

Ghostbusting in Detroit: Evidence on nonfilers from a controlled field experiment

Ben Meiselman (University of Michigan)

21 December 2016, 17:00-18:15, room 465, bld.72

Abstract: Many people who owe income tax fail to file a timely tax return. In communication with these “ghosts,” what messages from the tax authority are effective for eliciting a return? This is the first study to address message content in communication with income tax nonfilers. I assess the efficacy of messages related to *penalty salience*, *punishment probability*, *compliance cost*, and *civic pride* by evaluating the response to experimental mailings distributed to 7,142 Detroit residents. The *penalty salience* message was the most effective. Relative to a basic mailing that requested a return, *penalty salience* mailings that stated the statutory penalty for failing to file a return tripled response rates from 3% to 10%, increased the number of back-year returns filed per response from 0.08 to 0.27, and raised the fraction of filed returns that admitted tax due from 39% to 52%. *Compliance cost* mailings that enclosed a blank tax return and *punishment probability* mailings that stated the recipient’s federal income also raised response rates relative to the basic mailing, but *civic pride* mailings did not. Mailings were more effective in eliciting returns from older, higher-income, and first-time nonfilers. I investigate the impact of treatment mailings on the behavior of untreated neighbors and find no evidence of geographic network effects.

Ghostbusting in Detroit: Evidence on nonfilers from a controlled field experiment*

Ben Meiselman[†]
University of Michigan

September 2016

[\[Latest version\]](#)

Abstract

Many people who owe income tax fail to file a timely tax return. In communication with these “ghosts,” what messages from the tax authority are effective for eliciting a return? This is the first study to address message content in communication with income tax nonfilers. I assess the efficacy of messages related to *penalty salience*, *punishment probability*, *compliance cost*, and *civic pride* by evaluating the response to experimental mailings distributed to 7,142 Detroit residents. The *penalty salience* message was the most effective. Relative to a basic mailing that requested a return, *penalty salience* mailings that stated the statutory penalty for failing to file a return tripled response rates from 3% to 10%, increased the number of back-year returns filed per response from 0.08 to 0.27, and raised the fraction of filed returns that admitted tax due from 39% to 52%. *Compliance cost* mailings that enclosed a blank tax return and *punishment probability* mailings that stated the recipient’s federal income also raised response rates relative to the basic mailing, but *civic pride* mailings did not. Mailings were more effective in eliciting returns from older, higher-income, and first-time nonfilers. I investigate the impact of treatment mailings on the behavior of untreated neighbors and find no evidence of geographic network effects.

Keywords: Nonfiler, Tax Evasion, Income Tax

JEL Codes: H24, H26

*The data in this paper pertaining to the field experiment is subject to a non-disclosure agreement and was provided to the author from the City of Detroit. The content of the manuscript is also subject to disclosure approval from the City of Detroit.

[†]I am grateful for guidance from Joel Slemrod. I thank Odell Bailey, Will Boning, Charlie Brown, Donna Brown, Enda Hargaden, Jim Hines, Chris House, Ryan Kellogg, Jason Kerwin, Gaurav Khanna, Margaret Lay, Dayanand Manoli, Nimit Modi, Stefan Nagel, Carol O’Cleireacain, Debra Pospiech, Daniel Reck, David Szymanski, Ugo Troiano, Michelle Weston, and Eleanor Wilking. I gratefully acknowledge funding from the Office of Tax Policy Research, the Michigan Institute for Teaching and Research in Economics, and Rackham Graduate School at the University of Michigan. The views expressed here are my own and do not necessarily reflect the views of the City of Detroit.

1 Introduction

Tax authorities want to know what messages induce compliance from noncompliant taxpayers. Relative to other enforcement mechanisms like audits or site visits, the marginal cost of written communication is low. Even better, the marginal cost of making communication more effective is zero; the postage cost of mailing a letter that gets filed in the dustbin is the same as the postage cost of mailing a letter that induces timely compliance. Tax authorities want to send a message that works.

One common form of noncompliance is failure to file a tax return. For the federal individual income tax, Erard et al. (2014) estimate that 6.1% of required tax year 2012 returns were not filed on time. For Detroit's individual income tax, I estimate that 46% of required tax year 2014 returns were not filed on time. Controlled experiments are becoming more common in the tax compliance literature, but nonfilers have been the focus of only one such empirical paper, which examined the effect of repetition and reminders on filing rates (Guyton et al., 2016).

This paper provides the first evidence from a controlled experiment about message content in communication with income tax nonfilers. The experiment was designed and conducted by the author in collaboration with the City of Detroit. Detroit's income tax division sent mailings in April through June 2016 to 7,142 suspected ghosts—people who likely owed tax but did not file a tax year 2014 return. Each mailing contained one of several experimental messages, related variously to penalty salience, punishment probability, compliance cost, or civic pride. From the population of suspected ghosts with at least \$350 in estimated tax liability, nonfilers were randomly selected into experimental treatments and sent the same message in two mailings: a postcard, and then a letter one week later.

I compare the effectiveness of the various experimental messages for inducing taxpayer compliance. The main outcome of interest is the response rate, the rate at which mailings

elicited a tax return from taxpayers who had previously failed to file a return. I also evaluate response quality, including the amount of remittances, the likelihood of claiming a refund rather than admitting tax due, and the number of back-year returns accompanying the tax year 2014 return. I examine whether taxpayer behavior differed across treatments in ways that can be attributed only to messages on the postcard, such as the rate at which taxpayers accepted the letters, which required a signature for delivery. I identify taxpayer characteristics, including age and income, that were associated with higher response rates to the experimental messages. I investigate geographic network effects—the response rates of untreated neighbors to experimental mailings.

In communication with nonfilers, a simple penalty salience message is the most effective at inducing compliance. Mailings that stated the statutory penalty for failing to file elicited a tax return from 10.1% of intended recipients, more than triple the response rate to the contact-only control mailings and more than any other treatment mailings. Taxpayers in the penalty salience treatment were most likely to file back-year returns, most likely to admit tax due, and most likely to remit payment. Taxpayers responded more promptly to the penalty salience message, sometimes after receiving just the postcard, before the letter was even delivered.

The compliance cost mailings were also effective at raising response rates relative to a contact-only control, but the response quality was lower than the penalty salience mailings. Whereas all other treatment mailings differed only by one or two sentences in a prominent box on the postcard or letter, the compliance cost treatment letter also enclosed a blank tax form and return envelope. The response rate to the compliance cost mailings was 6.2%, double the response rate to the contact-only control mailings. However, the returns that were filed in response to the compliance cost mailings were more likely to claim refunds and less likely to admit tax due than the returns filed in response to the penalty salience mailings. Taxpayers also filed fewer back-year returns in response to compliance cost mailings.

Adding to the penalty salience message a punishment probability message that informed nonfilers that the city tax authority knew their 2014 federal income dampened response rates relative to the penalty salience message by itself. The idea behind the punishment probability message is that revealing the nonfiler’s federal income demonstrates that the tax authority has the ability to monitor taxpayer behavior and therefore raises the perceived probability of punishment. On its own, the punishment probability message raised response rates relative to the contact-only control. If the punishment probability and penalty salience messages both operated exclusively through their intended channels, we would expect that including both messages would raise response rates relative to one or the other by itself (Erard and Ho, 2001). However, when mailings included both the punishment probability message and the penalty salience message, the response rate was lower than the response rate to mailings with just the penalty salience message. This surprising result may be a consequence of limited taxpayer attention, supporting the conclusions from prior literature that simplicity is important in communication with taxpayers (Bhargava and Manoli, 2013).

This is the first controlled experiment to test the effectiveness of a civic pride message on taxpayers, and the response rate was statistically indistinguishable from the contact-only control. The civic pride message reminded taxpayers that the collection of taxes is essential to the successful resurgence of the City of Detroit. Normative appeals like the civic pride message have taken three forms in prior tax enforcement experiments: public service, fairness, and compliant majority messages. A typical “public service” message reminds taxpayers that tax revenue is used to pay for government services like education and law enforcement (Chirico et al., 2015; Blumenthal et al., 2001). A typical “fairness” message would ask taxpayers to comply with tax law in order to do their fair share (Torgler, 2004; Chirico et al., 2015). A typical “compliant majority” message would inform taxpayers of the compliance rate of their neighbors (Blumenthal et al., 2001; Fellner et al., 2013; Hallsworth et al., 2014; Hasseldine et al., 2007). Consistent with the results of this experiment, most

prior literature finds that normative appeals are not as effective as messages about the probability of being caught and the penalty if caught (Slemrod, 2015).

I find no evidence of geographic network effects. Network effects can be important even when per-neighbor effects are very small because treated individuals can have many neighbors. To investigate geographic network effects, I compute the distance between every treated nonfiler and every untreated taxpayer who filed a return within 90 days of the first postcard in the experiment. The effect of treatment mailings on filing rates of taxpayers within 100 meters of treated nonfilers was not statistically significant, and this finding was robust to alternative distances. If there are network effects from treatment, they are likely through family or coworkers rather than geographic neighbors.

I find that the penalty salience mailings improve welfare. I estimate that by applying the penalty salience treatment to the entire population of interest, the city could generate net revenue of \$266,000. With conservative assumptions about the social value of public spending, the penalty salience treatment generates more than enough revenue to offset the estimated compliance costs of taxpayers.

Heterogeneous response rates offer an opportunity to improve welfare even more. Taxpayers with higher income, older taxpayers, and taxpayers who were identified as nonfilers for the first time in TY2014 were more responsive to all treatments including penalty salience. The effects of age, income, filing history, and treatment status appear to be positive even when they are all at play. Relative to applying the penalty salience treatment to the entire population of interest, net welfare could be improved by focusing on a smaller population with higher response rates.

Section 2 gives background on the income tax system, the decision to file, and the estimated number of nonfilers in Detroit. Section 3 presents the design of a controlled field experiment. Section 4 presents the results of the field experiment. Section 5 discusses welfare, targeting criteria, and generalization of the findings. Section 6 concludes.

2 Background

2.1 Tax system

The City of Detroit levies an income tax on local residents and local workers. Regardless of where they work, residents owe 2.4% of income, with an exemption of \$600 for individuals and \$1,200 for joint filers. People who work in Detroit but reside elsewhere owe 1.2% of income with the same exemption levels. Detroit imposes other taxes such as property tax, but my focus here is on the income tax.

Whether the worker or the firm remits income tax to Detroit depends on worker classification and firm location. A firm must classify workers as either employees or contractors.¹ A firm located in the city must withhold from employees and remit income tax to Detroit. However, a firm located outside the city is not required to withhold Detroit income tax from employees, even if the employees owe Detroit income tax because they are Detroit residents.² A firm never remits income tax on behalf of contractors, regardless of the firm's location. City tax administrators believe one reason remittances by firms have fallen is that an increasing share of the workforce is classified as contractors.³

Reporting requirements also depend on worker classification and firm location. Firms issue forms that summarize annual income to all workers—a Form W2 for employees and a Form 1099 for contractors. A taxpayer must include a copy of W2s and 1099s she received when she files a tax return with the city. A firm located in Detroit must report the income information from 1099s to the city, and the income and withholding information from W2s

¹Generally, workers who receive benefits and over whom the firm has control are employees. The IRS has guidelines for distinguishing employees from contractors: <https://www.irs.gov/businesses/small-businesses-self-employed/independent-contractor-self-employed-or-employee>.

²State legislation may soon require firms outside the city to withhold and remit income tax on behalf of Detroit residents. See House Bill 4829 of 2015: <http://legislature.mi.gov/doc.aspx?2015-HB-4829>.

³“Detroit Chief Financial Officer John Hill blamed the [revenue] shortfall on a reduction in income tax withholding by employers in the city. In February, he told the Financial Review Commission, the state board overseeing Detroit’s fiscal affairs post-bankruptcy, that he believes a fair number of the new workers downtown are contract workers.” *Detroit Free Press*, March 7, 2015.

to the city. A firm located outside Detroit is not required to report income earned by Detroit residents.

Tax enforcement in Detroit is severely limited by administrative capacity. Detroit struggles just to process returns submitted on time by compliant taxpayers.⁴ Around the time of Detroit's bankruptcy in July 2013, lawyers for the city who wanted to sue taxpayers with known tax due were limited by the court, which had insufficient staff to process more than five such cases per week. Prior to tax year 2015, Detroit did not accept electronic returns; taxpayers were required to mail a paper return to a post office box or deliver a paper return in person to the municipal center.⁵

Within these limits, Detroit does audit tax returns, but not the same way as the IRS. For tax year 2014, Detroit contracted with Chase Bank to scan and manually key tax returns into a data file, which was then loaded into proprietary software called CityTax. City auditors can check information from returns in CityTax against information on federal income tax returns that are shared with Detroit by the IRS.⁶ Whereas IRS audits often independently verify information supplied by a taxpayer, the vast majority of Detroit audits currently go no further than comparing the information in the city return to the information in the federal return. Information on the federal return is treated as verification.

Michigan gives cities legal tools for income tax enforcement. A city tax authority is permitted to examine records that will help it to assess tax liability, including the tax liability of individuals who did not file a return but are believed to owe income tax. The

⁴"Taxpayers often wait months or even years before their refund checks arrive." *Detroit Free Press*, March 7, 2015.

⁵Detroit's tax administration is changing. In recognition of capacity constraints, Detroit turned over primary responsibility for processing city returns to the state beginning with tax year 2015. Even as Detroit ceded some responsibility to Michigan, the city maintained its own compliance and enforcement apparatus. The sample in this paper is for tax year 2014, for which the city retained full responsibility.

⁶The IRS shares federal tax information with state and local governments for the purpose of tax enforcement. Third party information reporting is an important mechanism of tax enforcement, as noted by, for example, Erard and Ho (2004) and Pomeranz (2015). This context is somewhat unusual because the "third party" is another level of government.

city does not have automatic subpoena power over records, but it can sue noncompliant individuals in court to compel documents. Willful failure to file a return, remit tax owed, or permit the tax authority to examine records is a misdemeanor.⁷

Detroit has two available pathways for pursuing identified individuals who have not filed a tax return. The first pathway is to send a “proposed assessment” to the taxpayer based on the city’s belief of what the taxpayer owes. If the taxpayer receives and does not dispute the proposed assessment, the tax debt becomes official. If the taxpayer then does not remit the tax debt, Detroit sends the debt to a collection agency. The second pathway is a criminal procedure. The city can charge an individual who fails to file a tax return with a misdemeanor. Prior to calendar year 2016, Detroit used the first pathway exclusively—issuing proposed assessments and forwarding unpaid tax debt to a collection agency. Detroit filed misdemeanor charges against noncompliant taxpayers in calendar year 2016 for the first time in many years.

As part of the proposed assessment pathway, the city must be able to prove that the taxpayer received the proposed assessment in order for the tax debt to become official. There is no such notification requirement for the city to charge taxpayers with a misdemeanor. To be courteous and reduce enforcement costs, city administrators prefer to communicate with taxpayers prior to charging them with a misdemeanor, but Detroit is under no legal obligation to do so. The city’s burden of ensuring the taxpayer is notified when it pursues the proposed assessment pathway may have led taxpayers to believe that they could avoid getting in trouble by refusing to accept the treatment letter, which was sent via certified mail and therefore required a signature for delivery.

