National Tax Journal, June 2020, 73 (2), 331–360 https://doi.org/10.17310/ntj.2020.2.02

COMMUNICATING TAX PENALTIES TO DELINQUENT TAXPAYERS: EVIDENCE FROM A FIELD EXPERIMENT

Taylor Cranor, Jacob Goldin, Tatiana Homonoff, and Lindsay Moore

We analyze a large field experiment conducted with the Colorado Department of Revenue to study the presentation of financial incentives and social norms in tax delinquency notices. We find that notices that highlight and provide information about financial penalties modestly raise the payment rate among delinquent taxpayers, with larger effects for notices providing greater detail. In contrast, we find no payment effect from highlighting social norms for timely payment. Our results suggest that attention to seemingly minor decisions about the wording of notices sent by tax authorities can increase tax payments and reduce administrative costs associated with taxpayer delinquency.

Keywords: tax compliance, communication outreach, field experiment

JEL: H26, D91

I. INTRODUCTION

Delinquent tax payments represent a central problem for tax authorities. In the United States, such payments represent approximately 25 percent of uncollected federal tax revenue (Perez-Truglia and Troiano, 2015). Virtually all government bodies that collect taxes — whether at the municipal, state, or federal level — send notices to those taxpayers who owe an outstanding tax liability. These notices typically include information on the taxpayer's balance due as well as information on the financial consequences of failing to make the required payment.

Traditional models of tax compliance assume that individuals decide whether to pay their balance due by weighing the cost of payment against the expected penalties for noncompliance (e.g., Allingham and Sandmo, 1972). Such models suggest that vary-

Taylor Cranor: Yale Law School, New Haven, CT, USA (Taylor.Cranor@yale.edu)

Jacob Goldin: Stanford Law School, Stanford, CA, USA and NBER, Cambridge, MA, USA (jsgoldin@law.stanford.edu)

Tatiana Homonoff: Robert F. Wagner School of Public Service, New York University, New York, NY, USA and NBER, Cambridge, MA, USA (tah297@nyu.edu)

Lindsay Moore: Behavioural Insights Team, Washington, DC, USA (lindsay.moore@bi.team)

ing the financial incentives for payment can affect taxpayer decisions about evasion or payment but do not focus on the possibility that variation in *how* those incentives are communicated to taxpayers might affect whether taxpayers decide to comply. In contrast, a growing literature in behavioral economics suggests that how the costs and benefits of payment are presented to taxpayers shapes how taxpayers respond (Slemrod, 2018; Pomeranz and Vila-Belda, 2019). If correct, this literature suggests that modifying the presentation of incentives described in delinquent tax notices can improve compliance for essentially no cost.

This article experimentally evaluates the effectiveness of a range of small modifications to a state's tax delinquency notice that varied the presentation of the incentives for timely tax payment. Specifically, we report results from a natural field experiment conducted in collaboration with the Colorado Department of Revenue (DOR). Our sample consists of the approximately 90,000 households that comprise the universe of delinquent taxpayers for the state of Colorado for tax year 2015. These households represent approximately 3.5 percent of Colorado income tax returns and collectively owe over $85 million in state income taxes.

The taxpayers in our sample were randomly assigned to receive one of several versions of a delinquency notice sent to taxpayers by the DOR. The first two treatment groups emphasized the financial penalty for noncompliance. Taxpayers assigned to these treatments received one of two notice variants: A detailed version listed the interest rate penalty associated with delayed payment and a general version emphasized the existence of a financial incentive for timely payment but did not provide details. A third experimental treatment emphasized social norms by including information on the high fraction of Colorado taxpayers who pay their tax bill on time. The final experimental group served as the control group and received the version of the notice sent in prior years. The notices associated with different treatment groups differed only with respect to a single sentence.

Despite the seemingly minor differences between the notices, we document nontrivial differences in their effect on taxpayer behavior. We find that the detailed penalty notice increased the fraction of taxpayers making a full payment before the statutory deadline or creating a payment plan by 1.6 percentage points (a 4.1 percent increase) relative to the control notice. The estimated effect of the general penalty notice was approximately half as large as that of the detailed penalty, a statistically significant difference. We find no evidence that the social norms notice was more effective than the control; we estimate its effect to be near zero and not statistically significant.

Interpreting the economic significance of our estimates requires understanding whether the treatments simply speed up the timing of payments or if they encourage payments from taxpayers who would not have paid had they not received the treatment message. The notices themselves emphasize the importance of taking action by the statutory deadline (30 days from the notice's receipt). To investigate the persistence of our observed treatment effects, we collected data on taxpayer payments for several

months after the statutory deadline. Focusing on payments made within 100 days, we find the effect of the penalty notice declines only slightly (by approximately one-third of its size for payments made by the statutory deadline). This persistence suggests that the treatment induces new payments rather than simply speeding up the timing of payments that would have otherwise been made within this time frame.

To understand the effect of our intervention on revenue collected, we examine treatment effect heterogeneity by taxpayer balance due. We observe differences in how the notices affect payment decisions along this margin: In response to the penalty notices, the effect of the detailed penalty on full payment was limited to taxpayers in the lowest and middle tertiles of balances due. In contrast, taxpayers in the highest tertile of balances due (those owing at least $433) were no more likely to make a full payment after receiving one of the penalty notices. On the other hand, the detailed penalty increased the rate at which taxpayers with high balances due enrolled in payment plans. Consistent with our earlier results, the same pattern is present, but attenuated, with respect to the general penalty, and we observe no effect of the social norms notice in any of the balance due categories. To the extent our observed payment effects are driven by taxpayers with lower balances due, the primary fiscal benefit of interventions such as the one we study may stem from reducing the cost of subsequent compliance efforts, rather than directly inducing taxpayers to remit large payments of outstanding tax liability. In particular, by reducing the number of taxpayer accounts that have not paid by the statutory deadline — the date at which additional intensive collection efforts are first imposed — the changes in message design we study can reduce costs to both the government and taxpayers associated with additional follow-up and more intensive enforcement activities such as imposition of tax liens and garnishment of taxpayer wages.

We contribute to a growing literature that experimentally evaluates communications sent by tax authorities with the goal of raising compliance. Much of this literature focuses on approaches to raising tax compliance that broadly fall within two categories: emphasizing financial incentives and social norms appeals. This literature shows consistent evidence that emphasizing financial incentives is effective at increasing timely payment of tax liability (for a review, see Pomeranz and Vila-Belda, 2019).[1] However, the literature that evaluates using social norms to increase tax compliance finds mixed results: Some studies find positive effects (Del Carpio, 2013; Kettle et al., 2016; Hallsworth et al., 2017; Bott et al., forthcoming; Larkin et al., 2019), whereas others do not (Blumenthal et al., 2001; Castro and Scartascini, 2015; Meiselman, 2018;

[1] Other notable papers on this topic include Fellner, Sausgruber, and Traxler (2013), Castro and Scartascini (2015), Chirico et al. (2019), Meiselman (2018), and De Neve et al. (2019). A related literature investigates the effects of communications from a tax authority that manipulate taxpayers' perceptions about the likelihood of an audit (Slemrod, Blumenthal, and Christian, 2001; Kleven et al., 2011; Dwenger et al., 2016; Castro and Scartascini, 2015; Gangl et al., 2014; Bérgolo et al., 2017; Bott et al., forthcoming).

John and Blume, 2018; Biddle, Fels, and Sinning, 2018). Our primary contribution is to provide new evidence on the effectiveness of these approaches within a large field experiment in a policy-relevant setting — income tax collection by U.S. state taxing authorities. An additional benefit of our approach relative to much (but not all) of the prior literature is that it permits comparison of social norms and penalty-focused treatments, holding fixed the characteristics of the population being studied. Finally, to our knowledge, prior work has not studied the effects of the level of detail at which penalties for noncompliance are described.[2]

Prior work that experimentally evaluates communication of tax penalties and social norms in shaping taxpayer compliance has primarily done so outside the United States, at other levels of government within the United States, and/or for other types of tax.[3] The insights obtained from these studies may not generalize to the U.S. state tax administration setting. For social norms, the motivational force of a norm may vary across cultures or based on the geographic scope of the comparison population. Similarly, with respect to communications that highlight penalties, the effect is likely to differ between the United States and other countries based on the system of tax payment (e.g., near-exact withholding with reconciliation in the United Kingdom versus frequent refunds and balances due in the United States), base level of compliance, and what types of penalties are legally available or used in practice by the taxing authority.

Finally, our paper contributes to a broader literature on the effectiveness of informational messages from government agencies at increasing civic responsibility and engagement with government programs. For example, communications from the Internal Revenue Service sent to low-income nonfilers emphasizing the availability of the Earned Income Tax Credit (EITC) have been shown to increase EITC take-up (Bhargava and Manoli, 2015; Guyton et al., 2016; Manoli and Turner, 2014). Similarly, government communications aimed at increasing take-up of benefit programs have been shown to be effective in a variety of policy areas such as disability insurance (Armour, 2018), public health insurance (Aizer, 2007), postsecondary education (Barr and Turner, 2018), and retirement savings (Goldin, Homonoff and Tucker-Ray, 2017).

The remainder of the paper proceeds as follows. Section II describes the institutional setting in which our experiment occurred. Section III provides details on our experi-

---

[2] In a new working paper, De Neve et al. (2019) study the effect of shortening and simplifying a tax compliance letter, which sheds light on the information overload — one mechanism through which the level of detail could affect the letter's effectiveness. In related work, John and Blume (2018) find that simplifying the information contained in a notice increased tax compliance.

[3] Within the United States, for example, Meiselman (2018) and Chirico et al. (2019) compare the effectiveness of communications emphasizing social norms and financial incentives sent by tax authorities in the cities of Detroit and Philadelphia (respectively). We find similar results — that highlighting financial penalties increases timely payments, whereas social norms appeals have no effect — in a different institutional policy setting (property taxes versus income taxes in the case of Chirico et al. (2019)) and a different population (nonfilers versus delinquent filers in the case of Meiselman (2018)). To our knowledge, the only prior papers to study such policies in the state income tax setting are Blumenthal et al. (2001), which focused exclusively on normative appeals (closely related to social norms), and Slemrod, Blumenthal, and Christian (2001), which studied interventions designed to shape taxpayer perceptions about audit probabilities

mental design. Section IV describes our administrative data. Section V provides the results of the experiment. Section VI discusses our findings in relation to the literature and concludes.

