters pass a net-revenue-increasing criterion. With a fixed tax authority budget, net revenue from one additional dollar of resources spent on in-person visits is slightly higher than net revenue from an additional dollar spent to send letters. With some additional assumptions, both treatments also easily pass a welfare-increasing criterion.

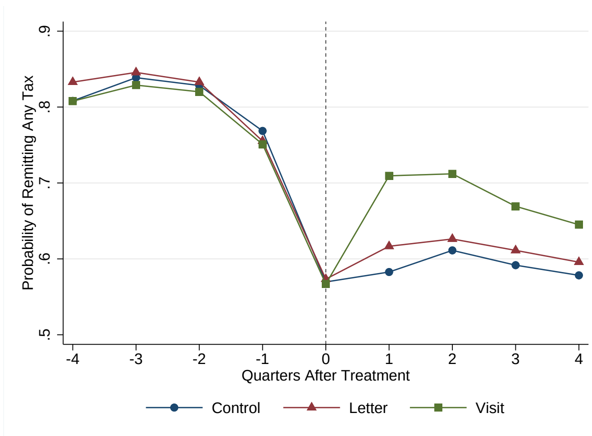
Figure 2.1: Outcome Means by Treatment Group

(a) Tax Remitted

(b) Probability of Remitting Any Tax



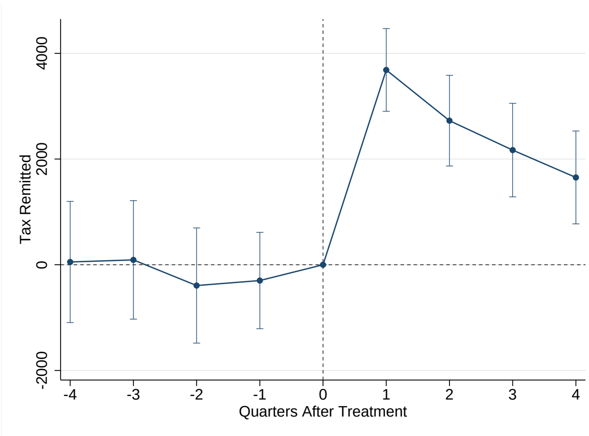
(c) Log(Tax Remitted)



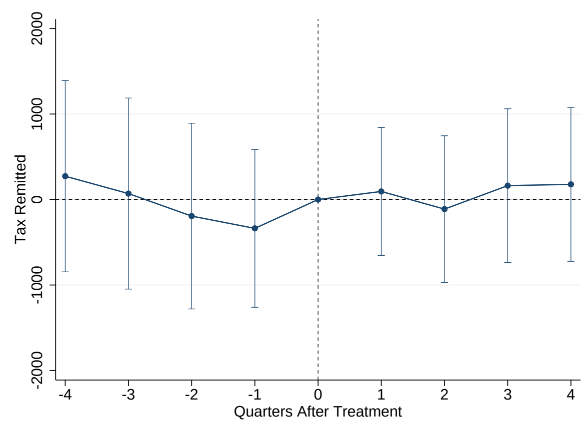
Notes: Tax remitted winsorized at the 98th percentile. Log(tax remitted) excludes firms remitting no tax.

Figure 2.2: Direct Effects

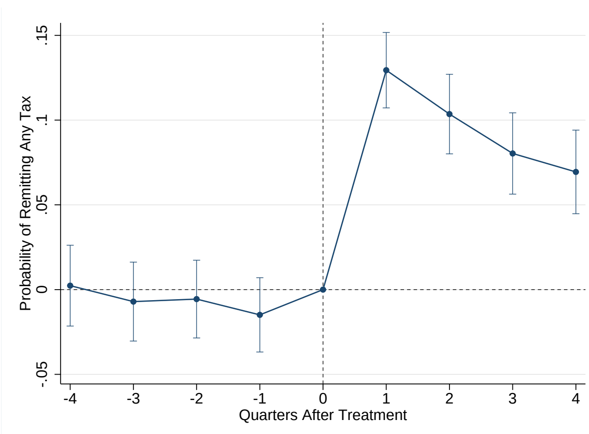
(a) Visit: Tax Remitted



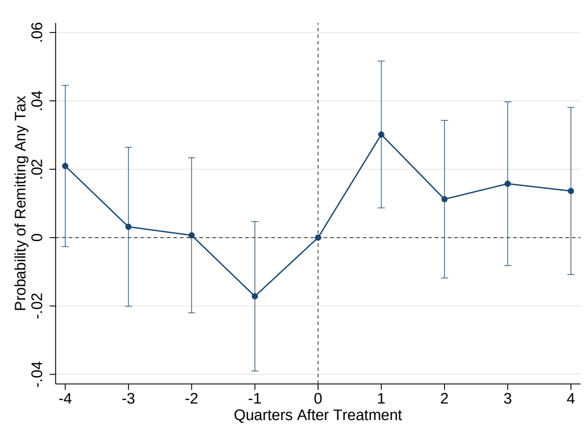
(b) Letter: Tax Remitted



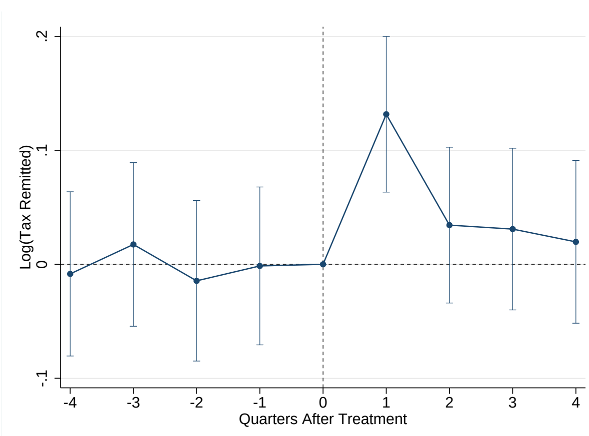
(c) Visit: Probability of Remitting Any Tax



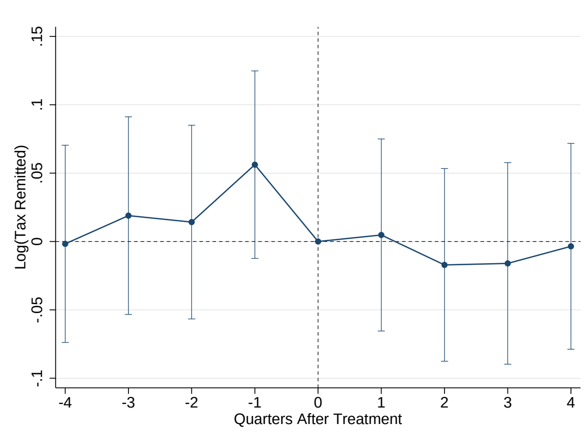
(d) Letter: Probability of Remitting Any Tax



(e) Visit: Log(Tax Remitted)



(f) Letter: Log(Tax Remitted)

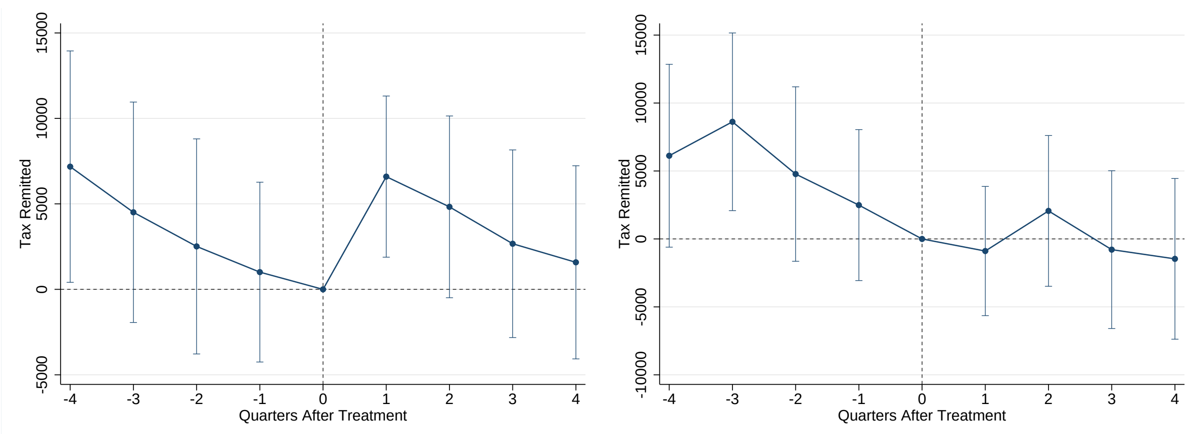


Notes: figures plot estimates and 95 percent confidence intervals. Tax remitted winsorized at the 98th percentile. Log(tax remitted) excludes firms remitting no tax.

Figure 2.3: Direct Effects: Top Ten Percent vs. Bottom Ten Percent by Size

(a) Visit: Tax Remitted

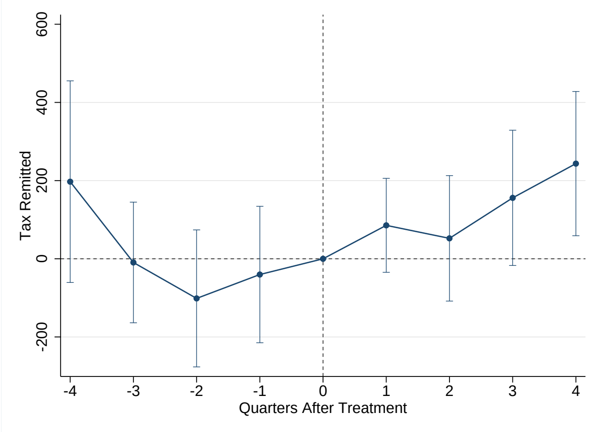
(b) Letter: Tax Remitted



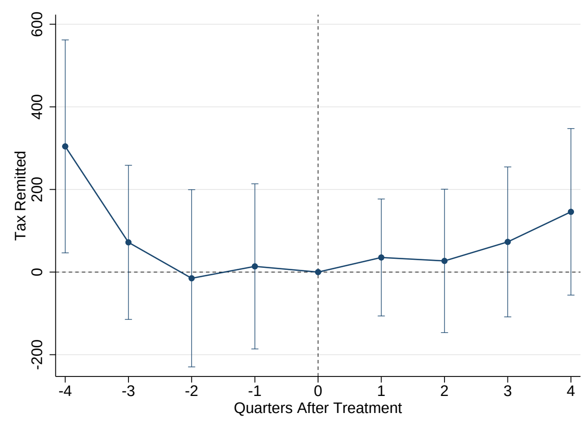
Notes: figures plot estimates and 95 percent confidence intervals. Tax remitted is winsorized at the 98th percentile. Log(tax remitted) excludes firms remitting no tax. Size is the number of W-2 employees in the year before treatment. The largest ten percent of firms have at least 67 employees, while the smallest ten percent of firms have at most two employees.

Figure 2.4: Tax Preparer Network Effects

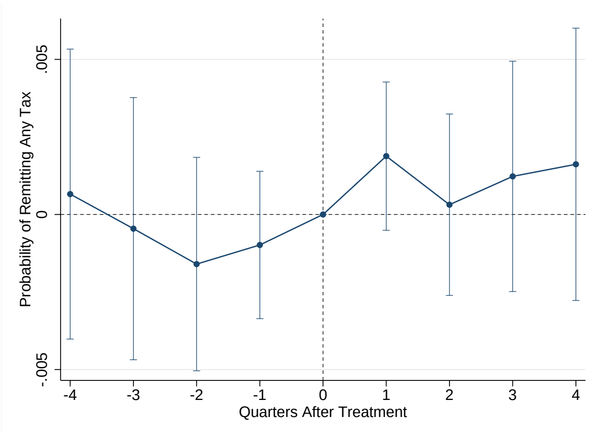
(a) Visit: Tax Remitted



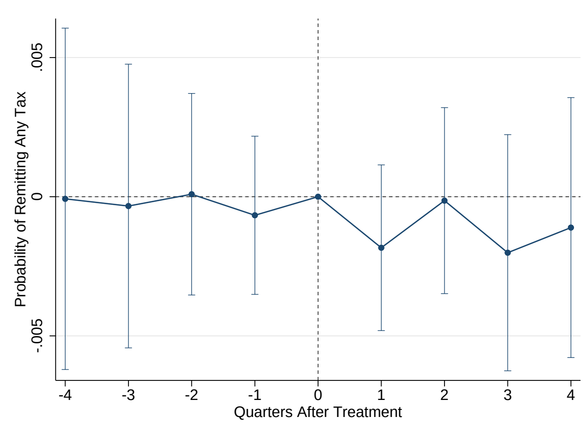
(b) Letter: Tax Remitted



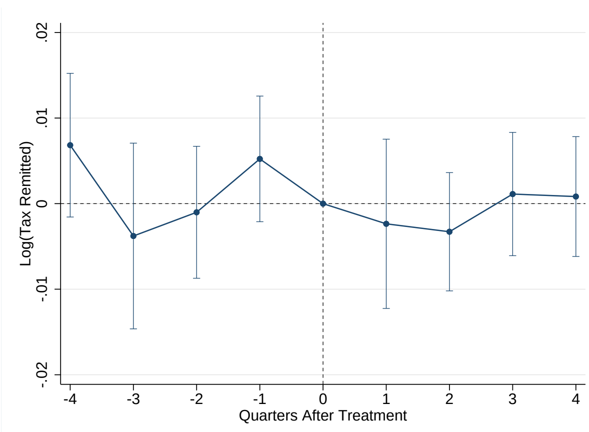
(c) Visit: Probability of Remitting Any Tax



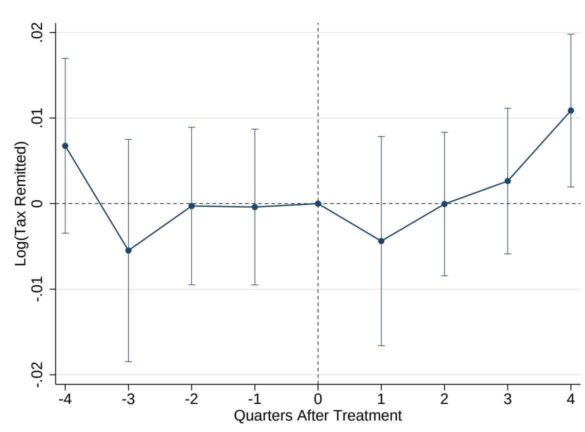
(d) Letter: Probability of Remitting Any Tax



(e) Visit: Log(Tax Remitted)



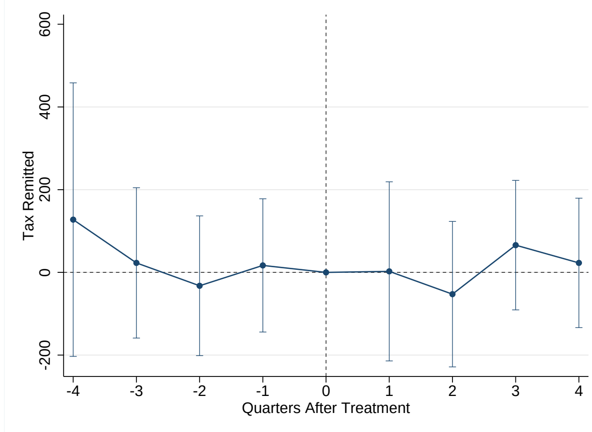
(f) Letter: Log(Tax Remitted)



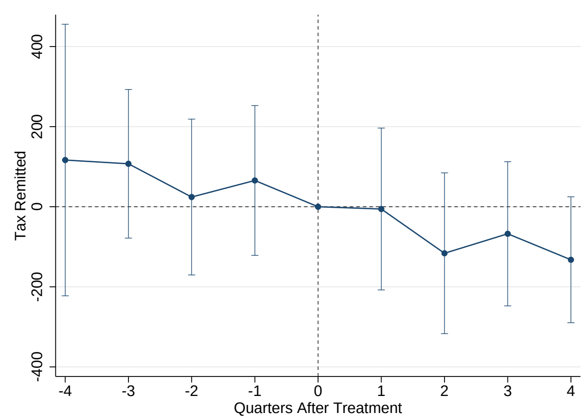
Notes: figures plot estimates and 95 percent confidence intervals. Standard errors clustered by preparer. Tax remitted is winsorized at the 98th percentile. Log(tax remitted) excludes firms remitting no tax.

Figure 2.5: Tax Preparation Firm Network Effects

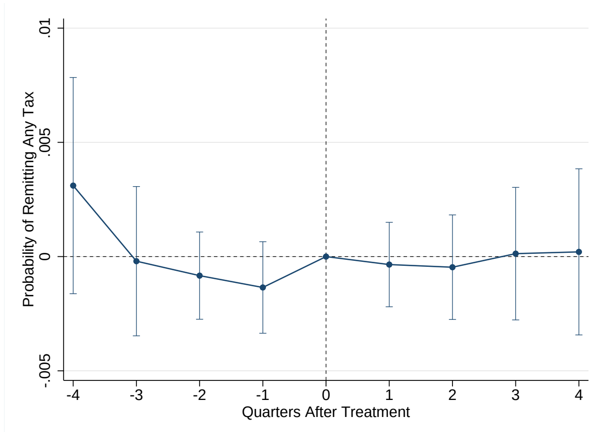
(a) Visit: Tax Remitted



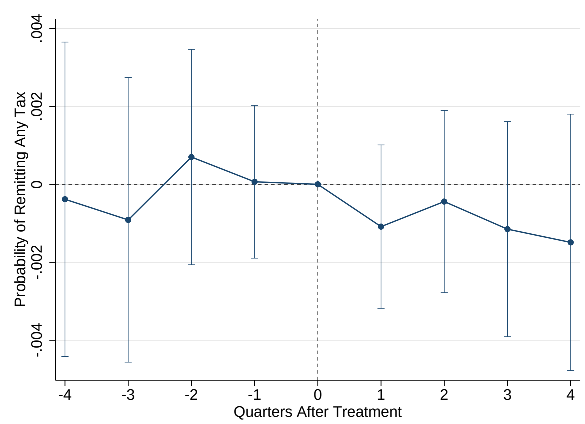
(b) Letter: Tax Remitted



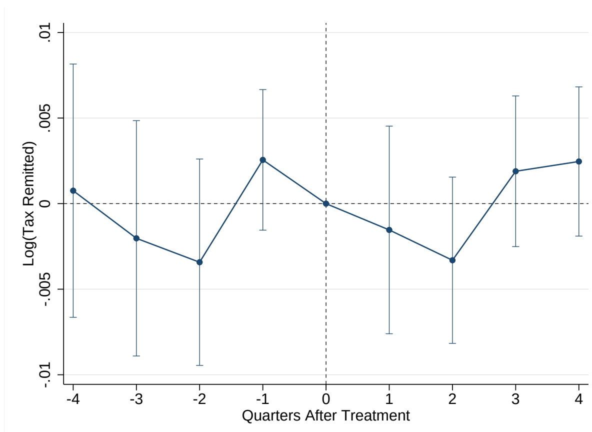
(c) Visit: Probability of Remitting Any Tax



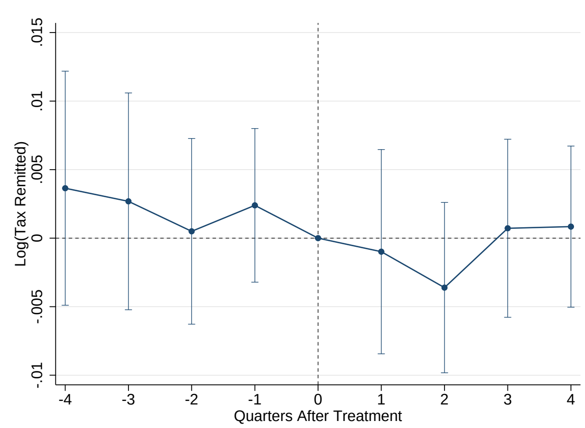
(d) Letter: Probability of Remitting Any Tax



(e) Visit: Log(Tax Remitted)



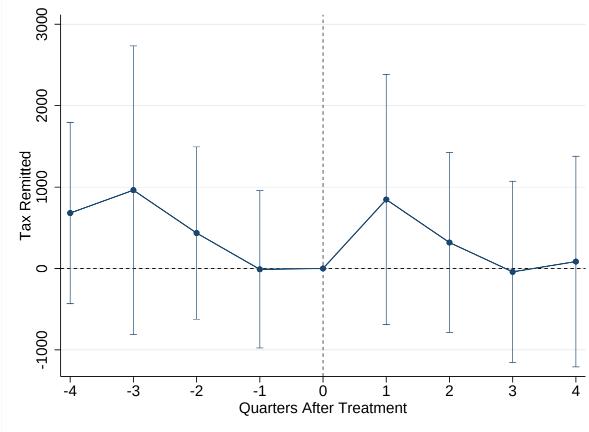
(f) Letter: Log(Tax Remitted)



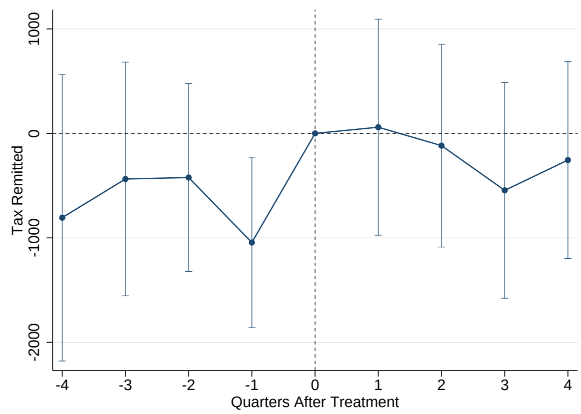
Notes: figures plot estimates and 95 percent confidence intervals. Standard errors clustered by tax preparation firm. Tax remitted is winsorized at the 98th percentile. Log(tax remitted) excludes firms remitting no tax.