⁷City Income Tax Act of 1964, Act 284 at 141.673 and 141.699: <https://legislature.mi.gov/documents/mcl/archive/2014/May/mcl-Act-284-of-1964.pdf>.

2.2 Filing decision

The logic of the standard model of income reporting can be naturally extended to the decision whether to file a return. In the standard model of [Allingham and Sandmo \(1972\)](#), taxpayer reports depend on the probability of audit and the penalty for a false report. In an extension by [Erard and Ho \(2001\)](#), taxpayer choice of whether to file a return depends on the probability of detection and the penalty for nonfiling. One suspects that Detroit residents and workers correctly perceive that the probability of punishing nonfilers is low. However, the statutory penalty for failing to file an income tax return is substantial: a \$500 fine and 90 days in jail.

The extended model of filing a return includes compliance costs, which appear to be important in Detroit. Many workers who are owed a refund from the city, because they have income tax withheld from their paychecks exceeding tax liability, still fail to file a return. The standard model cannot explain this behavior. It is possible that some of these workers decide not to claim a refund as a form of “donation” to the city, but it seems likely that compliance costs are more important. Compliance costs should be at least as large for taxpayers with tax due as it is for taxpayers who are owed a refund.

There may also be nontax reasons to avoid truthfully reporting residence. For instance, car insurance rates are particularly high in Detroit, higher than in districts immediately adjacent to the city.⁸ A resident of downtown Detroit would save money on car insurance by claiming residence at a suburban address. Workers may believe their true residence is more likely to be detected by the car insurance company if it is truthfully reported on an income tax return. So the decision not to file a return may be jointly determined by considerations of the probability of detection by the tax authority, the penalty for detection, compliance

⁸According to carinsurance.com, the average annual auto insurance rate was \$1,400 higher in central Detroit than in selected suburbs adjacent to Detroit. Reported auto insurance rates are averages by zip code for a 2014 Honda Accord for a single 40-year-old male with a clean record and good credit. The average rates in central Detroit were \$4,846 in Downtown (zip code 48226), \$5,025 in Midtown (48201), \$4,945 in New Center / North End (48202), \$4,827 in Downtown (48207), and \$4,636 in Corktown / Woodbridge (48216). The average rates in selected suburbs were \$3,491 in Southfield (48075), \$3,489 in Oak Park (48237), \$2,621 in Ferndale (48220), \$3,139 in Grosse Pointe (48230), and \$4,256 in Dearborn (48126).

costs, and nontax reasons for claiming residence elsewhere.

It is also possible that failure to file a return is not the result of a conscious decision or optimizing behavior. Some taxpayers may mistakenly believe that they filed a city return electronically. Many taxpayers file federal and state returns electronically, but Detroit did not accept city tax returns electronically prior to tax year 2015. If taxpayers use tax preparation software, they may think they are done with all of their federal, state, and local returns when they click the submit button, but that is not true if they owe Detroit income tax. Detroit only processes income tax returns that are mailed to a post office box or hand delivered. Furthermore, some Detroit residents and workers, especially those new to the area, may honestly be unaware that Detroit has an income tax.⁹

2.3 Nonfiler population size

In designing a tax enforcement approach to nonfilers, it would be helpful to know how many nonfilers there are. That would be easy for a tax administrator to calculate if she knew who is in the tax base and who filed tax returns. The identity of filers is known, but the identity and size of the tax base is unknown. Detroit's income tax base consists of residents and workers whose income exceeds the exemption amount.¹⁰

I estimate the number of people who owed Detroit income tax for tax year 2014 to be approximately 387,000. To calculate this figure, I use the Current Employment Statistics (CES) program of the Bureau of Labor Statistics to estimate the number of people who work in Wayne County and the number of employed Wayne County residents. I then utilize Longitudinal Employer-Household Dynamics (LEHD) Origin-Destination Employment Statistics from the Census Bureau to estimate Detroit's share of workers in Wayne County

⁹Awareness of Detroit's income tax seems comparable to awareness of city income tax in Ohio cities Cincinnati and Columbus, judging by an index of search interest from Google Trends. See Appendix Figure A.2. [Hoopes et al. \(2015\)](#) discuss tax enforcement with uninformed taxpayers.

¹⁰The exemption amount is \$600 for single filers and \$1,200 for joint filers.

Table 1: Estimated Detroit Tax Base

Year	Detroit residents who work in Detroit	Detroit residents who work elsewhere	Nonresidents who work in Detroit	Total
2012	68,970	191,878	121,542	382,389
2013	66,468	191,176	123,256	380,901
2014	67,562	194,144	125,398	387,103

Note: Estimates of Detroit resident-workers, nonresident workers, and worker nonresidents are obtained by applying LEHD Detroit shares of Wayne County to the CES workforce of Wayne County. Wayne County workforce is from Current Employment Statistics (CES) program of the Bureau of Labor Statistics. Detroit share of Wayne County workers and employed residents is from Longitudinal Employer-Household Dynamics (LEHD) Origin-Destination Employment Statistics from the Census Bureau.

and Detroit’s share of employed residents of Wayne County. Table 1 shows that the estimated income tax base of Detroit was 387,000 people when these shares from LEHD are applied to the workforce of Wayne County from CES.

For a given tax year, the population of nonfilers shrinks over time because many individuals file city tax returns months or years late. The population of nonfilers for a given tax year is thus a moving target. For the past several tax years, Detroit received around 150,000 returns on time and 30,000 over the subsequent year. Thousands of city returns continued to trickle in more than a year after the filing deadline. For the purpose of cross-year comparisons, it is therefore important to specify the date on which the population is being measured.

As of April 2016, when the field experiment began, I estimate the number of people who were Detroit nonfilers for tax year 2014 to be 179,000. The estimate comes from subtracting the number of people in the tax base who filed returns from the total number of people in the tax base. That estimate implies that 46% of individuals who were required to file Detroit tax returns failed to file a return. Assuming 40% of joint returns have two earners, as in Table 2, and that 17.3% of nonfilers would file joint returns, there were 167,000 missing returns, equal to 48% of required returns.¹¹

¹¹Erard et al. (2014) estimate there were 7.6 million federal individual income tax nonfilers in 2012 (6.1% of required returns).

A notable source of uncertainty is the number of joint filers who earned income. When all income is reported by third parties to the tax authority on W2s and 1099s, then the tax authority knows whether one or both individuals in a couple filing jointly are among the 387,000 individuals in the tax base. However, for income with no third-party reporting, there is no way to know whether each individual is in the tax base. Table 2 shows the computation of nonfilers by subtracting the number of people in the tax base who filed returns from the number of people in the tax base.

3 Design of a controlled field experiment

3.1 Sample

A sample of 9,523 individuals for the field experiment was randomly selected from a population of 42,704 suspected nonfilers. The population of nonfilers met the following criteria: (1) The IRS identified the individual as a federal taxpayer with a Detroit residence and income taxable to Detroit in tax year 2014, (2) Detroit had no record of the individual filing a 2014 city income tax return as of April 2016, (3) Detroit estimated the individual had 2014 tax due to the city of at least \$350, (4) the individual’s address appeared to be valid,¹² (5) Detroit had no record of the individual passing away or filing for bankruptcy.

Table 3 reports summary statistics for individuals who filed a federal return in tax year 2014 with a Detroit address by local filing status, sample eligibility, and sample selection. Among federal filers, individuals who failed to file a city return were younger on average and more likely to file as a head of household. Local nonfilers had lower income, and they were much more likely to have been identified by Detroit as a nonfiler for a tax year prior to 2014. Around 84% of nonfilers in the sample were also identified as a nonfiler for a prior year.

¹²To avoid pursuing individuals who were not actually Detroit residents, addresses were excluded if they had a zip code that is shared between Detroit and another city (e.g. Highland Park). To reduce the nondelivery rate, addresses were excluded if they had a street name that was not shared by other federal taxpayers, on the grounds that it was likely to be an erroneous address.

Table 2: Estimated Detroit Nonfilers

Year	Individuals in Detroit tax base	Returns filed	Joint returns	Joint returns with two earners	Nonfilers
2012	382,389	205,275	82,046	32,135	144,979
2013	380,901	193,455	77,084	31,133	156,313
2014	387,103	178,859	72,843	28,847	179,397

Note: Individuals in Detroit tax base (column 2) is author’s estimate explained in the text and Table 1. Returns filed (column 3) includes resident, non-resident, and partial-year returns. A return is considered to be a joint return (column 4) if it was marked as such by the taxpayer. A return is considered to be a joint return with two earners (column 5) if there was a W2 associated with the “secondary” social security number. Nonfilers (column 6) is equal to the size of the tax base (column 2) less the number of returns (columns 3) and the number of joint returns with two earners (column 5).

Table 3: Summary Statistics (TY 2014)

	Filer		Nonfiler		Population		Sample	
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Age	50.2	15.4	44.8	17.5	39.6	12.0	39.6	12.0
FS = single (%)	46.4	49.9	43.4	49.6	39.4	48.9	39.2	48.8
FS = married filing jointly (%)	26.8	44.3	17.3	37.8	10.6	30.8	10.8	31.0
FS = head of household (%)	24.3	42.9	37.9	48.5	48.6	50.0	48.6	50.0
Wage Income (\$ 000s)	44.4	61.3	27.4	80.2	31.4	23.3	31.3	23.7
Total Inc. (\$ 000s)	56.9	91.5	40.1	175.4	33.9	42.6	33.4	24.7
Adjusted Gross Inc. (\$ 000s)	56.5	90.1	39.6	172.0	33.6	42.5	33.2	24.5
Taxable Inc. (\$ 000s)	37.4	75.4	22.6	167.5	15.9	41.3	15.5	22.6
NF in 1-3 yrs before TY2014 (%)	0.3	5.2	20.9	40.7	40.5	49.1	40.6	49.1
NF in 4-8 yrs before TY2014 (%)	0.2	4.0	20.5	40.4	43.7	49.6	44.0	49.6
Log income	3.7	0.8	3.3	0.7	3.4	0.5	3.4	0.5
Observations	61,745		185,137		42,754		9,523	

Note: This table reports means of observable characteristics based on administrative tax data for Detroit residents who filed federal income tax returns for tax year 2014. Filers are Detroit residents who filed a federal return and a city return. Nonfilers (NF) are people who filed a federal return but no city return. Filing status (FS) is reported from the federal return.

Detroit estimates tax due from nonfilers using an algorithm that includes federal income information from the IRS and local withholding information from city employers. The city’s algorithm for estimating tax due is correct within \$15 of actual tax due for 70% of taxpayers who file both local and federal returns. Incomplete withholding information from employers causes discrepancies between Detroit’s estimation of tax due and actual tax due. Detroit’s estimation of tax due is too high for nonfilers with employers who did not submit W2s to the city electronically.¹³ Two sources of income—active duty military pay and pension

¹³Detroit accepts W2s from employers in electronic (online or CD) and paper format. Around 4% of the 12,700 employers who file an annual report with individual income tax withholding do so electronically. If an employer submitted a W2 electronically, then Detroit used the withholding amount for the nonfiler to estimate tax due. W2s that were submitted in paper form only were not digitized or used to estimate

income—also cause discrepancies between Detroit’s estimation of tax due and actual tax due. Detroit’s estimation of tax due is too high for nonfilers with these types of income.¹⁴

Detroit excluded taxpayers with addresses that were likely to be invalid. For prior tax years, Detroit sent tens of thousands of letters to nonfilers, thousands of which were returned as undeliverable. For tax year 2014, Detroit used a filter on addresses that marked about 7% of IRS addresses as likely to be invalid prior to sample selection. The United States Postal Service contracts with private vendors to offer paid address verification services, but Detroit does not pay for those services.

3.2 Experimental treatments

Taxpayers in the sample were sent two separate mailings in sequence, one week apart. The first mailing was a postcard, and the second mailing was a letter.¹⁵ The postcard listed the types of income that are taxable by Detroit and directed taxpayers where to find tax forms and filing instructions. The letter informed the nonfiler that Detroit believes they had taxable income and failed to file a city tax return for tax year 2014. Taxpayers were randomly assigned to a treatment status, which varied the content of a prominent box in both the postcard and the letter. Table 4 reports the message associated with each treatment status. Examples of postcards and letters are in Appendix Figure A.1.¹⁶

Penalty salience. One treatment status tested whether penalty salience affects tax com-

individual income tax due. By dollar value, around 20% of tax prepayments reported on city returns, including employer withholding and estimated payments from business income, are visible to the tax division and able to be connected to the taxpayer before receiving the city return. See Appendix Table A.2.

¹⁴Active duty military pay appears as wage (W2) income on a federal 1040. It is taxable income to the federal government, but it is not taxable income to Detroit. Detroit cannot systematically distinguish between active duty military pay and other wage income, although it can request that information for individual taxpayers. Similarly, pension income appears as other (1099-MISC) income on a federal 1040. It is taxable income to the federal government, but it is not taxable income to Detroit. As with military pay, Detroit cannot systematically distinguish between pension income and other income from a 1099-MISC, although it can request that information for individual taxpayers.

¹⁵To track delivery rates, the letters were sent via United States Postal Service certified mail.

¹⁶This study was submitted for approval to the University of Michigan Health Sciences and Behavioral Sciences Institutional Review Board. The IRB determined that this study had a status of “Not Regulated”.

Table 4: Experimental treatments

Treatment	Intervention	Message in prominent box on letter ¹
Penalty salience	Postcard and letter	Failure to file a tax return is a misdemeanor punishable by a fine of \$500 and 90 days in jail.
Punishment probability	Postcard and letter	Our records indicate you had federal total income of \$X for tax year 2014. ¹
Compliance cost	Postcard and letter, form and return envelope enclosed with letter	For your convenience, City Income Tax Form D-1040(R) is enclosed with this letter. ¹
Civic pride	Postcard and letter	Detroit’s rising is at hand. The collection of taxes is essential to our success.
Penalty salience × Punishment probability	Postcard and letter	Our records indicate you had federal total income of \$X for tax year 2014. ¹ Failure to file a tax return is a misdemeanor punishable by a fine of \$500 and 90 days in jail.
Contact-only (control)	Postcard and letter	None
No-contact (control)	None	N/A

Note: This table describes the experimental treatments. 1,200 taxpayers were assigned to each experimental treatment other than the no-contact control, to which 2,400 taxpayers were assigned.

¹ The boxed message was exactly the same on the postcard and the letter within each treatment other than the punishment probability treatments and the compliance cost treatment. In the punishment probability treatments, the boxed message on the postcard was, “The letter you receive will indicate how much taxable income you had in tax year 2014.” In the compliance cost treatment, the boxed message on the postcard was, “For your convenience, City Income Tax Form D-1040(R) will be enclosed with the letter.”

pliance. The boxed message stated that failure to file a tax return is a misdemeanor, and the statutory penalty for the misdemeanor is a fine of up to \$500 and 90 days in jail. Absent this treatment, the statutory penalty was almost certainly unknown by the vast majority of Detroit residents. The city had not prosecuted anyone under the misdemeanor provision for many years. The message in this treatment status was not phrased as a threat, but it is comparable to other field experiments that test “threats” of various sorts.¹⁷

¹⁷The “threat” treatment in [Chirico et al. \(2015\)](#) actually uses threatening language. Most other threat treatments are based on the threat of auditing a return, rather than the threat of punishment if no further action is taken.