II. INSTITUTIONAL BACKGROUND

Like other tax authorities for states and cities, the Colorado DOR sends letters to taxpayers who owe additional taxes beyond what they paid when filing their tax return. These "Notice of Deficiency" (NOD) letters inform taxpayers that they have an additional unpaid tax liability or that they are not entitled to the full refund they claimed on their return. The letters instruct taxpayers to pay the additional amount due, create a payment plan, or challenge the determination of additional liability by the statutory deadline — 30 days after the notice's receipt.

Colorado law creates several incentives for taxpayers receiving an NOD to pay the additional tax liability. First, unpaid tax liability accrues interest at a statutorily provided interest rate. The applicable rate of interest increases by 3 percentage points if payment is not made by the statutory deadline. At the time of our study, the annual interest rate was 3 percent for payments made within 30 days of the NOD's receipt and 6 percent for tax liability that remained unpaid as of that date.

The second incentive for taxpayers to pay the tax liability reported on the NOD is that Colorado law imposes financial penalties on delinquent taxpayers. This penalty applies to taxpayers who have not paid by 30 days after the NOD is received. The magnitude of the penalty is initially set at 5 percent of the outstanding tax liability and increases by 0.5 percentage points each month, up to a maximum penalty rate of 12 percent. In combination with the interest rate described above, the maximum annual penalty is, therefore, 18 percent of the tax liability owed, comparable to a credit card interest rate. Setting up a payment plan does not erase previously imposed penalties but stops the penalty rate from increasing.

Finally, taxpayers whose tax liability remains unpaid after the expiration of the 30-day period following the NOD receipt may face additional financial consequences such as garnishment of wages or bank accounts, referral to a collection agency, a lien or judgment against personal property, and even the sale or seizure of real property. Taxpayers may also avoid the imposition of these additional measures by setting up a payment plan for their unpaid tax liability. These consequences are detailed in the "Notice of Final Determination and Demand for Payment," which is sent immediately following the statutory deadline.

Approximately 100,000 NOD letters are mailed by the DOR each year. The letter population includes both Colorado taxpayers and out-of-state taxpayers who owe Colorado income taxes. As of July 2016, the total amount owed to the state by taxpayers who filed, but did not pay, on time was approximately $85 million. For context, the state budget office estimates that it will collect $220.9 million in taxes owed after one year (whether from taxpayers who filed but are delinquent in their payments or from individuals who owe tax but who failed to file a return) (Office of the State Controller, 2016).

III. EXPERIMENTAL DESIGN

In July 2016, the DOR conducted a randomized controlled trial to assess whether minor modifications to the NOD letters were associated with an increase in the fraction of taxpayers paying off their liability in full or creating payment plans. The sample consisted of 90,349 Colorado taxpayers who were identified by the DOR as delinquent and were slated to be mailed the NOD letter in July of 2016, following the close of the tax filing season in mid-April of 2016. The taxpayers included in the sample had each filed a return during the 2016 tax filing season (i.e., for tax year 2015) but had failed to pay the full amount reported due on the return.[4] Most of these taxpayers had not made any tax payments as of the NOD mailing date, though the sample includes taxpayers who paid an amount less than the full balance due as well. The taxpayers in this sample were randomly assigned to receive one of four versions of the NOD letter. We describe these letters below and summarize their content in Table 1.

Taxpayers assigned to the control group received the same version of the NOD letter that had been sent to all taxpayers in prior years (Appendix Figure 1). This letter describes the financial incentive for timely payment with the following sentence: "The Statement of Account reflects a 3% interest discount if paid within 30 days." Additionally, the standard NOD letter states that "Penalty and interest have been charged in accordance with Colorado tax law," but does not provide details about the existence of the delinquent taxpayer penalty other than a brief reference to the DOR's website.[5]

The "detailed penalty" treatment (Appendix Figure 2a) reflected a number of modifications to the control letter designed to bolster the effectiveness of the incentive for timely payment. First, motivated by the literature on loss aversion (Kahneman and Tversky, 1979), the incentive was reframed as a penalty for nonpayment instead of a benefit for payment before the deadline. Second, the letter stated that the change in interest rate was a function of the applicable law, which may have reduced inattention or conveyed to taxpayers that the penalty was more likely to be imposed than if it was discretionary. Third, the sentence presenting information about the penalty was bolded to increase the salience of the incentive. Finally, relative to the control, the detailed penalty treatment included additional detail about how the magnitude of the interest rate varies based on whether payment is timely. The modified sentence in the detailed penalty letter read: "By law, if you do not pay within 30 days, the interest rate on your account will double from 3% to 6%."

The "general penalty" treatment (Appendix Figure 2b) was similar to the detailed penalty treatment in most regards, but it provided information about the financial incentives for on-time payment at a higher level of generality. Like the detailed penalty,

[4] The sample excludes individuals who failed to file any return and individuals who filed a return for which the DOR adjusted the calculation of tax liability owed. Taxpayers in these groups are placed in a different pool by the DOR and receive a different set of communications.

[5] Specifically, the notice states: "For more information regarding penalties and interest, please see FYI General 11 at www.TaxColorado.com."

Table 1
Treatment Variants and Potential Mechanisms

Treatment	Key Text	Potential Mechanisms (Relative to Control)
Control	The Statement of Account reflects a 3% interest discount if paid within 30 days.	
Detailed penalty	By law, if you do not pay within 30 days, the interest rate on your account will double from 3% to 6%.	• Loss aversion • Inattention/salience • Concrete information/less abstract
General penalty	By law, if you do not pay within 30 days, any penalty associated with your account will increase for each month you do not pay (until the statutory maximum is reached).	• Loss aversion • Inattention/salience • Avoids information overload
Social norms	Nine out of ten people in Colorado pay their tax on time. You are currently in the very small minority of people who have not yet paid.	• Descriptive/minority social norms

The control, detailed penalty, and general penalty treatments do not include information about overall patterns of timely payment. The social norms treatment includes the same language about the financial penalty as the control.

the general penalty treatment framed the financial incentive as a loss, noted that the penalty occurred by operation of law, and bolded the sentence in the letter containing penalty information. Instead of providing information about the specific magnitude of the interest rate change, however, the general penalty letter simply noted the existence of a financial penalty that was increasing in the months elapsed before payment. The key sentence in the general penalty treatment read: "By law, if you do not pay within 30 days, any penalty associated with your account will increase for each month you do not pay (until the statutory maximum is reached)."

Theoretically, either the general or detailed penalty may be more effective. On the one hand, the general penalty letter may be less effective than the detailed penalty if providing specific information about the penalty makes it more concrete and easier

for the taxpayer to understand. On the other hand, it may be that taxpayers are equally averse to paying any penalty (at least within a broad range of amounts), in which case the specific information may contribute little extra motivational force. In addition, the extra information contained in the detailed penalty notice may contribute to information overload (e.g., Bhargava and Manoli, 2015), dampening the notice's motivational force relative to the general penalty treatment.

A limitation of our experimental design is that it does not allow us to disentangle the precise mechanisms by which the penalty treatments affect behavior. In particular, relative to the control, both the detailed and general penalty letters increase the salience of the penalty and frame it in terms of a loss.

Finally, the "social norm" treatment (Appendix Figure 2c) was motivated by recent studies suggesting that information about the prevalence of tax compliance can increase payment rates among delinquent taxpayers. Hallsworth et al. (2017) finds significant increases in tax compliance among U.K. taxpayers, highlighting that such information is most effective when the norm is descriptive (versus injunctive), tailored to the recipient's local area, and references the fact that the taxpayer is in the minority composed of noncompliant taxpayers. Along these lines, the letter associated with this treatment used the exact language from the most effective treatment message in Hallsworth et al. (2017), adapted to our geographic setting: "Nine out of ten people in Colorado pay their tax on time. You are currently in the very small minority of people who have not yet paid." This sentence was placed in its own paragraph, in bolded and italicized text. The social norms treatment did not modify the language describing the penalty from the control.

Taxpayers were assigned to each of the four treatment groups with equal probabilities. To improve the precision of the analysis, randomization was stratified based on the taxpayer's age, balance due, whether the taxpayer's return was filed before the due date, and whether the taxpayer was a Colorado resident.

IV. DATA

To analyze the results of the experiment, we used anonymized data on taxpayer payments and payment plan creation from the universe of delinquent taxpayers for the state of Colorado (101,068 taxpayers). These data were provided to us by the DOR. In addition to these outcome variables, the data included information on taxpayer age, ZIP code of residence, tax balance due, and whether the taxpayer's return was filed on time. We supplement the DOR data in our analysis with ZIP code level information from the American Communities Survey (ACS) on income and educational attainment.

Our main outcome variables relate to whether taxpayers made a payment on their account. Specifically, we track whether a taxpayer fully pays off the outstanding balance on their account, as well as whether a taxpayer makes a partial payment by creating a payment plan with DOR. The DOR categorizes taxpayers into three groups: those who paid off their balance in full by Final Determination (FD), the date corresponding (approximately) to the statutory deadline of 30 days from the taxpayer's receipt of the

NOD;[6] those who created a payment plan by FD; and those who took no action before FD. For taxpayers who made multiple payments but did not create a payment plan, we add up the payment amounts and use the last date of payment to determine if and when the taxpayer successfully paid off their balance. Taxpayers who had bills from before the December 31, 2015 filing period or who had multiple bills were excluded from the sample.[7] The final sample consisted of 90,349 taxpayers.

V. RESULTS

Table 2 provides descriptive statistics of our sample population. Column (1) describes the characteristics of the delinquent taxpayers comprising our sample. The mean taxpayer is approximately 44 years old and owes a balance due of $515. Figure 1 presents the full distribution of balances due. Not surprisingly, the vast majority of our sample are Colorado residents. Columns (2)–(5) investigate the balance of taxpayer characteristics across treat-

Table 2
Descriptive Statistics

	Treatment Groups					
	(1) Full Sample	(2) Detailed Penalty	(3) General Penalty	(4) Social Norms	(5) Control	(6) *p*-Value
Individual Characteristics						
Age	43.9	43.9	43.9	43.8	43.9	0.8467
Balance due ($)	515	514	518	515	514	0.9440
Return received on time (%)	77.9	77.9	77.8	77.9	78.0	0.9839
Colorado resident (%)	91.1	91.1	91.0	91.1	91.3	0.8137
Neighborhood Characteristics						
Median household income ($)	64,215	64,073	63,982	64,229	64,576	0.0171
Bachelor's degree or higher (%)	37.1	37.2	36.8	36.9	37.4	0.0008
N	90,349	22,571	22,625	22,613	22,540	

Source: Colorado DOR, 2016 (individual characteristics) and American Community Survey 2011–2015 (neighborhood characteristics). Local demographic data are reported at the ZIP code tabulation area (ZCTA) level and converted to the ZIP code level using UDS Mapper Cross-Walk. The *p*-value in Column (6) is associated with the F-test for equality across the four experimental groups.