Figure 2.6: Narrow Geographic (ZIP+4) Network Effects

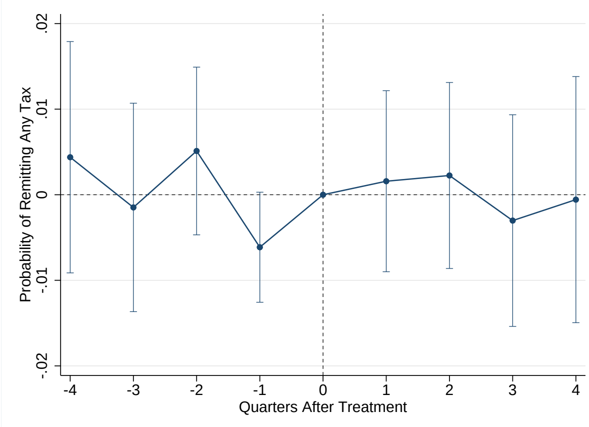
(a) Visit: Tax Remitted



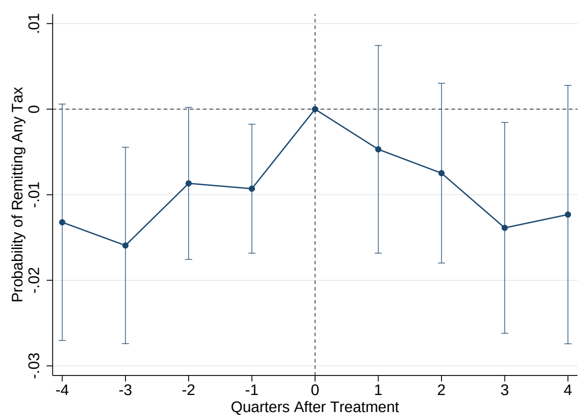
(b) Letter: Tax Remitted



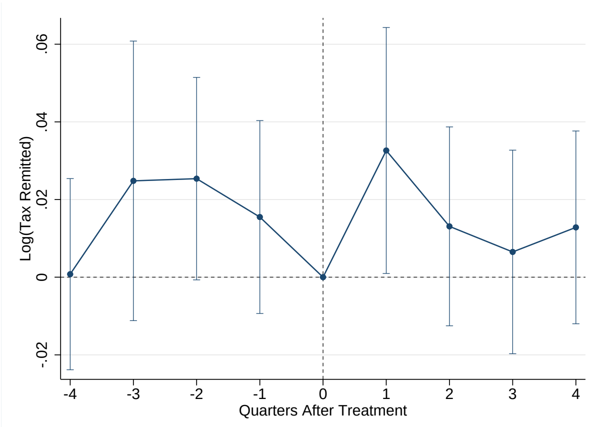
(c) Visit: Probability of Remitting Any Tax



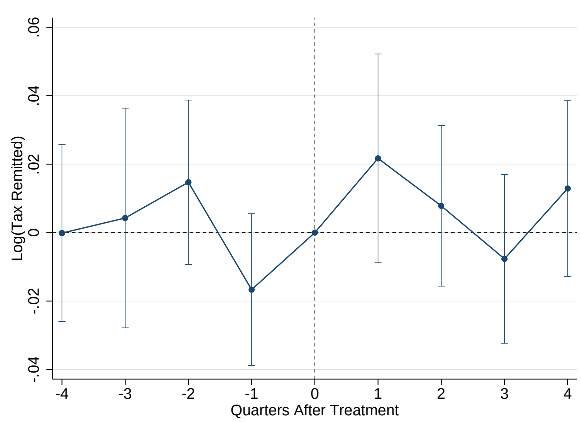
(d) Letter: Probability of Remitting Any Tax



(e) Visit: Log(Tax Remitted)



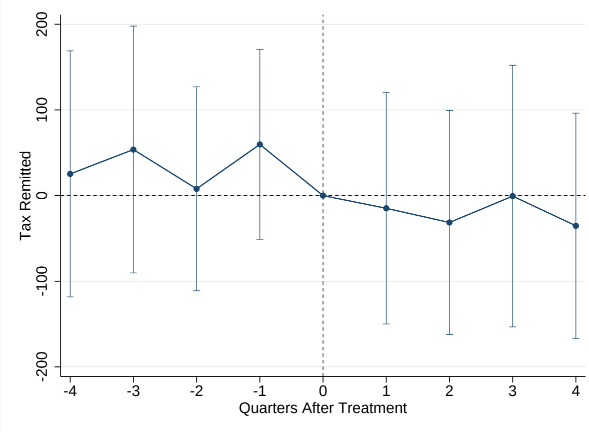
(f) Letter: Log(Tax Remitted)



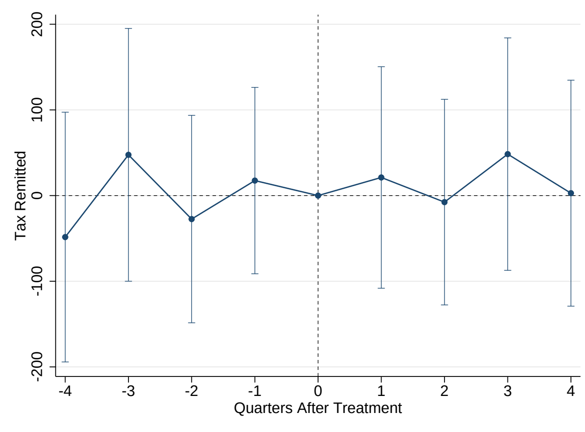
Notes: figures plot estimates and 95 percent confidence intervals. Standard errors clustered by ZIP+4. Tax remitted is winsorized at the 98th percentile. Log(tax remitted) excludes firms remitting no tax.

Figure 2.7: Geographic (ZIP Code) Network Effects

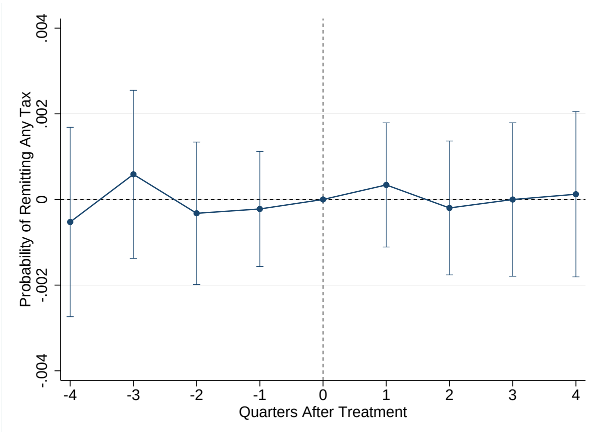
(a) Visit: Tax Remitted



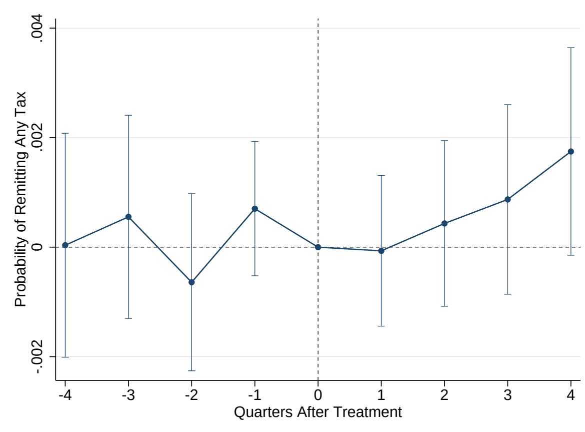
(b) Letter: Tax Remitted



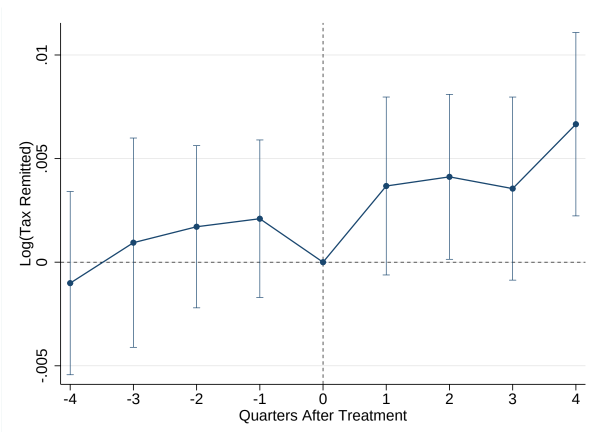
(c) Visit: Probability of Remitting Any Tax



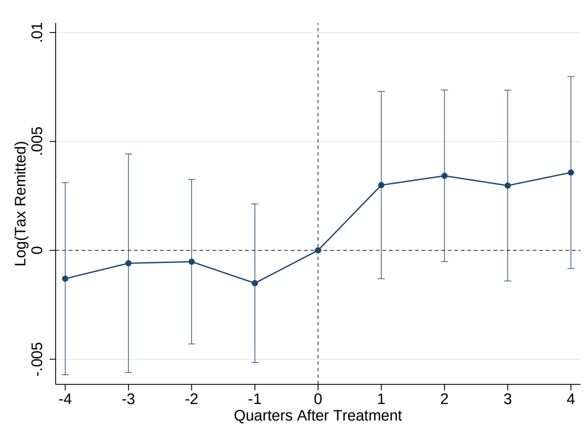
(d) Letter: Probability of Remitting Any Tax



(e) Visit: Log(Tax Remitted)

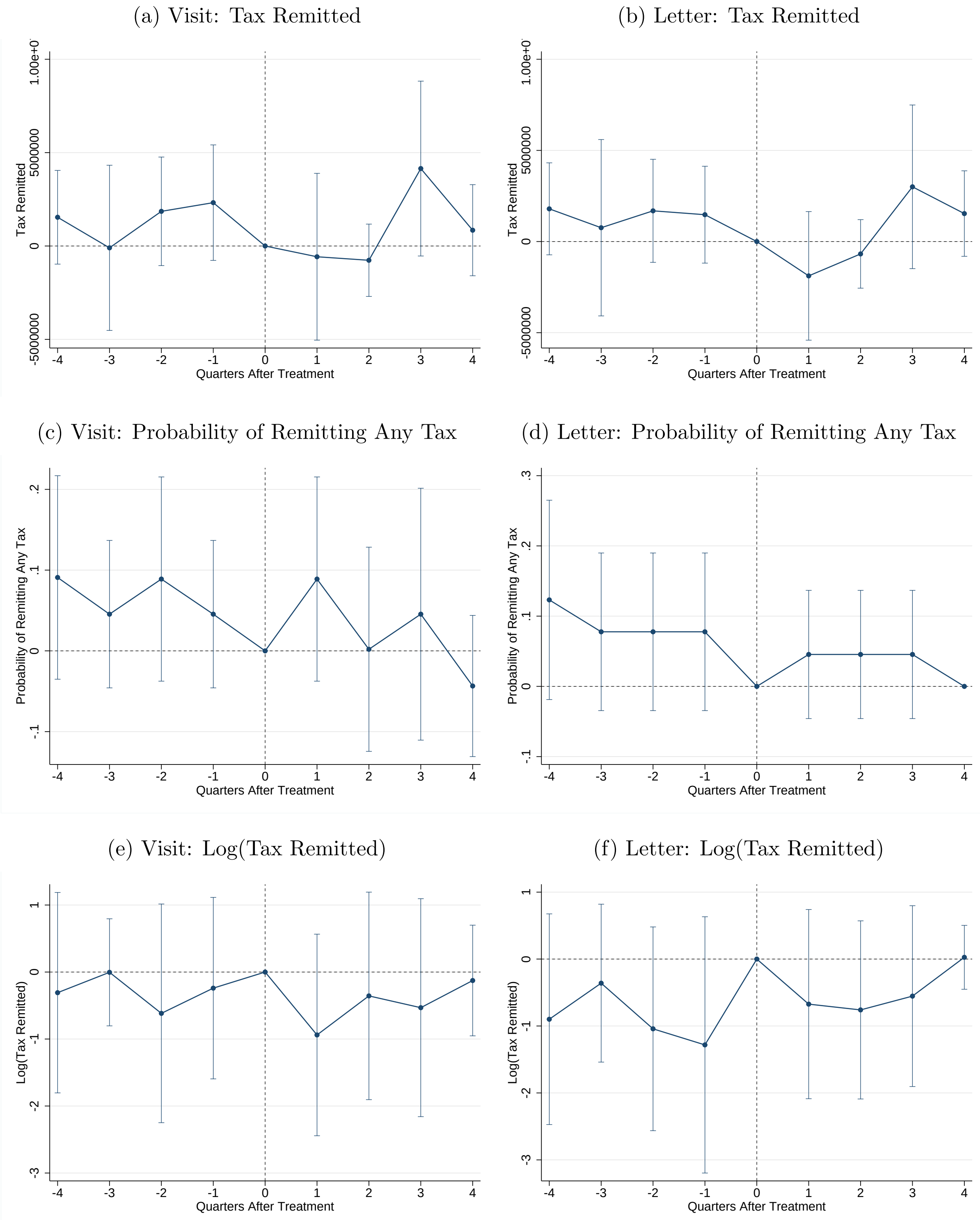


(f) Letter: Log(Tax Remitted)



Notes: figures plot estimates and 95 percent confidence intervals. Standard errors clustered by ZIP Code. Tax remitted is winsorized at the 98th percentile. Log(tax remitted) excludes firms remitting no tax.

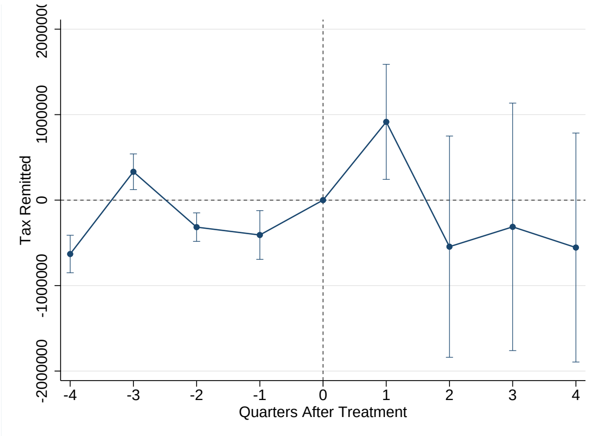
Figure 2.8: Effects on Parents of Treated Firms



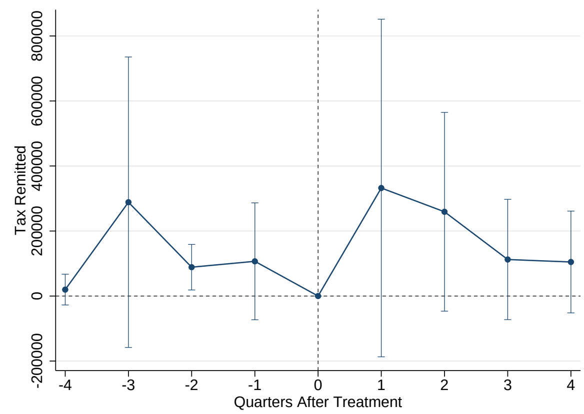
Notes: figures plot estimates and 95 percent confidence intervals. Standard errors clustered by parent firm. Tax remitted is winsorized at the 98th percentile. Log(tax remitted) excludes firms remitting no tax.

Figure 2.9: Effects on Subsidiaries of Treated Firms

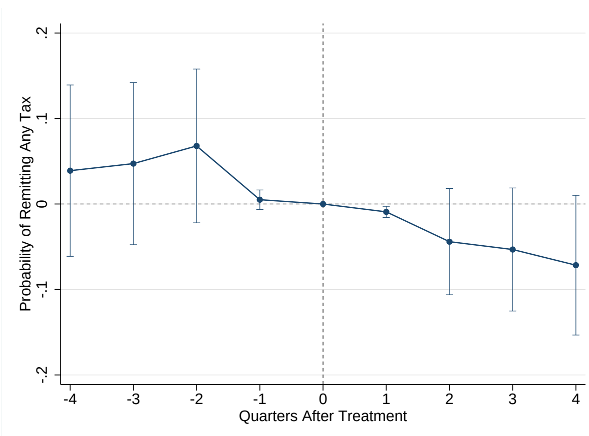
(a) Visit: Tax Remitted



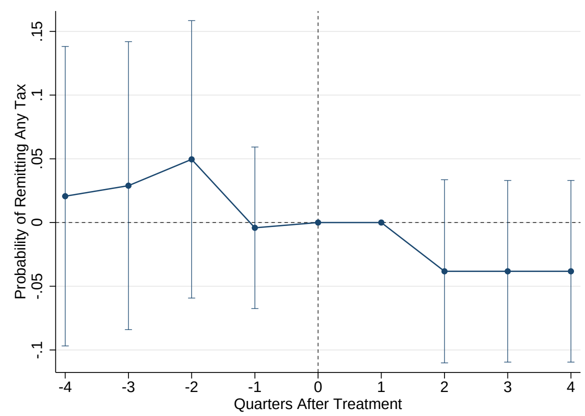
(b) Letter: Tax Remitted



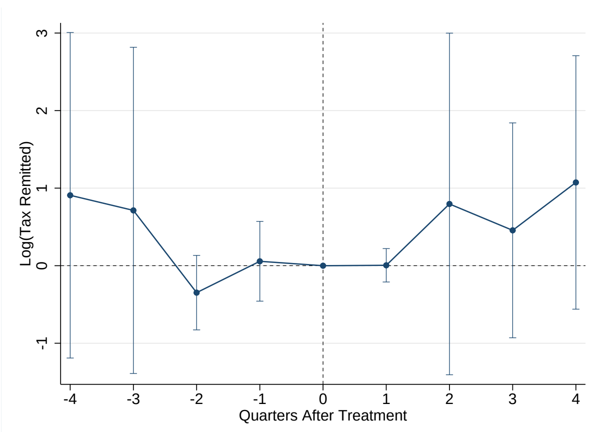
(c) Visit: Probability of Remitting Any Tax



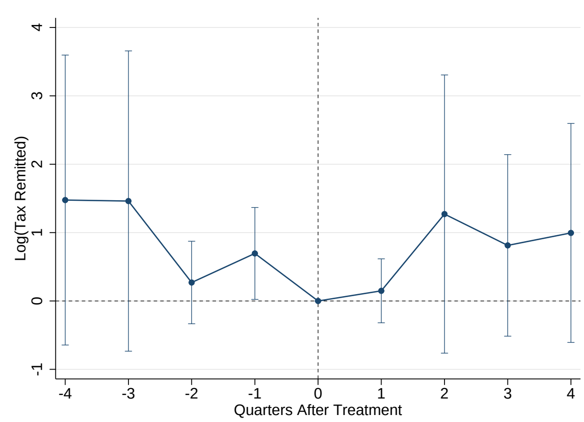
(d) Letter: Probability of Remitting Any Tax



(e) Visit: Log(Tax Remitted)



(f) Letter: Log(Tax Remitted)



Notes: figures plot estimates and 95 percent confidence intervals. Standard errors clustered by parent firm. Tax remitted is winsorized at the 98th percentile. Log(tax remitted) excludes firms remitting no tax.

Table 2.1: Descriptive Statistics One Quarter Before Treatment

	Form 941	Alert C	Control	Letter	Visit
Tax Remitted	21,604	10,683	10,499	11,024	10,523
	[57,279]	[20,554]	[20,022]	[21,265]	[20,342]
Any Tax Remitted	0.686	0.570	0.570	0.573	0.567
	[0.464]	[0.495]	[0.495]	[0.495]	[0.496]
Employees	15.0	27.6	27.5	27.4	27.9
	[32.6]	[39.1]	[39.3]	[38.4]	[39.5]
Median Tax Remitted	2,650	2,846	2,841	2,793	2,899
Median Employees	4	14	14	14	14
Number of Firms	6,489,930	12,172	3,894	4,069	4,209

Notes: Means reported except where otherwise indicated. Sample standard deviations in brackets. Form 941 statistics are from a ten percent random sample of all firms filing Form 941 at any point in the prior year. Employees is the number of Forms W-2 filed in the calendar year before treatment. Tax remitted and employees are winsorized at the 98th percentile.

Table 2.2: Treatment Timeline

By December 31, 2014	Q4 2014 Alert C status determined by algorithm, treatment groups randomly assigned.
January 1-15, 2015	Treatment carried out.
By March 31, 2015	Q1 2015 Alert C status determined by algorithm.
April 1-15, 2015	Firms receive a follow-up round of their assigned treatment if they have both Q4 2014 and Q1 2015 Alert C status.