Punishment probability. Another treatment status was intended to affect the perceived probability of punishment. The boxed message revealed that Detroit knew the recipient's total federal income, which is among the information provided by the IRS to Detroit. The rationale for this treatment is that a taxpayer will feel punishment is more likely if the tax authority reveals that it has relevant information. Revealing this information is intended to raise the perceived probability of punishment, relative to the letters that do not reveal that Detroit has information about the taxpayer other than name and address.

Compliance cost. The cost to the taxpayer of filing a return was reduced by a treatment status that enclosed a blank tax form and a return envelope. The enclosed return was for Detroit residents for tax year 2014, Form D-1040(R).¹⁸ The boxed message referred to the tax form as being provided for the convenience of the recipient. Although the monetary cost of the form and envelope is small, the nonmonetary cost could be substantial, including the time and effort to find the form online or retrieve it from the Coleman A. Young Municipal Center in downtown Detroit.

Civic pride. One set of mailings tested the effect of an appeal to civic pride. The boxed message proclaimed the importance of tax collection to the resurgence of Detroit. This is the first normative appeal of its kind, but it is not the only type of normative appeal that is potentially relevant in communication with taxpayers. In similar tax enforcement field experiments, normative appeals to taxpayers have (1) reminded taxpayers of services provided by tax dollars, (2) informed taxpayers about the compliance rate of their neighbors, and (3) referred to a general principle of equity or fairness.

Penalty salience \times punishment probability. The messages in the penalty salience treatment status and the punishment probability treatment status were combined in a separate treatment group. The boxed message stated the taxpayer's income first, then the penalty.

¹⁸The individual income tax form for nonresidents, Form D-1040(NR), was sent to some nonfilers in the experiment in place of the tax form for residents.

Standard theory about the decision to file suggests that the interaction between penalty salience and punishment probability should be important. If the other treatments are effective and operate through the intended channel of raising the perceived penalty and probability of punishment, then we would expect the interaction treatment to elicit a higher response than either by itself.

Control. Two groups of nonfilers were assigned to “control” groups. One group received no contact at all, and the other group was sent mailings with the prominent box omitted from both the postcard and the letter. There is considerable evidence that taxpayers respond to any kind of contact from the tax authority, probably because it alerts the taxpayer that the tax authority can monitor their behavior, so it is important to isolate the effect of the contact-only mailings from the effect of the particular messages in the other treatment groups.¹⁹ For the purpose of measuring responses, individuals in the no-contact group were assigned to batches as if they were being sent postcards and letters.

From the population of 42,704 nonfilers that met the sample selection criteria, 1,200 individuals were randomly selected for each of the 6 treatment groups that received letters (including the contact-only control group), and 2,400 individuals were randomly selected to be in a no-contact control group. To stay within the limits of the Detroit tax division’s administrative capacity, the postcards and letters were sent in staggered batches.²⁰ Each batch had an approximately equal number of nonfilers from each treatment group. There were 119, 581, 2,160, 2,160, and 2,160 individuals in batches one through five, respectively.²¹ The treatment groups are not exactly the same size because the city’s address filter was refined shortly before sending the second batch. Also, individuals were removed from the

¹⁹Chirico et al. (2015) and Fellner et al. (2013) describe field experiments that used similar contact-only letters to isolate the response to particular messages from the response to contact from the tax authority.

²⁰The tax division reports that it was unable to handle the phone calls that resulted from large batches (tens of thousands) of similar letters to nonfilers in past years. That likely dampened response rates and the effectiveness of contact. Therefore, in this field experiment, postcards and letters were dispersed in batches.

²¹Postcards were sent on April 18, May 2, May 16, June 1, and June 13-15. Letters were sent on April 25, May 9, May 24-26, June 9, and June 23.

sample if they filed a city tax return between the time the sample was selected and the time the postcards were mailed. Individuals removed from the sample were replaced with other individuals randomly selected from the population of nonfilers whenever possible. Appendix Table A.3 reports summary statistics by treatment status.

This field experiment was accompanied by another change that may have affected response rates. The State of Michigan took responsibility for processing individual City of Detroit income tax returns for tax year 2015. The state did not take any responsibility for past returns, so there is no direct impact on tax year 2014 returns. The shift to processing tax returns by the state was not directly related to this experiment, although both were motivated by a desire by city administrators to improve the efficiency of tax enforcement. There is no reason to think that nonfilers in one treatment status had a different level of exposure to this change than nonfilers with a different treatment status.

4 Results

Table 5 summarizes the response of nonfilers to mailings in the field experiment. Of the 7,142 taxpayers in the sample to whom mailings were sent, letters were delivered to 3,950 taxpayers (55.3%), and 420 taxpayers (5.9%) responded by filing a return within 75 days of the initial mailing. Even though the mailings only mentioned tax year 2014 specifically, many taxpayers filed returns for multiple years, such that the number of returns per filer was 1.15.²² The sum of refunds claimed by taxpayers who received mailings was \$13,109, the sum of tax due admitted was \$91,642, and the sum of payments remitted was \$42,712.

The overall response rate to sample mailings was 6.3%. Inclusion in the sample was conditional on the city estimating tax due above \$350, but 34% of returns nevertheless claimed a refund.²³ Of returns claiming refunds, the average refund size was \$75. About

²²When a taxpayer calls or visits the tax division, staff instruct the taxpayer to file returns for all missing years.

²³The most common discrepancy between estimated and actual tax due is withholding that Detroit did

Table 5: Summary of response

	Contact only	Penalty salience	Punishment probability	Compliance cost	Civic pride	Penalty x Punishment	All letters	No contact
Sample size	1,185	1,191	1,191	1,189	1,195	1,191	7,142	2,381
Letters delivered	661	670	658	621	643	697	3,950	0
Filers	36	120	58	74	46	116	450	7
Returns filed	39	153	69	83	50	129	523	7
Claiming refunds	16	44	19	34	21	41	175	5
Admitting tax due	15	80	37	30	18	62	242	2
Remitting payment	10	44	19	16	10	36	135	1
Total claimed (\$)	758	3,092	834	3,367	1,276	3,782	13,109	297
Total admitted (\$)	6,183	33,413	11,494	9,388	11,804	19,360	91,642	1,720
Total remitted (\$)	5,046	17,237	4,353	2,157	4,278	9,641	42,712	1,690
Delivered % of sample	55.8	56.3	55.2	52.2	53.8	58.5	55.3	0.0
Filed % of sample	3.0	10.1	4.9	6.2	3.8	9.7	6.3	0.3
Filed % of delivered	5.4	16.1	8.1	11.3	5.9	14.1	10.2	.
Filed % of nondelivered	0.0	2.3	0.9	0.7	1.4	3.6	1.5	0.3
Returns per filer	1.08	1.27	1.19	1.12	1.09	1.11	1.16	1.00
Refunds % of returns	41.0	28.8	27.5	41.0	42.0	31.8	33.5	71.4
Tax due % of returns	38.5	52.3	53.6	36.1	36.0	48.1	46.3	28.6
Payment % of tax due	66.7	55.0	51.4	53.3	55.6	58.1	55.8	50.0
Avg refund claimed (\$)	47.38	70.27	43.89	99.03	60.76	92.24	74.91	59.40
Avg due (\$)	412.20	417.66	310.65	312.92	655.78	312.26	378.68	860.00
Avg remittance (\$)	504.60	391.75	229.11	134.81	427.80	267.81	316.39	1690.00
Claim per letter (\$)	0.64	2.60	0.70	2.83	1.07	3.18	1.84	0.12
Due per letter (\$)	5.22	28.05	9.65	7.90	9.88	16.26	12.83	0.72
Remit per letter (\$)	4.26	14.47	3.65	1.81	3.58	8.09	5.98	0.71

Note: This table reports summary statistics for responses within 75 days of sending the postcard. It includes information from returns received through the Income Tax Division's post office box and returns processed by Chase Bank. Initially, 1,200 taxpayers were selected to be sent each of the treatment mailings. A few taxpayers were removed without being replaced because the city refined its address validity criteria, and a few taxpayers were removed without being replaced because they filed a tax return shortly before the postcard would have been sent.

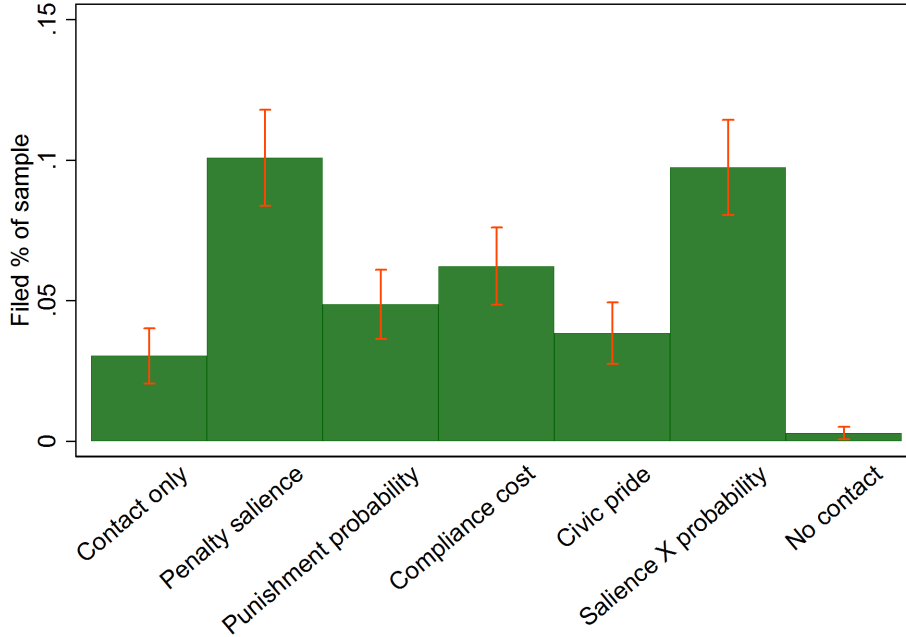
half of the returns that were filed admitted tax due, and on returns that admitted tax due the average due was \$379. Taxpayers are instructed to remit payment along with the return, but only 56% of returns that admitted tax due were accompanied by a remittance. The average remittance was \$316.

4.1 Response rates

Figure 1 shows response rates by experimental treatment. The penalty salience mailing elicited the highest response rate (10.1%), followed by penalty salience \times punishment prob-

not know about. However, many individuals claimed a refund without enclosing a W2 to prove withholding, and without a W2 the city does not issue a refund.

Figure 1: Response rates by treatment status



Note: This figure shows response rates by treatment status, where a response is filing a return within 75 days of the first mailing.

ability (9.7%), compliance cost (6.2%), punishment probability (4.9%), civic pride (3.8%), and contact-only (3.0%) mailings. The individuals in the no-contact control group, of course, did not receive a letter, and the “response” rate of filers as a percent of the no-contact sample was 0.3%. Each individual in the no-contact control was assigned to a batch of outgoing postcards, so a return from a no-contact individual is considered to be a response if it is received any time after the postcards were sent to that batch, just as if the individual had been sent a postcard.

Table 6 summarizes the results of linear probability model treatment effects using the following ordinary least squares regression specification:

$$\mathbb{P}[\text{response}_i = 1] = \alpha + \sum \beta^j \text{treatment}_i^j + \gamma X_i + e_i \quad (1)$$

where indicator variables denoting treatment status j ($treatment_i^j$) predict individual i 's binary response $response_i$, equal to one if the taxpayer filed a return and zero otherwise.²⁴ A vector of individual characteristics X_i includes income, the presence of nonwage income, filing status, batch, and filing history. Treatment effects are estimated relative to the excluded “no-contact” control condition, in which taxpayers were not sent any mailings.

Column 3 of Table 6 reports the response rate to delivered letters with the full set of controls (Local Average Treatment Effect). Some letters were not delivered because the address was invalid, so we should not expect those taxpayers to have received the letter or the treatment message. The response rate to delivered letters is therefore higher than the response rate to all letters. A delivered penalty salience letter raised response rates by 14.2 percentage points relative to the no-contact control, about three times the effect of the contact-only letter, which raised response rates by 4.7 percentage points. A delivered penalty salience \times punishment probability letter raised response rates by 13.0 percentage points, the compliance cost letter by 9.7 percentage points, the punishment probability letter by 6.9 percentage points, and the civic pride letter by 5.1 percentage points.

The coefficients on the control variables for filing history and income are significant. Filing history is a measure of how many times Detroit has identified the taxpayer as a nonfiler using the process of comparing federal returns to city returns. The filing history variables included in the regression are dummy indicators for (1) whether the taxpayer was identified using this method for the first time in tax year 2014 and (2) whether the taxpayer was identified using this method in 1-3 prior years. The omitted dummy indicator is for whether the taxpayer was identified using this method in 4 or more prior years. Relative to taxpayers who had been identified many times before, taxpayers who were identified in tax year 2014 for the first time were 3.5 percentage points more likely to respond to delivered letters.

²⁴Dependent variables are scaled by a factor of 100 so that coefficients can be read in percentage points.

