Online Appendix: For Online Publication Only

Tax Audits as Scarecrows: Evidence from a Large-Scale Field Experiment

Marcelo Bergolo, Rodrigo Ceni, Guillermo Cruces, Matias Giaccobasso, and Ricardo Perez-Truglia

04/11/21

A Letters and Survey

This appendix presents samples of the five types of letters sent in our experiment: the *baseline* letter (A.1), the *audit-statistics* letter (A.2), the *audit-threat* letter (A.3), the *audit-endogeneity* letter (A.4), and the *public goods* letter (A.5). Appendix A.6 presents a sample of the email invitation to complete the online survey sent by the IRS, and Appendix A.7 presents the questionnaire module about perceptions of audit probabilities and penalty rates included in the survey.

A.1 Sample Letter: *Baseline* Letter

DGI | Dirección General Impositiva

Montevideo, August 20th 2015

Mr../Ms. Taxpayer:

The DGI has the authority to perform inspections (see Art. 68 of the tax code) and routine audits of taxpayers on the basis of crosschecks and assessment of data compiled to detect oversights and inconsistency on tax returns as well as pending tax debts.

The aim of the DGI, and the primary challenge it faces, is to ensure the collection of revenue to sustain life in society. Additionally, its task is to generate a framework of fair and transparent competition where the failure of some to meet their obligations does not have an unfavorable impact on honest taxpayers. In order to meet these goals, inspections are performed in a routine fashion.

Your micro, small, or medium-sized business has been randomly selected to receive this information. *It is solely for your information and its receipt does not require you to present any documentation to the DGI offices.*

We ask you to comply with your tax obligations for the sake of the country we all want, a more and more developed Uruguay with greater and greater social cohesion.

Sincerely,

El Director General de Rentas

Lic. Joaquín Serra

Collection and Controls Division
Internal Revenues Services

A.2 Sample Letter: *Audit-Statistics* Letter

DGI | DIRECCIÓN GENERAL IMPOSITIVA

Montevideo, August 20th 2015

Mr../Ms. Taxpayer:

The DGI has the authority to perform inspections (see Art. 68 of the tax code) and routine audits of taxpayers on the basis of crosschecks and assessment of data compiled to detect oversights and inconsistency on tax returns as well as pending tax debts.

On the basis of historical information on similar businesses, there is a probability of p% that the tax returns you filed for this year will be audited in at least one of the coming three years. If, pursuant to that auditing, it is determined that tax evasion has occurred, you will be required to pay not only the amount previously unpaid, but also a fee of approximately *θ%* of that amount.

The aim of the DGI, and the primary challenge it faces, is to ensure the collection of revenue to sustain life in society. Additionally, its task is to generate a framework of fair and transparent competition where the failure of some to meet their obligations does not have an unfavorable impact on honest taxpayers. In order to meet these goals, inspections are performed in a routine fashion.

Your micro, small, or medium-sized business has been randomly selected to receive this information. *It is solely for your information and its receipt does not require you to present any documentation to the DGI offices.*

We ask you to comply with your tax obligations for the sake of the country we all want, a more and more developed Uruguay with greater and greater social cohesion.

Sincerely,

El Director General de Rentas

Lic. Joaquín Serra

Collection and Controls Division
Internal Revenues Services

A.3 Sample Letter: *Audit-Threat* Letter

DGI Dirección General Impositiva

Montevideo, August 20th 2015

Mr../Ms. Taxpayer:

The DGI has the authority to perform inspections (see Art. 68 of the tax code) and routine audits of taxpayers on the basis of crosschecks and assessment of data compiled to detect oversights and inconsistency on tax returns as well as pending tax debts.

We would like to inform you that the business you represent is one of a group of firms pre-selected for auditing in 2016. A p% of the firms in that group will then be randomly selected for auditing.

The aim of the DGI, and the primary challenge it faces, is to ensure the collection of revenue to sustain life in society. Additionally, its task is to generate a framework of fair and transparent competition where the failure of some to meet their obligations does not have an unfavorable impact on honest taxpayers. In order to meet these goals, inspections are performed in a routine fashion.

Your micro, small, or medium-sized business has been randomly selected to receive this information. *It is solely for your information and its receipt does not require you to present any documentation to the DGI offices.*

We ask you to comply with your tax obligations for the sake of the country we all want, a more and more developed Uruguay with greater and greater social cohesion.

Sincerely,

El Director General de Rentas

Lic. Joaquín Serra

Collection and Controls Division
Internal Revenues Services

A.4 Sample Letter: *Audit-Endogeneity* Letter

DGI Dirección General Impositiva

Montevideo, August 20th 2015

Mr../Ms. Taxpayer:

The DGI has the authority to perform inspections (see Art. 68 of the tax code) and routine audits of taxpayers on the basis of crosschecks and assessment of data compiled to detect oversights and inconsistency on tax returns as well as pending tax debts.

The DGI uses data on thousands of taxpayers to detect firms that may be evading taxes; most of its audits are aimed at those firms. Evading taxes, then, doubles your chances of being audited.

The aim of the DGI, and the primary challenge it faces, is to ensure the collection of revenue to sustain life in society. Additionally, its task is to generate a framework of fair and transparent competition where the failure of some to meet their obligations does not have an unfavorable impact on honest taxpayers. In order to meet these goals, inspections are performed in a routine fashion.

Your micro, small, or medium-sized business has been randomly selected to receive this information. *It is solely for your information and its receipt does not require you to present any documentation to the DGI offices.*

We ask you to comply with your tax obligations for the sake of the country we all want, a more and more developed Uruguay with greater and greater social cohesion.

Sincerely,

El Director General de Rentas

Lic. Joaquín Serra

Collection and Controls Division
Internal Revenues Services

A.5 Sample Letter: *Public-Goods* Letter

DGI Dirección General Impositiva

Montevideo, August 20th 2015

Mr../Ms. Taxpayer:

The DGI has the authority to perform inspections (see Art. 68 of the tax code) and routine audits of taxpayers on the basis of crosschecks and assessment of data compiled to detect oversights and inconsistency on tax returns as well as pending tax debts.

If those who currently evade their tax obligations were to evade 10% less, the additional revenue collected would enable all of the following: to supply 42,000 portable computers to school children; to build 4 high schools, 9 elementary schools, and 2 technical schools; to acquire 80 patrol cars and to hire 500 police officers; to add 87,000 hours of medical attention by doctors at public hospitals; to hire 660 teachers; to build 1,000 public housing units (50m2 per unit). There would be resources left over to reduce the fiscal burden. The tax behavior of each of us has direct effects on the lives of us all.

The aim of the DGI, and the primary challenge it faces, is to ensure the collection of revenue to sustain life in society. Additionally, its task is to generate a framework of fair and transparent competition where the failure of some to meet their obligations does not have an unfavorable impact on honest taxpayers. In order to meet these goals, inspections are performed in a routine fashion.

Your micro, small, or medium-sized business has been randomly selected to receive this information. *It is solely for your information and its receipt does not require you to present any documentation to the DGI offices.*

We ask you to comply with your tax obligations for the sake of the country we all want, a more and more developed Uruguay with greater and greater social cohesion.

Sincerely,

El Director General de Rentas

Lic. Joaquín Serra

Collection and Controls Division
Internal Revenues Services

A.6 Sample Letter: Invitation to the Online Survey

Dear Taxpayer:

The DGI's strategic objectives for this period include improving taxpayer services. In 2013, the first Survey on the Costs of Tax Compliance for Small and Medium-Sized Businesses was administered with the support of the Inter-American Center of Tax Administrations (CIAT) and the United Nations (UN). The DGI, in conjunction with a group of academics, has designed a new version of the survey (for more information, visit www.dgi.gub.uy). You can give us your answers on the website where you will find instructions on how to fill out the simple questionnaire; the entire process should take no more than fifteen minutes.

Respond to survey

To address these concerns, a random sample of taxpayers will receive a survey to be answered anonymously.

You are one of the randomly selected taxpayers, which is why you have received this communication. We are grateful for the time and effort you dedicate to assessing this questionnaire and to responding to it as precisely as possible.

Let me assure you that the survey is completely anonymous and the selection of recipients entirely random. The success of this project lies in the precision of your responses. It is on the basis of those responses and the real information they provide that the DGI will be able to hone the design, in the present and in the future, of its strategies to reduce the costs of compliance.

If you have any questions about this questionnaire, please send an e-mail to encuestas@cedlas.org.

We would like to thank you once again for your contribution to this project, which we are sure will benefit all taxpayers.

Sincerely,

Joaquín Serra
Director of the Income Tax Department

PS: If the "Respond to survey" link doesn't open, copy the following address in your browser:https://URL.

A.7 Survey Questionnaire

Introductory Text:

We would like you to respond to a survey about the costs of paying taxes. We hope you have the ten minutes that responding to the questionnaire will require. We are interested in your opinion and hope you will be frank in your responses, which are anonymous and used only for statistical purposes. We would like to thank you for your participation.

Questions Included in Main Module:

Q1) Have you been subject to a DGI audit (inspection or monitoring) at any point in the last three years?

□Yes.

□ No.

Q2)In your opinion,what is the probability that the tax returns filed by a company like yours be audited at least once in the next three years (from 0% to 100%)?

[ ] %

Q3) How sure are you of your response?

□ Not at all sure.

□ A little sure.

□Somewhat sure.

□ Very sure.

Q4)Let's imagine that a company like yours is audited and that tax evasion is detected. What, in your opinion, is the penalty (in %) as determined by law that the firm must pay in addition to the originally unpaid amount? For example, a fee of X% means that, for each $100 not paid, the firm would have to pay those original $100 plus $X in fees.

[ ] %

Q5) How sure are you of your response?

□ Not at all sure.

□ A little sure.

□Somewhat sure.

□ Very sure.

Q6) In your opinion, if a firm that evades taxes doubles the amount it is evading, what is the effect on its probability of being audited?

□ It would increase significantly.

□It would increase slightly.

□It would not change.

□It would diminish slightly.

□It would diminish significantly.

B Additional Results, Specifications, and Robustness Checks

B.1 Descriptive Statistics

Table B.1 reports firms' characteristics including VAT payments made in the three months before we sent the letters, the number of years each firm was registered with the IRS, the number of employees, and other basic variables. Column (2) provides these statistics for all firms in the main experimental sample. On average, firms in our sample had 4.8 employees and had been registered with the IRS for 15.2 years. Ten percent of the firms had been audited at least once sometime in the previous three years. For comparison, column (1) of Table B.1 shows the same statistics for the universe of all registered firms. By design, the firms in our experimental sample are small, both in terms of the number of employees and the level of VAT payments. Column (3) of Table B.1 provides statistical information on the secondary experimental sample (i.e., the *audit-threat* treatment arm). Despite some statistically significant differences between the two groups, the firms are broadly comparable in size. The main difference between firms in the two experimental samples is that the audit rates were four percentage points higher in the *audit-threat* sample, which is by design because the IRS selected firms classified as high risk for the *audit-threat* treatment arm, and such firms were more likely to have been targeted for audits in the past.

Column (4) reports the same descriptive statistics for the subsample of firms invited to answer the survey, and column (5) reports the p-value of the mean difference between firms that were invited and firms that were not. While the share of firms that paid VAT and the amount of VAT paid were not statistically different between the two groups of firms (p-values of 0.875 and 0.993 respectively), there were some statistically significant differences for other characteristics. Firms invited to the survey had been registered with the IRS for ten fewer months; they were two percentage points less likely to have been audited in the last three years than firms in the main experimental sample on average (p-values of 0.001 and <0.001 respectively). Firms in the survey sample had 1.6 more employees and were, on average, 11.5 percentage points less likely to have filed a comprehensive tax return in 2013 (p-value<0.001 for both). There are some additional statistically signficant differences in terms of sectors of operation. Firms in the service sector, as opposed to the retail goods sector, were overrepresented in the group of firms invited to participate in the survey (58.8% vs 49.6%, p-value<0.001). Overall, while differences in some characteristics are some statistically significant, most do not seem economically significant.