[6] In practice, the DOR assigns the date of FD as 45 days from the date the letter was queued to be mailed.

[7] A very small number of taxpayers received two NODs for the same period, resulting in multiple bills. This occurs if the taxpayer filed an original return that resulted in an NOD, and then subsequently filed an amended return that resulted in another NOD. We exclude these taxpayers from our analysis since they received multiple payment reminders from the DOR, though our results are not sensitive to their inclusion.

Figure 1
Distribution of Balances Due

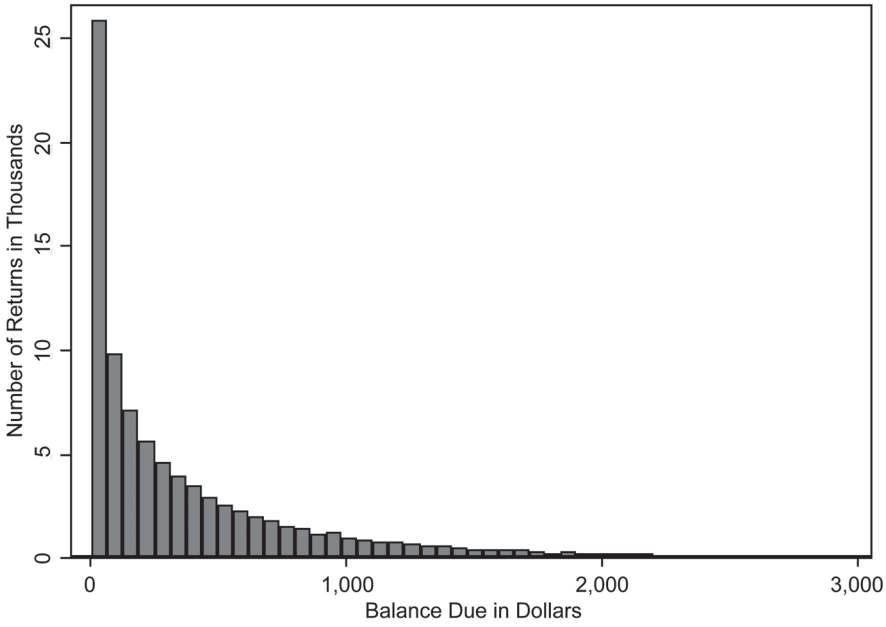


Figure 1 presents the distribution of balance due in the sample. The figure excludes the 2,309 returns with a balance due of over $3,000.

ment groups. Rows 1–4 show the balance across the available taxpayer-level characteristics: age, balance due, whether the return was received by the filing deadline, and whether the taxpayer was a Colorado resident. We also investigate the balance of local demographic characteristics of the taxpayer at the ZIP code level including median household income and bachelor's degree completion rates. In contrast to the taxpayer-level characteristics, we observe economically small but precisely estimated differences across treatment groups in these local demographic variables. To address this, we confirm that differences in geographic characteristics do not appear to be driving our results and we include either local demographics or ZIP code fixed effects in our empirical specifications below.

Our baseline empirical specification is a linear probability model that takes the form

$$y_i = \alpha + \beta_{DP} DP_i + \beta_{GP} GP_i + \beta_{SN} SN_i + \gamma x_i + \varepsilon_i$$

for taxpayer i, where y indicates the binary outcome of interest (paying a bill or making a payment plan); DP_i, GP_i, and SN_i are indicators for being assigned to the detailed penalty, general penalty, or social norms treatment group, respectively; and x_i is a vector of taxpayer-specific characteristics.[8]

[8] The results are qualitatively similar when estimated with a probit model.

A. Main Results

1. Full Payments

Our first analysis considers the effect of the notice variants on payments by taxpayers. The outcome we consider is an indicator for whether the taxpayer has paid off their balance in full by the statutory deadline (30 days from receipt of the NOD letter).[9] We focus on the statutory deadline, because after this stage, a delinquency moves to FD at which time the DOR begins to engage in more intensive collection activities such as wage garnishment, bank levies, the filing of judgments and liens, and the use of private collection agencies. These activities entail costs to the DOR, so policies that reduce these costs can help the department increase the net revenue it takes in.

Table 3 presents the results of this analysis. Among control group members, 34 percent of notified taxpayers made a full payment by the statutory deadline, suggesting substantial room for improvement in tax compliance. Column (1) presents our baseline specification, which includes indicators for treatment group assignments and no control variables. We estimate that receiving the detailed penalty treatment induces a 0.9 percentage point increase in the proportion of delinquent taxpayers who pay off their balance in full by the statutory deadline, relative to the control NOD letter. The estimated effect is statistically significant and represents a 2.6 percent increase relative to mean payment rate under the control NOD letter. The estimated effect of the general penalty is also positive, but the estimated coefficient is approximately half as large in magnitude as the detailed penalty and is not statistically different from zero. The estimated effect of the social norms penalty is near zero in magnitude and is not statistically significant.

Column (2) adds controls for the individual characteristics described in Table 2: taxpayer age, state residency, late filing status, and balance due.[10] The addition of these controls does not appreciably change the estimated coefficients, consistent with the observed balance of individual characteristics across treatment groups. In contrast, controlling for the local demographic characteristics from Table 2 — median income and college completion rates — in Column (3) slightly raises the magnitude of the estimated coefficients. Finally, adding ZIP code level fixed effects (Column (4)) yields our preferred specification. The estimated effect of the detailed penalty treatment on payments is 1.1 percentage points, representing a 3.2 percent increase over the control group.[11] The estimated effect of the general penalty treatment is 0.5 percentage points, and the estimated effect of the social norms treatment is approximately 0.2 percentage points, though neither of these effects is statistically significant. The analysis, thus, suggests that the detailed penalty treatment is the most effective and is associated with an economically modest increase in the fraction of delinquent taxpayers paying off their balance in full by the statutory deadline.

---

[9] Because we do not observe the exact date the letter was received, we follow DOR and use as our cutoff 45 days from the date the letter was queued to be mailed.

[10] For age and balance due, controls include continuous measures of age and balance due, the age and balance categories used in randomization stratification, and the interaction of the continuous and categorical variables. The specification also includes an indicator for individuals for whom age is missing (2,472 individuals).

[11] A randomization inference test yields an estimated p-value of approximately 0.007 (Appendix Figure 3).

Table 3

Effect of Notice on Payment Rate by Treatment

	(1)	(2)	(3)	(4)
Detailed penalty	0.888**	0.855**	0.961**	1.105***
	(0.449)	(0.404)	(0.397)	(0.406)
General penalty	0.460	0.464	0.665*	0.495
	(0.428)	(0.365)	(0.357)	(0.361)
Social norms	–0.198	0.010	0.155	0.179
	(0.460)	(0.404)	(0.406)	(0.415)
Test: detailed = general	0.365	0.350	0.472	0.149
Test: detailed = norms	0.015	0.038	0.047	0.026
Individual characteristics	No	Yes	Yes	Yes
Neighborhood characteristics	No	No	Yes	No
ZIP code fixed effects	No	No	No	Yes
Control group mean	33.940	33.940	33.940	33.940
N	90,349	90,349	90,349	90,349

Outcomes: An indicator for making a full payment by statutory deadline (within 45 days of the NOD). Individual characteristics include age controls, balance controls, Colorado residency, and whether the return was filed on time. Age controls include a continuous age variable, dummy variables for four age categories, and an interaction of the dummies with the continuous age variable. Similarly, the balance controls include a continuous variable for the log of the balance due, dummies for six balance categories, and an interaction of the dummies with the continuous balance variable. Neighborhood characteristics include ZCTA-level median household income and college attainment. The table reports *p*-values associated with the F-test for equality of treatment effects across the experimental groups (Detailed Penalty versus General Penalty and Detailed Penalty versus Social Norms, respectively). All columns report results from a linear probability model. Units are percentage points (0–100). The robust standard error is in parentheses. $*p < 0.1$, $**p < 0.05$, and $***p < 0.01$.

2. Payment Plan Creation

Table 4 turns to our next outcome of interest, a taxpayer's partial payment of their outstanding balance through the creation of a payment plan. As described in Section II, the initiation of a payment plan prevents the penalty rate from increasing and avoids the initiation of more intensive collection activities at FD.

For reference, Column (1) replicates our preferred specification on the likelihood of making a full payment from Table 2. Column (2) repeats that specification using payment plan creation as the outcome variable. We observe a similar pattern as with full payments. The results suggest that the detailed penalty treatment was the most effective at inducing taxpayers to create payment plans — it was associated with a statistically

Table 4

Effect of Notice on Payment and Payment Plans by Treatment

	(1) Paid in Full	(2) Plan	(3) Paid or Plan
Detailed penalty	1.105***	0.465***	1.570***
	(0.406)	(0.175)	(0.429)
General penalty	0.495	0.149	0.644
	(0.361)	(0.197)	(0.409)
Social norms	0.179	–0.250	–0.071
	(0.415)	(0.166)	(0.414)
Test: detailed = general	0.149	0.131	0.049
Test: detailed = norms	0.026	0.000	0.000
Control group mean	33.940	4.468	38.407
N	90,349	90,349	90,349

Outcomes: An indicator for making a full payment (Column (1)), creating a payment plan (Column (2)), and either making a full payment or creating a payment plan (Column (3)) by statutory deadline (within 45 days of the NOD). All regressions include age controls, balance controls, whether the return was filed on time, and ZIP code fixed effects. Age controls include a continuous age variable, dummy variables for four age categories, and an interaction of the dummies with the continuous age variable. Similarly, the balance controls include a continuous variable for the log of the balance due, dummies for six balance categories, and an interaction of the dummies with the continuous balance variable. The table reports *p*-values associated with the F-test for equality of treatment effects across the experimental groups (Detailed Penalty versus General Penalty and Detailed Penalty versus Social Norms, respectively). All columns report results from a linear probability model. The units are percentage points (0–100). Robust standard errors are in parentheses. $*p < 0.1$, $**p < 0.05$, and $***p < 0.01$.

significant 0.5 percentage point increase in payment plans, representing approximately an 11 percent increase relative to the control group mean of 4.5 percent. The effects of the other two treatments were near zero and not statistically significant.