Table 6: Response by experimental intervention, linear probability model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Filed	Filed	Filed	Filed	Delivered	Net due	Call/filed
Treatments							
Penalty salience	15.83*** (1.05)	15.31*** (1.03)	15.30*** (1.03)	9.74*** (0.74)	0.28 (2.03)	35.30*** (6.29)	18.96*** (1.14)
Punishment probability	7.76*** (1.05)	7.70*** (1.04)	7.65*** (1.04)	4.70*** (0.74)	-0.58 (2.03)	12.42* (6.34)	9.01*** (1.15)
Compliance cost	10.98*** (1.08)	10.72*** (1.06)	10.74*** (1.06)	5.88*** (0.74)	-3.60* (2.03)	8.48 (6.48)	12.31*** (1.17)
Civic pride	5.62*** (1.06)	5.23*** (1.05)	5.20*** (1.05)	3.54*** (0.74)	-2.13 (2.03)	12.56** (6.40)	7.32*** (1.16)
Penalty X punishment	13.77*** (1.03)	13.53*** (1.02)	13.57*** (1.02)	9.43*** (0.74)	2.65 (2.03)	15.68** (6.20)	18.46*** (1.12)
Contact only	5.15*** (1.05)	4.82*** (1.04)	4.80*** (1.04)	2.77*** (0.74)		7.65 (6.33)	7.00*** (1.14)
Other variables							
Filing jointly		3.31 (2.62)	2.21 (2.63)	2.69 (1.90)	14.07*** (5.36)	4.13 (16.03)	3.77 (2.90)
Nonfiler in TY2014 only		5.32*** (0.92)	5.20*** (0.92)	3.91*** (0.66)	1.84 (1.82)	19.78*** (5.61)	6.99*** (1.01)
Log income			2.71*** (0.68)	2.31*** (0.50)	-1.57 (1.39)	29.16*** (4.17)	2.75*** (0.75)
Nonwage income indicator			1.29* (0.67)	1.05** (0.49)	3.76*** (1.33)	-2.20 (4.09)	1.48** (0.74)
Constant	0.29 (0.49)	-11.05*** (3.63)	-38.02*** (7.71)	-32.51*** (5.43)	54.28*** (15.64)	-295.04*** (45.24)	-42.71*** (8.50)
Mean of dep var	6.48	6.49	6.49	4.81	55.31	10.55	8.18
p-val on pen=penxpun	0.11	0.16	0.18	0.71	0.24	0.01	0.72
p-val on pride=contact	0.73	0.76	0.76	0.37		0.54	0.83
R ²	0.05	0.08	0.09	0.06	0.01	0.02	0.10
Filing status dummies		X	X	X	X	X	X
Batch fixed effects		X	X	X	X	X	X
Observations	6,331	6,322	6,322	9,508	7,129	6,321	6,322
Delivered letters	X	X	X	X	X	X	X
Nondelivered letters				X	X		
No-contact control	X	X	X	X		X	X

Standard errors in parentheses, * p < 0.10, ** p < 0.05, *** p < 0.01

Note: This table estimates the response of nonfilers to experimental treatments using ordinary least squares regressions. The sample in all columns includes nonfilers to whom a letter was delivered. Columns 4 and 5 also include nonfilers to whom a letter was sent but not delivered. All columns except column 5 also include a no-contact control group. The dependent variable in columns 1 through 4 is a dummy indicator (scaled by 100) for whether the nonfiler filed a return. The dependent variable in column 5 is a dummy indicator for whether a letter was delivered. The dependent variable in column 6 is net due, which is equal to the sum of admitted tax due less the sum of refund claimed for all returns from a filer. The dependent variable in column 7 is a dummy indicator for whether the nonfiler called, visited the tax division in person, or filed a return.

Income was positively associated with response rates. For each point of log income, taxpayers were 4.2 percentage points more likely to respond. Taxpayers with nonzero nonwage income were 1.9 percentage points more likely to respond.

Column 4 of Table 6 reports the response rate to sent letters, many of which never reached the intended recipient (Intent to Treat).²⁵ If there were no selection into the treatment of a delivered letter and no effect from a nondelivered letter, then we would expect the estimates in column 4 to be smaller but proportional to the estimates in column 3. The basic patterns—the relative effects of the experimental treatments and the significance of the controls—are still present, but the estimates are not proportional. Penalty salience letters were more likely to be delivered, and taxpayers were more likely to respond to penalty salience mailings even if the letter was never delivered. Recall that the message in the letter was repeated from the postcard, so taxpayers may have filed a return on the basis of the message in the postcard even if they did not authorize delivery of the letter.

4.2 Response quality

The quality of responses also varied by experimental treatment. This section considers three dimensions of response quality: number of returns per filer, the propensity of the return to claim a refund or admit tax due, and the dollar amount of net tax due.

Table 5 shows that the penalty salience mailing elicited 1.29 returns per filer, whereas the other mailings elicited just 1.03 to 1.13 returns per filer. Nonfilers who called or visited the tax division were instructed by staff to file all delinquent returns including for tax years other than 2014.

Some taxpayers who did not file a return within 75 days were recorded in the tax division’s call and visit log. Taxpayers in the penalty salience and penalty salience \times punishment

²⁵Whether the letter was delivered depends on the validity of the address (many letters were returned by the USPS because they had invalid addresses) and a choice by the recipient whether to sign for delivery. The delivery rate could and indeed does vary by treatment, an issue I return to in Section 4.3.

probability treatments were more likely to file returns, and they were also more likely to call or visit the tax division without ever filing a return.

The penalty salience and penalty salience \times punishment probability mailings elicited high response rates of all kinds—returns with refunds, tax due, and remittances. The punishment probability mailings elicited a lower response rate than the penalty salience mailings, but the proportion of returns with refund and tax due was about the same. That was not the case for the other mailings. Contact-only, compliance cost and civic pride mailings elicited returns that claimed refunds relatively more frequently than they admitted tax due. The response rate to penalty salience mailings was only slightly better than the response rate to penalty salience \times punishment probability mailings, but it elicited 50% more tax due and remittance by dollar value.

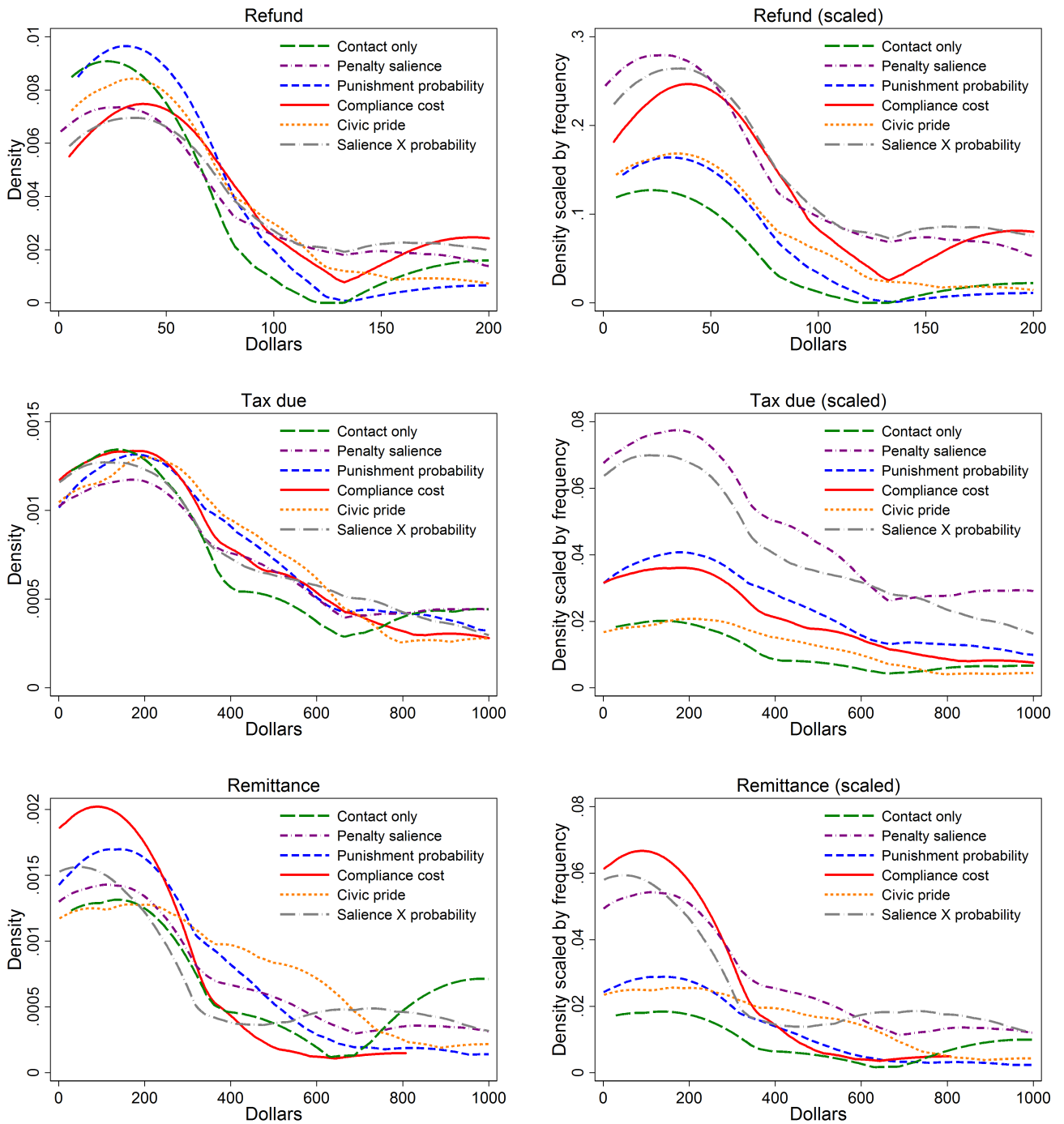
Because the totals and averages of refund claimed, tax due admitted, and payment remitted from the six experimental treatments are sensitive to extreme observations, Figure 2 shows the distribution of refund, tax due, and remittance with truncated distributions. One striking feature of this figure is that the scaled compliance cost density is very similar to the corresponding penalty salience and penalty salience \times punishment probability densities in terms of refund claimed, but it is far below them in terms of tax due admitted. For small remittances, the compliance cost sample was also similar to the penalty salience sample, but the compliance cost mailings did not elicit any large remittances.

4.3 Response to postcard

Although the purpose of the postcard was to raise response rates and deepen treatment through repetition, the postcards also influenced taxpayer behavior through other channels.

Figure 3 shows that some taxpayers filed a return after receiving the postcard and before receiving the letter. Penalty salience, penalty salience \times punishment probability, and civic pride all had similar distributions of response time relative to delivery date (left panel),

Figure 2: Distribution of refund, tax due, and remittance by treatment status

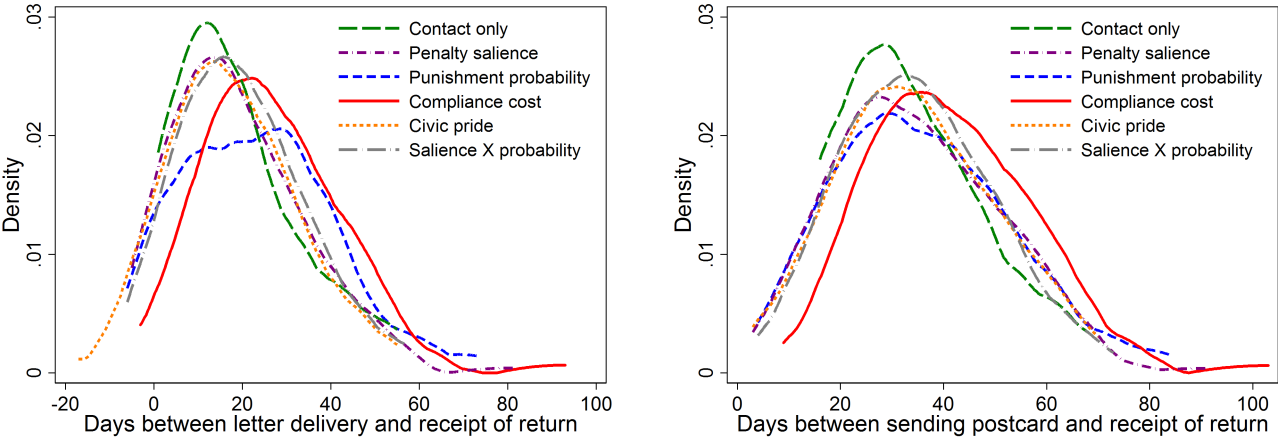


Note: This figure plots kernel density estimates for refund, tax due, and remittance. The graphs in the second column are scaled by frequency of response.

including some responses before the letter was delivered. The compliance cost distribution has a similar shape but is shifted to the right. Such a shift would be consistent with taxpayers who read the message on the postcard, which promised specifically that a tax form would be enclosed with the letter, and chose to wait until the letter arrived. The punishment probability postcard also made a specific promise, that the letter would report the taxpayer’s total income. The punishment probability distribution appears flattened and somewhat shifted to the right relative to the penalty salience and civic pride distributions. Recipients may have chosen to spend a bit more time verifying the information in the letter before responding to it.

The right panel of Figure 3 shows that the contact-only distribution has no mass until two weeks after the postcard was sent. This truncation suggests people may not have read or seen the postcard, because the only responses were after the letter arrived. That was the only treatment without a boxed message on the postcard and letter, so it is possible that the box itself, regardless of the content, attracted attention.

Figure 3: Response time by treatment status



Note: This figure shows a smoothed kernel density estimate of response time by treatment status. Response time is measured from some starting point until the date on which the city records a return as received, which is based on the postmark of the envelope. The starting point in the left panel is the delivery date of the letter according to USPS. The starting point in the right panel is the date on which postcards were sent.

Some taxpayers declined to receive the letter. The full-sample delivery rate was 55%, and an additional 26% of letters were unclaimed, meaning that the postal service believed it had a valid address but was unable to obtain a signature authorizing delivery after repeated attempts. Taxpayers who read the postcard and knew what was coming in the letter may have chosen whether to accept the letter, and indeed there are differences by treatment status that correspond roughly to response rates to delivered mail. Table 7 shows that the penalty salience and penalty salience \times punishment probability treatments, which had the highest response rates to delivered mail, also had the highest delivery rates. Although this is suggestive, column 5 of Table 6 shows that the only treatment letter with a delivery rate statistically different from the contact-only control at the 10% level was the compliance cost letter.

Table 7: Nonfiler letter delivery rates

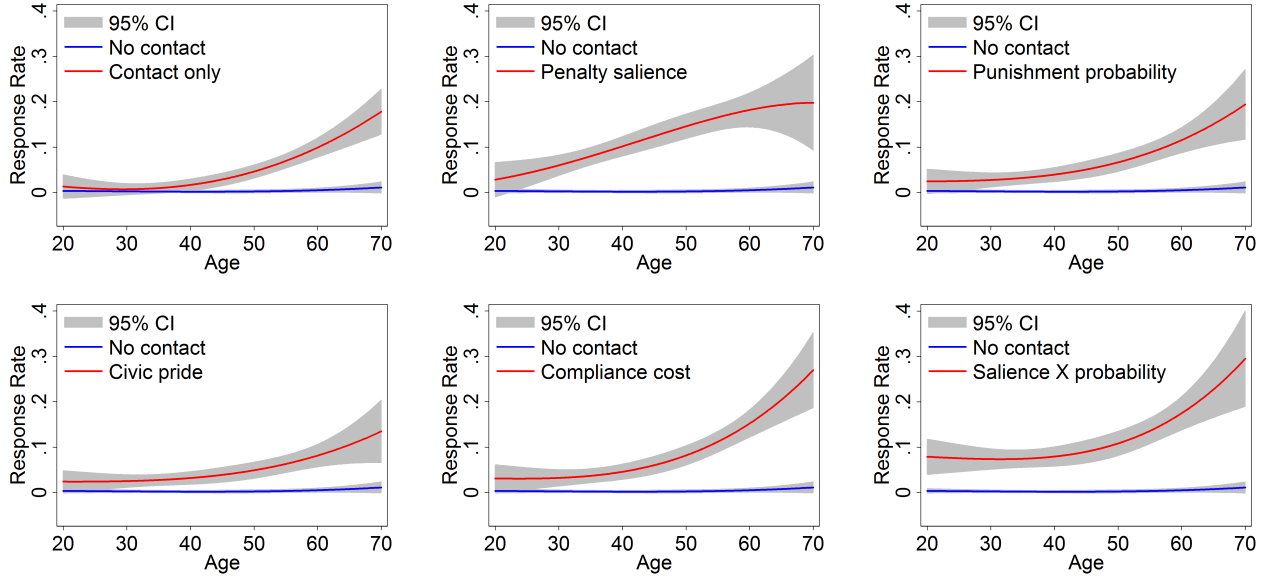
Treatment status	Delivery status						Total	
	Delivered		Unclaimed		Undeliverable			
Contact only	661	60.4%	285	26.1%	148	13.5%	1,094	100.0%
Penalty salience	670	61.0%	292	26.6%	136	12.4%	1,098	100.0%
Punishment probability	658	58.6%	333	29.7%	132	11.8%	1,123	100.0%
Compliance cost	621	56.5%	311	28.3%	167	15.2%	1,099	100.0%
Civic pride	643	57.8%	341	30.7%	128	11.5%	1,112	100.0%
Penalty X punishment	697	64.2%	279	25.7%	110	10.1%	1,086	100.0%
Total	3,950	59.7%	1,841	27.8%	821	12.4%	6,612	100.0%

Note: This table reports the delivery status of certified letters sent to nonfilers, according to the USPS tracking website. Relatively few penalty salience letters go unclaimed. Totals may not sum to number of letters sent if the USPS tracking website did not report the final status of a letter.