B.2 Delivery Status

An imbalance in the delivery rate of the letters across treatment arms could be a threat to the validity of our results. Since we used the postal service's certified delivery service, we have credible and complete information on the delivery status of each letter. To rule out the possibility of such an imblanace, Table B.2 reports the delivery status of the letters by treatment arm. Columns (1) through (4) report the distribution of delivery statuses by treatment arm for firms in the main sample, while column (5) reports the p-value of the joint equality test. In general, there are no differences by type of letter between the three treatment arms and the *baseline* letter. Only 19.9% of the letters sent to firms in the main sample were returned–a rate that varies little across treatment arms (p-value=0.290). When a letter's delivery status is "Returned" in the postal service dataset, the firm to which it was addressed is excluded from the analysis.

The only statistically significant difference between the treatment arms is the percentage of missing letters; it is smaller for firms that were selected for the *public-goods* message and the *audit-endogeneity* message that for firms selected for the *audit-statistics* message and the *baseline* message. Though statistically significant at a 5% level, these differences are not economically significant (6.7% and 6.2% for *public-goods* and *audit-endogeneity* respectively vs. 7.9% and 7.5% for *audit-statistics* and *baseline*).

Columns (6) through (8) provide an analogous breakdown for the secondary sample, where each delivery status is balanced across treatment arms. Compared to the main sample, the percentage of letters returned to the sender was higher (11.9% vs. 19.9%), perhaps because this sample was specifically selected by the IRS. However, there are no differences between the two treatment arms within the secondary sample.

B.3 Summary Statistics of Tax Payments

Table B.3 reports some descriptive statistics for the pre- and post-treatment periods on the outcome variables used throughout the paper for firms that received the *baseline* letter. The pre-treatment period covers the year immediately before the treatment (October 2014–September 2015) and the post-treatment period covers the twelve months immediately after the treatment (October 2015–September 2016). On average, the amount of VAT paid by a firm that received a *baseline* letter during the pre-treatment period was USD 7,770 and the median around USD 4,830, with a standard deviation of USD 8,210. In the subsequent year, the average amount of VAT paid was USD 6,470, the median USD 3,740, and the standard deviation USD 7,770. This represents a reduction of 16.7% in the average VAT payments between pre- and post-treatment periods. Since the group of firms analyzed consists mostly of small and medium-sized firms, this difference could be explained by a high turnover rate.

The bulk of total VAT payments is concurrent payments (about 95% of the total). Retroactive VAT payments represent only about 5% of total VAT payments made by the firms that received the *baseline* letter. The small share of retroactive VAT payments shows that most taxpayers do not make this type of payment. Indeed, the 75th percentile of this distribution is zero. The average amount of retroactive payments made by these firms during the pre-treatment period was USD 400, while during the post-treatment period it was about USD 300. The trends for both concurrent and retroactive payments are consistent with the trend observed for overall VAT payments–and the same holds true of other tax payments and amounts reported in VAT tax returns. Firms paid an average of USD 4,050 of other taxes in the pre-treatment period, and USD 3,300 in the post-treatment period. These amounts include other sales taxes, corporate income taxes, and others. The standard deviation was USD 8,540 in the pre-treatment period and USD 5,430 in the post-treatment period.

VAT tax returns provide additional information that firms report to the IRS regarding total VAT liabilities. The average final VAT liability for the *baseline* group is USD 7,790 in the pre-treatment period and USD 7,190 in the post-treatment period, and its evolution is similar to the sum of monthly VAT payments. The median VAT liability is USD 5,200 in the pre-treatment period and USD 3,840 in the post-treatment period, and the standard deviation is USD 7,910 and 9,930 respectively.

In addition to variables that capture the magnitude of VAT payments–variables more closely associated with intensive margin responses–Table B.3 reports some descriptive statistics about whether firms actually make payments and the number of payments they make. By construction, all the firms in the sample made at least one VAT payment in the pre-treatment period. Ninety-eight percent of the firms made more than three payments, and 89% made more than six payments. In the post-treatment period, 96% of the firms made at least one payment, 91% more than three, and 71% more than six. This pattern is also consistent with the evolution of the number of payments; it also suggests a high turnover rate for the firms in our sample.

This descriptive analysis indicates the importance of VAT to the Uruguayan tax structure. VAT payments are almost twice as high as other tax payments; they represent over 60% of total payments made by the firms in our sample.

B.4 Robustness Checks and Additional Results from Online Survey

B.4.1 Survey Results: Selection into the Survey

In this section, we report a series of additional results and robustness tests for the analysis conducted with the survey data.

Selective responsiveness is a potential concern about the survey data–a threat if the treatment itself induced differential response rates to the survey, or if particular groups of firms were more likely to participate in the survey. We present three pieces of evidence that contribute to our interpretation of the results.

First, we focus on the response rates to the survey and to the two questions about audit probability and penalty rate, that is, the ones relevant to our analysis. Table B.4 reports a series of statistics that shows the selection process from the experimental sample to the final subsample of firms used in the survey analysis, categorized by treatment arm. Invitations with the link to the survey were sent to all firms in our main and secondary samples that had reported their email address to the IRS. The total number of firms invited to the survey was 3,867, or about 23% of our main experimental sample.[75] The share of firms that started the survey (i.e., that answered at least the first two questions in the survey) was 24.5%, and those responses were balanced across treatments. The vast majority of the responses were from individuals who either identified themselves as owners or did not reply to the question about their role at the firm (about 76.5%). An individual who accessed the survey does not necessarily reach its relevant questions, even though they were placed at the beginning: 22.3% of the owners who started the survey did not answer the questions that collected information about prior beliefs of audit probability and penalty rate, which is comparable to the overall non-response rate for the survey. Furthermore, even if respondents reached that point in the survey (i.e., they reported a non-missing value in the previous question), 6.6% skipped the audit probability question and 8.6% the penalty question, a comparable average rate of skipping other questions in the survey (6.1%). We use these answers in our analysis, and in all cases the response rates, however measured, are balanced across treatment arms. We perform a more direct test of selective responsiveness by testing the effect of the treatment on the response rates to the two questions key to our analysis, that is, the ones on audit probability and penalty rate. To do this, we define a dummy variable that takes the value of one if the individual answered the audit probability or the penalty rate question and zero if they did not. We regress this dummy on an *audit-statistics* treatment indicator using

[75] While we sent some invitations to firms in the secondary sample, their responses are not used in our analysis because there were too few firms that satisfied both criteria.

our pooled control group as the reference group (*baseline* or *public-goods* letter recipients). Table B.5 reports the results of these estimates. The *audit-statistics* treatment does not affect the response rate to the audit probability question or the penalty rate question. To address possible differential drop-out rates, we also report the results restricting the analysis to individuals who answered all the survey questions up to the audit probability and penalty rate questions. The results remain the same. Both pieces of evidence suggest that there is no sign of differential behavior in answering the survey by treatment status.

Table B.6 complements this analysis with a balance test for some of the characteristics of the individual survey respondents and their firms who answered the audit probability and penalty rate questions. Column (1) presents information about the age, gender, and city of the respondent as well as some firm characteristics such as size, sector of operation, number of years registered with the IRS, number of locations, and number of employees for firms that received the *audit-statistics* letter. Column (2) does the same for firms that received the *public-goods* letter, column (3) for firms that received the *audit-endogeneity* letter, and column (4) for firms that received the *baseline* letter. Column (5) reports the p-value of the mean test for the four groups. All self-reported characteristics are balanced across treatment arms except for the age of the respondent, the number of years the firm has been registered with the IRS, and the percentage of firms with one employee. For those characteristics, the differences are economically irrelevant, though statistically significant at a 5% level.

Third, Figure B.1 reports a series of placebo tests where we replicate our estimation strategy to measure the effect of the *audit-statistics* message on questions that we do not expect to be affected by our treatment. Reported in other survey modules, these questions aimed to collect information on tax-compliance costs as explained in Section 2. We report the effects of the *audit-statistics* message on the answers to five placebo questions: 1) "On a scale from 1 to 5 where 1 is "Strongly disagree" and 5 is "Strongly agree," to what extent do you agree with the following statement: *Tax compliance generates non-pecuniary costs*"; 2) "On a scale from 1 to 5 where 1 is "Not at all stressful " and 5 is "Very stressful," please rate the level of stress generated by all the steps required to fulfill your tax obligations"; 3) "How much time do you spend informing yourself about your tax obligations?"; 4) "How much time do you spend registering all the transactions made by your firm?"; and 5) "What is the estimated monthly cost of all activities and supplies related to tax compliance?" The estimation strategy is identical to the one used for Figure 5 where the comparison group consists of firms that received the *baseline* letter and firms that received the *public-goods* letter. Figure B.2 contains a summary of the placebo tests compared to the results for the key variables of interest for our analysis (i.e., the ones reported in Figure 5) where, for comparison purposes, all estimates are expressed in standard deviations of the outcome variable for the control group. Figures B.1 and B.2 show that, as expected, the *audit-statistics* message had no effect on the responses to these placebo questions.

B.4.2 Survey Results: Robustness Checks

To analyze the survey, we pooled respondents from the *baseline* and the *public-goods* treatments to form a sufficiently large comparison group. The rationale was that neither of the two messages included information on audit probabilities or tax evasion penalty rates. In this appendix, we assess the robustness of the survey results in Section 5.1 to alternative definitions of the sample and the comparison group.

Panels (a) and (b) of Figure B.3 replicate the results in panels (a) and (b) of Figure 5. The shaded gray bars show the distribution of perceptions for the sixty-nine survey respondents that received the *baseline* letter only (Figure 5 relied on the 137 observations from the pooled *baseline* and *public-goods* groups). The dashed curve red line corresponds to the distribution of signals sent to the firms in the *audit-statistics* letters. Although slightly smaller than in the pooled control group, the average perceived audit probability for the *baseline* letter group (37.7%, in panel (a)) is still substantially higher than the 11.7% for the overall sample, a statistically significant difference. The results are also consistent with the main results when we look at the perceived penalty rate in panel (b). There are no statistically significant differences between the perceived penalty rate by the firms that received the *baseline* letter and our estimates from the overall data.

Second, we perform an additional robustness test of the results in Figure 5. To increase the likelihood that survey respondents were the ones who received our experimental messages, we replicate the analysis, this time restricting our sample to survey respondents who, in the survey, self-identified as firm owners. Presented in panels (c) and (d) of Figure B.3, the results with this restricted sample, which reduced the treatment group from 365 to 341 observations and the pooled control group from 137 to 125, are very similar to the results reported in the body of this paper. Our *audit-statistics* treatment significantly reduced the perceived average probability of audits, although it did not affect the average perception of penalty rates. Finally, panels (e) and (f) of Figure B.3 report the results of a similar analysis, but this time using all answers regardless of the respondent's self-reported role at the firm, that is, not only owners, but also managers, inhouse accountants, external tax advisors, and other employees. This modification increases the size of the treatment group from 365 to 465 and the size of the control group from 137 to 179. The results in terms of magnitude, direction, and statistical significance are the same.