Table 2.3: Status One Quarter After Treatment

	Alert A or B	Alert C	No Status
Visit (percent)	2	19	78
Letter (percent)	5	28	66
No Treatment (percent)	5	28	67

Notes: Alert A or B status reflects higher risk than firms in with Alert C status, and Alert C status reflects higher risk than the general population of Form 941 filers. All firms with Alert A or B status receive field contact as part of routine procedure. Firms that continued to have Alert C status one quarter after treatment received a follow-up dose of their initially assigned treatment. Firms with no status one quarter after treatment did not receive a follow-up dose of treatment. Source: Author calculations.

Table 2.4: Direct Effects

	Tax Remitted	Any Tax Remitted	Log(Tax Remitted)
Letter * One Quarter Post	94.4	0.0302***	0.00476
	(382)	(0.0110)	(0.0358)
Letter * Two Quarters Post	-112	0.0112	-0.0171
	(438)	(0.0118)	(0.0359)
Letter * Three Quarters Post	163	0.0158	-0.0160
	(459)	(0.0122)	(0.0376)
Letter * Four Quarters Post	177	0.0136	-0.00353
	(459)	(0.0125)	(0.0384)
Visit * One Quarter Post	3,686***	0.129***	0.132***
	(399)	(0.0113)	(0.0348)
Visit * Two Quarters Post	2,726***	0.104***	0.0344
	(438)	(0.0120)	(0.0349)
Visit * Three Quarters Post	2,169***	0.0803***	0.0309
	(451)	(0.0122)	(0.0362)
Visit * Four Quarters Post	1,652***	0.0694***	0.0197
	(448)	(0.0126)	(0.0364)
Letter * Post F-test P-value	0.912	0.0627	0.941
Visit * Post F-test P-value	<0.0001	<0.0001	0.000272
Quarter Fixed Effects	Yes	Yes	Yes
Number of Firm-Quarters	109,548	109,548	77,051
R-Squared	0.0281	0.0513	0.0220

Notes: Standard errors (in parentheses) clustered by firm. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Probability results from linear probability model.

Table 2.5: Direct Effects: Top vs. Bottom Ten Percent by Size

	Tax Remitted	Any Tax Remitted	Log(Tax Remitted)
Letter * Top vs Bottom 10 by Size * One Quarter Post	-889	-0.0300	-0.0297
	(2,425)	(0.0483)	(0.175)
Letter * Top vs Bottom 10 by Size * Two Quarters Post	2,063	0.00604	0.102
	(2,829)	(0.0513)	(0.178)
Letter * Top vs Bottom 10 by Size * Three Quarters Post	-786	-0.0486	-0.0462
	(2,961)	(0.0526)	(0.181)
Letter * Top vs Bottom 10 by Size * Four Quarters Post	-1,465	-0.0819	0.0386
	(3,016)	(0.0526)	(0.187)
Visit * Top vs Bottom 10 by Size * One Quarter Post	6,595***	-0.0311	0.147
	(2,404)	(0.0491)	(0.169)
Visit * Top vs Bottom 10 by Size * Two Quarters Post	4,826*	-0.0127	0.191
	(2,712)	(0.0507)	(0.174)
Visit * Top vs Bottom 10 by Size * Three Quarters Post	2,668	-0.0577	0.0624
	(2,799)	(0.0519)	(0.172)
Visit * Top vs Bottom 10 by Size * Four Quarters Post	1,582	-0.0586	0.190
	(2,881)	(0.0527)	(0.173)
Quarter Fixed Effects	Yes	Yes	Yes
Number of Firm-Quarters	24,687	24,687	16,593
R-Squared	0.222	0.0852	0.221

Notes: Standard errors (in parentheses) clustered by firm. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Probability results from linear probability model. Firms in the bottom 10 percent by size have at most two W-2 employees in the year before treatment, while firms in the top ten percent by size have at least 67 such employees.

Table 2.6: Preparer Network Effects

	Tax Remitted	Any Tax Remitted	Log(Tax Remitted)
Preparer Links to Letter Firms * One Quarter Post	35.3	-0.00183	-0.00437
	(72.2)	(0.00152)	(0.00624)
Preparer Links to Letter Firms * Two Quarters Post	27.0	-0.000140	-0.0000449
	(88.6)	(0.00170)	(0.00428)
Preparer Links to Letter Firms * Three Quarters Post	73.0	-0.00201	0.00264
	(92.6)	(0.00216)	(0.00434)
Preparer Links to Letter Firms * Four Quarters Post	146	-0.00111	0.0109**
	(103)	(0.00238)	(0.00455)
Preparer Links to Visit Firms * One Quarter Post	85.5	0.00188	-0.00236
	(61.3)	(0.00122)	(0.00505)
Preparer Links to Visit Firms * Two Quarters Post	52.3	0.000315	-0.00328
	(81.9)	(0.00149)	(0.00353)
Preparer Links to Visit Firms * Three Quarters Post	156*	0.00123	0.00112
	(88.3)	(0.00189)	(0.00368)
Preparer Links to Visit Firms * Four Quarters Post	243***	0.00162	0.000830
	(94.1)	(0.00224)	(0.00357)
Quarter Fixed Effects	Yes	Yes	Yes
Quarter * Total Preparer Links to Alert C Fixed Effects	Yes	Yes	Yes
Number of Preparer Clusters	10,219	10,219	9,357
Number of Firm-Quarters	1,796,994	1,796,994	1,193,501
R-Squared	0.00361	0.00500	0.0120

Notes: Standard errors (in parentheses) clustered by Preparer. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Probability results from linear probability model.

Table 2.7: Preparation Firm Network Effects

	Tax Remitted	Any Tax Remitted	Log(Tax Remitted)
Preparation Firm Links to Letter Firms * One Quarter Post	-5.67	-0.00109	-0.000987
	(103)	(0.00107)	(0.00380)
Preparation Firm Links to Letter Firms * Two Quarters Post	-116	-0.000442	-0.00361
	(102)	(0.00119)	(0.00317)
Preparation Firm Links to Letter Firms * Three Quarters Post	-67.5	-0.00115	0.000720
	(91.8)	(0.00141)	(0.00331)
Preparation Firm Links to Letter Firms * Four Quarters Post	-132*	-0.00149	0.000841
	(80.2)	(0.00168)	(0.00300)
Preparation Firm Links to Visit Firms * One Quarter Post	2.47	-0.000349	-0.00154
	(110)	(0.000943)	(0.00309)
Preparation Firm Links to Visit Firms * Two Quarters Post	-52.7	-0.000465	-0.00331
	(89.7)	(0.00117)	(0.00248)
Preparation Firm Links to Visit Firms * Three Quarters Post	65.9	0.000129	0.00189
	(79.9)	(0.00148)	(0.00225)
Preparation Firm Links to Visit Firms * Four Quarters Post	22.9	0.000206	0.00246
	(79.8)	(0.00185)	(0.00222)
Quarter Fixed Effects	Yes	Yes	Yes
Quarter * Total Preparation Firm Links to Alert C Fixed Effects	Yes	Yes	Yes
Number of Preparation Firm Clusters	9,759	9,759	9,053
Number of Firm-Quarters	3,563,361	3,563,361	2,468,149
R-Squared	0.00502	0.00620	0.0131

Notes: Standard errors (in parentheses) clustered by Preparation Firm. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Probability results from linear probability model.

Table 2.8: ZIP+4 Network Effects

	Tax Remitted	Any Tax Remitted	Log(Tax Remitted)
ZIP+4 Links to Letter Firms * One Quarter Post	59.5	-0.00469	0.0217
	(527)	(0.00619)	(0.0156)
ZIP+4 Links to Letter Firms * Two Quarters Post	-117	-0.00748	0.00782
	(495)	(0.00536)	(0.0120)
ZIP+4 Links to Letter Firms * Three Quarters Post	-545	-0.0139**	-0.00767
	(526)	(0.00628)	(0.0126)
ZIP+4 Links to Letter Firms * Four Quarters Post	-255	-0.0123	0.0129
	(481)	(0.00770)	(0.0131)
ZIP+4 Links to Visit Firms * One Quarter Post	847	0.00158	0.0326**
	(784)	(0.00540)	(0.0162)
ZIP+4 Links to Visit Firms * Two Quarters Post	319	0.00225	0.0131
	(563)	(0.00554)	(0.0131)
ZIP+4 Links to Visit Firms * Three Quarters Post	-41.5	-0.00302	0.00650
	(568)	(0.00631)	(0.0134)
ZIP+4 Links to Visit Firms * Four Quarters Post	84.9	-0.000572	0.0128
	(660)	(0.00734)	(0.0127)
Quarter Fixed Effects	Yes	Yes	Yes
Quarter * Total ZIP+4 Links to Alert C Fixed Effects	Yes	Yes	Yes
Number of ZIP+4 Clusters	5,916	5,916	5,476
Number of Firm-Quarters	290,745	290,745	201,828
R-Squared	0.00326	0.0104	0.00891

Notes: Standard errors (in parentheses) clustered by ZIP+4. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Probability results from linear probability model.

Table 2.9: ZIP Code Network Effects

	Tax Remitted	Any Tax Remitted	Log(Tax Remitted)
ZIP Code Links to Letter Firms * One Quarter Post	21.1	-0.0000667	0.00299
	(66.0)	(0.000702)	(0.00219)
ZIP Code Links to Letter Firms * Two Quarters Post	-7.65	0.000433	0.00342*
	(61.2)	(0.000772)	(0.00201)
ZIP Code Links to Letter Firms * Three Quarters Post	48.4	0.000872	0.00297
	(69.2)	(0.000883)	(0.00223)
ZIP Code Links to Letter Firms * Four Quarters Post	2.79	0.00175*	0.00357
	(67.3)	(0.000967)	(0.00225)
ZIP Code Links to Visit Firms * One Quarter Post	-14.9	0.000339	0.00368*
	(68.9)	(0.000740)	(0.00219)
ZIP Code Links to Visit Firms * Two Quarters Post	-31.4	-0.000198	0.00412**
	(66.7)	(0.000798)	(0.00203)
ZIP Code Links to Visit Firms * Three Quarters Post	-0.644	-0.00000113	0.00355
	(77.9)	(0.000914)	(0.00225)
ZIP Code Links to Visit Firms * Four Quarters Post	-35.3	0.000122	0.00666***
	(67.1)	(0.000984)	(0.00226)
Quarter Fixed Effects	Yes	Yes	Yes
Quarter * Total ZIP Code Links to Alert C Fixed Effects	Yes	Yes	Yes
Number of ZIP Code Clusters	7,046	7,046	7,008
Number of Firm-Quarters	3,181,959	3,181,959	2,159,992
R-Squared	0.00170	0.00624	0.00949

Notes: Standard errors (in parentheses) clustered by ZIP Code. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Probability results from linear probability model.

Table 2.10: Effects on Parents of Treated Firms

	Tax Remitted	Any Tax Remitted	Log(Tax Remitted)
Subsidiary Letter * One Quarter Post	-1,883,667	0.0455	-0.673
	(1,771,504)	(0.0459)	(0.693)
Subsidiary Letter * Two Quarters Post	-679,367	0.0455	-0.760
	(943,835)	(0.0459)	(0.652)
Subsidiary Letter * Three Quarters Post	3,002,550	0.0455	-0.554
	(2,254,423)	(0.0459)	(0.662)
Subsidiary Letter * Four Quarters Post	1,533,720	-1.05e-14	0.0261
	(1,177,288)	(6.19e-08)	(0.234)
Subsidiary Visit * One Quarter Post	-576,292	0.0889	-0.939
	(2,243,087)	(0.0635)	(0.738)
Subsidiary Visit * Two Quarters Post	-764,650	0.00198	-0.356
	(972,303)	(0.0635)	(0.759)
Subsidiary Visit * Three Quarters Post	4,147,976*	0.0455	-0.532
	(2,351,437)	(0.0783)	(0.798)
Subsidiary Visit * Four Quarters Post	845,508	-0.0435	-0.127
	(1,225,135)	(0.0439)	(0.405)
Quarter Fixed Effects	Yes	Yes	Yes
Parent Clusters	78	78	32
Observations	702	702	266
R-Squared	0.0301	0.0159	0.106

Notes: Standard errors (in parentheses) clustered by parent. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Probability results from linear probability model.

Table 2.11: Effects on Subsidiaries of Treated Firms

	Tax Remitted	Any Tax Remitted	Log(Tax Remitted)
Parent Letter * One Quarter Post	332,315	4.92e-15	0.148
	(258,302)	(1.17e-08)	(0.223)
Parent Letter * Two Quarters Post	259,164*	-0.0383	1.27
	(152,060)	(0.0357)	(0.968)
Parent Letter * Three Quarters Post	112,400	-0.0383	0.813
	(92,011)	(0.0354)	(0.632)
Parent Letter * Four Quarters Post	104,735	-0.0383	0.995
	(77,877)	(0.0354)	(0.762)
Parent Visit * One Quarter Post	915,375***	-0.00917***	0.00531
	(334,732)	(0.00323)	(0.103)
Parent Visit * Two Quarters Post	-544,889	-0.0441	0.796
	(643,782)	(0.0309)	(1.05)
Parent Visit * Three Quarters Post	-312,853	-0.0532	0.456
	(720,091)	(0.0358)	(0.659)
Parent Visit * Four Quarters Post	-554,952	-0.0716*	1.07
	(666,202)	(0.0406)	(0.778)
Quarter Fixed Effects	Yes	Yes	Yes
Parent Clusters	49	49	19
Number of Firm-Quarters	3,573	3,573	780
R-Squared	0.129	0.0805	0.237

Notes: Standard errors (in parentheses) clustered at the parent level. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Probability results from linear probability model.

Table 2.12: Dollar Values and Network Multipliers

	(1) Direct	(2) Preparer	(3) Prep Firm	(4) ZIP+4	(5) ZIP Code	(6) Parent	(7) Subsidiary
Links	1	22.8	98.1	2.84	659	0.00657	0.0326
Letter Effect per Link	94.4	35.3	-5.67	59.5	21.1	-1,883,667	332,315
(se)	382	72.2	103	527	66.0	1,771,504	258,302
Dollars per Letter	94.4	806	-556	169	13,938	-12,380	10,839
(se)	382	1,649	10,116	1,497	43,493	11,643	8,425
Letter Network Multiplier	1	8.53	-5.89	1.79	148	-131	115
Visit Effect per Link	3,686***	85.5	2.47	847	-14.9	-576,292	915,375***
(se)	399	61.3	110	784	68.9	2,243,087	334,732
Dollars per Visit	3,686***	1,952	242	2,404	-9,800	-3,788	29,856***
(se)	399	1,399	10,837	2,224	45,440	14,743	10,918
Visit Network Multiplier	1	0.530	0.0658	0.652	-2.66	-1.03	8.10

Notes: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$.

Table 2.13: Dollar Values and Network Multipliers: Four-Quarter Totals

	(1) Direct	(2) Preparer	(3) Prep Firm	(4) ZIP+4	(5) ZIP Code	(6) Parent	(7) Subsidiary
Links	1	22.8	98.1	2.84	659	0.00657	0.0326
Letter Effect per Link	322	281	-322	-857	64.7	1,973,237	808,615
(se)	1,483	273	296	1,759	215	2,512,362	522,807
Dollars per Letter	322	6,417	-31,580	-2,434	42,647	12,969	26,374
(se)	1,483	6,231	29,062	4,993	141,806	16,512	17,052
Letter Network Multiplier	1	19.9	-98.0	-7.55	132	40.3	81.9
Visit Effect per Link	10,233***	537***	38.6	1,209	-82.2	3,652,542	-497,319
(se)	1,495	250	273	2,046	233	2,846,314	2,316,868
Dollars per Visit	10,233***	12,258***	3,789	3,433	-54,207	24,006	-16,220
(se)	1,495	5,702	26,820	5,807	153,582	18,707	75,567
Visit Network Multiplier	1	1.20	0.370	0.335	-5.30	2.35	-1.59

Notes: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$.

CHAPTER III

Effects of Financing Constraints on Small Firms: Evidence From a Tax Enforcement Experiment

with Alex Turk

3.1 Introduction

Do financial constraints matter for the day-to-day operations of small firms? Links between financial markets and employment loom large after the 2008 financial crisis, which may have reduced employment, especially at smaller firms, in the recession that followed (Chodorow-Reich, 2013; Greenstone, Mas, and Nguyen, 2014). Maintaining firms' access to credit through financial markets has been discussed as a policy objective (Bernanke, 2010), as a means to mitigate the macroeconomic consequences of financial crises.

One way that government policy affects firms' financial constraints is through the tax system. Firm investment rises when accelerated depreciation provides immediate additional cash flows, but not when investment does not respond to additional future cash flows (Zwick and Mahon, 2017). Repatriations of foreign earnings facilitated by tax holidays may increase investment, although difference-in-difference studies using different control groups find conflicting results on this point (Dharmapala, Foley, and Forbes, 2011; Blouin and Krull, 2009; Faulkender and Petersen, 2012). Dobridge (2015) finds that in some recessions firms use larger refunds to finance investment, while in other recessions refunds are dedicated to paying down debt.

The U.S. tax system affects the finances of even the smallest employers by requiring that they regularly remit deposits of employment taxes, which include both social security and Medicare payroll taxes and withheld federal personal income taxes. These employment tax deposits are due quarterly, monthly, or semiweekly frequency, and in 2017 totaled \$2.38 trillion[1], 12 percent of GDP. Employers do not always comply with the quarterly deposit

[1] Payroll tax revenue was \$1.05 trillion and individual income tax withheld by employers was \$1.33 trillion, per Table 1 of the IRS Data Book 2017.

requirement, so making late deposits is in effect a loan (subject to substantial interest) from the government to firms. These loans parallel the government lending to financially constrained individuals that Andreoni (1992) models. Financially constrained firms could rely on cash flows from revenue to finance payroll and other obligations. Can enforcing the requirement that firms make deposits on time reduce available cash flows enough to reduce firm employment and operations?

Tax enforcement contact could reduce employment because firms are financially constrained. If frictions in debt and equity markets limit the other available means for small firms to finance their operations, they may rely on revenues or deferring tax deposits to meet payroll needs. Late deposits incur penalties and interest, but may have lower transaction costs than alternative sources of financing, and may not be visible to other creditors or potential creditors unless the IRS files a Notice of Federal Tax Lien, making the tax debt public. Enforcement contact about deposit requirements limits the firm's ability to use late deposits as a source of funds, leading firms to shift limited cash from payroll needs to tax deposits.

Tax enforcement could also reduce employment by raising the perceived effective tax rate on employee compensation. Enforcement contact can deter firms from attempting to avoid paying employment taxes, especially as the employment tax-related contact we study takes place before returns are filed and tax is assessed. In part, contact signals that other enforcement actions are more likely, raising the likelihood that tax debt will be collected even when the firm has limited liability. The IRS has several tools to collect unpaid employment taxes, including trust fund recovery penalties for which corporate officers can be personally liable. While the IRS typically has 10 years to pursue debts, trust fund recovery assessments against corporate officers must be made within 3 years. Early contact can signal that limited liability is less likely to shield the officers from the tax debt. To the extent that contact changes firms', and their officers', perceptions of the difficulty of avoiding paying employment taxes, contact raises the effective employment tax rate. Higher perceived effective tax rates on employment then provide an incentive to cut employment.

We investigate the effect of enforcing the requirement to make employment tax deposits on employment and operations using a randomized experiment. In this experiment, the IRS randomly assigned 12,172 firms whose tax deposits had been declining, an indicator of financial distress, to receive a letter, receive an in-person visit at their place of business from an IRS Revenue Officer, or to a control group. The text of the letter or topic of the visit was a reminder about deposit requirements and penalties for late payment, which raises the effective cost of this form of financing. The effects of contact are large; Boning, Guyton, Hodge, Slemrod, and Troiano (2018) find that in-person contact raised the tax remitted by firms visited in person by an average of ten thousand dollars over the year following contact.

The experiment provides us with a cleanly identified setting in which to study the consequences of tax enforcement that reduces the cash flows available to small, likely financially-constrained firms. Randomization is the gold standard for causal inference, and rules out many of the concerns that arise with identification strategies for observational data. The firms we study are at risk of falling behind on their tax payments, and absent randomization it would be difficult to construct a comparison group. IRS administrative data allow us to trace how firm employment responds quarter-by-quarter and to track the use of non-employee contractors. Data on payments firms receive through credit card and third-party payment networks are available from information returns, and annual tax filings provide a measure of whether firms continue to operate.

We find that in-person contact about employment tax deposits substantially reduces employment for a prolonged period. We estimate that employee headcount falls by five percent in the second year after an in-person visit, and that quarterly payroll is three to seven percent lower up to fourteen quarters after the visit. We do not detect employment or payroll effects of the letter, which has no detectable immediate effect on the amount of tax remitted. In the longer run, there is no evidence that visited firms remit less tax than control firms, and the letter recipients' tax remitted is detectably lower only slightly, in a single quarter. We find that both the letter and visit reduce the employment tax firms report is due in the two years following contact. We find no evidence that firms substitute independent contractors for employees following contact, that firms have lower sales as partially proxied by third-party-reported payments[2], or that firms close and cease filing after contact.