4.4 Heterogeneity of response

Older taxpayers responded to experimental mailings at higher rates than younger taxpayers. This was true across all treatments, and the gap was larger for the more effective mailings. Figure 4 plots response rates by age for each of the mailing treatments relative to the no-contact control group. Taxpayers under age 40 had a response rate below 10% for the penalty and penalty salience \times punishment probability treatments and below 5% for the other treatments. The plot of response rates looks convex in age, such that response rates

Figure 4: Response rate by age, treatments relative to no contact



Note: This figure estimates response rate by age using a fractional polynomial regression of response on age.

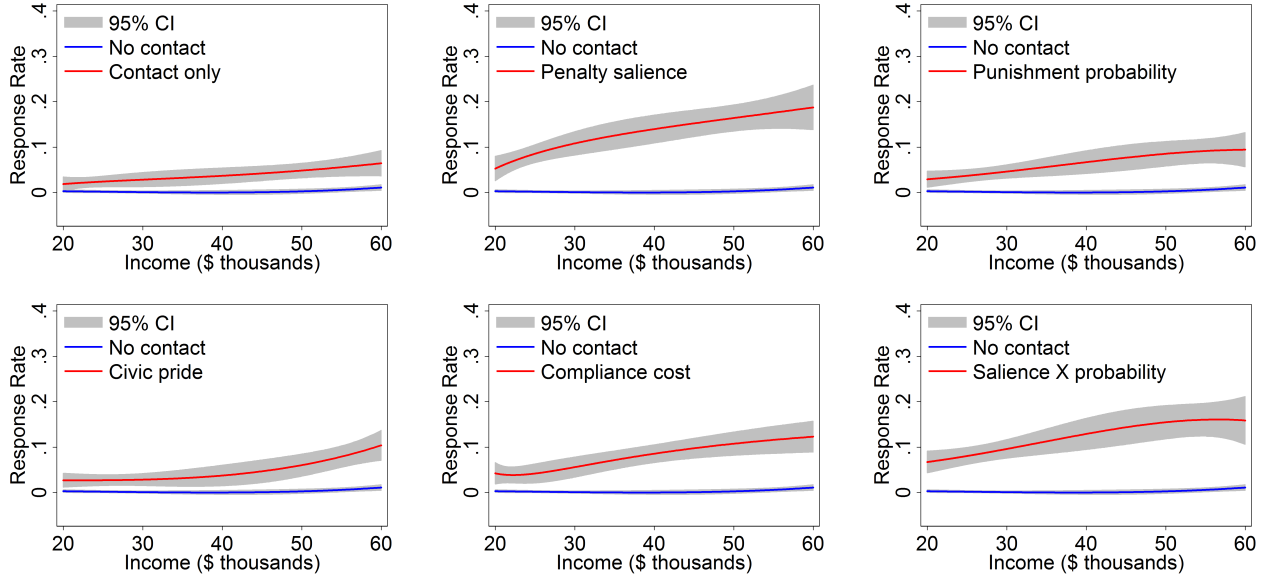
increase from age 40 to 50 and increase by even more from age 50 to age 60. By age 70, more than 20% of mailings elicit a return.²⁶

Income is highly correlated with age, so it is not surprising that response rates are higher for taxpayers with higher incomes. This is again true across treatment groups but more pronounced in the more effective treatments. Figure 5 plots response rates by income for each of the mailing treatments relative to the no-contact control group. The penalty and penalty salience \times punishment probability treatments elicited higher response rates even from taxpayers earning less than \$30K, whereas most of the gains from the compliance cost and punishment treatments came among taxpayers earning more than \$40K, and the civic pride treatment only raised response rates considerably above \$50K. The contact-only letter was not much more effective with higher-income taxpayers than with lower-income taxpayers.

Taxpayers who were identified as nonfilers from federal returns in more years were less likely to respond to experimental mailings. Table 8 shows response rates were higher for

²⁶Pension income is not taxable to Detroit. Some taxpayers over age 65 are pensioners, but others are among the highest active earners.

Figure 5: Response rate by income, treatments relative to no contact



Note: This figure estimates response rate by age using a fractional polynomial regression of response on age.

nonfilers identified for the first or second time than for other nonfilers in all large age-income bins. The response rates are less precisely measured in smaller bins, but one lesson from this table is that the effects of age, income, filing history, and treatment status appear to be positive even when they are all at play. The highest response rate at 39% was from taxpayers over age 50 with more than \$50K income who received the penalty salience mailing. However, the penalty salience message is not the most effective for all subgroups. Taxpayers under age 30 with less than \$30K income who had only been identified as nonfilers once were more responsive to the penalty salience \times punishment probability treatment.

4.5 Network effects

This section investigates behavioral responses of untreated taxpayers to the experimental mailings. The mailings could have influenced the behavior of untreated taxpayers if, for example, recipients of these mailings told their neighbors, relatives, or coworkers that they had been contacted by the Detroit Income Tax Division. Even a small effect per neighbor

Table 8: Response rate by age-income-filing history bins

Years nonfiler	Treatment status						Total
	Contact	Penalty	Punishment	Compliance	Civic pride	Pen X pun	
age <= 30, inc <= 30K							
1-2 years (N = 1036)	0.8%	5.5%	2.2%	2.3%	3.6%	9.0%	3.8%
3-5 years (N = 770)	0.0%	0.0%	4.1%	4.0%	0.0%	6.3%	2.4%
6-9 years (N = 231)	0.0%	2.7%	0.0%	9.1%	0.0%	0.0%	2.2%
Total	0.4%	3.1%	2.6%	3.8%	1.9%	6.8%	3.1%
age <= 30, 30K < inc <= 50K							
1-2 years (N = 214)	3.7%	14.8%	0.0%	3.7%	0.0%	12.5%	5.5%
3-5 years (N = 145)	4.3%	11.1%	0.0%	0.0%	0.0%	0.0%	2.8%
Total	4.0%	13.3%	0.0%	2.8%	0.0%	8.1%	4.4%
30 < age <= 50, 30K < inc <= 50K							
1-2 years (N = 311)	0.0%	18.2%	4.8%	10.9%	5.9%	17.4%	10.0%
3-5 years (N = 477)	0.0%	13.8%	3.1%	13.0%	2.6%	8.3%	6.5%
6-9 years (N = 590)	2.8%	6.8%	3.8%	2.7%	0.0%	4.5%	3.5%
Total	1.3%	11.5%	3.8%	8.0%	2.3%	9.3%	6.0%
30 < age <= 50, 50K < inc							
1-2 years (N = 221)	14.3%	14.7%	16.7%	8.3%	20.0%	10.0%	13.9%
3-5 years (N = 264)	0.0%	12.8%	8.0%	10.0%	6.5%	13.8%	8.6%
6-9 years (N = 221)	3.7%	0.0%	0.0%	15.6%	0.0%	3.1%	4.1%
Total	5.7%	9.8%	8.4%	11.2%	9.1%	8.6%	8.9%
50 < age, 30K < inc <= 50K							
1-2 years (N = 155)	8.3%	27.3%	18.8%	20.7%	11.8%	29.4%	20.4%
3-5 years (N = 209)	16.0%	23.1%	15.0%	0.0%	12.0%	22.2%	14.8%
6-9 years (N = 210)	4.5%	13.8%	7.4%	12.1%	0.0%	14.3%	9.5%
Total	10.2%	20.8%	12.7%	11.4%	8.2%	20.8%	14.3%
50 < age, 50K < inc							
1-2 years (N = 165)	19.0%	39.1%	17.6%	23.5%	30.0%	32.0%	27.6%
3-5 years (N = 174)	7.7%	25.9%	26.3%	8.3%	0.0%	23.5%	15.2%
6-9 years (N = 179)	0.0%	33.3%	0.0%	10.0%	8.7%	22.2%	10.7%
Total	8.7%	32.3%	12.7%	14.3%	11.9%	26.1%	17.7%

Note: This table shows response rate by age, income, and filing history. Only age-income-history cells with at least 100 observations are shown.

can add up to a substantial impact when one treated taxpayer has many neighbors. In other enforcement contexts, network effects like this appear to be important.²⁷

The question examined here is whether geographic neighbors of treated nonfilers filed tax returns in response to the experimental mailings. For each treated nonfiler, including the no-contact control group, I calculate the number of untreated neighbors within 100 meters who filed a return within 90 days of the date the first experimental mailing was sent. I geocoded the addresses of all treated nonfilers and all untreated taxpayers who filed a return

²⁷Drago et al. (2015) find that, when a sample of potential evaders of TV license fees were sent a letter, their untreated neighbors who did not receive a letter were more likely to comply with the fee. Boning et al. (2015) also present evidence on network effects in tax enforcement.

during the relevant time period, then computed the distance between every possible pair of treated nonfiler and untreated taxpayer. For each treated nonfiler, that gives a count of untreated taxpayers who filed during the relevant time period. I then regress the count of untreated taxpayers on treatment dummies.

I fail to find evidence of an effect of treatment mailings on the filing behavior of geographic neighbors. Table 9 shows that estimated coefficients on treatment dummies were not statistically different from zero. I repeated the procedure for a variety of distances, including 25, 50, 100, 200, 500, and 1000 meters.

Table 9: Untreated neighbor responses to treatment

	(1)	(2)	(3)	(4)	(5)	(6)
	<25m	<50m	<100m	<200m	<500m	<1km
Contact only	0.002 (0.021)	-0.002 (0.024)	-0.001 (0.029)	-0.055 (0.045)	-0.069 (0.095)	0.034 (0.254)
Penalty salience	-0.023 (0.019)	-0.034 (0.021)	-0.030 (0.026)	-0.071 (0.053)	-0.038 (0.107)	0.028 (0.219)
Punishment probability	-0.020 (0.016)	-0.026 (0.019)	-0.002 (0.026)	-0.055 (0.043)	-0.211** (0.102)	-0.268 (0.222)
Compliance cost	0.006 (0.016)	0.001 (0.019)	-0.003 (0.024)	-0.051 (0.043)	0.095 (0.103)	0.327 (0.284)
Civic pride	0.002 (0.018)	0.003 (0.020)	-0.003 (0.028)	-0.103** (0.047)	0.010 (0.118)	0.696 (0.510)
Penalty X punishment	-0.021 (0.014)	-0.029* (0.017)	-0.039 (0.025)	-0.126*** (0.047)	-0.097 (0.098)	-0.015 (0.256)
vconst	0.095*** (0.018)	0.142*** (0.020)	0.277*** (0.024)	0.839*** (0.050)	4.250*** (0.267)	14.717*** (1.343)
R^2	0.03	0.06	0.12	0.28	0.60	0.75
Observations	9,274	9,274	9,274	9,274	9,274	9,274

Standard errors in parentheses

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

Note: This table reports results from regressing the number of untreated neighbors who filed a return from an address within x meters of an individual in the sample on the treatment dummies. Standard errors are corrected for spatial correlation as described by [Conley \(1999\)](#) and implemented by [Hsiang \(2010\)](#).

5 Normative analysis

The frequency and quality of responses to the mailings in this field experiment can be understood as a combination of (1) how taxpayers respond to mailings and (2) how well taxpayers were targeted in the sample selection criteria. With better targeting mechanisms, the tax division could improve the yield per letter even relative to the most successful treatment condition. With that in mind, I undertake a cost-benefit and welfare analysis conditional on the exact selection criteria in the field experiment, then discuss targeting criteria, then rationalize and generalize the lessons from the field experiment.

5.1 Cost-benefit and welfare analysis

A tax authority that aims to maximize net revenue should send mailings to individuals for whom the expected marginal revenue exceeds the expected marginal cost. The marginal cost of mailings per nonfiler is estimated to be \$4.70. Appendix Table A.4 shows the components of the marginal cost of mailings, including materials, time, and postage. Postage for certification was more than two thirds of the marginal cost of mailings per nonfiler.²⁸

The revenue net of administrative costs per nonfiler is estimated in Table 10 separately for each experimental treatment. Remittances are a large and important part of net collection, but not the only part. They are offset by refunds, which are issued when withholding exceeds tax liability. Also, some tax debt which is not remitted with a tax return will eventually be recovered as a result of these mailings. Overall, net collections are equal to remittance plus recovered tax debt minus refunds issued. The administrative cost of processing returns is assumed to be one hour per taxpayer who files a return, with time valued at the salary of a

²⁸The revenue-maximizing condition is:

$$\gamma E[\tau Y_i - W_i - A] > MC \quad (2)$$

where γ is the response rate to a mailing, Y_i is income, W_i is withholding, A is the administrative cost of processing a response, and MC is the marginal cost of mailings per nonfiler. The left hand side of the equation is marginal revenue per mailing, and the right hand side of the equation is marginal cost per mailing.

Detroit tax examiner.²⁹

Marginal net revenue per nonfiler is positive in the penalty salience treatment and the penalty salience \times punishment probability treatment. The net revenue row of Table 10 subtracts the marginal cost of mailings per nonfiler from collections net of administrative costs. The net revenue is \$6.24 per nonfiler in the penalty salience treatment, \$0.38 per nonfiler in the penalty salience \times punishment probability treatment, and negative for the other treatments.

This net revenue analysis omits two potentially important channels. First, sampled taxpayers may comply at a higher rate in the future (specific deterrence). Second, other taxpayers may comply at a higher rate if they learn that Detroit is sending letters to nonfilers (general deterrence).

A tax authority that aims to maximize welfare should consider net revenue, the marginal social value of spending, and the private cost of compliance.³⁰ The optimal enforcement rule is to take an action (such as sending a letter) if the marginal social value of spending times net revenue exceeds the private cost of compliance. To estimate the private cost of compliance, I look at taxpayers in the sample who filed and were owed a refund. The average refund claimed was \$70. If we assume those taxpayers were rational and informed, \$70 should be a lower bound for the compliance cost of filing a return.³¹ In that case, the compliance

²⁹Staff time is valued at \$23.95 per hour, the hourly equivalent of the top annual salary of a Detroit tax examiner. *White Book, 2016-2017 Salary and Wage Adjustments*, March 2016, page 63, available at <http://www.detroitmi.gov/how-do-i/view-city-of-detroit-reports>.