An additional concern is that respondents may not have taken the survey seriously. One way of testing this is to analyze the robustness of our results without the respondents who indicate 50% as their perceived audit probability or penalty rate (Bruine de Bruin et al., 2002; Bruine de Bruin and Carman, 2012).[76] To address this concern we take two steps. First, we test whether the degree of certainty for individuals who indicate a 50% perceived audit probability or penalty rate is different from the degree of certainty for those who provide non-50% responses.[77] We present the results of this test in Figure B.4. Panel (a) of Figure B.4 reports the distributions of the degree of certainty in the answers provided to the audit probability question for 50% responses and non-50% responses. The average certainty for people with non-50% answers was 2.37 compared to a mean of 2.03 for people who answered 50% (measured on a scale from one to four, where one corresponds to an answer of "Not at all sure" and four to an answer of "Very sure"). This difference is statistically significant (p-value<0.001) and suggests that 50% answers indeed corresponded to individuals who did not feel sure about the answer they provided. The results are similar for Panel (b) of Figure B.4, which reports the results for certainty about the reported perceived penalty rate. Individuals with a 50% response feel less certain of their answers than individuals with a non-50% response. That does not necessarily mean, however, that individuals who answered 50% did not take the survey seriously. It simply reflects that they felt less certain about their answers. To rule out the possibility that our results could be driven by this type of response, Figure B.5 replicates our analysis reported in Figure 5 without individuals who provided 50% responses. Our findings hold, although our estimates are less precise due to the reduction in the number of answers included in the analysis.

If firms were rational, all these results would imply that firms would have paid less tax as a consequence of altering their beliefs. This is not, however, what we observe, which supports the hypothesis that the results are driven by the fear channel rather than by a rational re-optimization.

B.4.3 Survey Results: Beliefs About Audit Endogeneity

As in the case of the *audit-statistics* treatment arm, we conducted a survey of letter recipients that included a specific question to assess whether the information provided in the letter had an impact on beliefs about the endogeneity of audits:

> Perceived Audit Endogeneity: "In your opinion, if a firm that evades taxes doubles its evasion amount, the effect on its probability of being audited would..." The possible answers were: Increase significantly; Increase slightly; Not change;

[76] For simplicity's sake, in the remainder of this section we refer to such cases as 50% answers or 50% responses.

[77] We directly elicited the degree of certainty about perceived audit probability (in Q3) and penalty rate (in Q5) in the survey. These questions are reported in the survey questionnaire in Appendix A.7.

Diminish slightly; Diminish significantly.

The distribution of responses to this question is depicted in Figure B.6. The distribution of perceptions in the pooled control group–firms that received the *baseline* and *public goods* letters–suggests that firms were already aware of this endogeneity. In other words, there are no statistically significant differences in the distribution of perceptions for the *audit-endogeneity* group and the pooled control group group. On a scale from one to five, where one indicates that more evasion significantly increases the probability of being audited and five indicates that more evasion significantly diminishes the probability of being audited, the average belief was 1.45 in the pooled control group and 1.41 in the *audit-endogeneity* group (p-value of the difference=0.67).

B.4.4 Survey Results: Relation Between Signal and Self-Reported Perception

Figure B.7 shows the raw relation between the signal in the letter and the perceptions of audit probability and penalty rate reported in the survey. The x-axis depicts the signal in the letter and the y-axis the self-reported parameter. Panels (a) and (b) of Figure B.7 report the raw scatterplots for the audit probability and penalty rate respectively, where each dot represents a pair of values and the size of the dot is proportional to the number of individuals who received a given signal and their self-reported perception. While there is a positive correlation between the signal and the reported perception for the audit-probability parameter (depicted in panel (a)), individuals clearly overestimate the chances of being audited even after receiving the letter. Panel (b) also reports a positive corelation between the signal in the letter and the self-reported perception of penalty rate. In contrast to perceived audit probabilities, perceived penalty rates seem less disperse and closer to the actual value (about 30%). Panels (c) and (d) of Figure B.7 provide binned scatterplots that depict the results of regressing the self-reported perception of the parameter onto the signal in the letter for audit probability and penalty rate respectively. For the audit probability (depicted in panel (c)), individuals seem to have an adjustment rate of 40%. That means that for each additional percentage-point increase in the signal, individuals reported 0.4 additional percentage points in their perceived audit probability. The p-value associated with this coefficient does not allow us to reject the null hypothesis of no updating (p-value=0.169). For the penalty rate, the results depicted in panel (d) suggest that the information in the letter did not affect individuals' perceptions at all: for each additional percentage-point increase in the penalty rate as reported in the letters, individuals adjusted their perceived penalty rate by less than 10%. With a standard error of 0.20, the null hypothesis of the signal having no effect on perceptions cannot be rejected (p-value=0.965).

The information in the letter may also have altered individuals' degree of certainty about their perceptions of audit probability and penalty rates. Figure B.8 shows the effect of the *audit-statistics* letter on answers to Q3 and Q5, which ask directly how certain individuals are about their responses to the questions about perceived audit probability and penalty rate.[78] There are four possible answers for each question: "Not sure at all," "Somewhat sure," "Sure," and "Very sure." Figure B.8 suggests that, if anything, the information provided in the letter made firms less sure of their perceptions of p but did not affect their confidence about their perceived θ (p-values of 0.0688 and 0.8120 respectively).

B.5 Dynamics of the Effect-Raw Comparison Treatment vs Control

Figure B.9 plots the evolution of total VAT payments by treatment status, grouped by pairs of months. This raw visualization of the data allows for a very simple comparison of the treatment and comparison groups at each pair of months depicted. We group VAT payments by pairs of months because many firms in our sample are only required to make VAT payments on a bimonthly basis. Period zero, represented by the vertical dashed line, is defined as August–September 2015, that is, the period when the letters were delivered by the post office. In all panels, the lighter solid line represents the evolution of VAT payments for the comparison group, while the darker dashed line represents the evolution of VAT payments for the treated group. Panels (a) through (c) of Figure B.9 depict the treatment-control comparison in the main sample (i.e., *audit-statistics* vs. *baseline*, *audit-endogeneity* vs. *baseline,* and *public-goods* vs. *baseline* respectively). Panel (d) of Figure B.9 provides a comparison of the two treatment groups in the secondary sample: 50% audit probability threat vs. 25% audit probability threat. In all cases, we exclude firms whose letters were marked as "Returned" by the post office.

The results in Panels (a) through (c) of Figure B.9 indicate no difference in VAT payments between firms that received any of the treatment messages and firms that received the *baseline* letter before the letters were sent (this is formally contrasted by means of the balance tests reported in Table 1 and the falsification tests reported in Table 2). Immediately after receiving a letter, however, a wedge in VAT payments shows up between the two groups, and firms that were treated with any of the messages start to pay larger amounts of VAT compared to firms that received the *baseline* letter. These differences seem to be larger in the first twelve months after the experiment, although a smaller gap persists after the first post-treatment year. These patterns are consistent with the ones that can be derived

[78] These questions are found in the survey questionnaire in Appendix A.7.

from the difference-in-differences specification reported in Table 2. Panel (d) of Figure B.9 focuses on the secondary sample. The figure suggests that the evolution of VAT payments by treatment arm in this special sample is noisier than in the main sample; no clear pattern can be observed immediately after the letters were sent or in subsequent periods. This is consistent with the results reported in Table 3, where the results clearly indicate that the *audit-threat* letter did not have a differential effect on the group of firms that received the higher signal of probability of being audited.

B.6 Robustness Checks: Regression Analysis

B.6.1 Robustness Checks of Main Specification

To assess the robustness of the results for the main specification in Section 4, Table B.7 presents alternative estimates for the effects of the *audit-statistics*, *audit-endogeneity,* and *public-goods* treatments based on different specifications. The first column presents estimates of the treatment effects based only on the extensive margin of VAT payments: i.e., the outcome is coded as one if the firm made at least one payment in the post-treatment period, and zero otherwise. There is not much variation in the extensive margin: 96% of firms in the sample made positive payments in the post-treatment period. This is a direct byproduct of the selection of the subject pool: we excluded all firms that had not made at least three payments in the twelve months before the treatment assignment. To complement these results, columns (2) and (3) present the average treatment effects on alternative outcomes related to the number of payments made by the firms. Column (2) reports the average treatment effect on the probability of making at least three payments, and column (3) on the probability of making at least six payments in the same period.[79] The effects of the three different messages on the extensive margin and on the number of payments are close to zero and statistically insignificant.

The specifications in columns (4), (5), and (6) of Table B.7 use the number of VAT payments as the dependent variable. Column (4) corresponds to our main Poisson specification. Column (5) presents estimates based on OLS regressions, and column (6) presents estimates based on Tobit regressions. The Poisson model has a key advantage in this context: it deals naturally with the bunching of payments at exactly zero while allowing for the effects to be proportional. By contrast, the OLS specification does not deal with bunching at zero and does not allow for the effects on number of payments to be proportional. Because of the

[79] We do not include an outcome variable that reflects the treatment effect on the probability of making twelve payments during the first post-treatment year because, as mentioned, many firms in our sample are only required to make VAT payments on a bimonthly basis (only 25% of the sample makes payments every month). See Table B.3 for more details about the distribution of monthly payments.

nature of our sample, many firms are required to make VAT payments on a bimonthly basis. The Tobit specification is more appropriate than OLS since it takes into account the censored nature of the data at zero but does not allow for the effects to be proportional .

Columns (4), (5), and (6) of Table B.7 present estimates based on these specifications. They show the results to be identical in terms of sign and statistical significance of the coefficients, indicating robustness to these three alternative specifications. If anything, the effects are more statistically significant when using the OLS and Tobit models. Even though the results from the Poisson, OLS, and Tobit models are not directly comparable in terms of magnitude, they are roughly consistent. For instance, the Tobit model suggests that the *audit-statistics* message has an effect of USD 451 (p-value=0.003). Since the average outcome is USD 6,470, this Tobit coefficient amounts to an effect of about 6.9%, which is in the same order of magnitude as the Poisson model (it indicates an effect of 7.0% for the *audit-statistics* message (p-value=0.001)).

Finally, column (7) of Table B.7 reports the results of estimating our model on the final VAT liability calculated from an alternative administrative data source: annual tax returns. This outcome overcomes the possible concern that tax delinquency could be a driver of our results on the effect of retroactive payments specifically. The time frame for this outcome is completely different from the one for the monthly VAT payments used in our main specification, and hence the amounts (and thus the results) do not necessarily match. However, the results in column (7) of panel (a) of Table B.7 indicate that the effect of the *audit-statistics* message is 5.6% with this alternative measure of the outcome variable, which is similar in sign and magnitude to our main result. This effect is precisely estimated and statistically significant at the 5% level. The effect of the treatment in the falsification test is indistinguishable from zero.

The results in panel (b) of Table B.7 indicate that the effect of the *audit-endogeneity* message on the VAT reported in the annual tax return is similar to its effect on the VAT payments. The point estimate of the coefficient is 5.9%, which is smaller than the result for our main specification. In this case, however, the coefficient is less precisely estimated and statistically indistinguishable from zero at conventional levels. Finally, the results in panel (c) of Table B.7 indicate that the effect of the *public-goods* message on the annual VAT liability reported in the tax return is close to zero, a finding consistent with small effects observed immediately after receiving the letter that fade a few months later.

B.6.2 Alternative Specifications for the Effects of Signals on Audit Probabilities and Penalty Rates

We assess the robustness of the estimated effects of the signals about audit probabilities and penalty rates on post-treatment payments in two different ways.

First, panel (a) in Table B.8 replicates the analysis in Table 3 for the additional specifications we used in Table B.7. The results are essentially the same as the ones reported in Section 5: i.e., the specific information on audit probabilities and penalty rates in the *audit-statistics* message had no effect on compliance. At both the extensive and the intensive margin, all coefficients associated with the treatment variable, regardless of the specification used or the source of the dependent variable, are statistically insignificant. This is also true for the effects of the *audit-threat* message (panel (b) of Table B.8). Conditional on being treated, the information in the letter does not affect firms' compliance behavior.