Finally, because the treatments might cause taxpayers to switch between paying off their balance and payment plan creation, Column (3) of Table 4 investigates the effect of the treatment on the likelihood of taxpayers either making a full payment or creating a payment plan. The results suggest that the detailed penalty treatment increases the fraction of taxpayers making any payment by 1.6 percentage points, a 4.1 percent increase relative to the control group mean. The general penalty version was associated with an increase relative to the control of approximately half of this magnitude. The

difference in the estimated effect of the detailed and general penalty letters is statistically significant ($p < 0.05$). Again, we estimate the effect of the social norms letter to be near zero. The difference in the estimated effect of the detailed penalty and social norms letters is statistically significant as well ($p < 0.01$).[12]

B. Short- versus Longer-Term Outcomes

Interpreting the economic significance of our estimates in the prior subsection requires understanding whether the intervention increased payments from taxpayers who would not have paid had they not received the treatment message, or simply sped up payments by taxpayers who would have eventually paid. For example, Hallsworth et al. (2017) find large effects of the NOD letters emphasizing various types of social norms compared to the standard NOD on payments made within three weeks; however, these treatment effects fade and many are no longer significant when considering payment rates 70 days after the letter's mailing, suggesting that the treatment messages primarily increased the speed at which recipients paid their taxes rather than generating new payments. Our analyses thus far have focused on taxpayer actions taken by the statutory deadline using the DOR's definition of FD to be 45 days from the date the notice was queued to be mailed. This is the relevant cutoff for the imposition of financial penalties, but it is possible that the modifications to the NOD letter could have affected behavior apart from that specific margin as well, such as by inducing taxpayers to make their payments earlier than they would have otherwise.

Figure 2 presents the rate of full payments for each experimental group by day for the first 100 days after the letter's mailing. We focus only on full payments for this analysis because we lack data on the precise timing of payment plans created after the statutory deadline. Payment rates steadily increase over the first month for each group with roughly one-third of taxpayers making a full payment by the statutory deadline of 45 days and increasing to roughly 40 percent by 100 days. The figure shows that deviations by experimental group emerge around 30 days after the notice was sent and remain through the end of the follow-up period.

Table 5 estimates the effects of the treatments on taxpayer payments up to 100 days from the letter's mailing, varying the cutoff date across different time windows. The results show the effect is largest at the NOD threshold but suggest that the effects persist (albeit with a smaller magnitude) as far as 100 days from the letter's mailing. The results in the table, thus, provide suggestive evidence that the treatments cause some delinquent taxpayers to pay off their balance when they would not have otherwise done so in the months following receipt of the notice. The fact that the estimated effects of the treatments are smaller in magnitude when assessed at the later dates suggests that the treatments cause other taxpayers to move up the timing of the payments they would have otherwise made eventually.

[12] Appendix Table 1 presents the unadjusted means by experimental group for each of the three outcome variables in Table 4 as well as for the tax liability collected.

Figure 2

Payment Rate over Time by Experimental Group

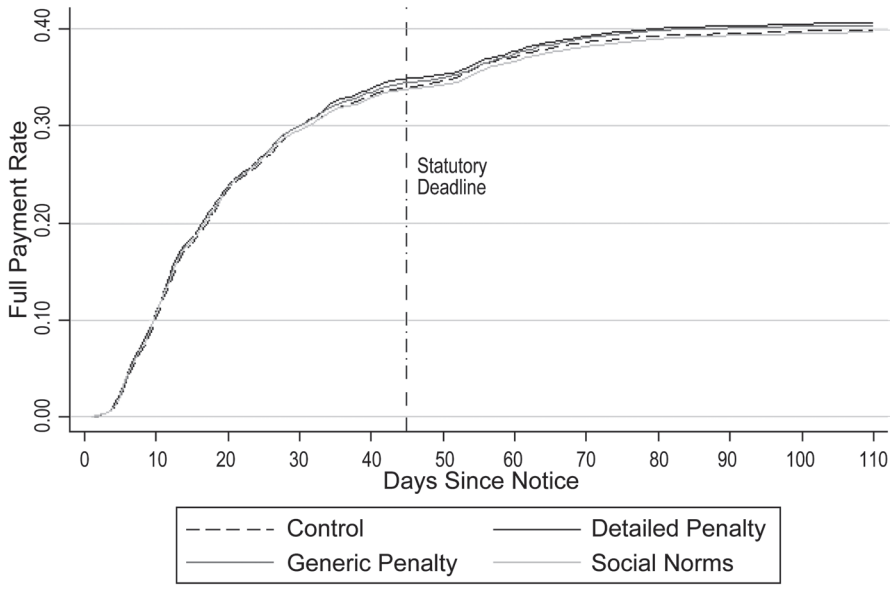


Figure 2 denotes the cumulative distribution function of payments made in full by day, separately by treatment group. The statutory deadline is defined as 45 days after NODs are sent.

C. Payment Effects by Balance Due

Our results so far suggest that the penalty-design notices affect the number of taxpayers who choose to pay off their balance due or set up a payment plan. However, understanding how this reduction in the number of delinquent taxpayer accounts maps into revenue requires understanding heterogeneity in the treatment effect by taxpayers' balance due.

We find strong evidence of heterogeneity along this margin. Table 6 considers our baseline specification separately by balance due tertile. For the lowest tertile (balances due of less than $95), we find that both penalty notices are effective. The detailed penalty increases the likelihood of full payment or payment plan creation by 1.8 percentage points (a 3.1 percent increase relative to the control) and the general penalty increases payments or payment plans by 1.2 percentage points (a 2.1 percent relative increase). Because the balance due is relatively low, it is perhaps not surprising that the treatment effect for taxpayers in this tertile is driven by full payments rather than the creation of payment plans. For taxpayers in the middle tertile (balances due between

Table 5
Effect of Notice on Payment and Payment Plans by Treatment

	(1) 45 Days	(2) 60 Days	(3) 75 Days	(4) 100 Days
Detailed penalty	1.105***	0.616	0.726*	0.751*
	(0.406)	(0.419)	(0.437)	(0.443)
General penalty	0.495	0.247	0.388	0.393
	(0.361)	(0.375)	(0.374)	(0.380)
Social norms	0.179	–0.143	–0.138	0.058
	(0.415)	(0.426)	(0.434)	(0.443)
Control group mean	33.940	37.116	38.935	39.712
N	90,349	90,349	90,349	90,349

Outcomes: An indicator for making a full payment within the specified number of days after the NOD is sent. DOR implements the statutory deadline as 45 days after the NOD is sent. All regressions include age controls, balance controls, whether the return was filed on time, and ZIP code fixed effects. Age controls include a continuous age variable, dummy variables for four age categories, and an interaction of the dummies with the continuous age variable. Similarly, the balance controls include a continuous variable for the log of the balance due, dummies for six balance categories, and an interaction of the dummies with the continuous balance variable. All columns report results from a linear probability model. Units are percentage points (0–100). Robust standard errors are in parentheses. *$p < 0.1$, **$p < 0.05$, and ***$p < 0.01$.

$95 and $433), the effect of the detailed penalty on full payments or plan creation is similar in percentage point terms (a 1.9 percentage point increase) but somewhat larger in relative terms (a 6.1 percent increase relative to the control). In contrast, we estimate the effect of the general penalty to be 0.3 percentage points and statistically insignificant. Finally, neither penalty treatment is statistically significant with respect to payments by taxpayers in the highest tertile (above $433), although we do observe a statistically significant 1.2 percentage point (12.3 percent) increase in payment plan creation associated with the detailed penalty for taxpayers in this group.[13] In each tertile, the effect of the social norms treatment is economically small and not statistically significant.

Using the estimates in Table 6, we conduct a back-of-the-envelope calculation for the total revenue effects from sending the various letters to all delinquent taxpayers in our sample. In particular, we estimate the overall effect of a notice on revenue by multiplying the estimated effect for each balance due tertile by the mean balance due among taxpayers in the corresponding tertile. Under the assumption that taxpayers who

[13] An important unknown for translating the effect of payment plan creation into revenue collection is the rate at which payment plans convert to full payments. Appendix Table 2 investigates the effect of the treatments by balance due on full payments at 100 days following letter receipt. The results are similar to those reported in Table 6, suggesting that payment plans do not convert into full payments within this (relatively short) time horizon.

Table 6
Treatment Effect Heterogeneity by Balance Due

	(1) Low Balance	(2) Medium Balance	(3) High Balance
Panel A. Paid			
Detailed penalty	1.9381**	1.7259**	–0.4349
	(0.8443)	(0.7757)	(0.5360)
General penalty	1.2705*	0.7159	–0.3129
	(0.7574)	(0.7263)	(0.5717)
Social norms	0.5318	0.6859	–0.6609
	(0.7550)	(0.8096)	(0.5501)
Control group mean	58.193	27.362	16.046
N	30,261	30,020	30,068
Panel B. Plan			
Detailed penalty	–0.119	0.171	1.209**
	(0.074)	(0.308)	(0.479)
General penalty	–0.027	–0.470	0.747
	(0.077)	(0.292)	(0.523)
Social norms	–0.017	–0.506	–0.368
	(0.079)	(0.321)	(0.445)
Control group mean	0.2119	3.4551	9.7911
N	30,261	30,020	30,068
Panel C. Paid or Plan			
Detailed penalty	1.820**	1.897**	0.774
	(0.852)	(0.758)	(0.664)
General penalty	1.244	0.246	0.434
	(0.763)	(0.771)	(0.763)
Social norms	0.515	0.180	–1.029
	(0.758)	(0.825)	(0.667)
Control group mean	58.405	30.817	25.837
N	30,261	30,020	30,068

Outcomes: An indicator for making a full payment by statutory deadline (Panel A), creating a payment plan (Panel B), and either making a full payment or creating a payment plan (Panel C). Columns (1)–(3) present results by the amount of the balance due, where low balance is less than $95, medium balance is between $95 and $433, and high balance is more than $433. All regressions include age controls, balance controls, whether the return was filed on time, and the ZIP code fixed effects. Age controls include a continuous age variable, dummy variables for four age categories, and an interaction of the dummies with the continuous age variable. Similarly, the balance controls include a continuous variable for the log of the balance due, dummies for six balance categories, and an interaction of the dummies with the continuous balance variable. All columns report results from a linear probability model. Units are percentage points (0–100). Robust standard errors are in parentheses. *$p < 0.1$, **$p < 0.05$, and ***$p < 0.01$.

enroll in a payment plan eventually pay off the full amount of their tax liability, our results imply an average per-person revenue effect of $1.46 from the detailed penalty, or $131,914 of additional tax liability collected within 45 days if the detailed penalty was sent to all delinquent taxpayers instead of the control.[14] Note that this estimate does not incorporate potential cost savings to the DOR or taxpayers from avoiding the intensive collection efforts imposed beginning after the statutory deadline.