³⁰Following Keen and Slemrod (2016), the welfare-maximizing condition is:

$$\nu' [MR - MC] > c \tag{3}$$

where ν' is the marginal social benefit of government spending and c is the compliance cost. The term inside the brackets is net collections per mailing, where marginal revenue MR and marginal cost MC correspond to equation 2. In this framework, a dollar in the hands of the tax authority is more valuable than a dollar in the hands of a taxpayer by a fixed factor ν' . The marginal compliance cost to taxpayers is considered by the tax authority, but the marginal cost to the tax authority MC receives a greater weight because that cost must be paid for with tax revenue.

³¹Benzarti (2015) estimates that the compliance cost to taxpayers of itemizing on a federal return is \$644, and that the cost of filing federal taxes is (an astounding) 1.25% of GDP.

Table 10: Net revenue and welfare

	Contact- only	Penalty saliency	Punishment probability	Compliance cost	Civic pride	Saliency × probability
Remit per letter [1]	4.26	14.47	3.65	1.81	3.58	8.09
Tax debt recovered [2]	0.19	2.72	1.20	1.22	1.26	1.63
Refund issued [3]	0.51	2.08	0.56	2.27	0.85	2.54
Net collection [4] = [1] + [2] - [3]	3.94	15.11	4.29	0.76	3.99	7.19
Cost of mailings [5]	4.70	4.70	4.70	4.70	4.70	4.70
Cost of processing responses [6]	0.73	2.41	1.17	1.49	0.92	2.33
Net revenue (NR) [7] = [4] - [5] - [6]	-1.49	8.00	-1.57	-5.43	-1.64	0.15
Social value of NR [8] = 1.2 × [7]	-1.79	9.60	-1.89	-6.51	-1.96	0.18
Compliance cost [9]	2.13	7.05	3.41	4.36	2.69	6.82
Net welfare [10] = [8] - [9]	-3.91	2.55	-5.30	-10.87	-4.66	-6.63

Note: All units are dollars per mailing. Refunds that are claimed are not always paid, e.g. if the taxpayer does not submit a W2, so “refund issued” is assumed to be 80% of claimed refund per letter. Similarly, admitted tax debt is not always collected, so “tax debt recovered” is assumed to be 20% of admitted due per letter that is not remitted with the return. Net collection is equal to remit per letter plus tax debt recovered minus refund issued. The cost of processing is assumed to be one hour (valued at \$23.95) per taxpayer who subsequently files. Collection net of administrative cost is equal to net collection minus cost of processing. Net revenue is equal to collection net of administrative cost minus the marginal cost of mailings per nonfiler in Table A.4 (\$4.70). Social value of spending is equal to 1.2 times net revenue. Compliance cost is assumed to be \$70 per taxpayer who files a return. Net welfare is social value minus compliance cost. This leaves out general and specific deterrence.

cost per letter in the penalty saliency treatment was \$6.58, and the net welfare per letter was \$0.90.

If the most effective treatment, the penalty saliency treatment, had been applied to the entire population of nonfilers that fit the sample selection criteria, then the city would have collected net revenue of \$266K and improved welfare by \$38K. This is inferred from a simple back-of-the envelope calculation multiplying the number of taxpayers who fit the sample selection criteria (42,704) by, first, the net revenue per letter (\$6.24) and, second, the net welfare per letter (\$0.90). However, there is potential for larger improvements in net revenue and net welfare by refining the selection criteria for which nonfilers are sent mailings.

5.2 Targeting criteria

The population examined by the field experiment included taxpayers who the city estimated owed at least \$350, without regard to age, income level, income composition, or filing history. My analysis suggests that the city could raise net revenue and net welfare by focusing on higher-yield demographics. Older taxpayers, higher-income taxpayers, and taxpayers who

have been identified fewer times as nonfilers had higher yields in the field experiment. These affects appear to operate individually, so that older taxpayers respond to mailings at higher rates at any income and with any filing history, and also to interact, such that taxpayers with all of the higher-response characteristics have particularly high response rates.

In Detroit, optimal enforcement strategy should make a distinction between wage and nonwage income. Most taxpayers in the sample who wound up claiming a refund had withholding from wage income that Detroit did not connect with the taxpayer when it estimated tax due. With this in mind, an optimal revenue-maximizing rule for contacting nonfilers should weight wage income using the probability that withholding was actually visible to the tax authority.³² Around 20% of withholding (by value) is visible to the tax authority before a city return is filed.

To illustrate the point that wage and nonwage income should be weighted differently, consider two taxpayers with total income of \$20,600 and no evidence of withholding. One has only wage income and the other has only nonwage income. With a \$600 exemption and a 2.4% tax rate, both have tax liability of \$480. However, if half of all wage earners for whom the city has no evidence of withholding actually did have exact withholding from wage income, then the expected tax due is only \$240 for the wage earner even though it is still \$480 for the nonwage earner. This issue can result in substantial refund claims for taxpayers with negative nonwage income. A taxpayer with the same total income of \$20,600, including \$40,600 wage income and negative \$20,000 nonwage income, would actually be owed a \$480 refund if tax was withheld from her wages as if her total income were \$40,600.

The city could improve targeting with access to certain services and information. With a better address filter and verification system, the city could send out fewer letters to bad

³²If wage and nonwage income are visible but withholding is sometimes invisible, the revenue-maximizing condition is:

$$\gamma[q(\tau Y_i^W - V_i) + \tau Y_i^{NW} - A] > MC \tag{4}$$

where Y_i^W is wage income, V_i is visible withholding, and Y_i^{NW} is nonwage income.

addresses and improve delivery rates. The selection of taxpayers would be better if the city could connect employer withholding to individual taxpayers more reliably.

5.3 Rationalizing and generalizing findings

For some taxpayers, the postcard was enough to elicit a response. The postcard also seems to have influenced the rate at which nonfilers claimed the letter; the messages associated with higher delivery rates were associated with higher filing rates. Repetition is important in communication with taxpayers, but some taxpayers do respond even to a simple postcard with a clear message.³³

The ghosts examined here are people who the city believes owe individual income tax to Detroit but do not file a return—*nonfilers*. Prior tax enforcement field experiments primarily examined taxpayer response to audits and communication with individual income tax *filers*, firms, and delinquents.³⁴ Fellner et al. (2013) examine the response of Austrian households, who must register and remit a fee for televisions and radios, to communication from a subsidiary of the Austrian public broadcasting company. The field experiment described by Fellner et al. (2013) is similar to the one described in this paper because it targets nonfilers—households that failed to file that registration. Guyton et al. (2016) examine the effect of reminders on filing rates of individual income tax nonfilers. Unlike Guyton et al. (2016), who focus on the frequency and volume of communication with nonfilers, this paper isolates the effect of message content.

The response of individual income tax nonfilers to increased penalty salience and punishment probability is consistent with prior evidence about other taxpayers and partially

³³Guyton et al. (2016) examine the impact of reminders on filing behavior.

³⁴For evidence on individual income tax *filers*, see Slemrod et al. (2001), Blumenthal et al. (2001), Hallsworth et al. (2014), Kleven et al. (2011), Bott et al. (2014), and Torgler (2004). For evidence on firms, see Ariel (2012), Gangl et al. (2014), Hasseldine et al. (2007), and Pomeranz (2015). For evidence on state tax delinquents, see Perez-Truglia and Troiano (2015). For evidence on property tax delinquents, see Chirico et al. (2015).

consistent with theory about nonfilers.³⁵ Detroit sent letters to suspected nonfilers in years prior to the field experiment described in this paper, but it had not enforced the statutory criminal penalty for failing to file a return for many years, so the probability of punishment was essentially zero.³⁶ The filing rate of nonfilers who received letters that revealed the city knew their federal income was 1.6 percentage points higher than the contact-only control, and the filing rate of nonfilers who received letters explaining the statutory penalty was 6.5 percentage points higher than the contact-only control. Theory predicts these two mechanisms should have a positive interaction, but the response to mailings with both messages was no more effective than the penalty salience message by itself.

Detroit is a somewhat unusual environment and might not be representative of taxpayer behavior in other contexts. When the city filed for bankruptcy in July 2013, it was the largest municipal bankruptcy in United States history. The city is still struggling to provide basic services, which may affect taxpayers' attitudes towards paying taxes and perceptions of the ability of Detroit to enforce tax law. These struggles also suggest that the marginal social value of spending may be particularly high in Detroit, such that the net welfare from undertaking enforcement efforts might be higher than estimated in the previous section. Detroit's population has low income relative to other cities, which suggests overall response rates should be even higher elsewhere for comparable messages because higher-income nonfilers responded to mailings at a higher rate.

³⁵Erard and Ho (2001) extend the theoretical framework of Allingham and Sandmo (1972) to include a decision whether to file a return, which depends on the probability of being caught and the penalty for nonfiling. Evidence of taxpayer response to changes in the probability of an audit were examined by Slemrod et al. (2001), Kleven et al. (2011), and Hasseldine et al. (2007).

³⁶Detroit has severe administrative capacity constraints, which makes developing countries a relevant comparison group. See Bird and Zolt (2005) for a discussion on the limits of individual income tax with limited administrative capacity.

6 Conclusion

A single sentence, strategically placed in mailings to attract attention, can have an economically meaningful impact on tax filing behavior. Informing taxpayers of a statutory penalty for failing to file a return elicited higher filing rates, more returns per filer, more admitted tax due, and more remittances. Even though both penalty salience and punishment probability were individually effective relative to the contact-only mailings, interacting these two treatments was no more effective, indeed less effective, than the penalty salience message by itself. This is inconsistent with the theoretical prediction that penalty salience and punishment probability should have a positive interaction. The interaction may have exhibited no improvement over penalty salience by itself because (1) the effectiveness of the penalty salience message depended on its simplicity, or (2) the penalty salience message had already exhausted the channel of affecting taxpayer behavior through perceived probability of punishment.

The results presented here add a new subgroup—nonfilers—to the evidence in the tax compliance literature that taxpayers respond more to incentive-based appeals than to normative appeals. The response rate to the civic pride treatment was not statistically different from the contact-only control group. Most studies that test normative appeals find they are not effective. Civic pride was a novel type of normative appeal, but it does not appear to be effective. Civic pride is inherently tailored to a locality, so it is possible that appeals to civic pride could work elsewhere, but results from Detroit are not promising. Perhaps people with whom a message about civic pride would succeed had already filed their tax returns.

Enclosing a blank tax form and return envelope was effective in eliciting higher response rates. The monetary cost to the taxpayer of a blank form and an envelope is trivial, so this suggests there are substantial nonmonetary costs to filing a return—e.g. the psychological toll of searching for tax form “D-1040(R)” on the city’s website.

The results presented here about a controlled experiment in Detroit relate to literature in several fields of economics. The civic pride message tested a purely behavioral response because it affected the framing of the decision of whether to file a return without affecting the expected penalty or cost of compliance. However, the other messages also tested behavioral responses. The penalty salience message, for example, tested the extent to which a particular form of communication has the ability to attract attention and influence perceived penalties, even though the actual penalties were no different among treatments.

This experiment showed that communication is a promising tool for fiscal authorities with limited capacity. Effective communication may improve enforcement not only for tax administration in the United States, but also more generally for law enforcement, especially with severe capacity constraints, which are often faced by policymakers in developing economies. The absence of evidence for geographic network effects in the context of tax enforcement on nonfilers is consistent with the claim that other networks such as families and coworkers are more relevant in urban environments.

References

- Michael G Allingham and Agnar Sandmo. Income tax evasion: A theoretical analysis. *Journal of Public Economics*, 1(3):323–338, 1972.
- Barak Ariel. Deterrence and moral persuasion effects on corporate tax compliance: Findings from a randomized controlled trial. *Criminology*, 50(1):27–69, 2012.
- Youssef Benzarti. How taxing is tax filing? Leaving money on the table because of compliance costs. Working Paper, University of California, Berkeley, 2015.
- Saurabh Bhargava and Dayanand Manoli. Why are benefits left on the table? Assessing the role of information, complexity, and stigma on take-up with an IRS field experiment. *The American Economic Review*, 2013.

- Richard M Bird and Eric M Zolt. The limited role of the personal income tax in developing countries. *Journal of Asian Economics*, 16(6):928–946, 2005.
- Marsha Blumenthal, Charles Christian, and Joel Slemrod. Do normative appeals affect tax compliance? Evidence from a controlled experiment in Minnesota. *National Tax Journal*, pages 125–138, 2001.
- Will Boning, John Guyton, Ron Hodge, Joel Slemrod, and Ugo Troiano. Direct and network effects of alternative collection enforcement treatments: Evidence from a randomized control experiment. 2015.
- Kristina Bott, Alexander W Cappelen, Erik Ø Sørensen, and Bertil Tungodden. You’ve got mail: A randomised field experiment on tax evasion. *NHH Norwegian School of Economics*, 2014.
- Michael Chirico, Robert P Inman, Charles Loeffler, John MacDonald, and Holger Sieg. An experimental evaluation of notification strategies to increase property tax compliance: Free-riding in the city of brotherly love. In *Tax Policy and the Economy, Volume 30*. University of Chicago Press, 2015.
- Timothy G Conley. Gmm estimation with cross sectional dependence. *Journal of econometrics*, 92(1):1–45, 1999.
- Francesco Drago, Friederike Mengel, and Christian Traxler. Compliance behavior in networks: Evidence from a field experiment. Discussion Paper 9443, IZA, 2015.
- Brian Erard and Chih-Chin Ho. Searching for ghosts: Who are the nonfilers and how much tax do they owe? *Journal of Public Economics*, 81(1):25–50, 2001.
- Brian Erard and Chin-Chin Ho. Mapping the U.S. tax compliance continuum. *Contributions to Economic Analysis*, 268:167–186, 2004.

- Brian Erard, Pat Langetieg, Mark Payne, and Alan Plumley. Missing returns vs. missing income: Estimating the extent of individual income tax filing noncompliance from IRS and Census data. Working Paper, B. Erard & Associates, 2014.
- Gerlinde Fellner, Rupert Sausgruber, and Christian Traxler. Testing enforcement strategies in the field: Threat, moral appeal and social information. *Journal of the European Economic Association*, 11(3):634–660, 2013.
- Katharina Gangl, Benno Torgler, Erich Kirchler, and Eva Hofmann. Effects of supervision on tax compliance: Evidence from a field experiment in Austria. *Economics Letters*, 123(3):378–382, 2014.
- John Guyton, Dayanand S Manoli, Brenda Schafer, and Michael Sebastiani. Reminders & recidivism: Evidence from tax filing & EITC participation among low-income nonfilers. Working Paper 21904, National Bureau of Economic Research, 2016.
- Michael Hallsworth, John List, Robert Metcalfe, and Ivo Vlaev. The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. Working Paper 20007, National Bureau of Economic Research, 2014.
- John Hasseldine, Peggy Hite, Simon James, and Marika Toumi. Persuasive communications: Tax compliance enforcement strategies for sole proprietors. *Contemporary Accounting Research*, 24(1):171–194, 2007.
- Jeffrey L Hoopes, Daniel H Reck, and Joel Slemrod. Taxpayer search for information: Implications for rational attention. *American Economic Journal: Economic Policy*, 7(3):177–208, 2015.
- Solomon M Hsiang. Temperatures and cyclones strongly associated with economic production in the caribbean and central america. *Proceedings of the National Academy of sciences*, 107(35):15367–15372, 2010.