Second, Table B.9 presents the results for an alternative specification of the elasticity estimation in Table 3. Instead of estimating the elasticities with respect to p and θ separately, we estimate the elasticity with respect to the product (expected penalty) $p * \theta$ in a regression of the form:

$$Y_i = \alpha + \tau_{p \cdot \theta} \cdot p_i \cdot \theta_i + \delta \cdot Post_t + \gamma_{p \cdot \theta} \left(p_i \cdot \theta_i \cdot Post_t \right) + \sum_{g=1}^{5} I_{\{i \in g\}} \left(\pi_g + \kappa_g \cdot Post_t \right) + \epsilon_i \quad \text{(B.1)}$$

Like in the model where p and θ were included separately, the elasticity computed with this alternative specification is statistically and economically insignificant.

Panel (a) of Figure B.10 provides additional evidence on the effect of the *audit-statistics* message (relative to the *baseline* letter). In the spirit of Figure 4, Figure B.10 summarizes the raw data, this time dividing the sample into two groups: firms that received letters with low signals of p ($p \leq$11.7%) and firms that received high signals of p ($p >$11.7%). Note that the evolution of the outcome variable is extremely similar for the two groups.

Panel (b) of Figure B.10 presents the equivalent analysis for firms that received messages with low and high values of θ (below and above the mean penalty rate in our sample, 30.6%). Again, the effects are very similar for the two groups.

The results in Table B.8, Table B.9, and Figure B.10 provide further evidence that what is at work is the fear channel rather than a rational re-optimization as a consequence of exposure to some information on tax enforcement mechanisms.

B.7 Additional Test: Breakdown of Other Taxes

In this section, we report additional and more detailed evidence about the treatment effects by analyzing the effects of the letters on different types of payments firms made regularly to the IRS. Regular payments by the firms to the IRS are comprised of VAT payments, corporate income tax payments, property tax payments, and personal income tax withholdings on behalf of employees. Table B.10 replicates the difference-in-differences estimates of the average effect of the different treatment arms presented in Table 2 but for different aggregations of payments for the first year after the experiment. The table reports both the treatment effect and the estimates from the falsification test comparing two pre-treatment periods, as does Table B.11 but for the results presented in Table 3. Column (1) reports the average effects on the total amount of tax paid by the firms, that is, the sum of VAT, corporate income tax, property tax, and personal income tax withholding payments. Columns (2) and (3) separate total payments into VAT and other taxes, while columns (4), (5), and (6) break down other taxes into their three components.

The results provided in Table B.10 show that the treatment effects caused by the *audit-statistics*, *audit-endogeneity,* and *public-goods* letters are not restricted to VAT payments; payments of other taxes increased as well, indeed by a slightly larger magnitude. When the "other taxes" variable is broken down into its components, the effects are too imprecise to identify any statistically significant effect on any single tax. Overall, all the coefficients are positive and of the right magnitude with no pre-treatment imbalances, but the effect of the *audit-statistics* message on corporate income tax is the only one that is statistically significant. The interpretation of the results reported in Table B.11 is similar, indicating that the information in the letters did not seem to have an effect on any tax payment except for personal income tax withholding. Though that effect is statistically significant in some specifications, there is no clear and consistent pattern across specifications.

B.8 Additional Test: Heterogeneity with Respect to Firm's Characteristics

To provide a more detailed description of how the messages affected different types of firms, we report here the treatment effects disaggregated by different firm characteristics. Because administrative records contain only a few observable firm characteristics, our analysis is based on four characteristics: number of employees, number of years the firms has been registered with the IRS, type of tax return filed, and activity sector. For number of employees and years registered with the IRS, we split the sample in two according to whether the observed value falls above or below the median (2.5 and 12.7 respectively). The IRS has different VAT schedules depending on firm size, and that is reflected in the type of tax return firms are required to file. Even for the small and medium-sized firms considered in this experiment, there are different types of tax returns. For a heterogeneity analysis, we divide firms into two groups: firms required to file a more comprehensive tax form (mostly larger firms) and firms required to file a simplified tax form according to special VAT regimes such as fixed VAT and professional independent VAT that have a separate form. As reported in Table 1, about 68% of firms filed a comprehensive VAT tax return in 2013. Finally, we also have information about the sector in which the firm operates. We split the sample into three groups: non-retail firms that sell goods (about 29%), retail firms (about 22%), and service providers (49%).

Figure B.11 depicts the average effect of the letters on VAT payments in the first year after the experiment (and the associated 95% confidence intervals) for each of the groups as defined above (similar to the "Post-Treatment" coefficient in column (1) of Table 2). Figure B.12 reports the same estimates but for the falsification test that compares treated and control firms in two pre-treatment periods (similar to the "Pre-Treatment" coefficient in column (1) of Table 2). Figure B.13 reports the estimated elasticities with respect to the audit probability and penalty rate in the letter for the same groups of firms, which is analogous to the results reported in Table 3. Figure B.14 reports the falsification test for the estimated elasticities. Figures B.11 and B.12 are based on four regressions. Each regression is an augmented version of equation (1) altered to include an additional interaction between the coefficient of interest and a dummy for the group of interest and for all the corresponding interaction terms. Figures B.13 and B.14 are similar, but based on equation (2) instead. This is our baseline specification for the elasticities of VAT payments to the parameters provided in the letters. Panel (a) of Figure B.11 compares the *audit-statistics* message with the *baseline* letter, while results in panels (b) and (c) replicate that comparison for *audit-endogeneity* and *public-goods* messages respectively. The notes in the figure report the p-value of the test that compares whether the treatment effects are different across groups within a single category.

The coefficients depicted in Figure B.11 show that while the letters have a positive effect on VAT payments overall, that effect is, it appears, homogeneous across firm types. Out of fifteen tests of equality of coefficients, only one is rejected at a 10% level (firms filing a more comprehensive tax return seem to respond slightly more than firms that do not file comprehensive tax returns). Figure B.13 is also consistent with homogeneous responses across groups of firms. Estimated elasticities are close to zero on average, as our main specification suggests, and there are no differences between groups based on any of the variables considered, except for firms in the service sector for which the test of equality of coefficients can be rejected at a 1% level (p-value=0.013). Falsification tests reported in Figures B.12 and B.14 are reassuring since they show that there are no differences in the "fake" treatment effect between firms in the pre-treatment period.[80] The general conclusion is, then, that all firms respond in a fairly similar way, although we do not have enough power to capture the differences that do exist.

B.9 Additional Test: Heterogeneity with Respect to Prior Beliefs

B.9.1 Measuring Prior Beliefs

In the *A&S* framework, firms with different prior beliefs about the probability of being audited should react differently to signals and information on this probability. To test this hypothesis, we need a measure for the prior beliefs of a particular firm. We construct a proxy of prior beliefs based on the firm's own audit history.

The intuition behind this approach is that, since there is little publicly available information on audit probabilities, firms likely form their beliefs based on their own audit experience. For instance, when a firm registers with the tax authority, its initial belief may follow the beta distribution with parameters $\{\alpha_0, \beta_0\}$. Assume that firm i has been registered for T_i years before our mailing campaign, and during this period it has experienced $N_i \leq T_i$ audits. If firm i is Bayesian, its belief about annual probability of being audited should follow a beta distribution with parameters $\{\alpha_1 = \alpha_0 + N_i,\ \beta_1 = \beta_0 + T_i - N_i\}$. The mean of that belief should be $\frac{\alpha_0+N_i}{\alpha_0+\beta_0+T_i}$. This implies a belief about the probability of being audited at least once in the following three years of $\hat{p}_i = 1 - \left(1 - \frac{\alpha_0+N_i}{\alpha_0+\beta_0+T_i}\right)^3$. In our main specification, we generate these proxies by setting $\{\alpha_0 = 0.13, \beta_0 = 1\}$. This baseline calibration generates an average belief that matches the actual average probability according to our administrative data, and we offer robustness tests using alternative calibrations.[81]

We can use the survey data to validate this proxy for prior belief. Among the 145 responses from the pooled control group, 10.3% of firms reported that they had been audited in the past, and these firms reported a higher average perceived probability of being audited at least once in the following three years (63.9%) relative to firms that had not been audited recently (38.1%), a large and statistically significant difference (p-value<0.001).[82] This ev-

[80] While in general the results from the falsification tests are reassuring and centered around zero, there is a common pattern of statistically significant "false" effects when the sample is split by type of tax return filed. This is due to an anomalous change in VAT payments in the *baseline* group for firms that did not file a comprehensive tax return. VAT payments for the *baseline* group that did not file a tax return fell significantly between the two years considered for the falsification test, which is why we observe a positive "false" treatment effect. This is observed across panels because the baseline group is used as the comparison group in panels (a) through (c).

[81] Since we only have information about audits for firms in our sample for the previous fifteen years, we set the maximum number of years registered with the IRS at fifteen.

[82] As reported in a follow-up paper (Bergolo et al., 2018), the indicator of recent audits is the single most important predictor of perceived audit probabilities among a host of different factors. Recent audit

idence suggests that, consistent with our proxy, firms are using their own audit history to form beliefs about the probability of being audited in the future.

B.9.2 Results

In the *A&S* framework, the effect of the *audit-statistics* letter on tax compliance should be larger for firms with relatively low priors for the audit probability ($\hat{p}$) compared to those with relatively high values. Since the *audit-statistics* message conveyed a signal of $\hat{p} = 11.7\%$ on average, its effect on compliance should have been greater for firms with $\hat{p} < 11.7\%$, since the signal should, on average, increase their perceived audit probability. The effect of the *audit-statistics* message should be negative for firms with $\hat{p}$>11.7%: on average, the signal should reduce their perceived probability of being audited.

Figure B.15 presents the results from this exercise. Panel (a) of Figure B.15 shows a binned scatterplot of the treatment effect of the *audit-statistics* letter for the four quartiles of $\hat{p}$.[83] This figure includes a vertical dashed line at 11.7%, the average audit probability conveyed in our letters. In contrast to the predictions from *A&S*, we fail to find a negative relationship between the effect of the *audit-statistics* message and the value of the prior belief: the slope is negative (-0.0006), but economically small and statistically insignificant (p-value of 0.525). Panel (b) of Figure B.15 presents a more direct test by combining the heterogeneity in prior beliefs with the heterogeneity in signals. Instead of grouping firms by prior belief, we group them by the difference between that prior belief and the specific signal sent to each firm in the personalized letter. The intuition is that the difference between the prior belief and the signal is the "surprise" conveyed by our information treatment. The figure includes a vertical dashed line at zero, the point where firms receive signals that are equal to their prior beliefs. The effect of the *audit-statistics* letter on compliance should be decreasing in $\hat{p}-p_{signal}$, positive for the group with $\hat{p}-p_{signal} < 0$ (i.e., those for whom the signal was higher than the prior belief), and negative for the group with $\hat{p} - p_{signal} > 0$ (i.e., those receiving a signal indicating that they were overestimating audit probability). The results in panel (b) are consistent with the results from panel (a): the slope of the relationship (indicated by the dashed red line) is zero (p-value=0.986).

As a robustness check, we provide an alternative calibration of the Bayesian model. In the above results, we selected values for the parameters α_0 and β_0 to "center" beliefs around

experience also has a positive effect on perceived penalty rate, but this effect is less significant: respondents from firms that were audited recently report an average perceived penalty rate of 40.0%, compared to 29.4% for respondents from firms that were not audited recently (this difference is statistically insignificant (p-value=0.201)).

[83] We base our analysis on quartiles of the probabilities because there is substantial bunching at zero, which forced us to divide the sample in four to get even-sized groups.

the true audit probability in our sample (11.7%). We present the results from an alternative calibration, which centers the perceived probability around the average value obtained from the control group in our post-treatment survey–an average perceived audit probability of 40.5% for the comparison group. The results are presented in Figure B.16. If anything, this alternative calibration makes *A&S* even less plausible: as shown in panel (b), we would expect almost everyone to adjust perceived probability downward, and thus reduce tax payments–exactly the opposite of what we find.