D. Potential Mechanisms

In this subsection, we explore channels through which our observed treatment effect might operate. We focus on several potential factors that shape whether an individual with unpaid tax liability will make a timely payment. First, the taxpayer may wish to avoid the cost of raising the funds needed for payment within the relevant time window (liquidity costs). All else equal, we expect liquidity costs to be increasing in the taxpayer's balance due (i.e., it is harder to come up with $500 than $50). Second, the taxpayer may view the task of making the payment to be unpleasant (e.g., because of the time or mental energy required) and seek to put off this task into the future (procrastination). As a first approximation, we expect procrastination effects to be largely invariant to the taxpayers' balance due, although we acknowledge it is possible that the taxpayers may perceive the process of paying a larger tax bill as more unpleasant than paying a smaller tax bill.[15] Against these benefits, taxpayers compare the costs from deferring tax payments that stem from interest and penalties assessed. Some taxpayers may be unaware of these financial incentives for payment or inattentive to them. By providing information about the incentives in an effective manner, the detailed and general penalty treatment letters may raise the perceived costs of nonpayment, thereby inducing previously unaware or inattentive taxpayers to make payments sooner than they otherwise would.

To study how the treatments interact with factors that potentially mediate timely payment behavior, we explore a range of heterogeneity analyses by taxpayer demographic characteristics. However, such characteristics may proxy for multiple factors that affect payment decisions. In addition, our analysis was not preregistered, which may generate concerns about selective reporting of results when subgroup analyses are used to test particular hypotheses. For these reasons, we view our analysis in this subsection as exploratory and the results as primarily suggestive. Because the demographic characteristics we observe are likely correlated, we consider them within a single multivariate

[14] We focus on revenue collected within 45 days because we do not observe payment plan creation after that time window. The corresponding estimates for the general penalty and social norms notice are $58,635 ($0.65/person) and –$99,443 (–$1.10/person), respectively.

[15] We assume that taxpayers with positive balance due will eventually pay in response to more intensive collection activities such as wage garnishment or imposition of a tax lien; hence, we do not treat permanent avoidance of tax as a benefit to failing to make a timely payment.

regression rather than in separate analyses. In particular, Table 7 reports results from the following specification:

$$y_i = \alpha + \beta T_i + \delta \mathrm{x}_i + \eta T_i x_i + \gamma z_i + \varepsilon_i,$$

where y_i denotes that i paid or created a payment plan within 45 days, T_i indicates assignment to the treatment, x_i is a vector of demographic characteristics for which we explore

Table 7

Treatment Effect Heterogeneity

	(1) Detailed Penalty	(2) General Penalty	(3) Social Norms
Treatment	1.919***	0.701	0.173
	(0.466)	(0.460)	(0.459)
Treatment × late filer	1.828*	0.053	0.166
	(1.108)	(1.128)	(1.064)
Treatment × non-CO resident	3.388	–0.618	2.050
	(2.518)	(2.818)	(2.949)
Treatment × log median income	-0.210	0.338	–3.415*
	(2.152)	(1.814)	(1.997)
Treatment × log median education	–0.569	–0.722	2.280*
	(1.351)	(1.114)	(1.190)
Treatment × age	–0.049**	–0.052*	–0.028
	(0.025)	(0.027)	(0.030)
Control group mean	33.940	33.940	33.940
N	43,896	43,940	43,909

Outcome: An indicator for making a full payment or creating a payment plan by statutory deadline. Each column separately compares one treatment arm (Detailed Penalty, General Penalty, and Social Norms) to the control group. Subgroup interactions include an indicator for whether the return was filed late, whether the taxpayer is not a Colorado resident, taxpayer age, log ZCTA-level median household income, and log ZCTA-level college attainment. All regressions include age controls, balance controls, whether the return was filed on time, and ZIP code fixed effects. Age controls include a continuous age variable, dummy variables for four age categories, and an interaction of the dummies with the continuous age variable. Similarly, the balance controls include a continuous variable for the log of the balance due, dummies for six balance categories, and an interaction of the dummies with the continuous balance variable. All columns report results from a linear probability model. Units are percentage points (0–100). Robust standard errors are in parentheses. *$p < 0.1$, **$p < 0.05$, and ***$p < 0.01$.

treatment effect heterogeneity, and z_i is a vector of the other demographic characteristics included in our earlier specifications. For ease of presentation, we estimate this model separately for each treatment arm.

We first consider how the treatment interacts with taxpayer procrastination related to making a payment. An initial piece of evidence comes from the prior subsection, where we show a relatively large treatment effect among taxpayers with low balances due. Recall that one-third of delinquent taxpayers in our sample had balances due of $95 or less. Suppose that liquidity costs increase convexly in balance due[16] so that liquidity costs tend to be quite small for taxpayers with very low balances due. Under this assumption, the fact that we observe so much nonpayment among taxpayers with low balances due suggests procrastination may play a role in explaining nonpayment decisions. In turn, the fact that such taxpayers respond to the treatment implies that those taxpayers who defer payment because of procrastination also tend to be the ones who, absent the treatment, would be unaware or inattentive to the financial costs for nontimely payment.

To further explore this hypothesis, we investigate treatment effect heterogeneity based on whether taxpayers were late in filing their annual tax return. Because procrastination seems a likely explanation for why a taxpayer would *file* a return after the deadline,[17] it may be that late filers tend to be the taxpayers most subject to procrastination. Hence, if the treatment tends to be particularly effective for taxpayers who do not pay because of procrastination, we might expect to observe larger treatment effects among late filers. Column (1) of Table 7 provides suggestive evidence that is consistent with this hypothesis: The estimated effect of the detailed penalty for late filers is 1.8 percentage points larger than that for filers who are not late, although the difference is only marginally statistically significant ($p \approx 0.10$). In contrast, for the general penalty and social norms treatments (Columns (2) and (3)), the estimated treatment effects are similar for late and on-time filers.

Next, we examine how the treatment effect varies based on factors that may be associated with misperceptions about the financial incentives for timely payment. To the extent the treatment operates by increasing the perceived costs of delaying payment, we would expect to observe larger treatment effects for taxpayers who, absent the treatment, would have been more likely to be inattentive to or unaware of the incentives for timely payment. To do so, we consider two demographic characteristics: the taxpayer's age and Colorado residency status. Younger taxpayers may be less aware of the penalties because they have had fewer years of experience interacting with the tax system. Similarly, although the majority of our sample reside in Colorado, approximately 10 percent do not. Out-of-state taxpayers may be less familiar with Colorado laws or

[16] This assumption implies that, on average, a taxpayer would have more than twice as hard of a time coming up with $2,000 relative to $1,000. This strikes us as plausible; someone who needs cash may have less costly sources from which to obtain it and more costly sources that must be turned to once the less costly sources have been exhausted.

[17] Recall that filing a return does not require paying the associated balance due at the same time.

unaware that the penalties for nonpayment apply to them. For both penalty treatments, we observe statistically significantly larger effects for younger taxpayers. In addition, we observe substantially larger effects for nonresidents than for residents for the detailed penalty treatment, but not for the general penalty treatment, although the difference is not statistically significant for either group.[18] Overall, these findings are consistent with the hypothesis that the treatment operates by correcting misperceptions about the incentives for timely payment.

Finally, we consider how the treatment effect varied based on taxpayer income and education. Although we cannot observe these characteristics at the individual level, we proxy for them using ZIP code level measures derived from the ACS. Income and education may mediate the effect of the treatment on taxpayer behavior in several ways. First, taxpayers with higher income and education may be more financially sophisticated and so may have more awareness of the penalties even absent receiving the letter. On the other hand, it is also possible that financial literacy affects the degree to which taxpayers absorb the information contained in the letters. Last, both education and income may proxy for lower liquidity costs by being associated with the ease at which taxpayers may be able to borrow funds from others in their networks. The results in Table 7 do not provide systematic evidence of treatment effect heterogeneity based on income and education.[19]

VI. DISCUSSION AND CONCLUSION

We document differences in the effectiveness of delinquency notices sent to taxpayers based on whether the notices make salient the details of financial incentives for timely payment. In contrast, we find that simply emphasizing the existence of a penalty, or emphasizing social norms against late payment, exerts smaller or near-zero effects. Because of our large sample, these effects are precisely estimated.

Our finding that small differences in notice wording can affect behavior underscores the importance of tax authorities paying attention to such issues when determining how to communicate with taxpayers. Reducing the total number of delinquent taxpayer accounts is valuable for taxing authorities because it reduces costs associated with further outreach to delinquent taxpayers such as additional mailings, phone calls, or

[18] Colorado residency could also matter with respect to the social norms treatment because the content of that communication focuses on the behavior of Colorado residents. To the extent that descriptive norms for one's own community are what matter for shaping behavior, we would predict that the direction of this interaction would be negative for nonresidents. However, Column (3) of Table 7 shows that we observe a larger effect of the social norms treatment for nonresidents compared to residents, although the difference is not statistically significant.