- Michael Keen and Joel Slemrod. Optimal tax administration. Working Paper 22408, National Bureau of Economic Research, 2016.
- Henrik Jacobsen Kleven, Martin B Knudsen, Claus Thustrup Kreiner, Søren Pedersen, and Emmanuel Saez. Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica*, 79(3):651–692, 2011.
- Ricardo Perez-Truglia and Ugo Troiano. Shaming tax delinquents: Theory and evidence from a field experiment in the United States. Working Paper 21264, National Bureau of Economic Research, 2015.
- Dina Pomeranz. No taxation without information: Deterrence and self-enforcement in the value added tax. *The American Economic Review*, 105(8):2539–2569, 2015.
- Joel Slemrod. Tax compliance and enforcement: New research and its policy implications. Working Paper, University of Michigan, 2015.
- Joel Slemrod, Marsha Blumenthal, and Charles Christian. Taxpayer response to an increased probability of audit: Evidence from a controlled experiment in Minnesota. *Journal of Public Economics*, 79(3):455–483, 2001.
- Benno Torgler. Moral suasion: An alternative tax policy strategy? Evidence from a controlled field experiment in Switzerland. *Economics of Governance*, 5(3):235–253, 2004.

A Appendix

Figure A.1: Example postcard and letter

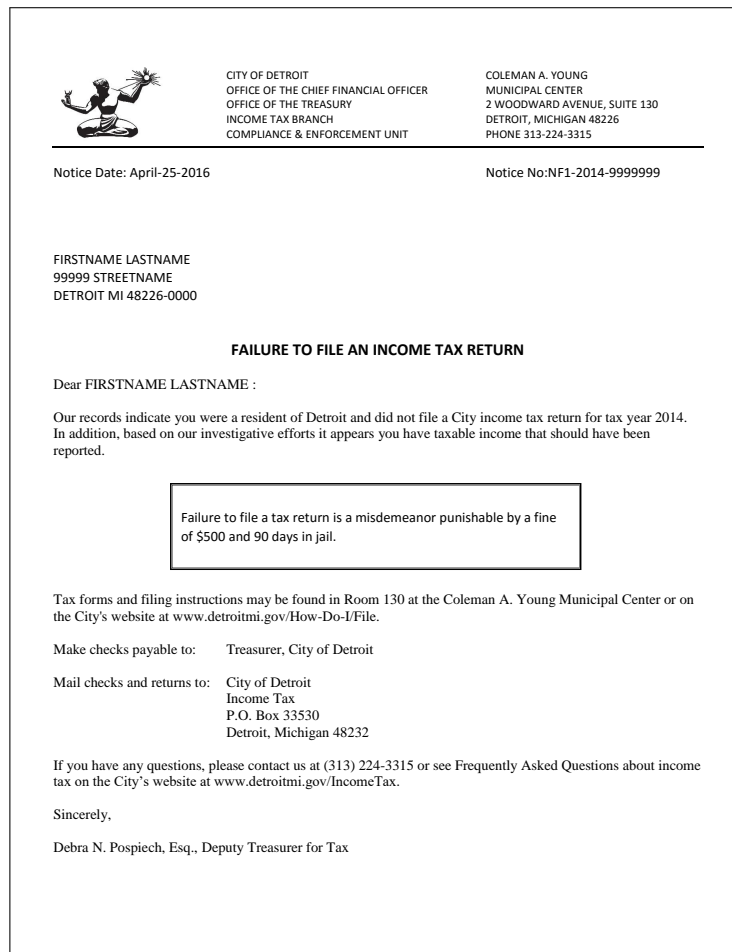
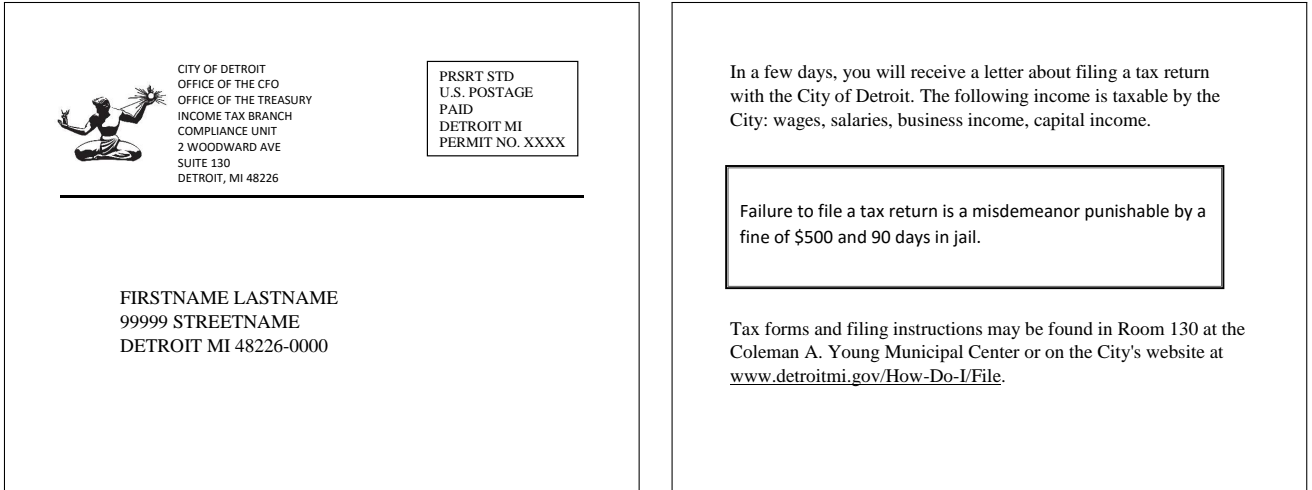
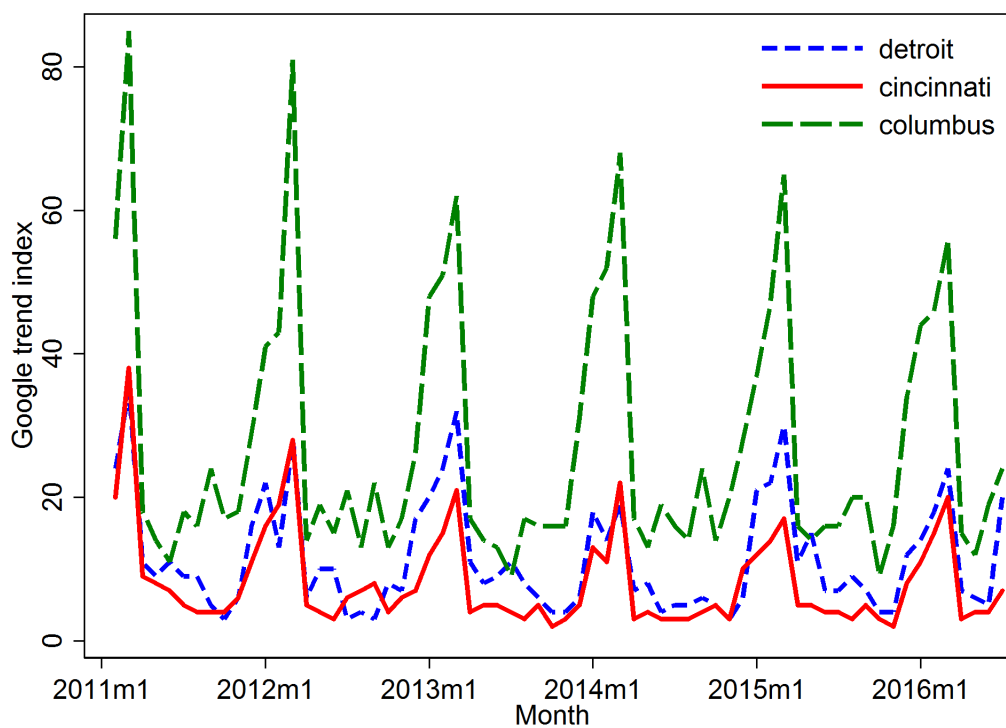


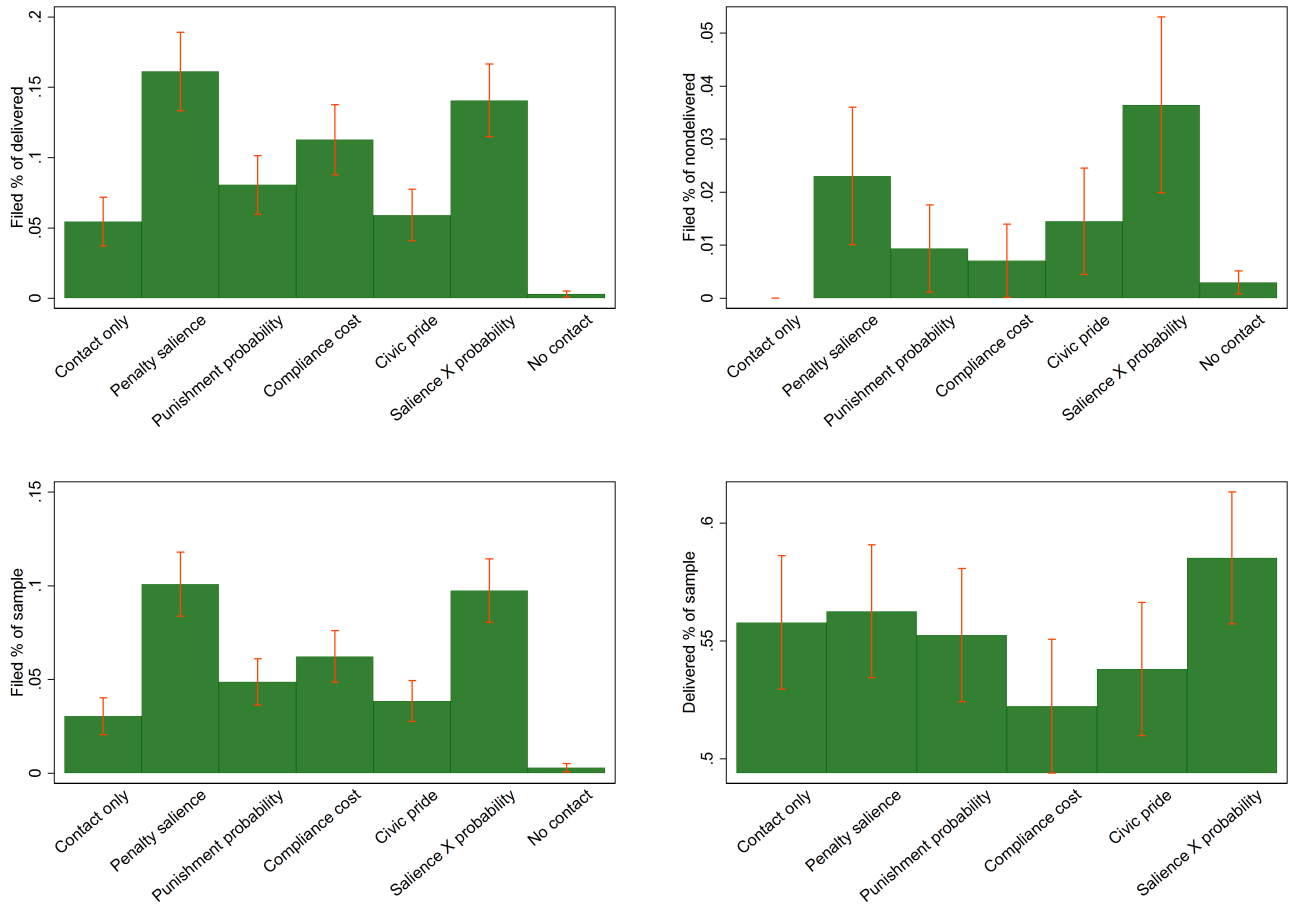
Figure A.2: Google search index for Detroit, Columbus, and Cincinnati income tax



Source: Google Trends.

Note: This figure compares search interest in “Detroit income tax” to corresponding search terms for Columbus and Cincinnati. Columbus has approximately the same population as Detroit but a much smaller metropolitan area. Cincinnati has a larger population in the city proper and about half of the population in the metropolitan area.

Figure A.3: Delivery and response rates by treatment status



Note: This figure shows response rates by treatment status.

Table A.1: States with local income tax

State	Localities	State	Localities
Alabama	4	Michigan	22
California	1	Missouri	2
Colorado	3	New Jersey	1
Delaware	1	New York	4
Indiana	91	Ohio	774
Iowa	297	Oregon	2
Kansas	535	Pennsylvania	2,961
Kentucky	218	West Virginia	3
Maryland	24		

Source: Tax Foundation

Note: The types of localities that levy income tax vary widely. In Michigan the localities that levy income tax are cities. In Maryland all 23 counties and one city, Baltimore, levy income tax. In Pennsylvania 2,492 municipalities and 469 school districts levy income tax.

Table A.2: Summary of resident filers

	Sum	Mean	25th percentile	Median	75th percentile
Wage income	2,126,720,935	32,708	7,663	24,027	45,037
Nonwage income	155,004,803	2,384	0	0	53
Liability	52,883,420	813	212	556	1,053
Visible payments	9,372,238	144	0	0	0
Invisible payments	39,461,816	607	0	343	875
Taxdue	7,573,737	116	0	0	59
Overpayments	3,524,371	54	0	5	31
Withheld (tax return)	45,636,852	898	323	684	1,186

Note: Visible payments include some withholdings and all estimated payments.