We consider the welfare implications of enforcing deposit requirements in light of these findings[3]. Employment effects *per se* do not affect welfare. Changes in tax remitted do affect welfare, but we find no evidence that visited firms reduce tax remitted in the long run, and only a slight dip in tax remitted by letter recipients. We consider the possibility that employment declines and the associated drop in employment taxes due eventually slightly reduces the trust fund taxes recovered from corporate officers. This does not alter the conclusion one would reach from the firms' tax deposits alone: that the visit improves welfare and the letter slightly reduces it. Absent from these welfare calculations are any changes in innovation spillovers from the affected small firms or efficiency gains from production shifting

[2]In future, we plan to investigate whether firms report lower sales or corporate income tax liability on their annual returns, which could reflect reductions in operations or substitution into income tax misreporting. We also plan to investigate whether firms report items on their corporate tax returns that suggest reductions in investment or in profits, other possible margins of response.

[3]Firms' welfare is not well-defined, so we assume a representative agent owns and works for the firms, and adopt the Keen and Slemrod (2017) framework to consider the welfare effects of tax administration.

to lower-cost firms under a uniform enforcement regime.

This paper contributes experimental evidence that employment responds to cash flows to the literature on business responses and financing constraints. Existing literature uses quasi-experimental strategies with naturally occurring variation in cash flow, including from the 2008 financial crisis (Greenstone, Mas, and Nguyen, 2014; Chodorow-Reich, 2013), from acceleration of payments to federal contractors (Barrot and Nanda, 2016), and from mandatory pension plan contributions (Rauh, 2006). Our evidence from a randomized experiment requires less stringent identification assumptions than these natural experiments.

We also demonstrate that effective tax enforcement can in fact simultaneously raise revenue and reduce the reported tax base. Our results show that as tax remitted rises, firms adjust their behavior, specifically employment, shrinking the tax base. We show that in this case firms do not use available avoidance strategies (independent contractors) as a substitute for evasion (failure to remit tax deposits). Our results leave open the possibility that firms substitute away from failure to remit deposits and towards misreporting employees or income. Firms could, for example, evade employment taxes by taking employees off the books altogether.

Our findings relate to the literature on optimal tax administration and production distortions. Diamond and Mirrlees (1971) show that optimal tax policy does not distort production, but their model omits administrative costs. Incorporating administrative costs can make tax instruments that distort production optimal in practice. Best, Brockmeyer, Kleven, Spinnewijn, and Waseem (2015), for example, show using data from Pakistan that a turnover tax produces revenue at much lower administrative cost than a profits tax, and thus improves welfare. We demonstrate that changes in enforcement can raise large amounts of revenue at little cost while distorting production, and find that this can be optimal in a new context.

The remainder of the paper includes a discussion of the experiment and background in Section 3.2, the method and results in Section 3.3, implications for welfare in Section 3.4, and concluding remarks in Section 3.5

3.2 Background, Experiment, and Data

Employers remit most of the U.S. federal personal income tax, which they withhold on behalf of their employees, as well as social security and medicare taxes. Collectively these taxes are a large share of U.S. federal revenue, and the IRS monitors firms' remittances to ensure compliance. The interest structure for failure to make these deposits on time includes penalties of two percent if the deposits are up to five days late, five percent if the deposits are 6-15 days late, ten percent if the deposits are more than fifteen days late, and fifteen percent if the deposits are not made within ten days of the IRS sending a bill, with interest charged at the federal short-term rate plus three percent. Enforcement can even pierce the corporate veil–if an employer does not make the required payments, corporate officers can be personally liable for them through the trust fund recovery penalty.

As part of the enforcement regime for employer tax deposits, the IRS assesses the risk that employers will fail to deposit the required tax each quarter. Firms with declining tax remittances are assigned to risk levels. Revenue Officers routinely visit the firms at the highest risk, which are designated Alert A or B. Firms at the next-highest risk level are designated Alert C, and may not be contacted.

We study a randomized experiment in which 12,172 firms at the moderately-high-risk Alert C level were assigned to receive either a letter or an in-person visit about their employment tax deposits in January-February 2015. Both the letter and visit include reminders about deposit responsibilities and penalties. The letter is included as Figure B.1. The experimental treatments include a follow-up rule for the second quarter after treatment: visited firms still in the Alert C risk category after one quarter received a second visit, and letter firms still in the Alert C risk category after one quarter received a second letter. This rule more closely mimics the policy regime for Alert A and B firms, which receive repeated visits as long as they remain in those categories, and thus mimics the policy regime.

We use IRS administrative data on return filings and tax payments to study the subsequent behavior of these firms. Our measure of tax remitted is employment tax deposits towards a calendar quarter made by the end of that calendar quarter[4]. Information on employment, wages, and employment tax due comes from the quarterly employment tax returns (Form 941), and from annual sums of employee payment information returns (Form W-2). Form 1099-MISC captures non-employment compensation the firms in our sample pay to independent contractors who may substitute for employees, while Form 1099-K allows us to observe payments to the firms we study from credit card companies and other third-party payment networks, a measure of firm sales. We also examine whether corporations file an annual return (Forms 1120 or 1120S), as the annual filing requirement for these forms implies that filing is a measure of whether the business continues to operate. Employees reported quarterly reflects employment in a pay period at the end of the calendar quarter, while the annual measure of employment captures the number of people the firm employs at any point during the year. We take as our measure of wages the wages, tips, and other compensation paid to employees in a quarter or year. Tax due to be remitted by the firm is the total of federal income tax withheld and social security and medicare taxes. To manage the influence

[4]This measure excludes late deposits.

of a few large firms that would otherwise dramatically inflate the variance of the estimates, we winsorize each of these variables at the 98th percentile of the pre-treatment distribution, excluding the filing indicator and 1099-MISC counts and payments, as fewer than one percent of firms report any such payments.

The firms we study are fairly small, as Table 3.1 illustrates. They have an average of 12.8 employees as of the end of the quarter before treatment, and employ 27.6 people at any point during the year before treatment. They paid an average of $126,100 in wages, tips, and other compensation in the quarter before treatment, and a total of $551,862 over the year before treatment. The total social security, medicare, and federal income taxes the firm is responsible for remitting in the quarter before treatment is $34,282, and in the year before treatment is $109,914. The average number of independent contractors paid is 3.08, with average contractor non-employment compensation of $133,279. On average, firms received $227,791 through credit card and similar networks. The firms are mostly corporations (including S corporations), with 74.4 percent filing a return in the year before treatment, and the remainder of the firms are presumably a mix of sole proprietorships, partnerships, etc.

3.3 Method and Results

3.3.1 Event Study Specification

We use an event-study regression design that captures the time path of the response to contact and accounts for both time-invariant differences across firms and time-period specific shifts. We estimate the regression specification

$$Y_{it} = \sum_j \sum_k \beta_{jk} * Treat_j * Period_k + \mu_t + \epsilon_{it}, \tag{3.1}$$

where Y_{it} is an outcome of interest, for example the number of employees or wages paid, $Treat_j$ is an indicator for whether firm i received treatment j, either a letter or a visit, $Period_k$ is an indicator equal to one if t is k periods before or after treatment, μ_t is a time fixed effect common to all firms, and ϵ_{it} is an error term. The effects of interest are the set of post-treatment β_{jk}, that is, the effect of each treatment in each period after treatment. We cluster standard errors by firm.

For β_{jk} to capture the causal effect of enforcement contact, it must be the case that there are no confounding contemporaneous changes that would affect the control and treatment groups differently, and that the groups would experience similar changes if no firms were treated. These assumptions are likely to be satisfied because treatment status is assigned randomly, and treatment status only affects IRS actions towards firms through the contacts we describe. Time-varying shocks are unlikely to be correlated with randomly assigned treatment status, and are thus unlikely to account for the results we find.

A placebo test supports the assumption that absent treatment the control, letter, and visit groups' outcomes would be similar. In our regression specification, we include all time periods in the fourteen years before treatment, and find that there are not significant differences between the three groups in this extensive pre-treatment period for the outcomes we study.

3.3.2 Effects on Tax Remitted

To begin, we place our results in context by referring to the results Boning, Guyton, Hodge, Slemrod, and Troiano (2018) find for treated firms' tax payments. These results are plotted in Figure 3.1. In the year after receiving a letter, firms remit similar amounts of tax to the control group. Visiting a firm in person, in contrast, strongly increases the tax that firm remits in the next four quarters. Visited firms remit an additional \$10,233 across the four quarters after treatment.

We extend the time series in Boning, Guyton, Hodge, Slemrod, and Troiano (2018) forward beyond one year after treatment to cover the four years after treatment. These results are reported in the first column of Tables 3.2 and 3.4. We find that in the ninth quarter after receiving a letter, firms in fact remit \$1,072 *less*, with $p < 0.05$. The reasons for this decline are explored below. The visit has no effects on firms' payments more than one year later.

3.3.3 Effects on Number of Employees and Compensation

Firms contacted about employment taxes may cut back on employees or wages for two reasons. Tax enforcement contact could cause firms to perceive a higher effective tax rate on labor inputs, leading them to employ less labor. If it is difficult to obtain additional cash, making immediate tax payments could reduce the budget available to compensate employees and purchase other inputs.

We begin with firms receiving a letter. We find that receiving a letter does not cause firms to detectably reduce employment or compensation. Relative to control firms, letter recipients neither cut back the number of employees nor wages, tips, and other compensation significantly in any quarter or year after treatment, as Tables 3.2 and 3.3 report. Figures 3.2 and 3.3 show the effects of both treatments on employment and compensation.

In contrast, we find that in-person visits cause firms to substantially reduce the number of employees, persistently reducing compensation. Estimates for visited firms are reported in greater detail in Tables 3.4 and 3.5. Visited firms employ 1.39 fewer workers two years post-visit (with $p < 0.05$), a reduction of five percent relative to the population mean of 27.6 employees in the year before treatment (from Table 3.1). Visited firms pay significantly ($p < 0.05$) lower wages per quarter, by amounts varying between \$3,771 and \$7,221 in each of the ten quarters following treatment as well as in the 14th quarter after treatment. These effects are economically meaningful - they are between 1.5 and 3 percent of the pre-treatment mean quarterly compensation of \$126,100. Summing across quarters in which the reduction in compensation is significant yields \$62,329, which is 6.1 times larger than the \$10,233 in additional tax visited firms remit over the year following the visit. Given the relative magnitudes, financing constraints and the additional tax remitted can fully explain the wage reductions only if there is a compounding effect over time, where reductions in wages in each quarter deprive the firm of additional revenue that would otherwise fund compensation in future quarters. As one would expect, wages paid annually also decline, although the decline is statistically significant only two years after contact, at \$27,771.

We also investigate whether firms substitute non-employee labor for employees, avoiding the payroll and income tax withholding obligations that come with the employee relationship, and do not detect such substitution in either the number of non-employee contractors or the amount of contractor compensation. Figures 3.4 and 3.5 show that firms' use of contractors is not detectably higher in any year after treatment.

Our finding that shocks to financing constraints have large employment effects on a group of firms likely to be constrained is consistent with existing literature using quasi-experiments. Chodorow-Reich (2013) finds access to lending affected employment growth during the 2007-2009 recession by up to five percentage points, and that effects were largest for the smallest firms. The payroll response we find is large compared to Barrot and Nanda (2016), who find that for every dollar of federal payments accelerated by 15 days, payroll increased by ten cents. The difference could be explained by the larger change in payment timing in our experiment, where firms may make payments multiple quarters earlier after contact, or by the high likelihood that the firms we study are constrained.

3.3.4 Effects on Employment Tax Due

Given that firms report lower employee compensation, one might expect them to report that less employment tax is due. Firms are responsible for remitting federal income tax and social security and medicare taxes, and when wages fall, the required tax deposits also fall. We find that both letter and visit recipients report significantly less tax due after treatment. Figure 3.6 illustrates these reductions in tax due. The coefficient on letter recipients' tax due is significantly ($p < 0.05$) negative one, two, four, and eight quarters after treatment, with point estimates varying between -$1,343 and -$1,720. This is consistent with a reduction in wages paid (although the effects on wages paid for letter recipients are not statistically significant, this is consistent with the negative point estimates we find for letter recipients) and explains the $1,072 reduction in letter firms' tax remitted nine quarters after treatment discussed above. The tax that visited firms report is due is significantly ($p < 0.05$) negative in each of the first ten quarters after treatment and thirteen and fourteen quarters after treatment. The point estimates vary between -$2,219 six quarters after treatment and -$1,207 one quarter after treatment.

Why do visited firms' reports of the tax due fall without a significant fall in the tax they remit? Even though less tax is due, visited firms remit a higher fraction of the tax that is due–the visit is intended to raise compliance with the requirement to remit tax. The higher fraction of tax due that is remitted offsets the lower quantity of tax that is due beginning five quarters after treatment, resulting in no significant effect on tax remitted.

3.3.5 Effects on Payments Received and Filing a Return

Firms could also respond to contact that tightens financial constraints in ways that reduce sales or increase the likelihood they close. Credit card companies and other payments processors report the payments they make to firms, a partial measure of firms' sales, on Form 1099-K. We test whether firms close using firms' annual corporate tax return filings. To the extent that firms in the population we study are corporations, their annual tax return filings provide evidence that they continue to operate. All corporations are required to submit an annual tax return on either Form 1120 or Form 1120S (if a subchapter S corporation), so firms ceasing to file such a return are presumably closed. As the summary statistics in Table 3.1 show, 74 percent of firms in the population we study filed a corporate tax return for the year before treatment. If contact makes firms more likely to close, fewer contacted firms would file tax returns in later years.

We find that IRS contact about employment taxes does not reduce credit card payments to employers or the likelihood that they continue to operate. Total payments from credit card and third-party networks, which are reported on Form 1099-K and shown in Figure 3.7, do not decline for either treatment group following treatment. Firms receiving a visit or letter are also not substantially less likely to file a corporate tax return in each of the four years after treatment, as Figure 3.8 shows. As we do not find that contact makes firms less likely to file, it is unlikely that the declines in employment and wages we find are due to firm closures.

3.4 Welfare Consequences

Does evidence that tax enforcement leads firms to reduce employment change the welfare consequences of tax enforcement? We apply the Keen and Slemrod (2017) framework and address the challenge that firm welfare is not well-defined by assuming that a representative agent both owns and works for the firm. In this framework, firm responses to policy changes have welfare consequences because of a fiscal externality; they affect the government budget.

Our framework omits externalities other than the fiscal externality, although they may be important in this context, because they are difficult to quantify. If small firms generate positive spillovers from innovation, for example, then if tax enforcement leads them to cut back hiring and reduce innovation it reduces welfare through a non-standard channel. The usual framework assumes that firms face the same effective tax rate, but allocative inefficiency can arise when some firms have better evasion opportunities than others, and changes in enforcement can correct this inefficiency, as Kopczuk, Marion, Muehlegger, and Slemrod (2016) explore. In our context, if tax enforcement causes firms whose competitive advantage arose from evasion to shrink, reducing hiring, then reduced hiring reflects resource reallocation toward more productive uses.

Limiting the analysis to changes in the tax employers remit, the hiring effects and accompanying changes in tax due we find do not alter the conclusion in Boning, Guyton, Hodge, Slemrod, and Troiano (2018) visits we study raised revenue and improved welfare, but we do find that letters affect tax remitted only through a reduction nine quarters later, and thus reduce welfare. We find no evidence that the reductions in hiring translate into reductions in tax remitted by visited firms, or that tax remitted more than one year after IRS contact changes.

Conducting welfare analysis solely on the basis of the taxes firms remit assumes that changes in tax due will not affect revenue collected from other sources. However, the IRS may collect taxes that firms do not remit from the corporate officers responsible for remitting taxes. Suppose that the IRS ultimately collects a fraction α of any change in taxes due from sources other than the firm's own deposits. The value of α is likely small, as employment taxes are largely collected directly from employers, so we choose $\alpha = 0.05$. Then, letting $r_{R,L}$ and $r_{R,V}$ be the revenue effects of the letter and visit through the firm's own remittances, and αt_L and αt_R be the revenue effects of collecting α times the change in tax due through from other sources, the total revenue effect of a letter or visit is $r_L = r_{R,L} + \alpha t_L$ or $r_V = r_{R,V} + \alpha t_V$.

Summing across quarters with statistically significant changes, we have

$$r_L = -\$1,072 - \alpha * \$5,952 = -\$1,369.6 \tag{3.2}$$

$$r_V = \$10,233 - \alpha * \$21,116 = \$9,177.2. \tag{3.3}$$

We then use the formula $\Delta W = (v' - 1)r - v'a - c$ from Keen and Slemrod (2017), where ΔW is the change in welfare, a is the administrative cost of the treatment, c is the additional compliance cost borne by the taxpayer, and v' is the marginal value of public funds, for which we take the value $v' = 1.17$ from Mayshar (1991). The administrative cost of an additional letter is \$4, and an additional visit costs \$220. We assume that compliance costs are twice the administrative costs, following Mayshar (1991). Then the welfare effects are

$$\Delta W_L = 0.17 * -1,369.65 - 1.17 * 4 - 8 = -\$245.512 \tag{3.4}$$

$$\Delta W_V = 0.17 * 9,177.2 - 1.17 * 220 - 440 = \$862.724. \tag{3.5}$$

By this metric, the letter reduces welfare, while the visit increases welfare. It is clear from these calculations that the value of α does not affect the sign of the letter's welfare effect, and that either α or the reduction in tax due would need to be much larger for the visit to reduce welfare. One caveat is that one might expect letters to cause small, immediate increases in tax remitted, which would improve their consequences for welfare but may simply be smaller than we are able to detect. These calculations are also necessarily limited by the caveats discussed above - there are likely additional terms that should be included in the welfare calculations, but that are difficult to measure.

3.5 Conclusion

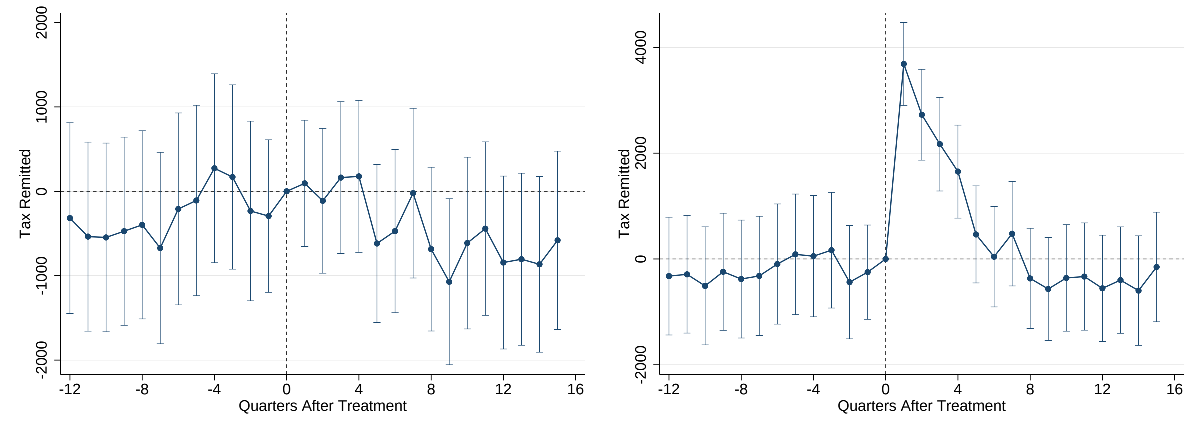
We find that randomly assigned in-person contact from IRS, which led firms to remit substantially more tax over the following year, also led them to substantially reduce the number of employees and total wages paid for more than two years after contact. On average across firms, the reduction in wages paid is more than twice as large as the increase in tax remitted, and both employment and wages fall by about five percent. In line with the reduction in wages paid, taxes due also fall. IRS contact does not lead firms to close, as contacted firms are no less likely to file annual corporate tax returns. If firms' reports are accurate, tax enforcement substantially changes how they do business.

Employment declines following tax enforcement are consistent with financial constraints, though changes in effective tax rates or underreporting employees could also explain our results. Enforcement contact tightens financial constraints, and could limit the funds available for firms to finance payroll, leading to employment declines. If instead visits change firms' expectations about the employment taxes they will need to remit, then employment declines could reflect adjustment to a regime of higher perceived effective tax rates. Employment declines could also reflect evasion by misreporting the number of employees.

What do changes in employment mean for the welfare effects of tax enforcement? In the simplest framework, only administrative and compliance costs and the tax remitted matter, not employment. The visit has large, positive welfare effects by this metric, while the only detectable effect of the letter on tax remitted is a slight, delayed dip, which implies a reduction in welfare. Assuming that five percent of the reduction in tax due associated with lower employment will be remitted by employees rather than the firm does not alter these welfare conclusions. The reduction in hiring might have other spillover effects, for example reducing welfare if it reduces the innovation spillovers firms generate or enhancing welfare if it reflects reallocation toward more productive firms that do not evade taxes.