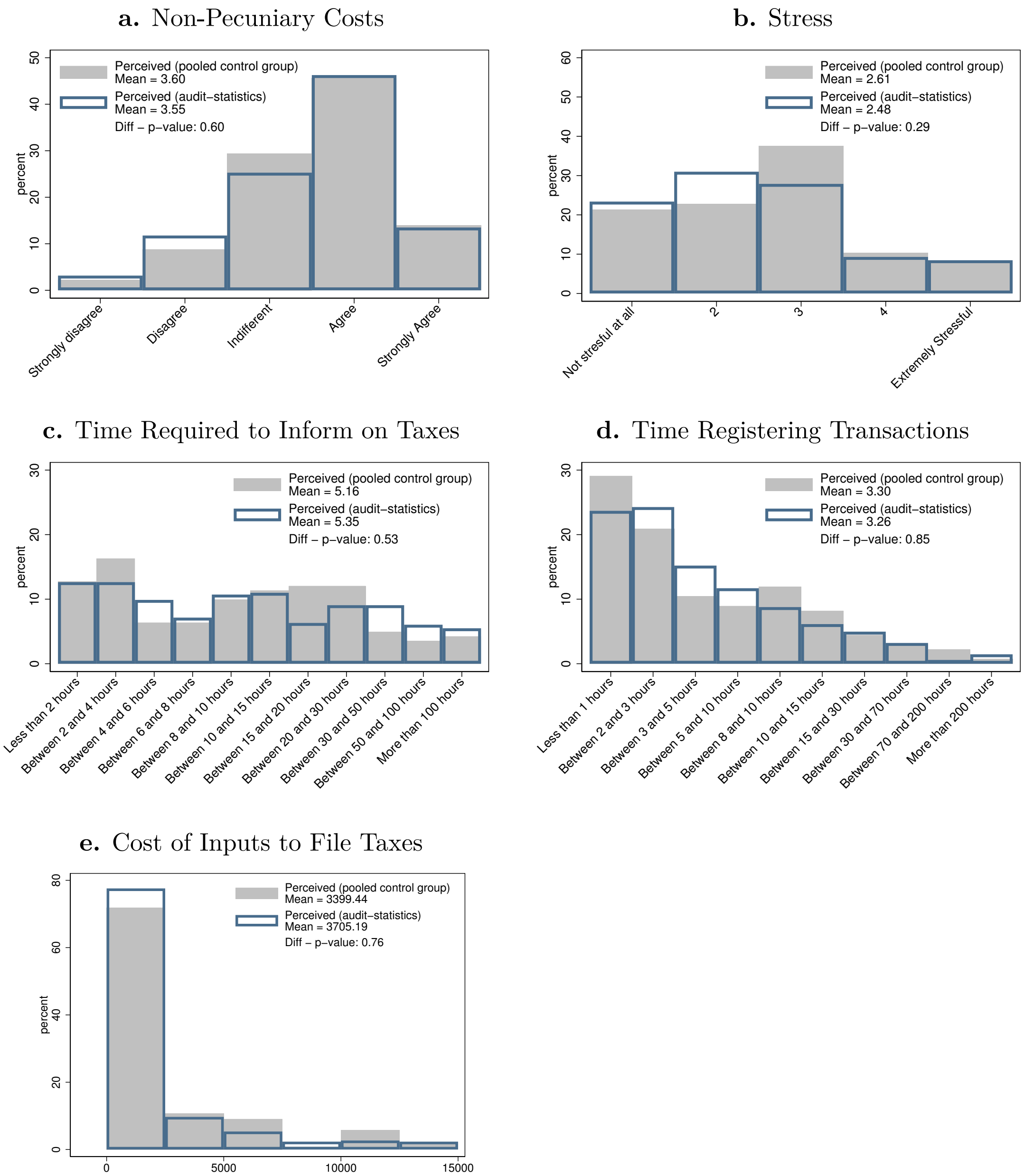
A more structured analysis is presented in Table B.12. For reference, column (1) of Table B.12 shows the baseline specification: i.e., in the whole sample, the *audit-statistics* message increased tax payments by 7.0%. We then divide the sample into two groups: firms with prior beliefs about the probability of being audited that were lower than the probability reported in the letter they received, and firms with prior beliefs about the probability of being audited that were equal to or higher than the probability reported in the letter they received. Since we divide the groups according to a characteristic of firms that received the *audit-statistics* message, the outcomes for both groups are compared to those of the firms that received the *baseline* letter. According to *A&S*, increasing taxpayers' perceived probability of being audited should result in higher tax payments, and the opposite should hold true when their perceived probability of being audited is reduced. The expected effect of the *audit-statistics* message is, therefore, positive on firms with $\hat{p} < p$ and negative on firms with $\hat{p} \geq p$. Column (2) in Table B.12 shows that the average effect of the *audit-statistics* message on VAT payments in the first year after receiving the letter is 6.1% for taxpayers with relatively low prior beliefs. Column (3) reports that the average effect for taxpayers with relatively high beliefs is 8.6%. The effect on firms with relatively low prior beliefs is positive but lower in magnitude than the effect on firms with relatively high prior beliefs, and the sign of the effect on the latter is the opposite of *A&S*-based predictions. Furthermore, the differences in magnitude are statistically and economically insignificant (p-value=0.347). If these results provide any evidence, it counters the *A&S* predictions.

We can reproduce a similar exercise without a Bayesian learning model. We compare the responses of firms with and without prior audit experience. In columns (4) and (5) of Table B.12, we compare the effects of the *audit-statistics* message on firms that were audited in the recent past (23.7% of the sample was audited at some point in the fifteen years before the treatment) and firms that were not audited during that period (the remaining 76.3% of the sample).[84] The intuition is that firms that were audited in the recent past believe the probability of being audited is higher than those that were not audited. The null hypothesis is thus that the effect of our *audit-statistics* messages is stronger for firms that were not

[84] That time frame reflects how far back the available IRS administrative records reach.

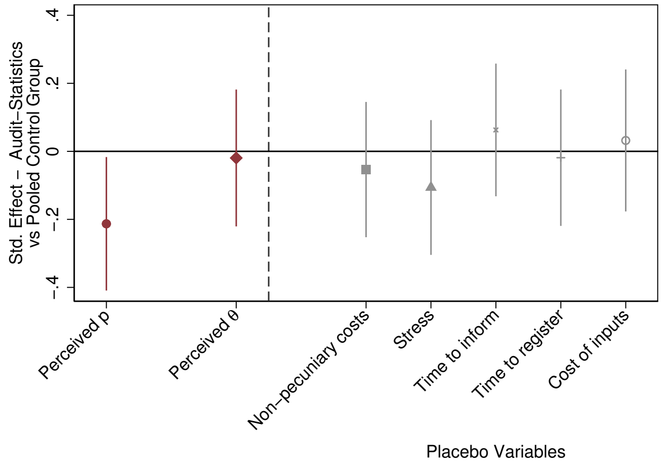
audited, since such firms would have had lower prior beliefs about audit probabilities and thus increase their perceived probabilities in response to the *audit-statistics* messages. We could even expect to see negative effects on compliance for firms that were audited in the past, because they were likely to have had high prior beliefs about audit probabilities that our information treatment should have reduced. The results in Table B.12 indicate that there are no heterogeneous effects with respect to recent audit experience. The difference in treatment effects for the two groups is small and statistically insignificant: the effect is 8.2% for the group of firms that was previously audited, and 6.6% for those with no prior audit experience (p-value of difference=0.729).

Figure B.1: Survey Results: Placebo Variables by Treatment Group



Notes: Placebo tests where we estimate the effect of receiving the *audit-statistics* letter on questions that we do not expect to be affected by our treatment. These questions were reported in other modules of the survey and were aimed to collect information about tax compliance costs. Panel (a) reports the effects on the answers to a question that asked: "On a scale from 1 to 5 where 1 is 'Strongly disagree' and 5 is 'Strongly agree,' to what extent do you agree with the following statement: *Tax compliance generates non-pecuniary costs*?" Panel (b) presents the responses to the following question: "On a scale from 1 to 5 where 1 is "Not stressful at all" and 5 is "Very stressful," please rate the level of stress created by all the steps required to fulfill your tax obligations." Panel (c) presents the responses to the following question: "How much time do you spend informing yourself about your tax obligations?" Panel (d) presents the responses to the following question: "How much time do you spend registering all the transactions made by your firm?" Panel (e) presents the responses to the following question: "What is the estimated monthly cost of all the inputs that you use in activities related to tax compliance?" The estimation strategy is exactly the same as the one used for Figure 5 (see corresponding notes for more detail). "Perceived (pooled control group)" refers to survey respondents who received the *baseline* letter or the *public-goods* letter during the experimental stage. "Perceived (*audit-statistics*)" refers to respondents who received *audit-statistics* letters. We report the mean responses and the p-value of the difference between the two groups. Analysis is restricted to individuals who are owners (or who did not answer the ownership question) and who answered the questions on perceived audit probability and penalty rate.

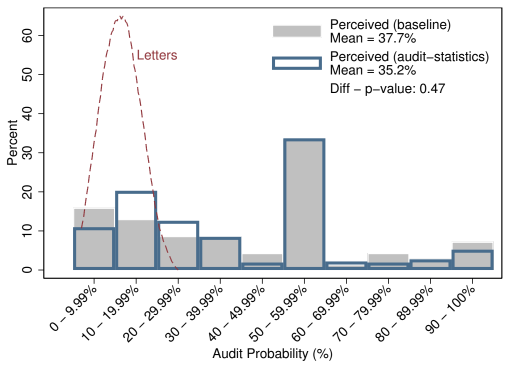
Figure B.2: Survey Results: Placebo Variables, *Audit-Statistics* vs. *Baseline* and *Public-Goods*



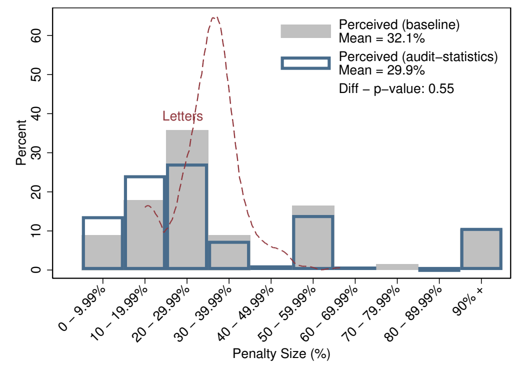
Notes: In this figure, we compare the treatment effects of the *audit-statistics* letter on perceptions of audit probability and penalty rate (coefficients reported to the left of the vertical dashed line) to the placebo outcomes reported in Figure B.1 (see corresponding notes for more details about the placebo outcomes). For all of the outcomes reported in the figure, we used the same estimation strategy as in Figure 5 (see corresponding notes for more details). The comparison group is comprised of respondents who received the *baseline* letter or the *public-goods* letter during the experimental stage. The treatment group is comprised of respondents who received *audit-statistics* letters. For comparison purposes, all outcome variables are expressed in terms of the standard deviations of the control group. In all cases, the analysis is restricted to individuals who are owners (or who did not answer the ownership question) and who answered the two questions that were relevant to our analysis (the ones about perceived audit probability and penalty rate).