[19] Column (3) of Table 7 shows marginally significant effects of income (negative) and education (positive) on the effectiveness of the social norms treatment. Because income and education are so highly correlated, we also consider an F-test of the null hypothesis that the coefficients on the income and education interactions are jointly zero; we are unable to reject this null hypothesis at conventional levels of statistical significance ($p = 0.14$).

enforcement actions. For many taxpayers, the effects of the presentational differences we study are modest — an increase in the payment rate of, at most, a few percent. However, delinquency notices of the type we study are already routinely sent to taxpayers; there is essentially *no additional cost* to designing the notice to have a more effective presentation rather than a less effective one. And, to the extent a tax authority is focusing on reaching one of the groups of taxpayers we observe to be particularly responsive (such as younger taxpayers), the benefits of the more effective presentation are likely to be greater.

Although we find evidence that the detailed penalty treatment positively affected the rate of on-time payments, a limitation of our experimental design is that our results do not speak to which aspects of that treatment are responsible for the increase in effectiveness. In particular, the detailed penalty treatment combines several changes relative to the control version of the letter. At the same time, comparing the effectiveness of the detailed and general penalty treatments does provide suggestive evidence about one aspect of the letter design — the choice of the level of generality at which the penalty is described. There, our finding that the point estimates associated with the detailed penalty letter are larger in magnitude than those associated with the general penalty letter suggests that additional precision can be helpful and need not contribute to information overload. Additionally, our experiment tests the effectiveness of emphasizing penalties and social norms separately; however, the interaction of the two message types may prove to be more or less effective than one message in isolation (Haynes et al., 2013). Future research could test the effect of interacting these two types of treatments or explore the relative importance of the elements of the detailed penalty treatment by considering intermediate versions of that treatment, varying which elements are included and which are left out.

Comparing our results to prior findings from behavioral public finance, our results are qualitatively consistent with the prior literature that has focused on U.S. taxing authorities. For example, Chirico et al. (2019) and Meiselman (2018) investigate the effect of communications by tax authorities in U.S. municipalities on property tax payment and income tax filing, respectively, and find that notices that emphasize financial penalties result in higher payment rates than notices emphasizing social norms. Chirico et al. (2019) find effects of substantially larger magnitudes than ours, although for the most part the confidence interval for their estimated effects overlaps with ours.[20] One possible explanation for the smaller effect we observe could be that a larger number of taxpayers in our sample were already aware that the state could impose a penalty compared to fewer taxpayers aware of the financial consequences of nonpayment of property taxes to a municipality (such as a property tax lien). Alternatively, it could be that taxpayers were more concerned with avoiding the penalties for nonpayment of property taxes that Chirico et al. (2019) emphasized in their letters (such as imposition of a lien) relative to the relatively modest financial penalties described in the letter by the DOR.

[20] Meiselman (2018) also finds large effects, but the outcome he studies is tax filing, not delinquent payments.

ACKNOWLEDGMENTS AND DISCLAIMER

We thank Christopher Muntean, James Viano, Kevin Furman, Heidi Humphreys, and Barbara Brohl at the Colorado DOR and David Padrino and David Demres at the Colorado Governor's Office for their efforts in data collection and implementation of this study. We are grateful to Ingrid Ellen, Lauren Jones, Jonathan Meer, Ben Meiselman, and Katherine Michelmore for conversations and suggestions that have greatly improved the quality of this project. Sarah Kotb and Vedant Vohra provided excellent research assistance. The views expressed here are our own and do not necessarily reflect the views of the State of Colorado.

DISCLOSURE

The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

REFERENCES

Aizer, Anna, 2007. "Public Health Insurance, Program Take-Up, and Child Health." *Review of Economics and Statistics* 89 (3), 400–415.

Allingham, Michael G., and Agnar Sandmo, 1972. "Income Tax Evasion: A Theoretical Analysis." *Journal of Public Economics* 1 (3–4), 323–338.

Armour, Philip, 2018. "The Role of Information in Disability Insurance Application: An Analysis of the Social Security Statement Phase-In." *American Economic Journal: Economic Policy* 10 (3), 1–41.

Barr, Andrew, and Sarah Turner, 2018. "A Letter and Encouragement: Does Information Increase Postsecondary Enrollment of UI Recipients?" *American Economic Journal: Economic Policy* 10 (3), 42–68.

Bérgolo, Marcelo L., Rodrigo Ceni, Guillermo Cruces, Matias Giaccobasso, and Ricardo Perez-Truglia, 2017. "Tax Audits as Scarecrows: Evidence from a Large-Scale Field Experiment." NBER Working Paper 23631. National Bureau of Economic Research, Cambridge, MA.

Bhargava, Saurabh, and Dayanand Manoli, 2015. "Psychological Frictions and the Incomplete Take-up of Social Benefits: Evidence from an IRS Field Experiment." *American Economic Review* 105 (11), 3489–3529.

Biddle, Nicholas, Katja M. Fels, and Mathias Sinning, 2018. "Behavioral Insights on Business Taxation: Evidence from Two Natural Field Experiments." *Journal of Behavioral and Experimental Finance* 18, 30–49.

Blumenthal, Marsha, Charles Christian, Joel Slemrod, and Matthew G. Smith, 2001. "Do Normative Appeals Affect Tax Compliance? Evidence from a Controlled Experiment in Minnesota." *National Tax Journal* 54 (1), 125–138.

Bott, Kristina M., Alexander W. Cappelen, Erik Ø. Sørensen, and Bertil Tungodden, forthcoming. "You've Got Mail: A Randomized Field Experiment on Tax Evasion." *Management Science*.

Castro, Lucio, and Carlos Scartascini, 2015. "Tax Compliance and Enforcement in the Pampas: Evidence from a Field Experiment." *Journal of Economic Behavior & Organization* 116, 65–82.

Chirico, Michael, Robert Inman, Charles Loeffler, John MacDonald, and Holger Sieg, 2019. "Deterring Property Tax Delinquency in Philadelphia: An Experimental Evaluation of Nudge Strategies." *National Tax Journal* 72 (3), 479–506.

De Neve, Jan-Emmanuel, Clement Imbert, Johannes Spinnewijn, Teodora Tsankova, and Maarten Luts, 2019. "How to Improve Tax Compliance? Evidence from Population-Wide Experiments in Belgium." Saïd Business School Working Paper 7. Saïd Business School, Oxford, UK.

Del Carpio, Lucia, 2013. "Are the Neighbors Cheating? Evidence from a Social Norm Experiment on Property Taxes in Peru." Working Paper.

Dwenger, Nadja, Henrik Kleven, Imran Rasul, and Johannes Rincke, 2016. "Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany." *American Economic Journal: Economic Policy* 8 (3), 203–32.

Fellner, Gerlinde, Rupert Sausgruber, and Christian Traxler, 2013. "Testing Enforcement Strategies in the Field: Threat, Moral Appeal and Social Information." *Journal of the European Economic Association* 11 (3), 634–660.

Gangl, Katharina, Benno Torgler, Erich Kirchler, and Eva Hofmann, 2014. "Effects of Supervision on Tax Compliance: Evidence from a Field Experiment in Austria." *Economics Letters* 123 (3), 378–382.

Goldin, Jacob, Tatiana Homonoff, and Will Tucker-Ray, 2017. "Retirement Contribution Rate Nudges and Plan Participation: Evidence from a Field Experiment." *American Economic Review* 107 (5), 456–461.

Guyton, John, Dayanand S. Manoli, Brenda Schafer, and Michael Sebastiani, 2016. "Reminders & Recidivism: Evidence from Tax Filing & EITC Participation among Low-Income Nonfilers." NBER Working Paper No. 21904. National Bureau of Economic Research, Cambridge, MA.

Hallsworth, Michael, John A. List, Robert D. Metcalfe, and Ivo Vlaev, 2017. "The Behavioralist as Tax Collector: Using Natural Field Experiments to Enhance Tax Compliance." *Journal of Public Economics* 148, 14–31.

Haynes, Laura C., Donald P. Green, Rory Gallagher, Peter John, and David J. Torgerson, 2013. "Collection of Delinquent Fines: An Adaptive Randomized Trial to Assess the Effectiveness of Alternative Text Messages." *Journal of Policy Analysis and Management* 32 (4), 718–730.

John, Peter, and Toby Blume, 2018. "How Best to Nudge Taxpayers? The Impact of Message Simplification and Descriptive Social Norms on Payment Rates in a Central London Local Authority." *Journal of Behavioral Public Administration* 1 (1), 1–11.

Kahneman, Daniel, and Amos Tversky, 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47 (2), 263–292.

Kettle, Stewart, Marco Hernandez, Simon Ruda, and Michael A Sanderson, 2016. "Behavioral Interventions in Tax Compliance: Evidence from Guatemala." World Bank Group Working Paper 7690. World Bank Group, Washington, DC.

Kleven, Henrik Jacobsen, Martin B Knudsen, Claus Thustrup Kreiner, Søren Pedersen, and Emmanuel Saez, 2011. "Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark." *Econometrica* 79 (3), 651–692.

Larkin, Christopher, Michael Sanders, Isabelle Andresen, and Felicity Algate, 2019. "Testing Local Descriptive Norms and Salience of Enforcement Action: A Field Experiment to Increase Tax Collection." *Journal of Behavioral Public Administration* 2 (1), 1–11.

Manoli, Dayanand S., and Nicholas Turner, 2014. "Nudges and Learning: Evidence from Informational Interventions for Low-Income Taxpayers." NBER Working Paper No. 20718. National Bureau of Economic Research, Cambridge, MA.

Meiselman, Ben S., 2018. "Ghostbusting in Detroit: Evidence on Nonfilers from a Controlled Field Experiment." *Journal of Public Economics* 158, 180–193.

Office of the State Controller, Department of Personnel & Administration, 2016. "Colorado Comprehensive Annual Financial Report." Available online at https://www.colorado.gov/pacific/sites/default/files/CAFR16.pdf.

Perez-Truglia, Ricardo, and Ugo Troiano, 2015. "Shaming Tax Delinquents: Theory and Evidence from a Field Experiment in the United States." NBER Working Paper No. 21264. National Bureau of Economic Research, Cambridge, MA.

Pomeranz, Dina, and José Vila-Belda, 2019. "Taking State-Capacity Research to the Field: Insights from Collaborations with Tax Authorities." *Annual Review of Economics* 11, 755–781.

Slemrod, Joel, 2018. "Tax Compliance and Enforcement: New Research and Its Policy Implications." NBER Working Paper 24799. National Bureau of Economic Research, Cambridge, MA.