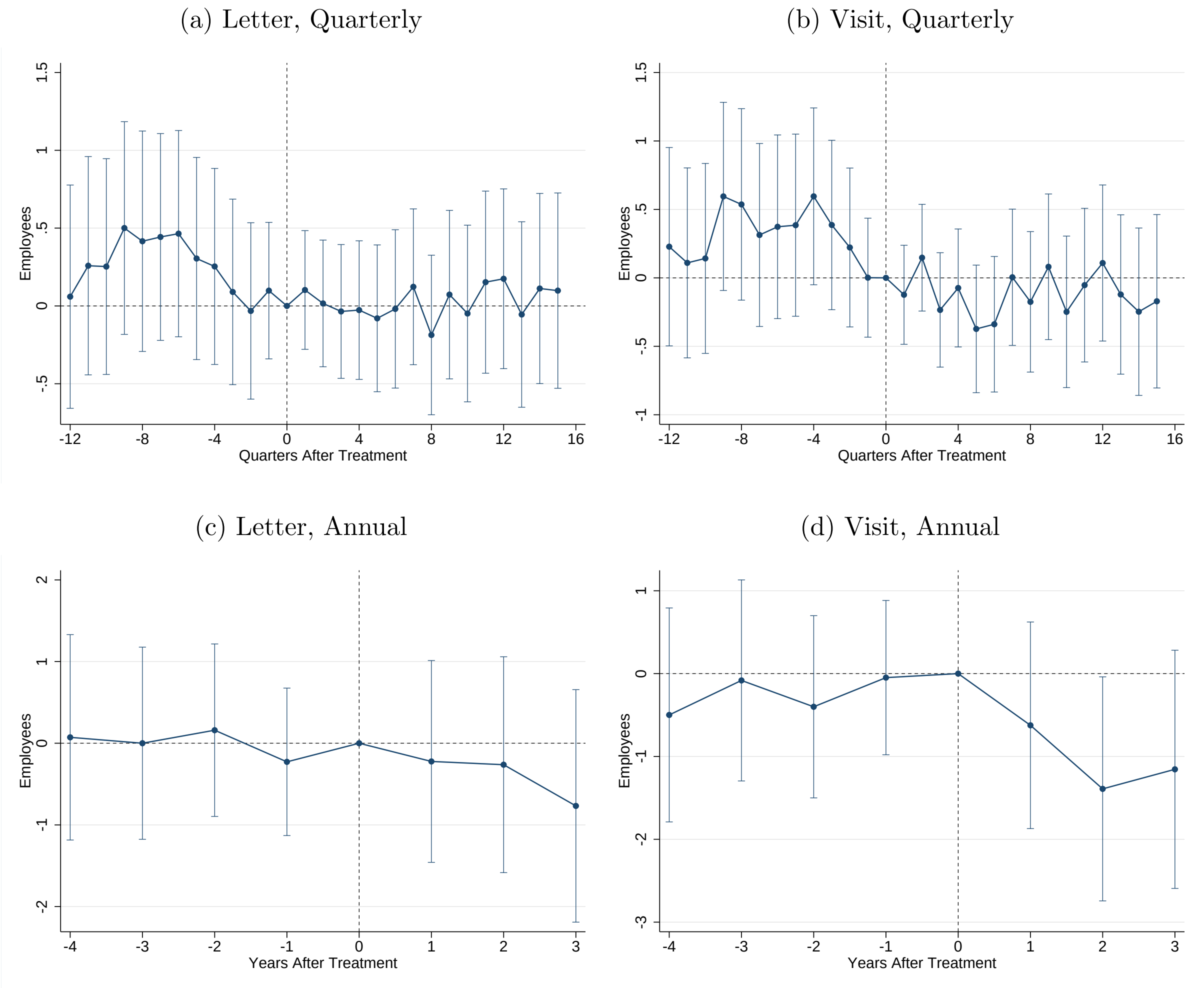
Figure 3.1: Effects of Treatments on Tax Remitted

(a) Letter (b) Visit



Notes: Plots 95 percent confidence intervals from standard errors clustered by firm. Winsorized at the 98th percentile. From IRS data on payments received.

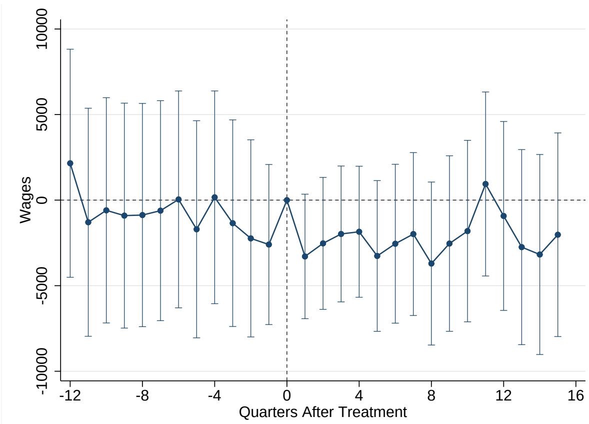
Figure 3.2: Effects of Treatments on Number of Employees



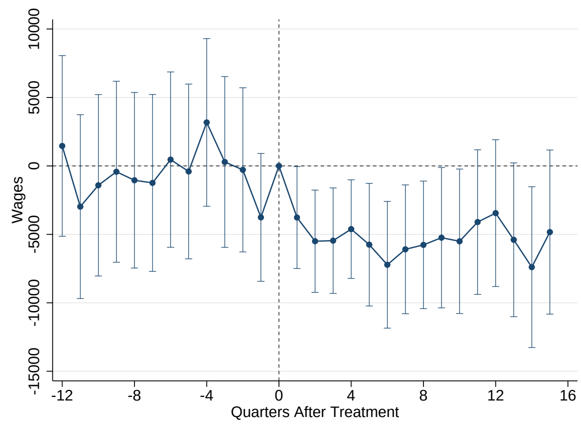
Notes: Plots 95 percent confidence intervals from standard errors clustered by firm. Winsorized at the 98th percentile. Quarterly values from Form 941 and annual values from total of all Forms W-2.

Figure 3.3: Effects of Treatments on Wages, Tips, and Other Compensation

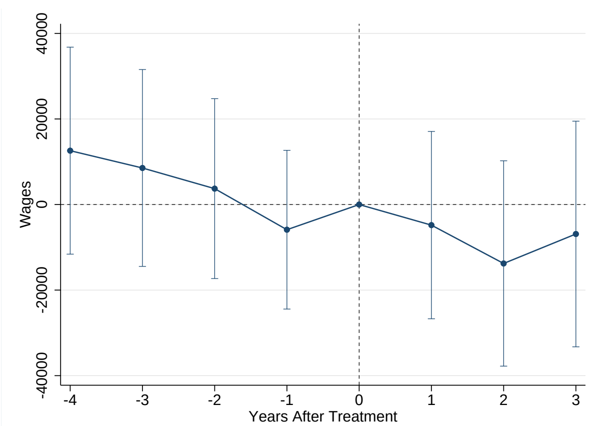
(a) Letter, Quarterly



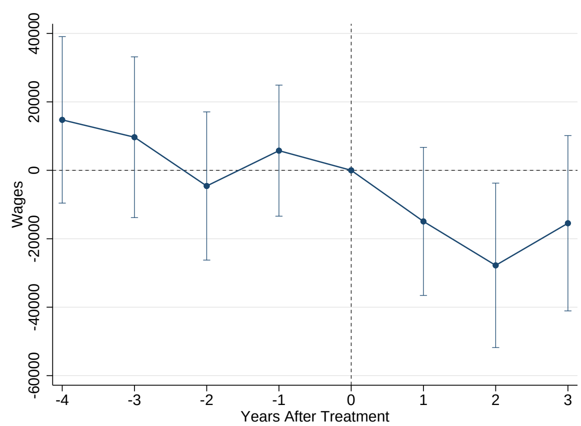
(b) Visit, Quarterly



(c) Letter, Annual

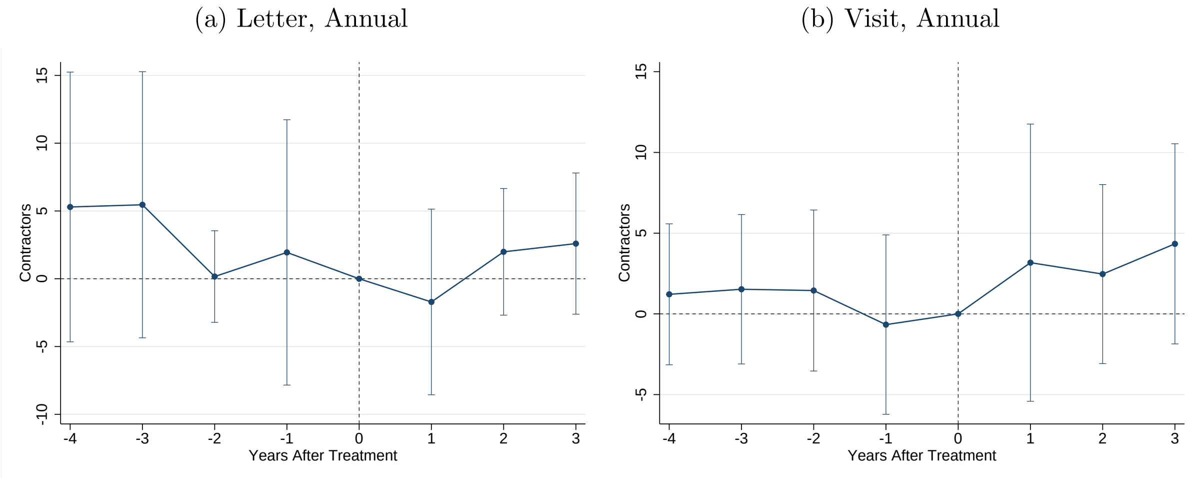


(d) Visit, Annual



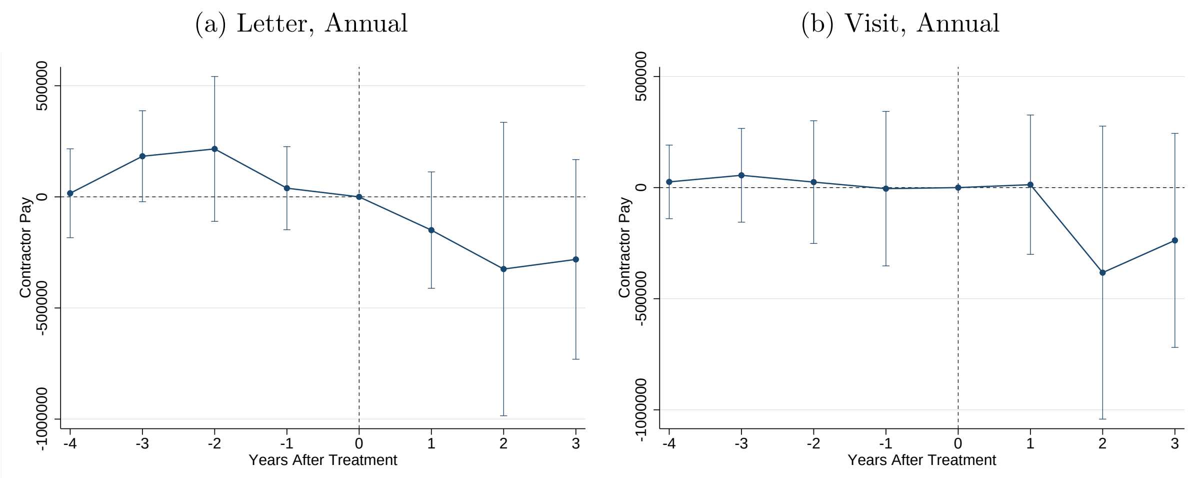
Notes: Plots 95 percent confidence intervals from standard errors clustered by firm. Winsorized at the 98th percentile. Quarterly values from Form 941 and annual values from total of all Forms W-2.

Figure 3.4: Effects of Treatments on Contractor Count



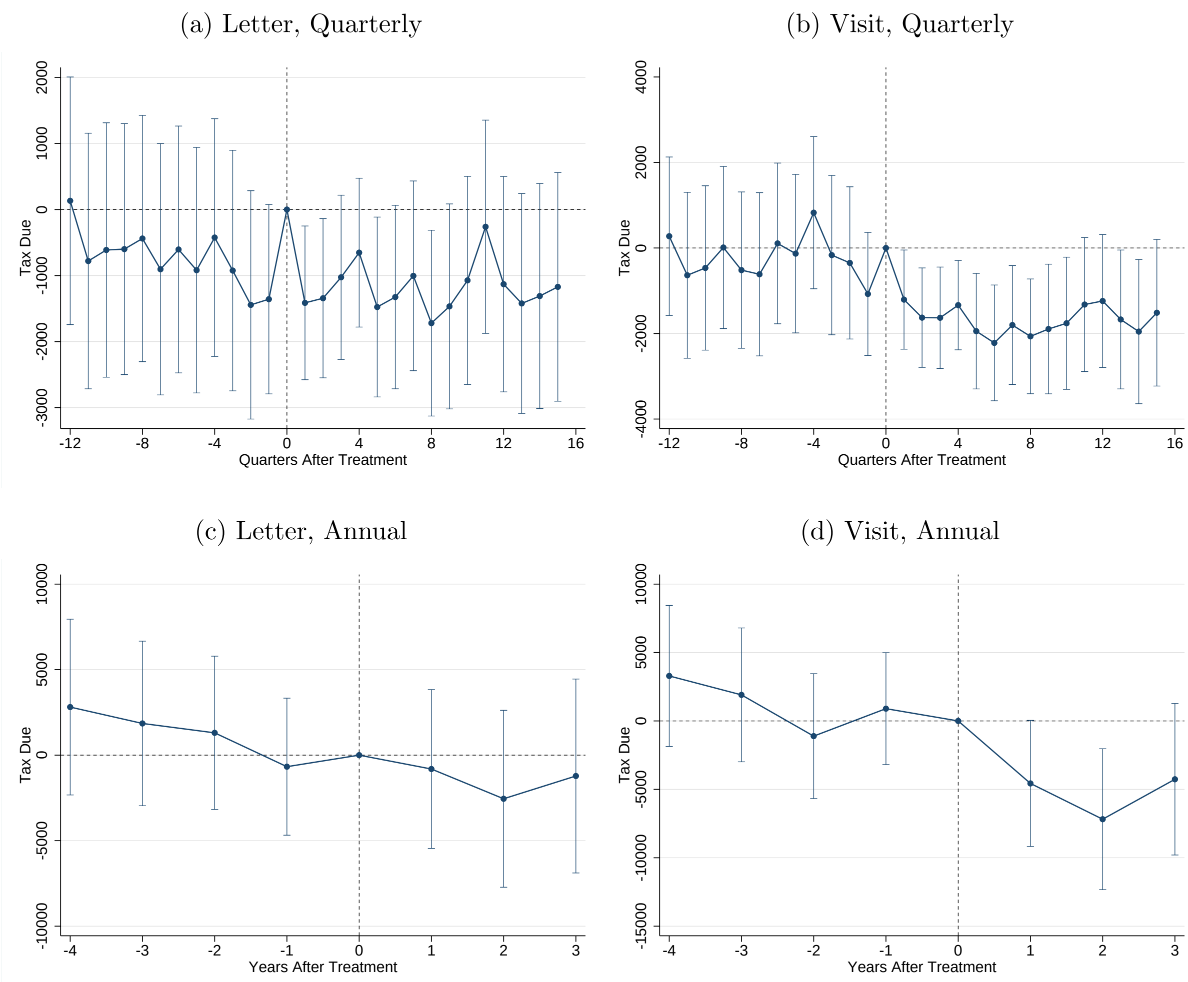
Notes: Plots 95 percent confidence intervals from standard errors clustered by firm. Winsorized at the 98th percentile. Values sum over all Forms 1099-MISC.

Figure 3.5: Effects of Treatments on Contractor Compensation



Notes: Plots 95 percent confidence intervals from standard errors clustered by firm. Winsorized at the 98th percentile. Values sum over all Forms 1099-MISC.

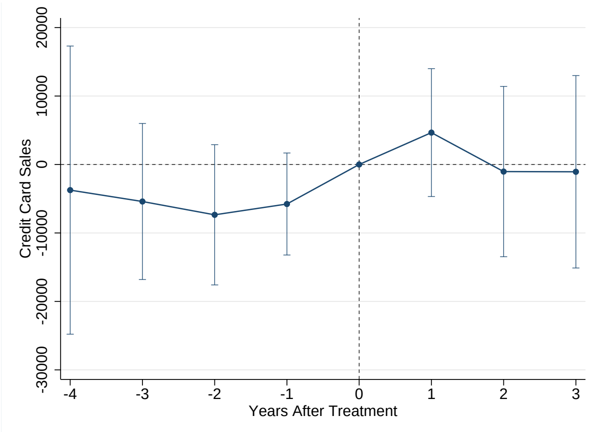
Figure 3.6: Effects of Treatments on Tax Due



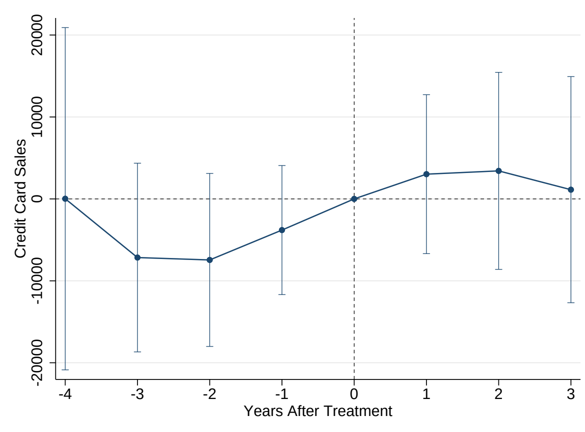
Notes: Plots 95 percent confidence intervals from standard errors clustered by firm. Winsorized at the 98th percentile. Quarterly values from Form 941 and annual values from total of all Forms W-2.

Figure 3.7: Effects of Treatments on Credit Card Payments

(a) Letter, Annual



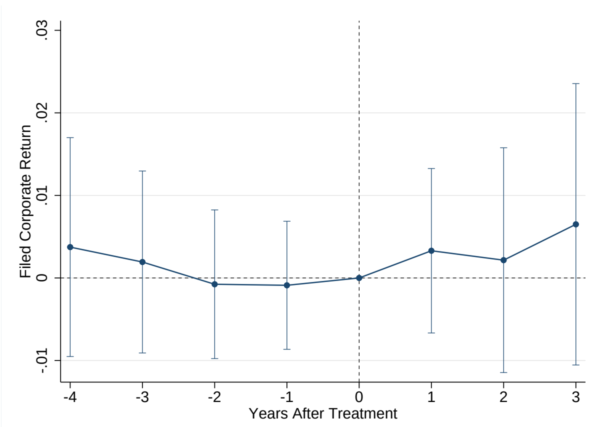
(b) Visit, Annual



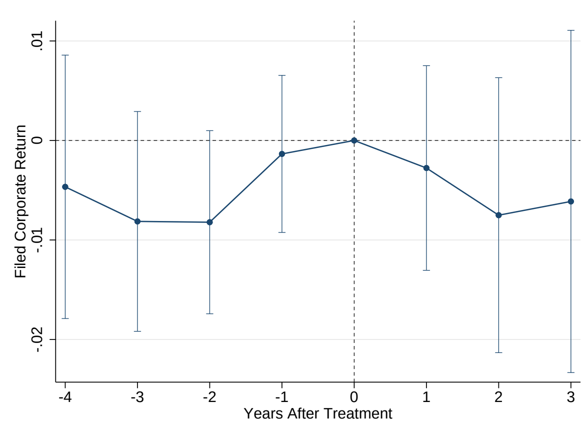
Notes: Plots 95 percent confidence intervals from standard errors clustered by firm. Winsorized at the 98th percentile. Values sum over all Forms 1099-K.

Figure 3.8: Effects of Treatments on Return Filing

(a) Letter



(b) Visit



Notes: Plots 95 percent confidence intervals from standard errors clustered by firm. Outcome is an indicator for filing either Form 1120, corporation tax return, or Form 1120S, S-corporation tax return.

Table 3.1: Descriptive Statistics One Period Before Treatment

	Alert C	Control	Letter	Visit
Employees	12.8	12.9	12.8	12.7
(Quarterly)	[16.6]	[16.8]	[16.5]	[16.6]
Employees	27.6	27.5	27.4	27.9
(Annual)	[39.1]	[39.3]	[38.4]	[39.5]
Wages, Tips, Other Compensation	126,100	125,161	127,714	125,409
(Quarterly)	[159,789]	[159,418]	[161,711]	[158,282]
Wages, Tips, Other Compensation	551,862	550,483	559,217	546,027
(Annual)	[706,291]	[713,070]	[718,536]	[687,914]
Tax Due	34,282	33,883	35,114	33,846
(Quarterly)	[46,353]	[45,911]	[47,532]	[45,599]
Tax Due	109,914	109,871	111,146	108,762
(Annual)	[153,208]	[154,772]	[155,265]	[149,740]
Contractors	3.08	5.40	2.71	1.31
(Annual)	[139]	[202]	[129]	[41.9]
Contractor Pay	133,279	153,039	139,993	108,506
(Annual)	[7,806,308]	[8,638,999]	[8,334,713]	[6,328,694]
Credit Card Sales	227,791	226,548	230,291	226,526
(Annual)	[480,229]	[476,135]	[481,839]	[482,540]
Filed Corporate Return	0.744	0.747	0.742	0.743
(Annual)	[0.436]	[0.435]	[0.437]	[0.437]
Number of Firms	12,172	3,894	4,069	4,209

Notes: Tax remitted prior to treatment is not reported to avoid disclosing IRS methods and procedures. Table reports means with sample standard deviations in brackets. Wages, tips, and other compensation, employees, and federal income tax to withhold are from Form 940 filings (quarterly) and totals of all Forms W-2 filed (annual). Tax due (quarterly) includes federal income tax and social security and medicare taxes due as reported on Form 941, and (annual) includes these taxes withheld as reported on Forms W-2. Contractor count and compensation sums non-employee compensation reported on Forms 1099-MISC. Payments by credit card sums payments by credit card and through third party networks reported on Forms 1099-K. Filing variable counts filing either Form 1120 or Form 1120S. Values are winsorized at the 98th percentile, except for filing indicator and contractors and pay (which are non-zero for less than one percent of the sample).