Figure B.3: Survey Results: Perceived p and θ, Alternative Samples and Comparison Group

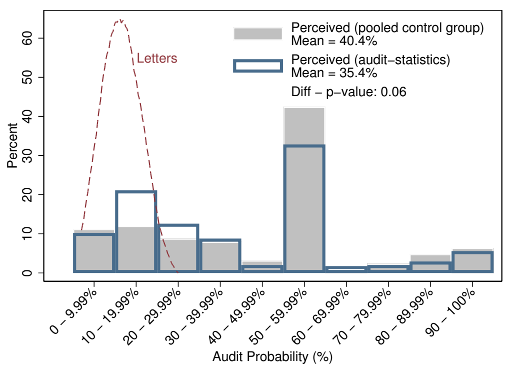
a. *p: Audit-Statistics* vs. *Baseline*



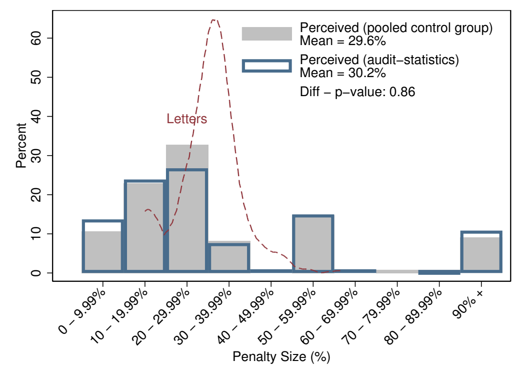
b. θ: *Audit-Statistics* vs. *Baseline*



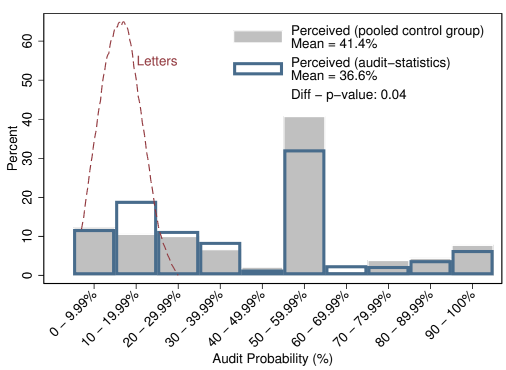
c. *p: Audit-Statistics* vs. *Baseline* and *Public-Goods*, Owners Only



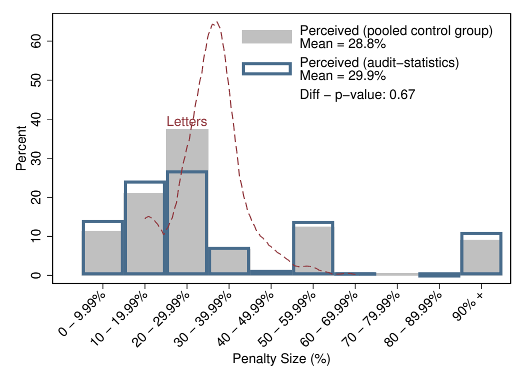
d. θ: *Audit-Statistics* vs. *Baseline* and *Public-Goods*, Owners Only



e. *p: Audit-Statistics* vs. *Baseline* and *Public-Goods*, Full Sample



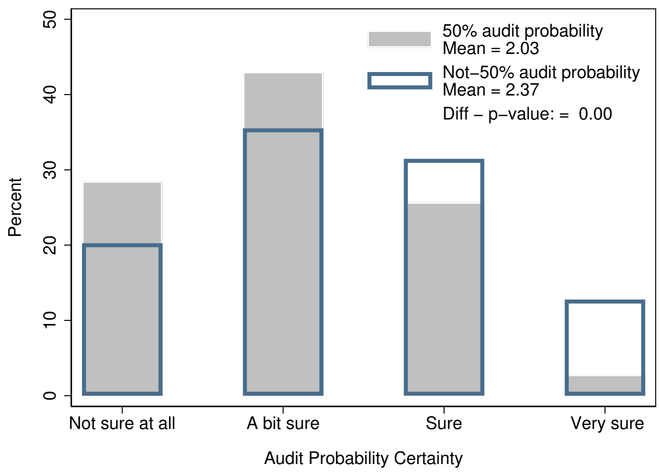
f. θ: *Audit-Statistics* vs. *Baseline* and *Public-Goods*, Full Sample



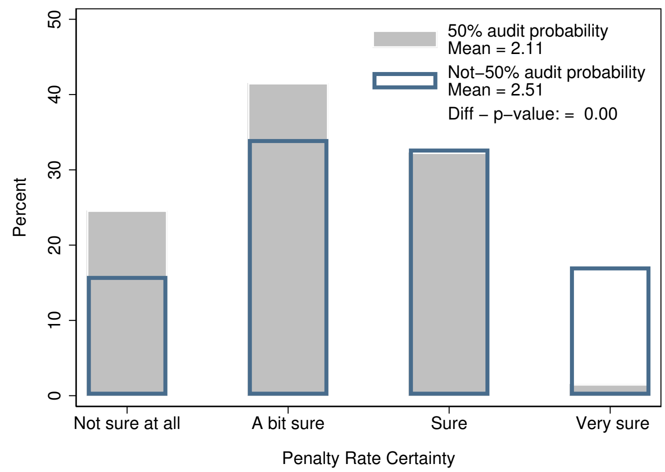
Notes: In panels (a) and (b), "Perceived (*baseline*)" (N=69) refers to survey respondents who received the *baseline* letter during the experimental stage while "Perceived (*audit-statistics*)" (N=365) refers to respondents who received the *audit-statistics* letter. In panels (c) and (d), we use a pooled control group comprised of recipients of the *baseline* and *public-goods* letters, but we restrict the sample to survey respondents who self-identified as owners (N of *baseline* group = 61, N of *public-goods* group = 64, N of *audit-statistics* group = 341). Panels (e) and (f) use the full sample regardless of the self-reported occupation in the survey, that is, it includes owners, managers, inhouse and external accountants, and other employees (N of *baseline* group = 89, N of *public-goods* group = 90, N of *audit-statistics* group = 465). We also report the mean of the perceptions for each parameter and the p-value of the difference between the groups in each panel. The answers correspond to survey Q2 and Q4 (see full survey questionnaire in Appendix A.7). In panels (a), (c), and (e), the x-axis represents the probability of being audited; in panels (b), (d), and (f), it represents the average penalty rate. The red line represents the density function of the information displayed in the *audit-statistics* letters, measured in the y-axis on the right (hidden for the sake clarity). In all cases, the analysis is restricted to letters that were not returned by the postal service.

Figure B.4: Certainty in Perceptions Reported by Having Answered 50%

a. Audit Probability (p)



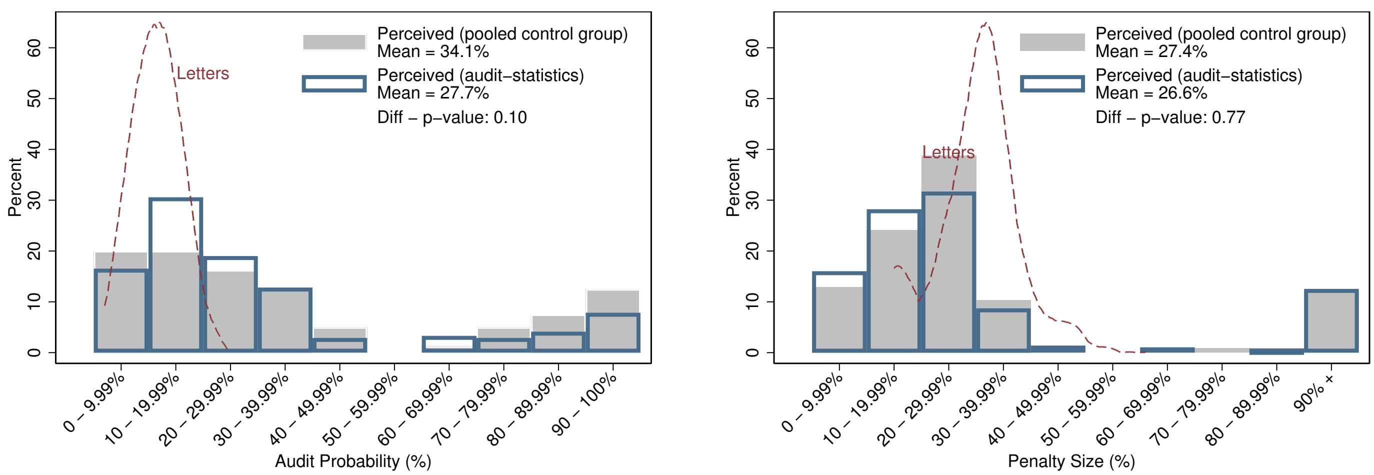
b. Penalty Rate (θ)



Notes: The histograms present the distributions of certainty about perceived audit probability and penalty rate as elicited in the survey. Panel (a) reports the distribution of answers to survey Q3 (see full survey questionnaire in Appendix A.7), splitting the sample between individuals who reported a perceived audit probability of 50% (N=180) and individuals who reported a different audit probability (N=322). Panel (b) presents a similar breakdown for survey question Q5, that is, it also splits the sample between individuals who reported a perceived penalty rate of 50% (N = 65) and individuals who reported a different perceived penalty rate (N = 397). In both panels, the sample is restricted to individuals who belong to our *audit-statistics* treatment arm or to the pooled control group (i.e., those who received either the *baseline* or the *public-goods* letter). We report the mean responses for the various groups, measured on a four-point scale where one corresponds to "Not at all sure" and four to "Very sure." We also report the p-value of the difference between the two groups.

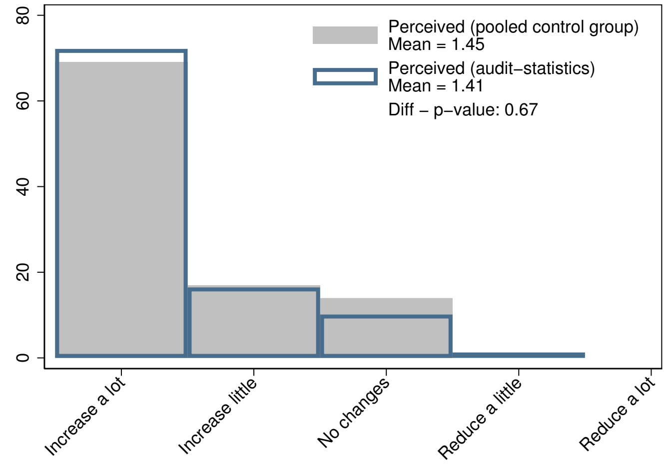
Figure B.5: Survey Results: Perception of Audit Probabilities and of Tax Evasion Penalty Rates by Treatment Group - Excluding 50% Responses

a. Audit Probability (p) **b.** Penalty Rate (θ)



Notes: The histograms are based on the survey responses of those who, in the post-treatment survey, either self-identified as owners or did not answer the question regarding their occupation in the firm and those who did not answer 50% to the audit probability and penalty rate questions. “Perceived (pooled control group)” (N = 80) refers to survey respondents who received either the *baseline* (N = 46) or the *public-goods* (N = 34) letter during the experimental stage (neither of the two letters contained any information on audit probabilities or penalty rates). “Perceived (*audit–statistics*)” refers to respondents who received *audit-statistics* letters (N = 242). In panel (a), the x-axis represents the probability of being audited; in panel (b), it represents the average penalty rate. We report the mean responses and the p-value of the difference between the two groups. The answers correspond to Q2 and Q4 of the survey (see full survey questionnaire in Appendix A.7). The red line represents the density function of the information displayed in the *audit-statistics* letters, measured in the y-axis on the right (hidden for the sake clarity).