Table A.3: Balance test

	Contact	Penalty	Punishment	Compliance	Civic pride	Pen X pun	No contact
Age	39.5 (11.9)	39.6 (11.9)	39.0 (11.7)	39.8 (11.9)	39.0 (11.9)	39.9 (11.8)	39.9 (12.2)
Single (%)	40.9 (49.2)	38.6 (48.7)	40.2 (49.1)	40.8 (49.2)	40.6 (49.1)	37.9 (48.5)	37.4 (48.4)
Filing jointly (%)	9.5 (29.3)	12.0 (32.5)	10.8 (31.1)	9.9 (29.8)	11.2 (31.6)	10.7 (30.9)	10.9 (31.2)
Head of household (%)	48.6 (50.0)	48.2 (50.0)	47.4 (50.0)	47.4 (50.0)	47.0 (49.9)	50.1 (50.0)	50.0 (50.0)
Wage Inc (\$ 000s)	30.5 (20.9)	31.3 (21.1)	31.0 (20.4)	31.0 (20.0)	31.7 (29.0)	31.3 (21.4)	31.9 (27.5)
Total Inc (\$ 000s)	32.9 (22.7)	33.7 (22.8)	33.1 (21.4)	33.1 (23.2)	33.7 (28.5)	33.0 (23.3)	33.9 (27.2)
AGI (\$ 000s)	32.7 (22.7)	33.5 (22.8)	32.8 (21.3)	32.9 (23.2)	33.4 (28.1)	32.8 (23.2)	33.7 (27.1)
Taxable Inc (\$ 000s)	15.3 (20.8)	15.7 (21.1)	14.8 (17.9)	15.5 (21.2)	16.0 (26.0)	14.9 (20.9)	15.9 (25.6)
NF in TY2014 only (%)	15.4 (36.1)	15.2 (35.9)	16.0 (36.7)	16.4 (37.1)	16.4 (37.1)	14.4 (35.1)	15.0 (35.7)
priorf24	40.8 (49.2)	41.4 (49.3)	38.4 (48.7)	40.1 (49.0)	42.5 (49.5)	42.3 (49.4)	39.4 (48.9)
priorf59	43.8 (49.6)	43.4 (49.6)	45.6 (49.8)	43.5 (49.6)	41.0 (49.2)	43.3 (49.6)	45.6 (49.8)
linc	10.3 (0.5)	10.3 (0.5)	10.3 (0.5)	10.3 (0.5)	10.3 (0.5)	10.3 (0.5)	10.3 (0.5)
Observations	1,182	1,188	1,189	1,187	1,194	1,189	2,379

Note: This table reports means of observable characteristics based on administrative tax data for individuals in the sample by treatment status. Standard deviations in parentheses.

Table A.4: Marginal cost of mailings per nonfiler

	Dollars	Source / Description
Materials		
Card stock	0.017	$\frac{\$17}{250 \text{ sheets}} \times \frac{1 \text{ sheet}}{4 \text{ postcards}}$
Envelopes	0.044	$\frac{\$22}{500 \text{ envelopes}}$
Ink	0.040	pcworld.com estimate
Paper	0.014	$\frac{\$7}{500 \text{ sheets}}$
Time		
Printing	0.033	$\frac{3 \text{ hours}}{2,160 \text{ letters}} \times \frac{\$23.95}{\text{hour}}$
Stuffing	0.033	$\frac{3 \text{ hours}}{2,160 \text{ letters}} \times \frac{\$23.95}{\text{hour}}$
Certifying	0.444	$\frac{40 \text{ hours}}{2,160 \text{ letters}} \times \frac{\$23.95}{\text{hour}}$
Applying postage	0.033	$\frac{3 \text{ hours}}{2,160 \text{ letters}} \times \frac{\$23.95}{\text{hour}}$
Postage		
Postcard	0.270	USPS permit imprint
Letter	0.465	USPS metered postage
Certification	3.300	USPS metered postage
Total	4.698	

Note: Staff time is valued at \$23.95 per hour, the hourly equivalent of the top annual salary of a Detroit tax examiner. *White Book, 2016-2017 Salary and Wage Adjustments*, March 2016, page 63, available at <http://www.detroitmi.gov/how-do-i/view-city-of-detroit-reports>. Marginal cost of mailings per nonfiler was a bit higher for the compliance cost group because the compliance cost letters enclosed a blank tax form and a return envelope. Also, the stuffing machine was less likely to stuff the outgoing envelope successfully, which required staff time to correct.

Table A.5: Response by experimental intervention, probit model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Penalty salience	0.20*** (0.02)	0.19*** (0.02)	0.19*** (0.02)	0.14*** (0.01)	0.00 (0.02)	0.15*** (0.01)	0.24*** (0.02)
Punishment probability	0.15*** (0.02)	0.15*** (0.02)	0.15*** (0.02)	0.10*** (0.01)	-0.01 (0.02)	0.11*** (0.01)	0.19*** (0.02)
Compliance cost	0.17*** (0.02)	0.17*** (0.02)	0.17*** (0.02)	0.11*** (0.01)	-0.04* (0.02)	0.13*** (0.01)	0.21*** (0.02)
Civic pride	0.13*** (0.02)	0.13*** (0.02)	0.13*** (0.02)	0.09*** (0.01)	-0.02 (0.02)	0.09*** (0.01)	0.17*** (0.02)
Penalty X punishment	0.19*** (0.02)	0.18*** (0.02)	0.18*** (0.02)	0.13*** (0.01)	0.03 (0.02)	0.14*** (0.01)	0.24*** (0.02)
Contact only	0.13*** (0.02)	0.12*** (0.02)	0.12*** (0.02)	0.08*** (0.01)		0.09*** (0.01)	0.17*** (0.02)
Filing jointly		0.02 (0.02)	0.01 (0.02)	0.01 (0.02)	0.14*** (0.05)	0.04 (0.02)	0.02 (0.03)
Nonfiler in TY2014 only		0.05*** (0.01)	0.05*** (0.01)	0.03*** (0.01)	0.02 (0.02)	0.04*** (0.01)	0.06*** (0.01)
Log income			0.02*** (0.01)	0.02*** (0.00)	-0.02 (0.01)	0.02*** (0.01)	0.02*** (0.01)
Nonwage income indicator			0.01* (0.01)	0.01* (0.00)	0.04*** (0.01)	0.01 (0.01)	0.01** (0.01)
R^2							
Filing status dummies		X	X	X	X	X	X
Batch fixed effects		X	X	X	X	X	X
Observations	6,331	6,322	6,322	9,508	7,129	6,321	6,322
Delivered letters	X	X	X	X	X	X	X
Nondelivered letters				X	X		
No-contact control	X	X	X	X		X	X

Standard errors in parentheses

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

Note: This table estimates the response of nonfilers to experimental treatments using probit regressions. The sample in all columns includes nonfilers to whom a letter was delivered. Columns 5 and 6 also include nonfilers to whom a letter was sent but not delivered. All columns except column 5 also include a no-contact control group. The dependent variable in columns 1 through 5 is a dummy indicator for whether the nonfiler filed a return. The dependent variable in column 6 is a dummy indicator for whether a letter was delivered. The dependent variable in column 7 is net due, which is equal to the sum of admitted tax due less the sum of refund claimed for all returns from a filer. The dependent variable in column 8 is a dummy indicator for whether the nonfiler called, visited the tax division in person, or filed a return.

Table A.6: Response by experimental intervention, by age and income bins, linear probability model

	(1)	(2)	(3)	(4)
	Filed	Filed	Filed	Filed
Penalty salience	7.30*** (1.46)	14.94*** (2.21)	16.98*** (2.36)	22.57*** (2.18)
Punishment probability	5.86*** (1.43)	2.37 (2.14)	7.35*** (2.43)	13.32*** (2.29)
Compliance cost	4.30*** (1.49)	6.08*** (2.23)	10.32*** (2.47)	19.92*** (2.28)
Civic pride	3.29** (1.42)	6.26*** (2.26)	3.12 (2.42)	9.09*** (2.30)
Penalty X punishment	9.81*** (1.42)	13.77*** (2.24)	13.43*** (2.19)	17.59*** (2.21)
Contact only	0.37 (1.44)	2.81 (2.26)	7.03*** (2.21)	9.78*** (2.32)
Constant	0.14 (0.67)	0.64 (1.00)	0.21 (1.08)	0.27 (1.07)
R^2	0.04	0.06	0.06	0.08
Age	Younger	Younger	Older	Older
Income	Lower	Higher	Lower	Higher
N	1,863	1,194	1,264	2,010

Standard errors in parentheses

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

Note: This table estimates the response of nonfilers to experimental treatments using ordinary least squares regressions. The sample in column 1 is restricted to nonfilers to whom a letter was delivered and the no-contact control group below the median age. The sample in column 2 is above the median age, the sample in column 3 is below the median income, and the sample in column 4 is above the median income. The dependent variable is a dummy indicator (scaled by 100) for whether the nonfiler filed a return.

Table A.7: Heterogeneity of response

	Treatment status							Total
	Contact	Penalty	Punishment	Compliance	Civic pride	Pen X pun	No contact	
Years nonfiler								
1 year	3.3%	12.2%	5.8%	7.2%	7.7%	14.6%	0.3%	6.4%
2 years	4.8%	15.5%	8.3%	6.8%	7.2%	13.7%	0.0%	7.1%
3 years	4.6%	11.5%	5.5%	7.9%	2.3%	11.9%	1.3%	5.7%
4 years	1.4%	10.7%	6.5%	4.4%	3.9%	8.3%	0.4%	4.7%
5 years	1.4%	9.0%	3.5%	5.0%	2.2%	5.5%	0.4%	3.2%
6 years	3.4%	7.5%	3.6%	7.1%	1.6%	8.2%	0.0%	4.0%
7 years	1.7%	5.6%	1.7%	4.8%	2.3%	3.7%	0.0%	2.4%
8 years	4.7%	3.3%	1.8%	8.1%	1.1%	10.5%	0.0%	3.8%
9 years	0.0%	8.0%	3.4%	1.5%	0.0%	1.7%	0.0%	2.0%
Total	3.0%	10.1%	4.9%	6.2%	3.8%	9.7%	0.3%	4.8%
Age								
Age <= 20	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
20 < Age <= 30	1.0%	4.8%	2.1%	3.9%	1.8%	7.0%	0.2%	2.6%
30 < Age <= 40	0.9%	8.3%	3.9%	2.7%	3.5%	8.3%	0.5%	3.6%
40 < Age <= 50	3.5%	10.0%	4.8%	7.9%	4.0%	8.2%	0.2%	4.9%
50 < Age <= 60	5.1%	20.4%	8.6%	8.2%	6.9%	14.3%	0.0%	7.8%
60 < Age	15.9%	16.7%	15.2%	23.3%	7.8%	23.1%	1.4%	13.0%
Total	3.0%	10.1%	4.9%	6.2%	3.9%	9.8%	0.3%	4.8%
Income								
Income <= 20	2.2%	4.3%	1.7%	5.2%	1.9%	5.9%	0.3%	2.7%
20 < Income <= 30	1.9%	8.5%	5.2%	3.3%	3.6%	8.6%	0.1%	3.9%
30 < Income <= 40	4.2%	13.6%	2.9%	8.5%	2.4%	9.7%	0.2%	5.1%
40 < Income <= 50	2.6%	14.9%	9.4%	8.6%	4.2%	15.0%	0.0%	6.9%
50 < Income <= 60	6.3%	11.8%	10.7%	11.5%	10.0%	16.4%	0.0%	8.3%
60 < Income	6.7%	21.7%	8.8%	11.5%	10.4%	16.5%	1.3%	9.6%
Total	3.0%	10.1%	4.9%	6.2%	3.8%	9.7%	0.3%	4.8%
Filing status								
Other	7.7%	0.0%	11.1%	8.7%	0.0%	13.3%	2.4%	5.8%
Single	4.5%	9.6%	5.0%	6.6%	5.4%	13.3%	0.1%	5.6%
Joint	4.4%	23.1%	13.8%	11.1%	3.7%	15.6%	0.4%	9.3%
Head of Household	1.4%	7.5%	2.5%	4.8%	2.7%	5.7%	0.3%	3.1%
Total	3.0%	10.1%	4.9%	6.2%	3.8%	9.7%	0.3%	4.8%
Treatment batch								
Batch 1	0.0%	15.0%	0.0%	15.0%	0.0%	10.0%	0.0%	5.1%
Batch 2	2.1%	16.5%	11.2%	7.1%	3.1%	11.6%	0.5%	6.6%
Batch 3	2.8%	9.5%	6.7%	7.0%	4.2%	8.4%	0.0%	4.8%
Batch 4	3.6%	11.5%	3.4%	5.0%	3.9%	10.3%	0.6%	4.8%
Batch 5	3.1%	7.3%	3.1%	5.9%	3.9%	10.1%	0.3%	4.2%
Total	3.0%	10.1%	4.9%	6.2%	3.8%	9.7%	0.3%	4.8%

Note: This table shows heterogeneity in response.

Table A.8: Network effects by 9-digit zip code

	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	Nonsampled	Nonsampled	Nonsampled	Nonsampled
Letters sent in sample	0.033*** (0.004)	0.023*** (0.004)	0.006* (0.004)	-0.004 (0.004)		
Federal returns		0.004*** (0.001)		0.004*** (0.000)		0.004*** (0.000)
Contact only					0.013 (0.009)	0.003 (0.009)
Penalty salience					0.002 (0.009)	-0.008 (0.009)
Punishment probability					0.013 (0.009)	0.002 (0.009)
Compliance cost					-0.003 (0.009)	-0.014 (0.009)
Civic pride					0.015* (0.009)	0.006 (0.009)
Penalty x punishment					-0.001 (0.009)	-0.011 (0.009)
No contact					0.006 (0.006)	-0.004 (0.006)
Constant	0.047*** (0.003)	0.025*** (0.004)	0.046*** (0.002)	0.024*** (0.003)	0.046*** (0.003)	0.024*** (0.003)
R^2	0.00	0.01	0.00	0.00	0.00	0.00
Observations	21,376	21,376	21,376	21,376	21,376	21,376

Standard errors in parentheses

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

Note: This table estimates geographic network effects. Observations are 9-digit zip codes. The dependent variable in columns 1 and 2 is the number of returns for tax years 2014 and earlier filed by all taxpayers in May 2016 through July 2016. The dependent variable in columns 3 through 6 is the number of returns for tax years 2014 and earlier filed by untreated taxpayers in May 2016 through July 2016. “Letters sent in sample” is a count of the number of sampled taxpayers in that add-on code. “Federal returns” is a count of the number of taxpayers identified by the IRS as filing a federal return in tax year 2014 who should have filed a Detroit return, which is meant to control for the population within the add-on code. Columns 5 and 6 regress the count of untreated taxpayers who filed returns on counts of the number of individuals who received each experimental mailing.

Table A.9: Tax due

	(1)	(2)	(3)	(4)	(5)
	Tax due	Tax due	Tax due	Tax due	Tax due
Total income	0.0054*** (0.0006)				
Wage income		0.0073*** (0.0007)	0.0073*** (0.0007)	0.0074*** (0.0007)	0.0074*** (0.0007)
Nonwage income		0.0011 (0.0011)	0.0021* (0.0012)	0.0118*** (0.0021)	0.0121*** (0.0022)
Dummy age >= 62			-174.8** (77.7)		
Age dummy * wage income				-0.0017 (0.0012)	-0.0014 (0.0013)
Age dummy * nonwage income				-0.0135*** (0.0024)	-0.0134*** (0.0025)
Age					-1.6 (1.8)
FS = married filing jointly					-37.9 (59.2)
FS = head of household					-49.7 (49.3)
Constant	-38.5 (34.9)	-88.1** (35.7)	-78.4** (35.8)	-100.9*** (34.8)	-8.2 (89.0)
R^2	0.154	0.194	0.203	0.255	0.258
Observations	459	459	459	459	459
Admitted tax due, mean from filed returns (\$)	203.40	203.40	203.40	203.40	203.40
Predicted due, sum of population (\$)	6,190,151	6,082,965	6,415,784	6,187,799	6,325,517

Standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: This table estimates admitted tax due on filed city returns in the sample, conditional on characteristics observable from federal returns.