Table 3.2: Effects of Letter, Quarterly Outcomes

	Tax Remitted	Employees	Wages	Tax Due
1 Quarter Post × Letter	94.4	0.103	-3,290	-1,413*
	(382)	(0.195)	(1,856)	(594)
2 Quarters Post × Letter	-112	0.0160	-2,527	-1,343*
	(438)	(0.208)	(1,967)	(615)
3 Quarters Post × Letter	163	-0.0354	-1,976	-1,026
	(459)	(0.219)	(2,025)	(634)
4 Quarters Post × Letter	177	-0.0267	-1,848	-652
	(459)	(0.227)	(1,953)	(575)
5 Quarters Post × Letter	-618	-0.0797	-3,262	-1,476*
	(477)	(0.241)	(2,247)	(695)
6 Quarters Post × Letter	-472	-0.0192	-2,547	-1,325
	(493)	(0.260)	(2,369)	(709)
7 Quarters Post × Letter	-21.5	0.123	-1,982	-1,004
	(513)	(0.255)	(2,429)	(733)
8 Quarters Post × Letter	-685	-0.187	-3,705	-1,720*
	(495)	(0.262)	(2,430)	(718)
9 Quarters Post × Letter	-1,072*	0.0727	-2,537	-1,467
	(502)	(0.276)	(2,616)	(792)
10 Quarters Post × Letter	-613	-0.0487	-1,808	-1,073
	(519)	(0.289)	(2,705)	(803)
11 Quarters Post × Letter	-442	0.153	944	-261
	(524)	(0.298)	(2,743)	(823)
12 Quarters Post × Letter	-844	0.175	-925	-1,130
	(523)	(0.295)	(2,816)	(832)
13 Quarters Post × Letter	-805	-0.0551	-2,745	-1,422
	(520)	(0.304)	(2,908)	(849)
14 Quarters Post × Letter	-865	0.112	-3,177	-1,310
	(531)	(0.312)	(2,981)	(869)
15 Quarters Post × Letter	-581	0.0983	-2,021	-1,170
	(539)	(0.320)	(3,036)	(883)
Time Fixed Effects	Yes	Yes	Yes	Yes
Number of Firms	12,172	12,172	12,172	12,172
R-Squared	0.0573	0.112	0.0431	0.0372

Notes: Standard errors (in parentheses) clustered by firm. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$. See notes to Table 3.1 for data sources.

Table 3.3: Effects of Letter, Annual Outcomes

	Employees	Wages	Tax Due	Contractors	Contractor Pay	Credit Card Sales	Filed Corp. Return
1 Year Post × Letter	-0.223	-4,811	-811	-1.71	-149,627	4,658	0.00330
	(0.630)	(11,170)	(2,368)	(3.49)	(133,613)	(4,763)	(0.00508)
2 Years Post × Letter	-0.263	-13,776	-2,550	1.99	-324,787	-1,035	0.00216
	(0.674)	(12,244)	(2,639)	(2.38)	(336,803)	(6,347)	(0.00695)
3 Years Post × Letter	-0.767	-6,887	-1,220	2.59	-281,384	-1,065	0.00650
	(0.726)	(13,453)	(2,892)	(2.66)	(229,160)	(7,170)	(0.00869)
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Firms	12,172	12,172	12,172	12,172	12,172	12,172	12,172
R-Squared	0.0126	0.0240	0.0170	0.000188	0.000130	0.112	0.0482

Notes: Standard errors (in parentheses) clustered by firm. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$. See notes to Table 3.1 for data sources.

Table 3.4: Effects of Visit, Quarterly Outcomes

	Tax Remitted	Employees	Wages	Tax Due
1 Quarter Post × Visit	3,686***	-0.124	-3,771*	-1,207*
	(399)	(0.185)	(1,899)	(592)
2 Quarters Post × Visit	2,726***	0.147	-5,501**	-1,629**
	(438)	(0.199)	(1,905)	(593)
3 Quarters Post × Visit	2,169***	-0.234	-5,459**	-1,632**
	(451)	(0.213)	(1,966)	(605)
4 Quarters Post × Visit	1,652***	-0.0738	-4,617*	-1,336*
	(448)	(0.220)	(1,836)	(533)
5 Quarters Post × Visit	464	-0.373	-5,753*	-1,945**
	(468)	(0.238)	(2,285)	(690)
6 Quarters Post × Visit	42.4	-0.339	-7,221**	-2,219**
	(484)	(0.253)	(2,361)	(691)
7 Quarters Post × Visit	478	0.00437	-6,092*	-1,801*
	(504)	(0.254)	(2,401)	(709)
8 Quarters Post × Visit	-367	-0.175	-5,769*	-2,066**
	(484)	(0.262)	(2,377)	(685)
9 Quarters Post × Visit	-567	0.0808	-5,244*	-1,894*
	(496)	(0.271)	(2,615)	(774)
10 Quarters Post × Visit	-358	-0.248	-5,506*	-1,760*
	(513)	(0.282)	(2,692)	(788)
11 Quarters Post × Visit	-332	-0.0531	-4,101	-1,322
	(517)	(0.286)	(2,693)	(800)
12 Quarters Post × Visit	-555	0.108	-3,448	-1,239
	(513)	(0.291)	(2,735)	(793)
13 Quarters Post × Visit	-399	-0.122	-5,398	-1,672*
	(513)	(0.297)	(2,867)	(828)
14 Quarters Post × Visit	-598	-0.247	-7,396*	-1,955*
	(528)	(0.312)	(2,996)	(861)
15 Quarters Post × Visit	-151	-0.171	-4,832	-1,514
	(529)	(0.323)	(3,058)	(875)
Time Fixed Effects	Yes	Yes	Yes	Yes
Number of Firms	12,172	12,172	12,172	12,172
R-Squared	0.0573	0.112	0.0431	0.0372

Notes: Standard errors (in parentheses) clustered by firm. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$. See notes to Table 3.1 for data sources.

Table 3.5: Effects of Visit, Annual Outcomes

	Employees	Wages	Tax Due	Contractors	Contractor Pay	Credit Card Sales	Filed Corp. Return
1 Year Post × Visit	-0.624	-14,931	-4,567	3.17	13,080	3,019	-0.00277
	(0.636)	(11,035)	(2,351)	(4.38)	(159,979)	(4,948)	(0.00525)
2 Years Post × Visit	-1.39*	-27,771*	-7,182**	2.47	-382,669	3,421	-0.00751
	(0.690)	(12,254)	(2,627)	(2.83)	(336,315)	(6,134)	(0.00705)
3 Years Post × Visit	-1.16	-15,467	-4,263	4.34	-237,639	1,126	-0.00613
	(0.733)	(13,065)	(2,823)	(3.16)	(245,712)	(7,041)	(0.00877)
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Firms	12,172	12,172	12,172	12,172	12,172	12,172	12,172
R-Squared	0.0126	0.0240	0.0170	0.000188	0.000130	0.112	0.0482

Notes: Standard errors (in parentheses) clustered by firm. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$. See notes to Table 3.1 for data sources.

APPENDICES

APPENDIX A

Paying Taxes Automatically: Behavioral Effects of Withholding Income Tax

Table A.1: Difference-in-Difference Estimates of Effects of $250 Withholding Reduction: Married Filing Jointly Subsample

	(1) Fully Paid (pp)	(2) Payment Net of Balance Due ($)	(3) Percentage of Balance Due Paid
2000 * Treated	0.00761		
	(0.0711)		
2001 * Treated	-0.0215		
	(0.0682)		
2002 * Treated	-0.191***		
	(0.0676)		
2003 * Treated	-0.0131	-12.10***	-0.0894
	(0.0640)	(3.870)	(0.0575)
2004 * Treated	-0.106*	-11.02***	-0.134**
	(0.0625)	(3.966)	(0.0563)
2005 * Treated	-0.0158	-4.831	-0.0376
	(0.0618)	(4.081)	(0.0559)
2006 * Treated	-0.0436	-3.441	-0.0729
	(0.0595)	(4.008)	(0.0538)
2007 * Treated	-0.172***	-1.461	-0.108**
	(0.0540)	(3.839)	(0.0500)
2008 * Treated	(omitted)	(omitted)	(omitted)
2009 * Treated	-1.677***	-16.80***	-0.766***
	(0.0599)	(3.765)	(0.0516)
2010 * Treated	0.0851	2.893	0.0591
	(0.0604)	(4.142)	(0.0547)
2011 * Treated	-0.0721	8.382*	-0.0732
	(0.0631)	(4.417)	(0.0575)
2012 * Treated	-0.0476	11.87**	-0.0442
	(0.0656)	(4.650)	(0.0599)
2013 * Treated	0.0180	14.53***	-0.00545
	(0.0670)	(4.837)	(0.0612)
F-stat Pre-2008	3.548	3.023	1.755
p-value Pre-2008	0.000407	0.00988	0.118
F-stat Post-2008	234.9	12.39	68.40
p-value Post-2008	<0.0001	<0.0001	<0.0001
Tax Unit Fixed Effects	Yes	Yes	Yes
Year FE * Demographics	Yes	Yes	Yes
Years	14	11	11
Tax Unit-Years	9200603	7243010	7243010
R-Squared	0.000150	0.0000181	0.0000536

Notes: standard errors (in parentheses) clustered by household. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Includes interactions between year fixed effects and 2008 age fixed effects, spouse's age, and dependents. Payment amount data begin in 2003. Payment net of balance due winsorized at the 1st and 99th percentiles. F-statistics report joint statistical significance of all pre-2008 or post-2008 coefficients.

Table A.2: Difference-in-Difference Estimates of Effects of $250 Withholding Reduction: Subsample with Fewer Than Three Dependents in 2008

	(1) Fully Paid (pp)	(2) Payment Net of Balance Due ($)	(3) Percentage of Balance Due Paid
2000 * Treated	0.154**		
	(0.0609)		
2001 * Treated	0.141**		
	(0.0589)		
2002 * Treated	0.00841		
	(0.0585)		
2003 * Treated	0.132**	-2.098	0.0653
	(0.0556)	(2.819)	(0.0509)
2004 * Treated	0.129**	-1.061	0.0783
	(0.0546)	(2.897)	(0.0501)
2005 * Treated	0.148***	1.685	0.111**
	(0.0539)	(2.970)	(0.0495)
2006 * Treated	0.0938*	2.445	0.0594
	(0.0517)	(2.892)	(0.0475)
2007 * Treated	-0.00738	2.433	-0.000388
	(0.0468)	(2.728)	(0.0439)
2008 * Treated	(omitted)	(omitted)	(omitted)
2009 * Treated	-1.456***	-11.26***	-0.711***
	(0.0513)	(2.692)	(0.0453)
2010 * Treated	0.242***	6.880**	0.206***
	(0.0519)	(2.950)	(0.0477)
2011 * Treated	0.195***	10.95***	0.174***
	(0.0538)	(3.151)	(0.0497)
2012 * Treated	0.235***	15.57***	0.210***
	(0.0557)	(3.349)	(0.0515)
2013 * Treated	0.329***	19.14***	0.288***
	(0.0570)	(3.514)	(0.0526)
F-stat Pre-2008	3.096	0.919	1.421
p-value Pre-2008	0.00170	0.468	0.213
F-stat Post-2008	298.2	21.09	112.6
p-value Post-2008	<0.001	<0.001	<0.001
Tax Unit Fixed Effects	Yes	Yes	Yes
Year FE * Demographics	Yes	Yes	Yes
Years	14	11	11
Tax Unit-Years	12,916,840	10,219,254	10,219,254
R-Squared	0.000149	0.0000174	0.0000768

Notes: standard errors (in parentheses) clustered by household. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Includes interactions between year fixed effects and 2008 age fixed effects, marital status, spouse's age, and dependents. Payment amount data begin in 2003. Payment net of balance due winsorized at the 1st and 99th percentiles. F-statistics report joint statistical significance of all pre-2008 or post-2008 coefficients.

Table A.3: Difference-in-Difference Estimates of Effects of $250 Withholding Reduction: Subsample without Social Security Income in 2008

	(1) Fully Paid (pp)	(2) Payment Net of Balance Due ($)	(3) Percentage of Balance Due Paid
2000 * Treated	0.0520		
	(0.0830)		
2001 * Treated	-0.0310		
	(0.0804)		
2002 * Treated	-0.385***		
	(0.0805)		
2003 * Treated	-0.0718	-3.995	-0.125*
	(0.0761)	(3.659)	(0.0704)
2004 * Treated	-0.0694	-6.151	-0.101
	(0.0745)	(3.800)	(0.0691)
2005 * Treated	-0.102	-3.343	-0.125*
	(0.0731)	(3.888)	(0.0678)
2006 * Treated	-0.133*	-2.047	-0.161**
	(0.0701)	(3.795)	(0.0650)
2007 * Treated	-0.202***	-2.505	-0.179***
	(0.0631)	(3.560)	(0.0599)
2008 * Treated	(omitted)	(omitted)	(omitted)
2009 * Treated	-3.325***	-38.52***	-1.704***
	(0.0768)	(3.607)	(0.0652)
2010 * Treated	-0.271***	-10.32***	-0.235***
	(0.0723)	(3.989)	(0.0670)
2011 * Treated	-0.0678	2.246	-0.0664
	(0.0747)	(4.227)	(0.0696)
2012 * Treated	0.0636	9.487**	0.0620
	(0.0776)	(4.529)	(0.0722)
2013 * Treated	0.216***	12.88***	0.185**
	(0.0795)	(4.686)	(0.0739)
F-stat Pre-2008	5.989	0.569	2.045
p-value Pre-2008	<0.0001	0.724	0.0691
F-stat Post-2008	491.0	37.89	188.2
p-value Post-2008	<0.0001	<0.0001	<0.0001
Tax Unit Fixed Effects	Yes	Yes	Yes
Year FE * Demographics	Yes	Yes	Yes
Years	14	11	11
Tax Unit-Years	8,765,752	6,935,877	6,935,877
R-Squared	0.000474	0.0000307	0.000200

Notes: standard errors (in parentheses) clustered by household. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Includes interactions between year fixed effects and 2008 age fixed effects, marital status, spouse's age, and dependents. Payment amount data begin in 2003. Payment net of balance due winsorized at the 1st and 99th percentiles. F-statistics report joint statistical significance of all pre-2008 or post-2008 coefficients.

Table A.4: Difference-in-Difference Estimates of Effects of $250 Withholding Reduction: Heterogeneity by Interest Income over $500 in 2008

	(1) Fully Paid (pp)	(2) Payment Net of Balance Due ($)	(3) Percentage of Balance Due Paid
2000 * Treated * Interest	0.155		
	(0.162)		
2001 * Treated * Interest	0.189		
	(0.153)		
2002 * Treated * Interest	0.183		
	(0.150)		
2003 * Treated * Interest	0.212	-8.788	0.150
	(0.139)	(9.978)	(0.119)
2004 * Treated * Interest	0.214	-16.20	0.126
	(0.136)	(10.19)	(0.117)
2005 * Treated * Interest	0.0356	-18.15*	0.0534
	(0.138)	(10.68)	(0.120)
2006 * Treated * Interest	0.0342	-18.51*	-0.0629
	(0.134)	(10.70)	(0.116)
2007 * Treated * Interest	0.452***	-4.948	0.219**
	(0.120)	(10.39)	(0.107)
2008 * Treated * Interest	(omitted)	(omitted)	(omitted)
2009 * Treated * Interest	-0.925***	-33.17***	0.0739
	(0.144)	(9.915)	(0.112)
2010 * Treated * Interest	-0.139	-34.29***	-0.156
	(0.130)	(10.87)	(0.113)
2011 * Treated * Interest	0.0392	-18.74*	-0.0337
	(0.135)	(11.17)	(0.119)
2012 * Treated * Interest	0.0369	-24.69**	-0.0644
	(0.141)	(11.52)	(0.124)
2013 * Treated * Interest	0.0517	-12.54	-0.0132
	(0.144)	(11.83)	(0.127)
F-stat Pre-2008	2.757	1.024	1.811
p-value Pre-2008	0.00482	0.401	0.107
F-stat Post-2008	11.49	3.084	0.914
p-value Post-2008	<0.0001	0.00872	0.471
Tax Unit Fixed Effects	Yes	Yes	Yes
Year FE * Demographics	Yes	Yes	Yes
Years	14	11	11
Tax Unit-Years	13,327,835	10,703,607	10,703,607
R-Squared	0.00277	0.00114	0.00183

Notes: standard errors (in parentheses) clustered by tax unit. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Interest income includes both taxable and tax-exempt interest. Includes interactions between year fixed effects, interest indicator, and 2008 age fixed effects, marital status, spouse's age, and dependents. Payment amount data begin in 2003. Payment net of balance due winsorized at the 1st and 99th percentiles. F-statistics report joint statistical significance of all pre-2008 or post-2008 coefficients.

Table A.5: Difference-in-Difference Estimates of Effects of $250 Withholding Reduction: Heterogeneity by Positive 2008 Interest Income

	(1) Fully Paid (pp)	(2) Payment Net of Balance Due ($)	(3) Percentage of Balance Due Paid
2000 * Treated * Interest	0.288*		
	(0.155)		
2001 * Treated * Interest	0.420***		
	(0.148)		
2002 * Treated * Interest	0.177		
	(0.146)		
2003 * Treated * Interest	0.416***	-5.375	0.361***
	(0.138)	(6.994)	(0.127)
2004 * Treated * Interest	0.382***	-0.928	0.356***
	(0.134)	(7.116)	(0.124)
2005 * Treated * Interest	0.315**	-2.642	0.329***
	(0.131)	(7.184)	(0.121)
2006 * Treated * Interest	0.395***	3.763	0.326***
	(0.125)	(6.996)	(0.116)
2007 * Treated * Interest	0.713***	1.048	0.307***
	(0.112)	(6.520)	(0.106)
2008 * Treated * Interest	(omitted)	(omitted)	(omitted)
2009 * Treated * Interest	-1.123***	-16.14**	-0.114
	(0.124)	(6.618)	(0.110)
2010 * Treated * Interest	-0.381***	-14.68**	-0.395***
	(0.126)	(7.244)	(0.117)
2011 * Treated * Interest	-0.142	-14.09*	-0.200
	(0.132)	(7.616)	(0.123)
2012 * Treated * Interest	-0.0525	-8.404	-0.0615
	(0.138)	(8.097)	(0.128)
2013 * Treated * Interest	0.0407	-11.73	-0.00418
	(0.142)	(8.572)	(0.133)
F-stat Pre-2008	5.845	0.435	2.608
p-value Pre-2008	<0.0001	0.824	0.0230
F-stat Post-2008	20.49	1.456	3.047
p-value Post-2008	<0.0001	0.201	0.00941
Tax Unit Fixed Effects	Yes	Yes	Yes
Year FE * Demographics	Yes	Yes	Yes
Years	14	11	11
Tax Unit-Years	13,327,835	10,703,607	10,703,607
R-Squared	0.00292	0.00107	0.00203

Notes: standard errors (in parentheses) clustered by tax unit. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Interest income includes both taxable and tax-exempt interest. Includes interactions between year fixed effects, interest indicator, and 2008 age fixed effects, marital status, spouse's age, and dependents. Payment amount data begin in 2003. Payment net of balance due winsorized at the 1st and 99th percentiles. F-statistics report joint statistical significance of all pre-2008 or post-2008 coefficients.