Slemrod, Joel, Marsha Blumenthal, and Charles Christian, 2001. "Taxpayer Response to an Increased Probability of Audit: Evidence from a Controlled Experiment in Minnesota." *Journal of Public Economics* 79 (3), 455–483.

APPENDIX

Appendix Table 1
Outcome Variable Means

	Treatment Groups			
	Detailed Penalty	General Penalty	Social Norms	Control
Paid in full (%)	34.8	34.4	33.7	33.9
Payment plan (%)	4.9	4.6	4.2	4.5
Paid or payment plan (%)	39.8	39.0	38.0	38.4
Tax liability collected ($)	150	147	138	146
N	22,571	22,625	22,613	22,540

Outcomes: indicators for making a full payment, creating a payment plan, and either making a full payment or creating a payment plan by statutory deadline. The tax liability collected is equal to a taxpayer's balance due if the taxpayer made a full payment or created a payment plan by statutory deadline.

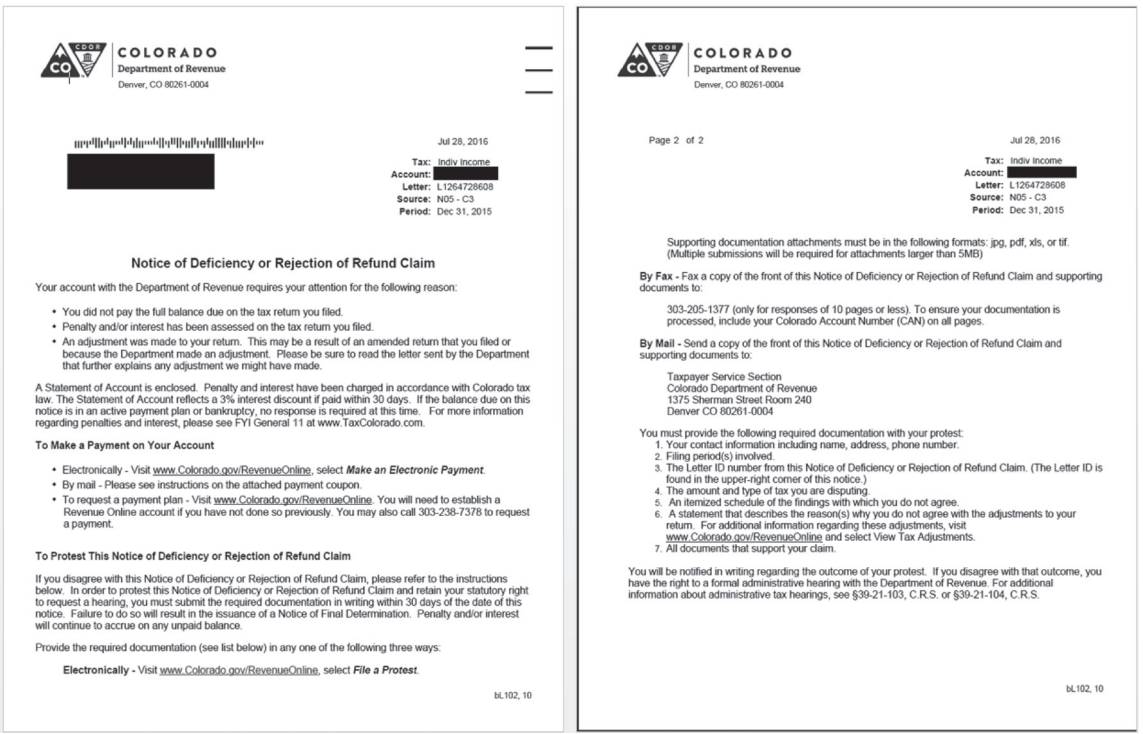
Appendix Table 2
Treatment Effect Heterogeneity by Balance Due—Paid within 100 Days

	(1) Low Balance	(2) Medium Balance	(3) High Balance
Detailed penalty	1.471*	1.426*	–0.633
	(0.834)	(0.824)	(0.590)
General penalty	1.244*	0.819	–0.653
	(0.700)	(0.754)	(0.618)
Social norms	0.987	0.539	–1.166*
	(0.739)	(0.810)	(0.634)
Control group mean	65.943	33.010	19.944
N	30,261	30,020	30,068

Outcomes: An indicator for making a full payment within 100 days of the NOD. Columns (1)–(3) present results by the amount of the balance due, where low balance is less than $95, medium balance is between $95 and $433, and high balance is more than $433. All regressions include age controls, balance controls, whether the return was filed on time, and ZIP code fixed effects. Age controls include a continuous age variable, dummy variables for four age categories, and an interaction of the dummies with the continuous age variable. Similarly, the balance controls include a continuous variable for the log of the balance due, dummies for six balance categories, and an interaction of the dummies with the continuous balance variable. All columns report results from a linear probability model. Units are percentage points (0–100). Robust standard errors are in parentheses. *$p < 0.1$, **$p < 0.05$, and ***$p < 0.01$.

Appendix Figure 1

Control Letter



COLORADO
Department of Revenue
Denver, CO 80261-0004

[redacted]

Jul 28, 2016

Tax: Indiv Income
Account: [redacted]
Letter: L1264728608
Source: N05 - C3
Period: Dec 31, 2015

Notice of Deficiency or Rejection of Refund Claim

Your account with the Department of Revenue requires your attention for the following reason:

- You did not pay the full balance due on the tax return you filed.
- Penalty and/or interest has been assessed on the tax return you filed.
- An adjustment was made to your return. This may be a result of an amended return that you filed or because the Department made an adjustment. Please be sure to read the letter sent by the Department that further explains any adjustment we might have made.

A Statement of Account is enclosed. Penalty and interest have been charged in accordance with Colorado tax law. The Statement of Account reflects a 3% interest discount if paid within 30 days. If the balance due on this notice is in an active payment plan or bankruptcy, no response is required at this time. For more information regarding penalties and interest, please see FYI General 11 at www.TaxColorado.com.

To Make a Payment on Your Account

- Electronically - Visit www.Colorado.gov/RevenueOnline, select ***Make an Electronic Payment***.
- By mail - Please see instructions on the attached payment coupon.
- To request a payment plan - Visit www.Colorado.gov/RevenueOnline. You will need to establish a Revenue Online account if you have not done so previously. You may also call 303-238-7378 to request a payment.

To Protest This Notice of Deficiency or Rejection of Refund Claim

If you disagree with this Notice of Deficiency or Rejection of Refund Claim, please refer to the instructions below. In order to protest this Notice of Deficiency or Rejection of Refund Claim and retain your statutory right to request a hearing, you must submit the required documentation in writing within 30 days of the date of this notice. Failure to do so will result in the issuance of a Notice of Final Determination. Penalty and/or interest will continue to accrue on any unpaid balance.

Provide the required documentation (see list below) in any one of the following three ways:

Electronically - Visit www.Colorado.gov/RevenueOnline, select ***File a Protest***.

bL102, 10

COLORADO
Department of Revenue
Denver, CO 80261-0004

Page 2 of 2

Jul 28, 2016

Tax: Indiv Income
Account: [redacted]
Letter: L1264728608
Source: N05 - C3
Period: Dec 31, 2015

Supporting documentation attachments must be in the following formats: jpg, pdf, xls, or tif. (Multiple submissions will be required for attachments larger than 5MB)

By Fax - Fax a copy of the front of this Notice of Deficiency or Rejection of Refund Claim and supporting documents to:

303-205-1377 (only for responses of 10 pages or less). To ensure your documentation is processed, include your Colorado Account Number (CAN) on all pages.

By Mail - Send a copy of the front of this Notice of Deficiency or Rejection of Refund Claim and supporting documents to:

Taxpayer Service Section
Colorado Department of Revenue
1375 Sherman Street Room 240
Denver CO 80261-0004

You must provide the following required documentation with your protest:

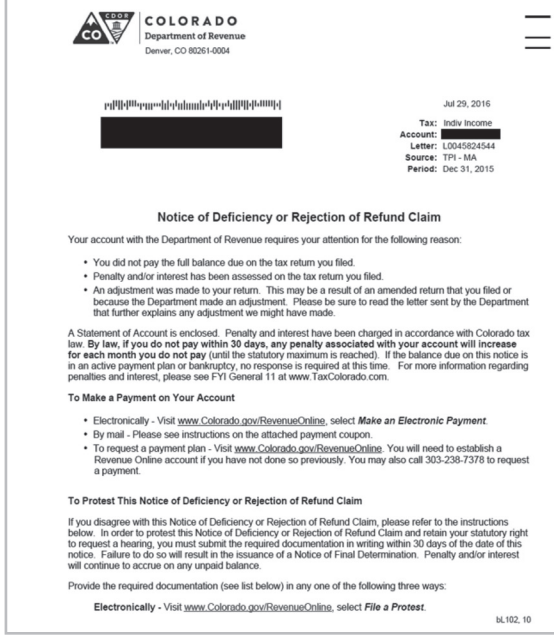
1. Your contact information including name, address, phone number.
2. Filing period(s) involved.
3. The Letter ID number from this Notice of Deficiency or Rejection of Refund Claim. (The Letter ID is found in the upper-right corner of this notice.)
4. The amount and type of tax you are disputing.
5. An itemized schedule of the findings with which you do not agree.
6. A statement that describes the reason(s) why you do not agree with the adjustments to your return. For additional information regarding these adjustments, visit www.Colorado.gov/RevenueOnline and select View Tax Adjustments.
7. All documents that support your claim.

You will be notified in writing regarding the outcome of your protest. If you disagree with that outcome, you have the right to a formal administrative hearing with the Department of Revenue. For additional information about administrative tax hearings, see §39-21-103, C.R.S. or §39-21-104, C.R.S.

bL102, 10

Appendix Figure 2a

Detailed Penalty Letter



COLORADO
Department of Revenue
Denver, CO 80261-0004

[redacted]

Jul 29, 2016

Tax: Indiv Income
Account: [redacted]
Letter: L0045824544
Source: TPI - MA
Period: Dec 31, 2015

Notice of Deficiency or Rejection of Refund Claim

Your account with the Department of Revenue requires your attention for the following reason:

- You did not pay the full balance due on the tax return you filed.
- Penalty and/or interest has been assessed on the tax return you filed.
- An adjustment was made to your return. This may be a result of an amended return that you filed or because the Department made an adjustment. Please be sure to read the letter sent by the Department that further explains any adjustment we might have made.