Figure B.6: Perception of Endogeneity of Audits: *Audit-Endogeneity* vs. *Baseline* and *Public-Goods*

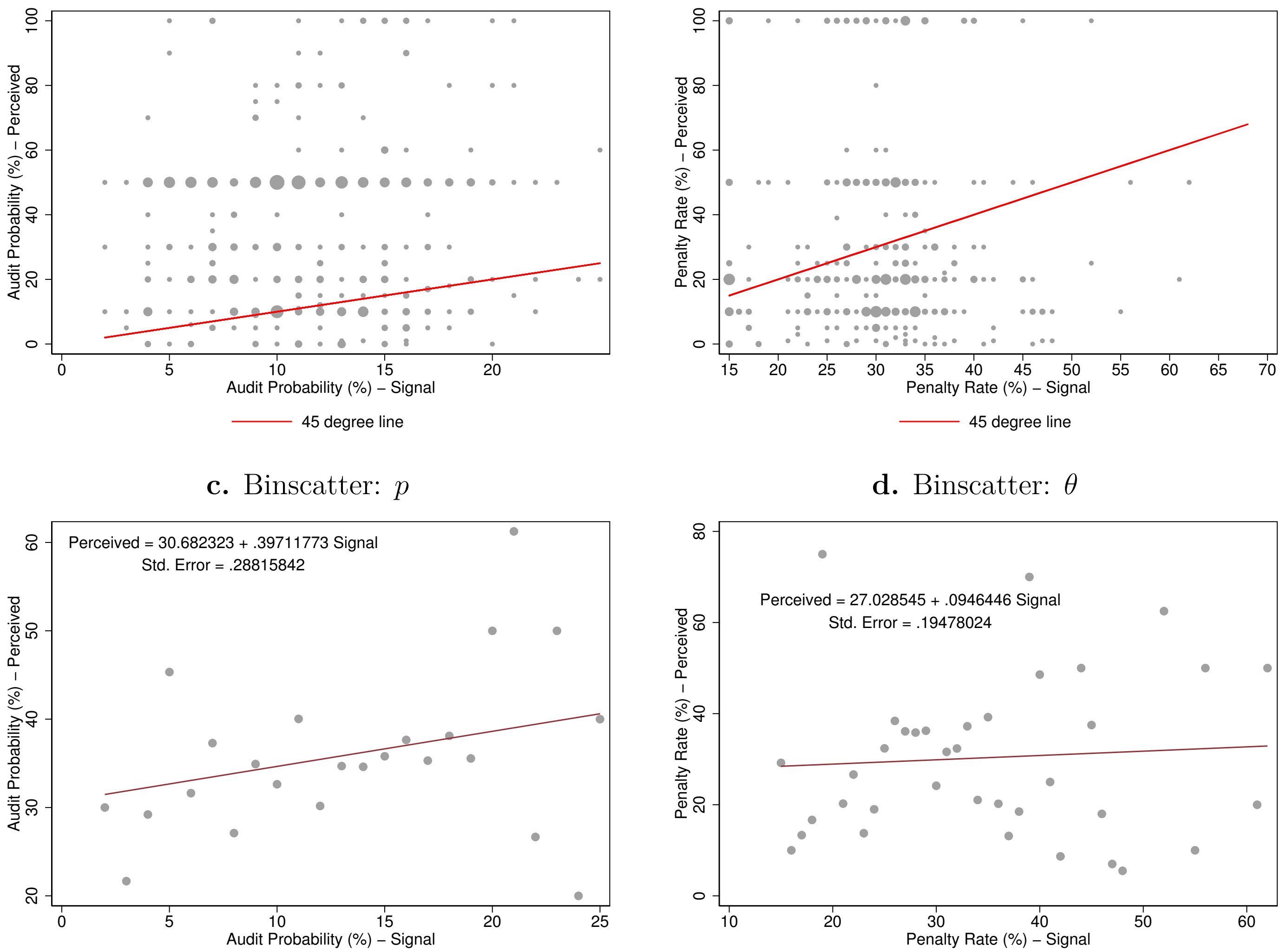


Notes: "Perceived (pooled control group)" (N = 137) refers to respondents who received either the *baseline* or the *public-goods* letter during the experimental stage, while "Perceived Endogeneity" (N = 79) refers to respondents who received *audit-endogeneity* letters. These answers correspond to Q6 of the survey (see the full questionnaire in Appendix A.7). The x-axis represents the different categories presented as survey options. We report the mean responses for the two groups, measured on a five-point scale where one corresponds to "Increase a lot" and five to "Reduce a lot." We also report the p-value of the difference between the two groups.

Figure B.7: Relation Between Signal Received and Self-Reported Perception

a. Scatterplot: p **b.** Scatterplot: θ

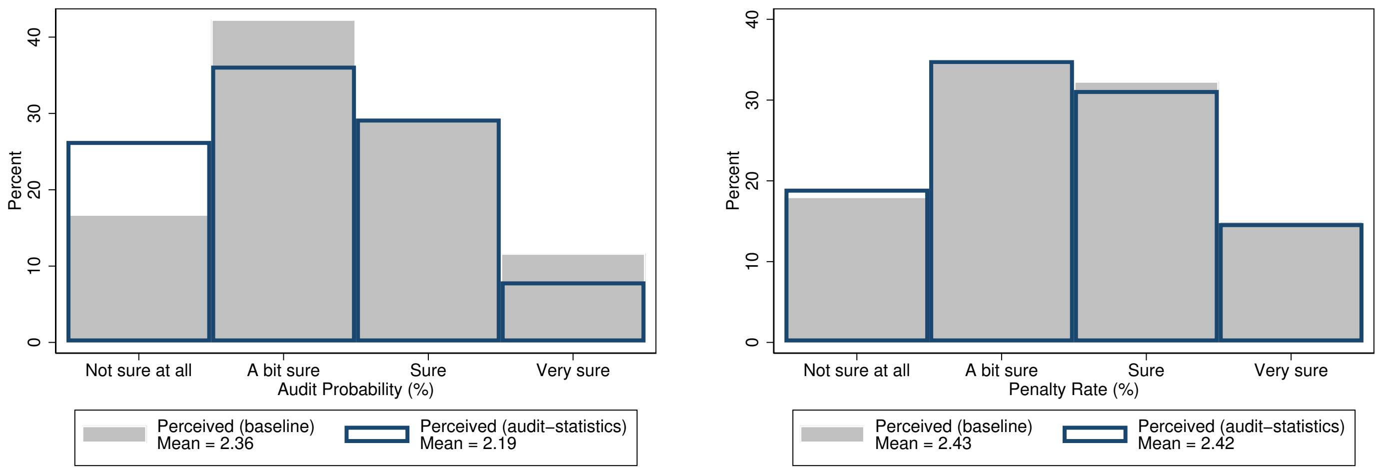
c. Binscatter: p **d.** Binscatter: θ



Notes: Figure B.7 shows the raw relations between the signal included in the letter and perceptions of audit probability and penalty rate reported in the survey for firms that received the *audit-statistics* letter (N = 365). The x-axis depicts the signal included in the letter and the y-axis the self-reported parameter. Panels (a) and (b) report the raw scatterplot where each dot represents a pair of values (signal, perception) and the size of the dot is proportional to the number of individuals. A solid 45-degree line is included in each panel. Panels (c) and (d) report the binned scatterplots of the same variable and include a line corresponding to the adjusted regression of perceptions over signals. Both the full equation predicted by the regression and the standard error of the coefficient on the signal are included as well. In all cases, the analysis is restricted to letters that were not returned by the postal service.

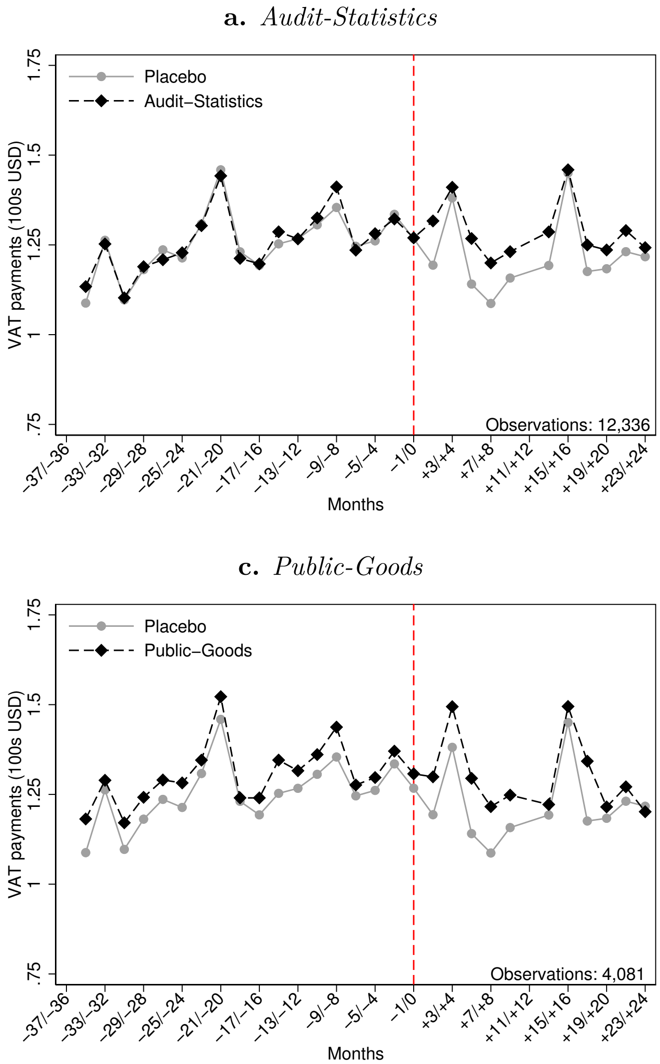
Figure B.8: Relation Between Signal Received and Self-Reported Certainty About Perception

a. Certainty about Self-Reported p **b.** Certainty about Self-Reported θ



The histograms are based on the survey responses to Q3 and Q5 of those who, in the post-treatment survey, either self-identified as owners or did not answer the question regarding their occupation in the firm (see full survey questionnaire in Appendix A.7). "Perceived (*baseline)*" (N = 137) refers to the survey respondents who received the *baseline* letter (N = 68) or the *public-goods* letter (N = 69) during the experimental stage (neither of the two letters contained any information on audit probabilities or penalty rates). "Perceived (*audit–statistics*)" refers to respondents who received *audit-statistics* letters (N = 364). In both panels, the x-axis represents a four-point scale that captures individuals' degree of certainty with respect to perceived audit probability (panel (a)) and perceived penalty rate (panel (b)). We also report the mean responses and the p-value of the difference between the two groups, calculated on a scale from one (corresponding to the "Not at All Sure" response) to four (corresponding to the "Very Sure" response).