Table A.6: Difference-in-Difference Estimates of Effects of $250 Withholding Reduction: Heterogeneity by Positive Interest Income in All Years 2005-2008

	(1) Fully Paid (pp)	(2) Payment Net of Balance Due ($)	(3) Percentage of Balance Due Paid
2000 * Treated * Interest	0.212		
	(0.145)		
2001 * Treated * Interest	0.350**		
	(0.138)		
2002 * Treated * Interest	0.0966		
	(0.136)		
2003 * Treated * Interest	0.206	-13.50*	0.220*
	(0.128)	(7.278)	(0.115)
2004 * Treated * Interest	0.250**	-5.477	0.210*
	(0.124)	(7.454)	(0.112)
2005 * Treated * Interest	0.208*	-7.566	0.206*
	(0.123)	(7.556)	(0.111)
2006 * Treated * Interest	0.390***	-0.498	0.338***
	(0.119)	(7.454)	(0.108)
2007 * Treated * Interest	0.585***	-8.862	0.298***
	(0.107)	(7.103)	(0.0998)
2008 * Treated * Interest	(omitted)	(omitted)	(omitted)
2009 * Treated * Interest	-1.177***	-19.66***	-0.141
	(0.126)	(7.258)	(0.107)
2010 * Treated * Interest	-0.521***	-33.46***	-0.531***
	(0.122)	(7.899)	(0.111)
2011 * Treated * Interest	-0.259**	-26.45***	-0.310***
	(0.127)	(8.213)	(0.116)
2012 * Treated * Interest	-0.161	-24.47***	-0.172
	(0.132)	(8.671)	(0.121)
2013 * Treated * Interest	-0.244*	-30.63***	-0.272**
	(0.136)	(9.093)	(0.125)
F-stat Pre-2008	4.810	1.241	2.459
p-value Pre-2008	0.00000615	0.287	0.0309
F-stat Post-2008	19.47	4.233	5.190
p-value Post-2008	<0.0001	0.000754	<0.0001
Tax Unit Fixed Effects	Yes	Yes	Yes
Year FE * Demographics	Yes	Yes	Yes
Years	14	11	11
Tax Unit-Years	12,800,881	10,184,534	10,184,534
R-Squared	0.00303	0.00111	0.00204

Notes: standard errors (in parentheses) clustered by tax unit. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Interest income includes both taxable and tax-exempt interest. Includes interactions between year fixed effects, interest indicator, and 2008 age fixed effects, marital status, spouse's age, and dependents. Payment amount data begin in 2003. Payment net of balance due winsorized at the 1st and 99th percentiles. F-statistics report joint statistical significance of all pre-2008 or post-2008 coefficients.

Table A.7: Difference-in-Difference Estimates of Effects of $250 Withholding Reduction: Heterogeneity by Dividend Income over $100 in 2008

	(1) Fully Paid (pp)	(2) Payment Net of Balance Due ($)	(3) Percentage of Balance Due Paid
2000 * Treated * Dividends	0.303*		
	(0.169)		
2001 * Treated * Dividends	0.222		
	(0.156)		
2002 * Treated * Dividends	-0.0121		
	(0.152)		
2003 * Treated * Dividends	-0.0347	1.051	0.0670
	(0.143)	(9.830)	(0.121)
2004 * Treated * Dividends	-0.0396	-0.496	0.0128
	(0.139)	(10.11)	(0.120)
2005 * Treated * Dividends	-0.0124	-7.701	0.0416
	(0.141)	(10.12)	(0.123)
2006 * Treated * Dividends	0.133	-8.011	0.114
	(0.137)	(10.01)	(0.119)
2007 * Treated * Dividends	0.216*	-7.685	0.0439
	(0.123)	(9.687)	(0.110)
2008 * Treated * Dividends	(omitted)	(omitted)	(omitted)
2009 * Treated * Dividends	-1.022***	-6.986	0.0289
	(0.150)	(9.868)	(0.116)
2010 * Treated * Dividends	-0.257*	-14.83	-0.284**
	(0.136)	(10.87)	(0.118)
2011 * Treated * Dividends	-0.192	-12.14	-0.176
	(0.140)	(11.29)	(0.123)
2012 * Treated * Dividends	-0.129	-11.40	-0.0919
	(0.145)	(11.67)	(0.129)
2013 * Treated * Dividends	-0.119	-6.943	-0.101
	(0.149)	(12.01)	(0.132)
F-stat Pre-2008	1.501	0.383	0.248
p-value Pre-2008	0.151	0.861	0.941
F-stat Post-2008	10.32	0.435	1.817
p-value Post-2008	<0.0001	0.825	0.106
Tax Unit Fixed Effects	Yes	Yes	Yes
Year FE * Demographics	Yes	Yes	Yes
Years	14	11	11
Tax Unit-Years	13,327,835	10,703,607	10,703,607
R-Squared	0.00275	0.00108	0.00181

Notes: standard errors (in parentheses) clustered by tax unit. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Dividend income is taxable dividend income. Includes interactions between year fixed effects, interest indicator, and 2008 age fixed effects, marital status, spouse's age, and dependents. Payment amount data begin in 2003. Payment net of balance due winsorized at the 1st and 99th percentiles. F-statistics report joint statistical significance of all pre-2008 or post-2008 coefficients.

APPENDIX B

Heard it Through the Grapevine: The Direct and Network Effects of a Tax Enforcement Field Experiment on Firms

Table B.1: Direct Effect of Letter: All Quarters

	Tax Remitted	Any Tax Remitted	Log(Tax Remitted)
Letter * Four Quarters Pre	273 (571)	0.0210* (0.0120)	-0.00170 (0.0368)
Letter * Three Quarters Pre	69.7 (570)	0.00317 (0.0119)	0.0189 (0.0369)
Letter * Two Quarters Pre	-194 (554)	0.000663 (0.0116)	0.0142 (0.0361)
Letter * One Quarter Pre	-337 (471)	-0.0172 (0.0112)	0.0562 (0.0350)
Letter * One Quarter Post	94.4 (382)	0.0302*** (0.0110)	0.00476 (0.0358)
Letter * Two Quarters Post	-112 (438)	0.0112 (0.0118)	-0.0171 (0.0359)
Letter * Three Quarters Post	163 (459)	0.0158 (0.0122)	-0.0160 (0.0376)
Letter * Four Quarters Post	177 (459)	0.0136 (0.0125)	-0.00353 (0.0384)
P-value from F-test of Letter in Pre Quarters	0.761	0.00948	0.239
P-value from F-test of Letter in Post Quarters	0.912	0.0627	0.941
Quarter Fixed Effects	Yes	Yes	Yes
Number of Firm-Quarters	109,548	109,548	77,051
R-Squared	0.0281	0.0513	0.0220

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.2: Direct Effect of Visit: All Quarters

	Tax Remitted	Any Tax Remitted	Log(Tax Remitted)
Visit * Four Quarters Pre	52.7 (585)	0.00234 (0.0122)	-0.00842 (0.0368)
Visit * Three Quarters Pre	92.1 (572)	-0.00707 (0.0119)	0.0174 (0.0366)
Visit * Two Quarters Pre	-393 (556)	-0.00559 (0.0117)	-0.0145 (0.0359)
Visit * One Quarter Pre	-299 (465)	-0.0149 (0.0112)	-0.00144 (0.0353)
Visit * One Quarter Post	3,686*** (399)	0.129*** (0.0113)	0.132*** (0.0348)
Visit * Two Quarters Post	2,726*** (438)	0.104*** (0.0120)	0.0344 (0.0349)
Visit * Three Quarters Post	2,169*** (451)	0.0803*** (0.0122)	0.0309 (0.0362)
Visit * Four Quarters Post	1,652*** (448)	0.0694*** (0.0126)	0.0197 (0.0364)
P-value from F-test of Visit in Pre Quarters	0.707	0.456	0.572
P-value from F-test of Visit in Post Quarters	8.32e-18	8.71e-29	0.000272
Quarter Fixed Effects	Yes	Yes	Yes
Number of Firm-Quarters	109,548	109,548	77,051
R-Squared	0.0281	0.0513	0.0220

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.3: Preparer Letter Network Effects with Pre-Treatment Quarters as Placebo Test

	Tax Remitted	Any Tax Remitted	Log(Tax Remitted)
Preparer Links to Letter Firms * Four Quarters Pre	304**	-0.0000760	0.00676
	(132)	(0.00313)	(0.00521)
Preparer Links to Letter Firms * Three Quarters Pre	71.9	-0.000335	-0.00548
	(95.2)	(0.00260)	(0.00663)
Preparer Links to Letter Firms * Two Quarters Pre	-15.0	0.0000893	-0.000276
	(109)	(0.00185)	(0.00469)
Preparer Links to Letter Firms * One Quarter Pre	13.8	-0.000667	-0.000396
	(102)	(0.00145)	(0.00464)
Preparer Links to Letter Firms * One Quarter Post	35.3	-0.00183	-0.00437
	(72.2)	(0.00152)	(0.00624)
Preparer Links to Letter Firms * Two Quarters Post	27.0	-0.000140	-0.0000449
	(88.6)	(0.00170)	(0.00428)
Preparer Links to Letter Firms * Three Quarters Post	73.0	-0.00201	0.00264
	(92.6)	(0.00216)	(0.00434)
Preparer Links to Letter Firms * Four Quarters Post	146	-0.00111	0.0109**
	(103)	(0.00238)	(0.00455)
Firm Fixed Effects	Yes	Yes	Yes
Quarter Fixed Effects	Yes	Yes	Yes
Quarter * Total Preparer Links to Alert C Fixed Effects	Yes	Yes	Yes
Preparer Clusters	10,219	10,219	9,357
Number of Firm-Quarters	1,796,994	1,796,994	1,193,501
R-Squared	0.00361	0.00500	0.0120

Notes: Standard errors (in parentheses) clustered by Preparer. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Probability results from linear probability model.

Table B.4: Preparer Visit Network Effects with Pre-Treatment Quarters as Placebo Test

	Tax Remitted	Any Tax Remitted	Log(Tax Remitted)
Preparer Links to Visit Firms * Four Quarters Pre	197	0.000657	0.00683
	(132)	(0.00238)	(0.00428)
Preparer Links to Visit Firms * Three Quarters Pre	-9.41	-0.000457	-0.00378
	(78.7)	(0.00216)	(0.00554)
Preparer Links to Visit Firms * Two Quarters Pre	-101	-0.00160	-0.00101
	(89.4)	(0.00176)	(0.00393)
Preparer Links to Visit Firms * One Quarter Pre	-40.3	-0.000983	0.00523
	(89.0)	(0.00121)	(0.00375)
Preparer Links to Visit Firms * One Quarter Post	85.5	0.00188	-0.00236
	(61.3)	(0.00122)	(0.00505)
Preparer Links to Visit Firms * Two Quarters Post	52.3	0.000315	-0.00328
	(81.9)	(0.00149)	(0.00353)
Preparer Links to Visit Firms * Three Quarters Post	156*	0.00123	0.00112
	(88.3)	(0.00189)	(0.00368)
Preparer Links to Visit Firms * Four Quarters Post	243***	0.00162	0.000830
	(94.1)	(0.00224)	(0.00357)
Firm Fixed Effects	Yes	Yes	Yes
Quarter Fixed Effects	Yes	Yes	Yes
Quarter * Total Preparer Links to Alert C Fixed Effects	Yes	Yes	Yes
Preparer Clusters	10,219	10,219	9,357
Number of Firm-Quarters	1,796,994	1,796,994	1,193,501
R-Squared	0.00361	0.00500	0.0120

Notes: Standard errors (in parentheses) clustered by Preparer. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Probability results from linear probability model.

Table B.5: Preparation Firm Letter Network Effects with Pre-Treatment Quarters as Placebo Test

	Tax Remitted	Any Tax Remitted	Log(Tax Remitted)
Preparation Firm Links to Letter Firms * Four Quarters Pre	117	-0.000384	0.00364
	(173)	(0.00206)	(0.00436)
Preparation Firm Links to Letter Firms * Three Quarters Pre	107	-0.000913	0.00269
	(94.7)	(0.00186)	(0.00403)
Preparation Firm Links to Letter Firms * Two Quarters Pre	24.2	0.000699	0.000498
	(99.3)	(0.00141)	(0.00346)
Preparation Firm Links to Letter Firms * One Quarter Pre	65.6	0.0000650	0.00239
	(95.5)	(0.001000)	(0.00286)
Preparation Firm Links to Letter Firms * One Quarter Post	-5.67	-0.00109	-0.000987
	(103)	(0.00107)	(0.00380)
Preparation Firm Links to Letter Firms * Two Quarters Post	-116	-0.000442	-0.00361
	(102)	(0.00119)	(0.00317)
Preparation Firm Links to Letter Firms * Three Quarters Post	-67.5	-0.00115	0.000720
	(91.8)	(0.00141)	(0.00331)
Preparation Firm Links to Letter Firms * Four Quarters Post	-132*	-0.00149	0.000841
	(80.2)	(0.00168)	(0.00300)
Firm Fixed Effects	Yes	Yes	Yes
Quarter Fixed Effects	Yes	Yes	Yes
Quarter * Total Preparation Firm Links to Alert C Fixed Effects	Yes	Yes	Yes
Preparation Firm Clusters	9,759	9,759	9,053
Number of Firm-Quarters	3,563,361	3,563,361	2,468,149
R-Squared	0.00502	0.00620	0.0131

Notes: Standard errors (in parentheses) clustered by Preparation Firm. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Probability results from linear probability model.

Table B.6: Preparation Firm Visit Network Effects with Pre-Treatment Quarters as Placebo Test

	Tax Remitted	Any Tax Remitted	Log(Tax Remitted)
Preparation Firm Links to Visit Firms * Four Quarters Pre	128	0.00311	0.000758
	(169)	(0.00241)	(0.00377)
Preparation Firm Links to Visit Firms * Three Quarters Pre	22.9	-0.000201	-0.00203
	(92.7)	(0.00167)	(0.00351)
Preparation Firm Links to Visit Firms * Two Quarters Pre	-32.4	-0.000834	-0.00342
	(86.2)	(0.000974)	(0.00308)
Preparation Firm Links to Visit Firms * One Quarter Pre	16.9	-0.00135	0.00256
	(82.2)	(0.00102)	(0.00210)
Preparation Firm Links to Visit Firms * One Quarter Post	2.47	-0.000349	-0.00154
	(110)	(0.000943)	(0.00309)
Preparation Firm Links to Visit Firms * Two Quarters Post	-52.7	-0.000465	-0.00331
	(89.7)	(0.00117)	(0.00248)
Preparation Firm Links to Visit Firms * Three Quarters Post	65.9	0.000129	0.00189
	(79.9)	(0.00148)	(0.00225)
Preparation Firm Links to Visit Firms * Four Quarters Post	22.9	0.000206	0.00246
	(79.8)	(0.00185)	(0.00222)
Firm Fixed Effects	Yes	Yes	Yes
Quarter Fixed Effects	Yes	Yes	Yes
Quarter * Total Preparation Firm Links to Alert C Fixed Effects	Yes	Yes	Yes
Preparation Firm Clusters	9,759	9,759	9,053
Number of Firm-Quarters	3,563,361	3,563,361	2,468,149
R-Squared	0.00502	0.00620	0.0131

Notes: Standard errors (in parentheses) clustered by Preparation Firm. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Probability results from linear probability model.

Table B.7: ZIP+4 Letter Network Effects with Pre-Treatment Quarters as Placebo Test

	Tax Remitted	Any Tax Remitted	Log(Tax Remitted)
ZIP+4 Links to Letter Firms * Four Quarters Pre	-806	-0.0132*	-0.000138
	(700)	(0.00705)	(0.0132)
ZIP+4 Links to Letter Firms * Three Quarters Pre	-436	-0.0159***	0.00427
	(570)	(0.00585)	(0.0164)
ZIP+4 Links to Letter Firms * Two Quarters Pre	-422	-0.00867*	0.0147
	(459)	(0.00453)	(0.0122)
ZIP+4 Links to Letter Firms * One Quarter Pre	-1,044**	-0.00929**	-0.0167
	(417)	(0.00384)	(0.0113)
ZIP+4 Links to Letter Firms * One Quarter Post	59.5	-0.00469	0.0217
	(527)	(0.00619)	(0.0156)
ZIP+4 Links to Letter Firms * Two Quarters Post	-117	-0.00748	0.00782
	(495)	(0.00536)	(0.0120)
ZIP+4 Links to Letter Firms * Three Quarters Post	-545	-0.0139**	-0.00767
	(526)	(0.00628)	(0.0126)
ZIP+4 Links to Letter Firms * Four Quarters Post	-255	-0.0123	0.0129
	(481)	(0.00770)	(0.0131)
Firm Fixed Effects	Yes	Yes	Yes
Quarter Fixed Effects	Yes	Yes	Yes
Quarter * Total ZIP+4 Links to Alert C Fixed Effects	Yes	Yes	Yes
ZIP+4 Clusters	5,916	5,916	5,476
Number of Firm-Quarters	290,745	290,745	201,828
R-Squared	0.00326	0.0104	0.00891

Notes: Standard errors (in parentheses) clustered by ZIP+4. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Probability results from linear probability model.

Table B.8: ZIP+4 Visit Network Effects with Pre-Treatment Quarters as Placebo Test

	Tax Remitted	Any Tax Remitted	Log(Tax Remitted)
ZIP+4 Links to Visit Firms * Four Quarters Pre	681	0.00438	0.000777
	(568)	(0.00689)	(0.0126)
ZIP+4 Links to Visit Firms * Three Quarters Pre	962	-0.00148	0.0248
	(904)	(0.00621)	(0.0184)
ZIP+4 Links to Visit Firms * Two Quarters Pre	435	0.00511	0.0254*
	(540)	(0.00500)	(0.0133)
ZIP+4 Links to Visit Firms * One Quarter Pre	-10.4	-0.00614*	0.0155
	(493)	(0.00328)	(0.0127)
ZIP+4 Links to Visit Firms * One Quarter Post	847	0.00158	0.0326**
	(784)	(0.00540)	(0.0162)
ZIP+4 Links to Visit Firms * Two Quarters Post	319	0.00225	0.0131
	(563)	(0.00554)	(0.0131)
ZIP+4 Links to Visit Firms * Three Quarters Post	-41.5	-0.00302	0.00650
	(568)	(0.00631)	(0.0134)
ZIP+4 Links to Visit Firms * Four Quarters Post	84.9	-0.000572	0.0128
	(660)	(0.00734)	(0.0127)
Firm Fixed Effects	Yes	Yes	Yes
Quarter Fixed Effects	Yes	Yes	Yes
Quarter * Total ZIP+4 Links to Alert C Fixed Effects	Yes	Yes	Yes
ZIP+4 Clusters	5,916	5,916	5,476
Number of Firm-Quarters	290,745	290,745	201,828
R-Squared	0.00326	0.0104	0.00891

Notes: Standard errors (in parentheses) clustered by ZIP+4. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Probability results from linear probability model.

Table B.9: ZIP Code Letter Network Effects with Pre-Treatment Quarters as Placebo Test

	Tax Remitted	Any Tax Remitted	Log(Tax Remitted)
ZIP Code Links to Letter Firms * Four Quarters Pre	-48.4	0.0000352	-0.00130
	(74.4)	(0.00104)	(0.00225)
ZIP Code Links to Letter Firms * Three Quarters Pre	47.6	0.000554	-0.000593
	(75.3)	(0.000946)	(0.00256)
ZIP Code Links to Letter Firms * Two Quarters Pre	-27.4	-0.000641	-0.000523
	(61.8)	(0.000825)	(0.00193)
ZIP Code Links to Letter Firms * One Quarter Pre	17.5	0.000703	-0.00151
	(55.5)	(0.000625)	(0.00186)
ZIP Code Links to Letter Firms * One Quarter Post	21.1	-0.0000667	0.00299
	(66.0)	(0.000702)	(0.00219)
ZIP Code Links to Letter Firms * Two Quarters Post	-7.65	0.000433	0.00342*
	(61.2)	(0.000772)	(0.00201)
ZIP Code Links to Letter Firms * Three Quarters Post	48.4	0.000872	0.00297
	(69.2)	(0.000883)	(0.00223)
ZIP Code Links to Letter Firms * Four Quarters Post	2.79	0.00175*	0.00357
	(67.3)	(0.000967)	(0.00225)
Firm Fixed Effects	Yes	Yes	Yes
Quarter Fixed Effects	Yes	Yes	Yes
Quarter * Total ZIP Code Links to Alert C Fixed Effects	Yes	Yes	Yes
ZIP Code Clusters	7,046	7,046	7,008
Number of Firm-Quarters	3,181,959	3,181,959	2,159,992
R-Squared	0.00170	0.00624	0.00949

Notes: Standard errors (in parentheses) clustered by ZIP Code. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Probability results from linear probability model.

Table B.10: ZIP Code Visit Network Effects with Pre-Treatment Quarters as Placebo Test

	Tax Remitted	Any Tax Remitted	Log(Tax Remitted)
ZIP Code Links to Visit Firms * Four Quarters Pre	25.3	-0.000526	-0.00101
	(73.3)	(0.00113)	(0.00226)
ZIP Code Links to Visit Firms * Three Quarters Pre	53.7	0.000587	0.000942
	(73.5)	(0.00100)	(0.00258)
ZIP Code Links to Visit Firms * Two Quarters Pre	7.88	-0.000323	0.00171
	(60.7)	(0.000849)	(0.00200)
ZIP Code Links to Visit Firms * One Quarter Pre	59.7	-0.000222	0.00210
	(56.5)	(0.000686)	(0.00194)
ZIP Code Links to Visit Firms * One Quarter Post	-14.9	0.000339	0.00368*
	(68.9)	(0.000740)	(0.00219)
ZIP Code Links to Visit Firms * Two Quarters Post	-31.4	-0.000198	0.00412**
	(66.7)	(0.000798)	(0.00203)
ZIP Code Links to Visit Firms * Three Quarters Post	-0.644	-0.00000113	0.00355
	(77.9)	(0.000914)	(0.00225)
ZIP Code Links to Visit Firms * Four Quarters Post	-35.3	0.000122	0.00666***
	(67.1)	(0.000984)	(0.00226)
Firm Fixed Effects	Yes	Yes	Yes
Quarter Fixed Effects	Yes	Yes	Yes
Quarter * Total ZIP Code Links to Alert C Fixed Effects	Yes	Yes	Yes
ZIP Code Clusters	7,046	7,046	7,008
Number of Firm-Quarters	3,181,959	3,181,959	2,159,992
R-Squared	0.00170	0.00624	0.00949

Notes: Standard errors (in parentheses) clustered by ZIP Code. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Probability results from linear probability model.