A Statement of Account is enclosed. Penalty and interest have been charged in accordance with Colorado tax law. **By law, if you do not pay within 30 days, any penalty associated with your account will increase for each month you do not pay** (until the statutory maximum is reached). If the balance due on this notice is in an active payment plan or bankruptcy, no response is required at this time. For more information regarding penalties and interest, please see FYI General 11 at www.TaxColorado.com.

To Make a Payment on Your Account

- Electronically - Visit www.Colorado.gov/RevenueOnline, select ***Make an Electronic Payment***.
- By mail - Please see instructions on the attached payment coupon.
- To request a payment plan - Visit www.Colorado.gov/RevenueOnline. You will need to establish a Revenue Online account if you have not done so previously. You may also call 303-238-7378 to request a payment.

To Protest This Notice of Deficiency or Rejection of Refund Claim

If you disagree with this Notice of Deficiency or Rejection of Refund Claim, please refer to the instructions below. In order to protest this Notice of Deficiency or Rejection of Refund Claim and retain your statutory right to request a hearing, you must submit the required documentation in writing within 30 days of the date of this notice. Failure to do so will result in the issuance of a Notice of Final Determination. Penalty and/or interest will continue to accrue on any unpaid balance.

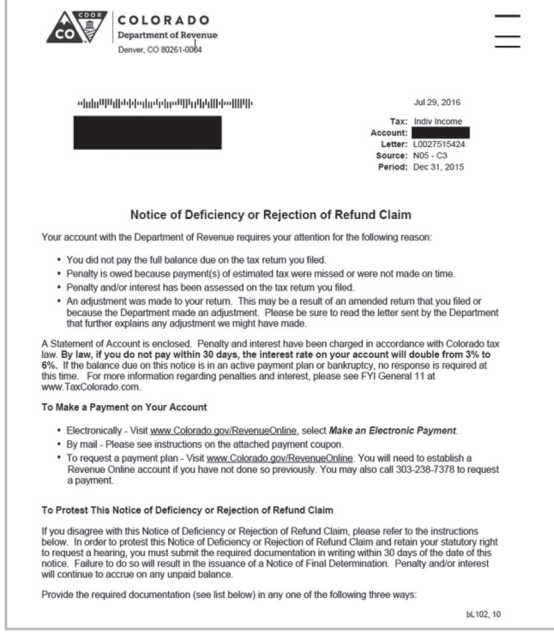
Provide the required documentation (see list below) in any one of the following three ways:

Electronically - Visit www.Colorado.gov/RevenueOnline, select ***File a Protest***.

bL102, 10

Appendix Figure 2b

General Penalty Letter



COLORADO
Department of Revenue
Denver, CO 80261-0004

[redacted]

Jul 29, 2016

Tax: Indiv Income
Account: [redacted]
Letter: L0027515424
Source: N05 - C3
Period: Dec 31, 2015

Notice of Deficiency or Rejection of Refund Claim

Your account with the Department of Revenue requires your attention for the following reason:

- You did not pay the full balance due on the tax return you filed.
- Penalty is owed because payment(s) of estimated tax were missed or were not made on time.
- Penalty and/or interest has been assessed on the tax return you filed.
- An adjustment was made to your return. This may be a result of an amended return that you filed or because the Department made an adjustment. Please be sure to read the letter sent by the Department that further explains any adjustment we might have made.

A Statement of Account is enclosed. Penalty and interest have been charged in accordance with Colorado tax law. **By law, if you do not pay within 30 days, the interest rate on your account will double from 3% to 6%.** If the balance due on this notice is in an active payment plan or bankruptcy, no response is required at this time. For more information regarding penalties and interest, please see FYI General 11 at www.TaxColorado.com.

To Make a Payment on Your Account

- Electronically - Visit www.Colorado.gov/RevenueOnline, select ***Make an Electronic Payment***.
- By mail - Please see instructions on the attached payment coupon.
- To request a payment plan - Visit www.Colorado.gov/RevenueOnline. You will need to establish a Revenue Online account if you have not done so previously. You may also call 303-238-7378 to request a payment.

To Protest This Notice of Deficiency or Rejection of Refund Claim

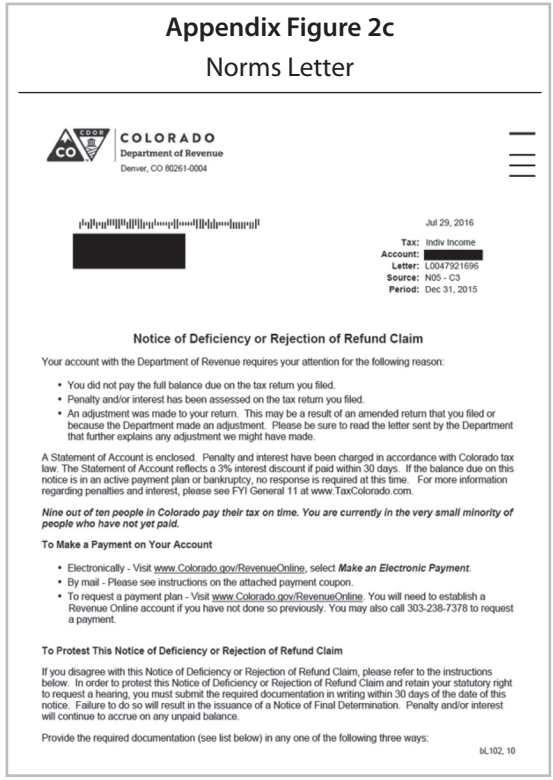
If you disagree with this Notice of Deficiency or Rejection of Refund Claim, please refer to the instructions below. In order to protest this Notice of Deficiency or Rejection of Refund Claim and retain your statutory right to request a hearing, you must submit the required documentation in writing within 30 days of the date of this notice. Failure to do so will result in the issuance of a Notice of Final Determination. Penalty and/or interest will continue to accrue on any unpaid balance.

Provide the required documentation (see list below) in any one of the following three ways:

bL102, 10

Appendix Figure 2c

Norms Letter



COLORADO
Department of Revenue
Denver, CO 80261-0004

[redacted]

Jul 29, 2016

Tax: Indiv Income
Account: [redacted]
Letter: L0047921696
Source: N05 - C3
Period: Dec 31, 2015

Notice of Deficiency or Rejection of Refund Claim

Your account with the Department of Revenue requires your attention for the following reason:

- You did not pay the full balance due on the tax return you filed.
- Penalty and/or interest has been assessed on the tax return you filed.
- An adjustment was made to your return. This may be a result of an amended return that you filed or because the Department made an adjustment. Please be sure to read the letter sent by the Department that further explains any adjustment we might have made.

A Statement of Account is enclosed. Penalty and interest have been charged in accordance with Colorado tax law. The Statement of Account reflects a 3% interest discount if paid within 30 days. If the balance due on this notice is in an active payment plan or bankruptcy, no response is required at this time. For more information regarding penalties and interest, please see FYI General 11 at www.TaxColorado.com.

Nine out of ten people in Colorado pay their tax on time. You are currently in the very small minority of people who have not yet paid.

To Make a Payment on Your Account

- Electronically - Visit www.Colorado.gov/RevenueOnline, select ***Make an Electronic Payment***.
- By mail - Please see instructions on the attached payment coupon.
- To request a payment plan - Visit www.Colorado.gov/RevenueOnline. You will need to establish a Revenue Online account if you have not done so previously. You may also call 303-238-7378 to request a payment.

To Protest This Notice of Deficiency or Rejection of Refund Claim

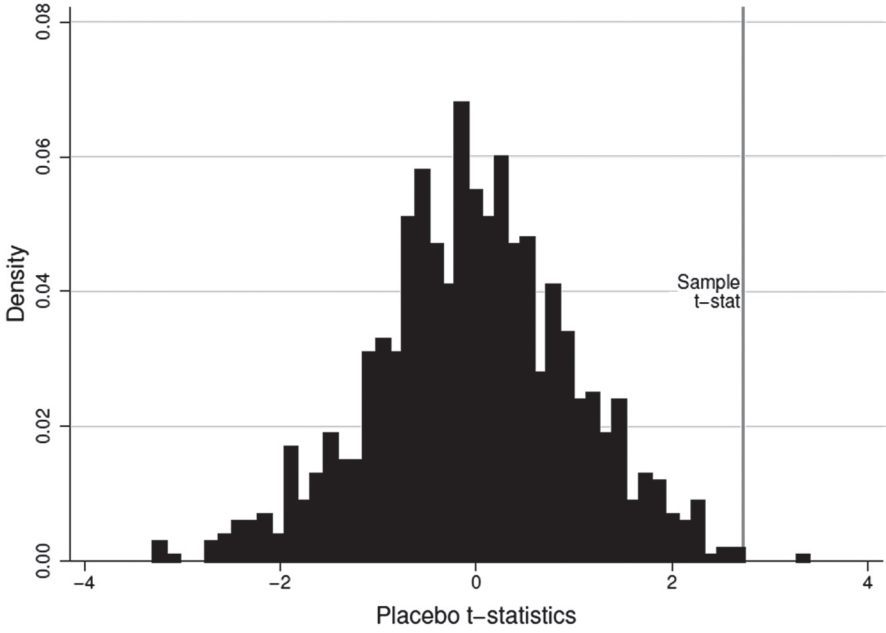
If you disagree with this Notice of Deficiency or Rejection of Refund Claim, please refer to the instructions below. In order to protest this Notice of Deficiency or Rejection of Refund Claim and retain your statutory right to request a hearing, you must submit the required documentation in writing within 30 days of the date of this notice. Failure to do so will result in the issuance of a Notice of Final Determination. Penalty and/or interest will continue to accrue on any unpaid balance.

Provide the required documentation (see list below) in any one of the following three ways:

bl.102, 10

Appendix Figure 3

Randomization Inference



This figure plots the distribution of treatment effect t-statistics for the detailed penalty generated from 1,000 random reassignments of the treatment across taxpayers in the sample population. The estimated t-statistics correspond to the specification reported in Column (1) of Table 3. The vertical line denotes the t-statistic estimated using the actual sample population. The sample is restricted to individuals in the detailed penalty treatment or control group.