Table B.11: Effect of Letter on Parent: All Quarters

	Tax Remitted	Any Tax Remitted	Log(Tax Remitted)
Subsidiary Letter * Four Quarters Pre	1,794,258	0.123*	-0.899
	(1,267,413)	(0.0713)	(0.772)
Subsidiary Letter * Three Quarters Pre	758,015	0.0777	-0.360
	(2,429,302)	(0.0564)	(0.578)
Subsidiary Letter * Two Quarters Pre	1,684,730	0.0777	-1.04
	(1,419,499)	(0.0564)	(0.746)
Subsidiary Letter * One Quarter Pre	1,472,714	0.0777	-1.28
	(1,334,352)	(0.0564)	(0.939)
Subsidiary Letter * One Quarter Post	-1,883,667	0.0455	-0.673
	(1,771,504)	(0.0459)	(0.693)
Subsidiary Letter * Two Quarters Post	-679,367	0.0455	-0.760
	(943,835)	(0.0459)	(0.652)
Subsidiary Letter * Three Quarters Post	3,002,550	0.0455	-0.554
	(2,254,423)	(0.0459)	(0.662)
Subsidiary Letter * Four Quarters Post	1,533,720	-1.05e-14	0.0261
	(1,177,288)	(6.19e-08)	(0.234)
Quarter Fixed Effects	Yes	Yes	Yes
Parent Clusters	78	78	32
Observations	702	702	266
R-Squared	0.0301	0.0159	0.106

Notes: Standard errors (in parentheses) clustered by parent. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Probability results from linear probability model.

Table B.12: Effect of Visit on Parent: All Quarters

	Tax Remitted	Any Tax Remitted	Log(Tax Remitted)
Subsidiary Visit * Four Quarters Pre	1,539,170	0.0909	-0.308
	(1,259,938)	(0.0633)	(0.734)
Subsidiary Visit * Three Quarters Pre	-98,195	0.0455	-0.00430
	(2,221,784)	(0.0459)	(0.392)
Subsidiary Visit * Two Quarters Pre	1,856,585	0.0889	-0.617
	(1,460,265)	(0.0635)	(0.800)
Subsidiary Visit * One Quarter Pre	2,321,260	0.0455	-0.240
	(1,552,670)	(0.0459)	(0.664)
Subsidiary Visit * One Quarter Post	-576,292	0.0889	-0.939
	(2,243,087)	(0.0635)	(0.738)
Subsidiary Visit * Two Quarters Post	-764,650	0.00198	-0.356
	(972,303)	(0.0635)	(0.759)
Subsidiary Visit * Three Quarters Post	4,147,976*	0.0455	-0.532
	(2,351,437)	(0.0783)	(0.798)
Subsidiary Visit * Four Quarters Post	845,508	-0.0435	-0.127
	(1,225,135)	(0.0439)	(0.405)
Quarter Fixed Effects	Yes	Yes	Yes
Parent Clusters	78	78	32
Observations	702	702	266
R-Squared	0.0301	0.0159	0.106

Notes: Standard errors (in parentheses) clustered by parent. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Probability results from linear probability model.

Table B.13: Effect of Parent Letter on Subsidiary: All Quarters

	Tax Remitted	Any Tax Remitted	Log(Tax Remitted)
Parent Letter * Four Quarters Pre	19,824	0.0207	1.48
	(23,511)	(0.0584)	(1.01)
Parent Letter * Three Quarters Pre	288,604	0.0289	1.46
	(222,215)	(0.0562)	(1.05)
Parent Letter * Two Quarters Pre	88,733**	0.0496	0.270
	(34,841)	(0.0541)	(0.287)
Parent Letter * One Quarter Pre	106,786	-0.00413	0.695**
	(89,427)	(0.0315)	(0.320)
Parent Letter * One Quarter Post	332,315	4.92e-15	0.148
	(258,302)	(1.17e-08)	(0.223)
Parent Letter * Two Quarters Post	259,164*	-0.0383	1.27
	(152,060)	(0.0357)	(0.968)
Parent Letter * Three Quarters Post	112,400	-0.0383	0.813
	(92,011)	(0.0354)	(0.632)
Parent Letter * Four Quarters Post	104,735	-0.0383	0.995
	(77,877)	(0.0354)	(0.762)
Quarter Fixed Effects	Yes	Yes	Yes
Parent Clusters	49	49	19
Number of Firm-Quarters	3,573	3,573	780
R-Squared	0.129	0.0805	0.237

Notes: Standard errors (in parentheses) clustered at the parent level. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Probability results from linear probability model.

Table B.14: Effect of Parent Visit on Subsidiary: All Quarters

	Tax Remitted	Any Tax Remitted	Log(Tax Remitted)
Parent Visit * Four Quarters Pre	-630,580***	0.0390	0.908
	(109,109)	(0.0498)	(0.999)
Parent Visit * Three Quarters Pre	331,858***	0.0473	0.713
	(103,805)	(0.0472)	(1.00)
Parent Visit * Two Quarters Pre	-316,185***	0.0679	-0.347
	(82,964)	(0.0447)	(0.229)
Parent Visit * One Quarter Pre	-408,154***	0.00504	0.0570
	(141,810)	(0.00565)	(0.245)
Parent Visit * One Quarter Post	915,375***	-0.00917***	0.00531
	(334,732)	(0.00323)	(0.103)
Parent Visit * Two Quarters Post	-544,889	-0.0441	0.796
	(643,782)	(0.0309)	(1.05)
Parent Visit * Three Quarters Post	-312,853	-0.0532	0.456
	(720,091)	(0.0358)	(0.659)
Parent Visit * Four Quarters Post	-554,952	-0.0716*	1.07
	(666,202)	(0.0406)	(0.778)
Quarter Fixed Effects	Yes	Yes	Yes
Parent Clusters	49	49	19
Number of Firm-Quarters	3,573	3,573	780
R-Squared	0.129	0.0805	0.237

Notes: Standard errors (in parentheses) clustered at the parent level. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Probability results from linear probability model.

Figure B.1: Letter

Department of the Treasury
Internal Revenue Service

IRS

Date:

Dear

Your federal tax deposits
We understand federal tax deposit requirements may be confusing and the resulting penalties can be significant. With this in mind, we reviewed your federal tax deposit history and your deposits appear to have decreased. This may be due to a change in your payroll, because you are a new business owner and are not familiar with deposit requirements, or it may be due to other factors.

Your responsibility as an employer
You, as the employer, have the responsibility of withholding trust fund taxes from employees' paychecks. Trust fund tax is money withheld, by an employer, from employees' wages for FICA (social security and Medicare tax) and income tax held in trust until paid to the Department of Treasury. This money must be paid periodically to the Treasury by making federal tax deposits.

What you need to do
Please tell us about the decrease in your deposits so that your account can be updated. You may do one of the following:

- Call the IRS at 1-866-897-4289 Monday through Friday, 8 AM to 8 PM eastern time, or
- Complete and return the enclosed Form 14143, *Reason for Decrease to Federal Tax Deposit.*

Penalty for failing to pay
Individuals who are required to account for and pay these taxes for the business may be personally liable for a penalty if the business fails to pay trust fund taxes. The penalty is equal to the amount of the unpaid trust fund taxes that the business owes the Treasury. For additional information, see the enclosed Notice 784, *Could You be Personally Liable for Certain Unpaid Federal Taxes?*

Penalty for failing to pay timely
If you do not pay these taxes on time or you do not include the required payment with your Form 941, *Employer's Quarterly Federal Tax Return*, interest and penalties will be assessed on any unpaid balance. Additionally, penalties of up to 15% of the amount not deposited may also be assessed, depending on the number of days the federal tax deposits are late.

Letter 4594 (Rev. 10-2013)
Catalog Number 54939M

Penalty for failing to file your return timely
In the event you are unable to pay your taxes timely, it is imperative to file your Form 941 Employer's Quarterly Federal Tax Return timely. If the return is filed after the due date, the law provides penalties for filing late unless there is a reasonable cause for the delay.

Additional information
For further information, please see Publication 15, *Circular E, Employer's Tax Guide*, or the Internal Revenue Service's small business employment tax section. Both are available at www.irs.gov. The employment tax section of the small business web page can be accessed by selecting "Businesses" at the home page, then selecting "Employment Taxes" under Business Topics.

Thank you for taking the time to keep up with your employment tax obligations.

Program Manager
Centralized Processing Operation
Philadelphia Compliance Services

Enclosures:
Form 14143
Notice 784

Letter 4594 (Rev. 10-2013)
Catalog Number 54939M

BIBLIOGRAPHY

Allingham, Michael G and Agnar Sandmo. 1972. Income tax evasion: a theoretical analysis. *Journal of Public Economics* 1, no. 3–4:323–338.

Alstadster, Annette, Wojciech Kopczuk, and Kjetil Telle. 2018. Social networks and tax avoidance: evidence from a well-defined norwegian tax shelter. Working Paper 25191, National Bureau of Economic Research. URL http://www.nber.org/papers/w25191.

Andrei, Amanda, Kevin Comer, and Matthew Koehler. 2014. An agent-based model of network effects on tax compliance and evasion. *Journal of Economic Psychology* 40:119–133.

Andreoni, James. 1992. IRS as loan shark: tax compliance with borrowing constraints. *Journal of Public Economics* 49, no. 1:35–46. URL http://www.sciencedirect.com/science/article/pii/004727279290062K.

Barrot, Jean-Noel and Ramana Nanda. 2016. Can paying firms quicker affect aggregate employment? Tech. rep., National Bureau of Economic Research.

Bernanke, Ben S. 2010. Restoring the flow of credit to small businesses: a speech at the Federal Reserve Meeting Series: "Addressing the Financing Needs of Small Businesses," Washington, D.C., July 12, 2010. Speech 534, Board of Governors of the Federal Reserve System (U.S.). URL https://ideas.repec.org/p/fip/fedgsq/534.html.

Best, Michael Carlos, Anne Brockmeyer, Henrik Jacobsen Kleven, Johannes Spinnewijn, and Mazhar Waseem. 2015. Production versus revenue efficiency with limited tax capacity: theory and evidence from pakistan. *Journal of Political Economy* 123, no. 6:1311–1355.

Bhargava, Saurabh and Dayanand Manoli. 2015. Psychological frictions and the incomplete take-up of social benefits: evidence from an IRS field experiment. *American Economic Review* 105, no. 11:3489–3529.

Blank, Rebecca M and David E Card. 1991. Recent trends in insured and uninsured unemployment: is there an explanation? *The Quarterly Journal of Economics* 106, no. 4:1157–1189.

Blouin, Jennifer and Linda Krull. 2009. Bringing it home: a study of the incentives surrounding the repatriation of foreign earnings under the American Jobs Creation Act of 2004. *Journal of Accounting Research* 47, no. 4:1027–1059.

Blumenthal, Marsha, Charles Christian, and Joel Slemrod. 2001. Do normative appeals affect tax compliance? evidence from a controlled experiment in Minnesota. *National Tax Journal* :125–138.

Board of Governors of the Federal Reserve System. 2018. Commercial bank interest rate on credit card plans, accounts assessed interest [TERMCBCCINTNS]. Retrieved from FRED, Federal Reserve Bank of St. Louis, https://fred.stlouisfed.org/series/TERMCBCCINTNS.

Boning, William C, John Guyton, Ronald H Hodge, Joel Slemrod, and Ugo Troiano. 2018. Heard it through the grapevine: direct and network effects of a tax enforcement field experiment. Tech. rep., National Bureau of Economic Research Working Paper.

Brockmeyer, Anne and Marco Hernandez. 2016. Taxation, information, and withholding: evidence from Costa Rica. Tech. rep., The World Bank.

Chang, Otto H and Joseph J Schultz. 1990. The income tax withholding phenomenon: evidence from TCMP data. *Journal of the American Taxation Association* 12, no. 1:88–93.

Chetty, Raj and Emmanuel Saez. 2013. Teaching the tax code: earnings responses to an experiment with EITC recipients. *American Economic Journal: Applied Economics* 5, no. 1:1–31.

Chodorow-Reich, Gabriel. 2013. The employment effects of credit market disruptions: firm-level evidence from the 2008–9 financial crisis. *The Quarterly Journal of Economics* 129, no. 1:1–59.

degl'Innocenti, Duccio Gamannossi and Matthew D Rablen. 2019. Tax evasion on a social network. Tech. rep., University of Sheffield.

Dharmapala, Dhammika, C Fritz Foley, and Kristin J Forbes. 2011. Watch what I do, not what I say: the unintended consequences of the Homeland Investment Act. *The Journal of Finance* 66, no. 3:753–787.

Diamond, Peter A and James A Mirrlees. 1971. Optimal taxation and public production I: production efficiency. *The American Economic Review* 61, no. 1:8–27.

Diehl, Volker, Harald Stein, Michael Hummel, Raphael Zollinger, and Joseph M Connors. 2003. Hodgkins lymphoma: biology and treatment strategies for primary, refractory, and relapsed disease. *ASH Education Program Book* 2003, no. 1:225–247.

Dobridge, Christine L. 2015. Fiscal stimulus and firms: a tale of two recessions. In *Proceedings. annual conference on taxation and minutes of the annual meeting of the national tax association*, vol. 108. JSTOR, 1–61.

Drago, Francesco, Friederike Mengel, and Christian Traxler. 2015. Compliance behavior in networks: evidence from a field experiment. Tech. rep., IZA Discussion Paper.

Dusek, Libor and Sutirtha Bagchi. 2017. Third-party reporting, tax collections, and the size of government: evidence from withholding. Tech. rep., SSRN.

Engström, Per, Katarina Nordblom, Henry Ohlsson, and Annika Persson. 2015. Tax compliance and loss aversion. *American Economic Journal: Economic Policy* 7, no. 4:132–64.

Faulkender, Michael and Mitchell Petersen. 2012. Investment and capital constraints: repatriations under the American Jobs Creation Act. *The Review of Financial Studies* 25, no. 11:3351–3388.

Feenberg, Daniel and Jonathan Skinner. 1989. Sources of IRA saving. *Tax policy and the economy* 3:25–46.

Feldman, Naomi E. 2010. Mental accounting effects of income tax shifting. *The Review of Economics and Statistics* 92, no. 1:70–86.

Greenstone, Michael, Alexandre Mas, and Hoai-Luu Nguyen. 2014. Do credit market shocks affect the real economy? quasi-experimental evidence from the Great Recession and 'normal' economic times. Tech. rep., National Bureau of Economic Research.

Guyton, John, Pat Langetieg, Day Manoli, Mark Payne, Brenda Schafer, and Michael Sebastiani. 2017. Reminders and recidivism: using administrative data to characterize nonfilers and conduct EITC outreach. *American Economic Review* 107, no. 5:471–75.

Hallsworth, Michael. 2014. The use of field experiments to increase tax compliance. *Oxford Review of Economic Policy* 30, no. 4:658–679.

Hallsworth, Michael, John A List, Robert D Metcalfe, and Ivo Vlaev. 2017. The behavioralist as tax collector: using natural field experiments to enhance tax compliance. *Journal of Public Economics* 148:14–31.

Heckman, James J and Jeffrey A Smith. 1999. The pre-programme earnings dip and the determinants of participation in a social programme. implications for simple programme evaluation strategies. *Economic Journal* 109, no. 457:313–348.

Internal Revenue Service. 2016. Tax gap estimates for tax years 2008-2010. https://www.irs.gov/newsroom/the-tax-gap. Accessed February 13, 2018.

———. 2018. *Internal revenue service data book, 2017. publication 55b*. Washington: Internal Revenue Service.

Jensen, Anders. 2019. Employment structure and the rise of the modern tax system. Working Paper 25502, National Bureau of Economic Research. URL http://www.nber.org/papers/w25502.

Jones, Damon. 2012. Inertia and overwithholding: explaining the prevalence of income tax refunds. *American Economic Journal: Economic Policy* 4, no. 1:158–85.

Keen, Michael and Joel Slemrod. 2017. Optimal tax administration. *Journal of Public Economics* 152:133–142.

Klassen, Kenneth J, Petro Lisowsky, and Devan Mescall. 2015. The role of auditors, non-auditors, and internal tax departments in corporate tax aggressiveness. *Accounting Review* 91, no. 1:179–205.

Klepper, Steven, Mark Mazur, and Daniel Nagin. 1991. Expert intermediaries and legal compliance: the case of tax preparers. *Journal of Law and Economics* 34, no. 1:205–229.

Kleven, Henrik J, Martin B Knudsen, Claus T Kreiner, Sren Pedersen, and Emmanuel Saez. 2010. Unwilling or unable to cheat? evidence from a randomized tax audit experiment in Denmark. Working Paper 15769, National Bureau of Economic Research. URL http://www.nber.org/papers/w15769.

Kopczuk, Wojciech, Justin Marion, Erich Muehlegger, and Joel Slemrod. 2016. Does tax-collection invariance hold? evasion and the pass-through of state diesel taxes. *American Economic Journal: Economic Policy* 8, no. 2:251–286.

Mayshar, Joram. 1991. Taxation with costly administration. *Scandinavian Journal of Economics* 93:75–88.

Meiselman, Ben S. 2018. Ghostbusting in Detroit: evidence on nonfilers from a controlled field experiment. *Journal of Public Economics* 158:180–193. URL http://www.sciencedirect.com/science/article/pii/S0047272718300057.

Mortenson, Jacob A and Andrew Whitten. 2016. How sensitive are taxpayers to marginal tax rates? evidence from income bunching in the United States. Tech. rep., Georgetown University.

OECD. 2015. Tax administration comparative information series. https://qdd.oecd.org/subject.aspx?Subject=TAS. Accessed May 11, 2018.

Ortega, Daniel and Carlos Scartascini. 2015. Who's calling: the effect of phone calls as a deterrence mechanism. Tech. rep., Inter-American Development Bank.

———. 2018. Don't blame the messenger: a field experiment on delivery methods for increasing tax compliance. Tech. rep., Inter-American Development Bank.

Perez-Truglia, Ricardo and Ugo Troiano. 2018. Shaming tax delinquents. *Journal of Public Economics* 167:120–137. URL http://www.sciencedirect.com/science/article/pii/S0047272718301762.

Pomeranz, Dina. 2015. No taxation without information: deterrence and self-enforcement in the value added tax. *American Economic Review* 105, no. 8:2539–2569.

Powell, David. 2015. Do payroll taxes in the United States create bunching at kink points? Tech. rep., Michigan Retirement Research Center Working Paper.

Rauh, Joshua D. 2006. Investment and financing constraints: evidence from the funding of corporate pension plans. *The Journal of Finance* 61, no. 1:33–71.

Rees-Jones, Alex. 2017. Quantifying loss-averse tax manipulation. *The Review of Economic Studies* 85, no. 2:1251–78.

Rincke, Johannes and Christian Traxler. 2011. Enforcement spillovers. *Review of Economics and Statistics* 93:1224–1234.

Saez, Emmanuel. 2009. Details matter: The impact of presentation and information on the take-up of financial incentives for retirement saving. *American Economic Journal: Economic Policy* 1, no. 1:204–28.

Sahm, Claudia R, Matthew D Shapiro, and Joel Slemrod. 2012. Check in the mail or more in the paycheck: does the effectiveness of fiscal stimulus depend on how it is delivered? *American Economic Journal: Economic Policy* 4, no. 3:216–50.

Shapiro, Matthew D and Joel Slemrod. 1995. Consumer response to the timing of income: evidence from a change in tax withholding. *American Economic Review* 85, no. 1:274–283.

Slemrod, Joel. forthcoming. Tax compliance and enforcement. *Journal of Economic Literature* .

Slemrod, Joel, Marsha Blumenthal, and Charles Christian. 2001. Taxpayer response to an increased probability of audit: evidence from a controlled experiment in Minnesota. *Journal of Public Economics* 79, no. 3:455–483. URL http://www.sciencedirect.com/science/article/pii/S0047272799001073.

Slemrod, Joel and Shlomo Yitzhaki. 1987. The optimal size of a tax collection agency. *Scandinavian Journal of Economics* 89:183–192.

USPS. 2016. Faq. https://faq.usps.com/. Accessed October 30, 2016.

Zonder, Jeffrey A, Pamela Pemberton, Helen Brandt, Anwar N Mohamed, and Charles A Schiffer. 2003. The effect of dose increase of imatinib mesylate in patients with chronic or accelerated phase chronic myelogenous leukemia with inadequate hematologic or cytogenetic response to initial treatment. *Clinical Cancer Research* 9, no. 6:2092–2097.

Zwick, Eric and James Mahon. 2017. Tax policy and heterogeneous investment behavior. *American Economic Review* 107, no. 1:217–48.