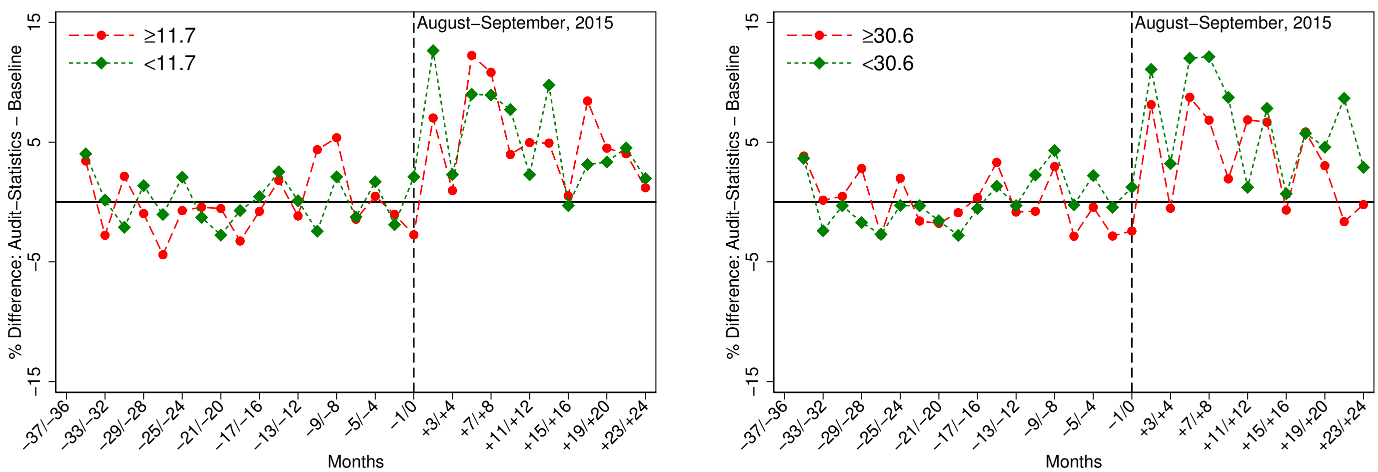
Figure B.9: Bimonthly VAT Payments, by Letter Type



Notes: These figures plot the bimonthly total VAT payments separated by treatment and comparison groups. Similar figures for the difference between these two groups are reported in Figure 4. The data used for this figure correspond to the period from October 2012 to September 2017. The months of August and September 2015, when most of the letters were delivered, are defined as the reference bimonthly period and marked with the dashed vertical line. Each figure plots the time series of total VAT payments for the treatment arm and the comparison group separately. Panel (a) (N = 12,336) presents the evolution of the main outcome variable for firms that received the *audit-statistics* message and the *baseline* message. Panel (b) (N = 4,103) provides the results for the *audit-endogeneity* message, and panel (c) (N = 4,081) for the *public-goods* message. Panel (d) (N = 4,048) presents the two treatment groups selected from the secondary sample. In this case, the two lines depict the two *audit-threat* messages that contain different probabilities of being audited: 50% vs 25%. In all cases, analysis is restricted to letters that were not returned by the postal service. For each pair of months, VAT payments are top-coded at the 99.99th percentile to avoid contamination of the results by outliers. In all panels, we omit the +11/+12 pair of months (August 2016–September 2016) because the information provided by the IRS for it is incomplete.

Figure B.10: Effects of *Audit-Statistics* by Level of the Signal

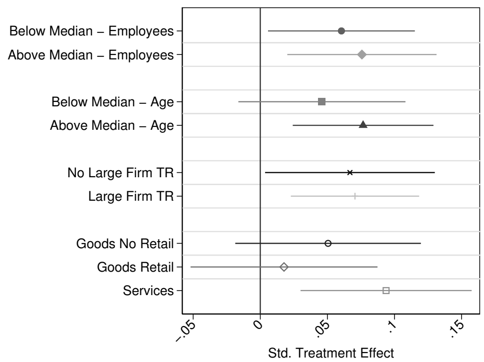
a. Effect of *audit-statistics* vs. *baseline*, by level of p **b.** Effect of *audit-statistics* vs. *baseline*, by level of θ



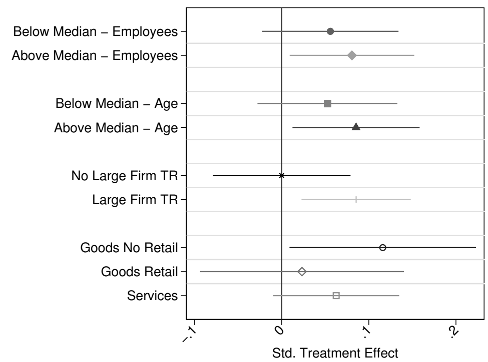
Notes: These figures plot the percentage difference in bimonthly total VAT payments between treatment and control groups, normalized by the average pre-treatment percentage difference (i.e., between months -35 and 0) for the same outcome, dividing the sample into firms that fall above and below the mean value of the corresponding signal. The red dots in panels (a) and (b) represent the effect for *audit-statistics* letter recipients with signals on p or θ above the mean respectively. Green dots represent the same effect but for those with signals on p and θ below the mean respectively. The data used in the figure cover the period October 2012–September 2017. The period of August 2015–September 2015, when most of the letters were delivered, is defined as the reference pair of months (it is indicated by the dashed vertical line). Each figure plots the difference between each treatment arm and the *baseline* letter (N = 12,336). In all cases, the analysis is restricted to letters that were not returned by the postal service. For each pair of months, VAT payments are top-coded at the 99.99th percentile to avoid the contamination of the results by outliers.

Figure B.11: Average Effects of *Audit-Statistics*, *Audit-Endogeneity,* and *Public-Goods* Letters on VAT Payments by Type of Firm

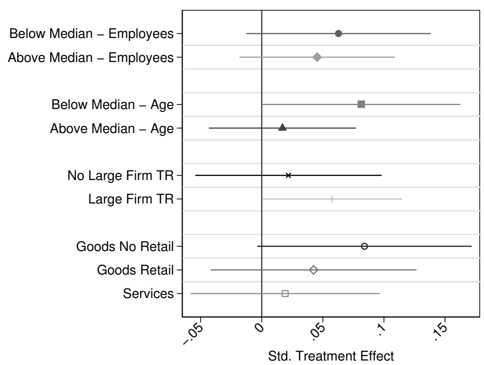
a. *Audit-Statistics*



b. *Audit-Endogeneity*



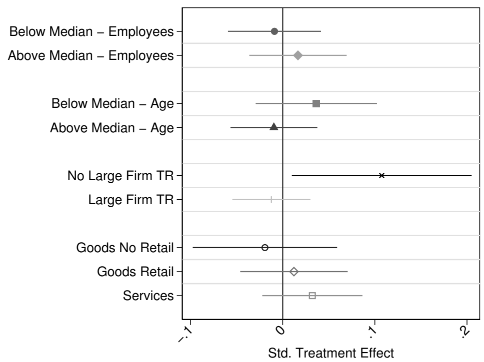
c. *Public-Goods*



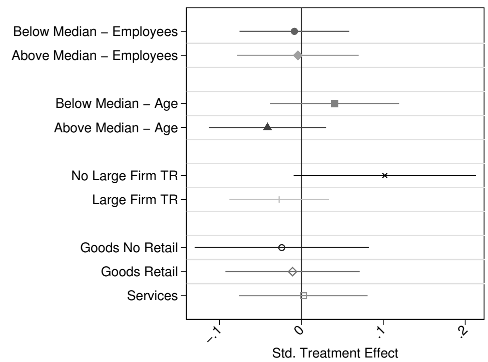
Notes: * significant at the 10% level, ** at the 5% level, *** at the 1% level. Standard errors are clustered at the firm level. Treatment effects are estimated for the first year post-treatment (October 2015–September 2016 vs. October 2014–September 2015) and based on an augmented version of the difference-in-differences specification reported in equation (1), which compares treated firms to control firms and pre-treatment to post-treatment periods using yearly aggregated variables. To capture the differential effects by group, we include a dummy variable for the interaction between the group of interest and the coefficient that captures the treatment effect. We also include all additional corresponding interactions with the time and treatment variables. More details about the baseline specification can be found in Table 2 and Section 4. More details about the augmented model can be found in Section B.8. Panel (a) compares the *audit-statistics* message with the *baseline* letter, while panels (b) and (c) replicate the comparison for the *audit-endogeneity* and *public-goods* messages respectively. The heterogeneity analysis is based on four characteristics: number of employees, age of firm (i.e., number of years registered with the IRS), type of tax return filed, and sector of operation. The median number of employees in the experimental sample is 2.5, while the median firm age is 12.7 years. Sixty-eight percent of firms filed a comprehensive VAT tax return in 2013 while the remaining 32% filed a simplified VAT tax return. Regarding the distribution of sector of operation, 29% of firms are non-retail sellers of goods , 22% are retail sellers, and 49% are service providers . In all cases, analysis is restricted to letters that were not returned by the postal service.

Figure B.12: Average Effects of *Audit-Statistics*, *Audit-Endogeneity,* and *Public-Goods* Letters on VAT Payments by Type of Firm, Pre-Treatment Falsification Test

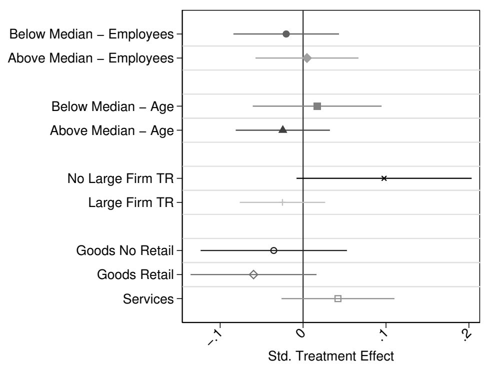
a. *Audit-Statistics*



b. *Audit-Endogeneity*



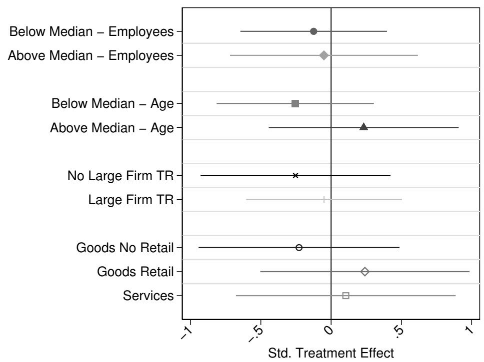
c. *Public-Goods*



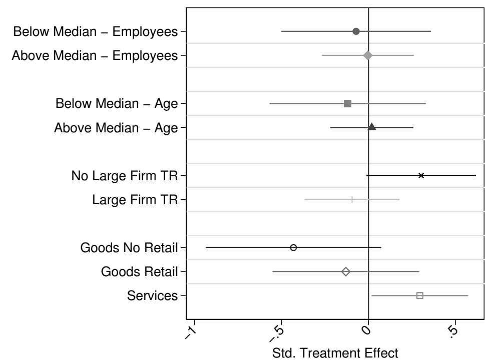
Notes: * significant at the 10% level, ** at the 5% level, *** at the 1% level. Standard errors are clustered at the firm level. Estimates presented in the figure correspond to falsification tests that compare two pre-treatment periods (October 2014–September 2015 vs. October 2013–September 2014) and are based on an augmented version of the difference-in-differences specification reported in equation (1), which compares treated firms to control firms using yearly aggregated variables. To capture the differential effects by group, we include a dummy variable for the interaction between the group of interest and the coefficient that captures the treatment effect. We include as well all additional corresponding interactions with the time and treatment variables. More details about the baseline specification can be found in Table 2 and Section 4. More details about the augmented model can be found in Section B.8. Panel (a) compares the *audit-statistics* message and the *baseline* letter, while panels (b) and (c) replicate the comparison for the *audit-endogeneity* and *public-goods* messages respectively. The heterogeneity analysis is based on four characteristics: number of employees, age of firm (i.e., number of years registered with the IRS), type of tax return filed, and sector of operation. The median number of employees in the experimental sample is 2.5, while the median age of the firms is 12.7 years. Sixty-eight percent of firms filed a comprehensive VAT tax return in 2013 while the remaining 32% filed a simplified VAT tax return. Regarding the distribution of sector of operation, 29% of firms are non-retail sellers of goods, 22% are retail sellers, and 49% are service providers. In all cases, analysis is restricted to letters that were not returned by the postal service.

Figure B.13: Elasticities of Tax Payments with Respect to Audit Probability and Penalty Rate in *Audit-Statistics* and *Audit-Threat* Sub-Treatments by Type of Firm

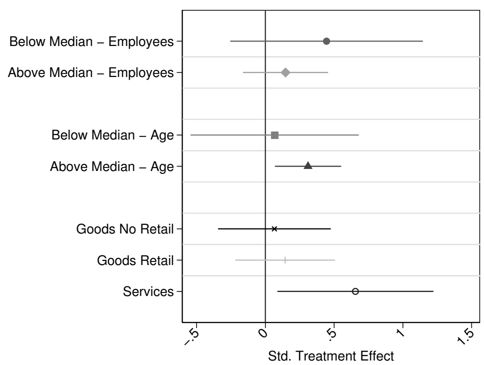
a. *Audit Probability*



b. *Penalty Rate*

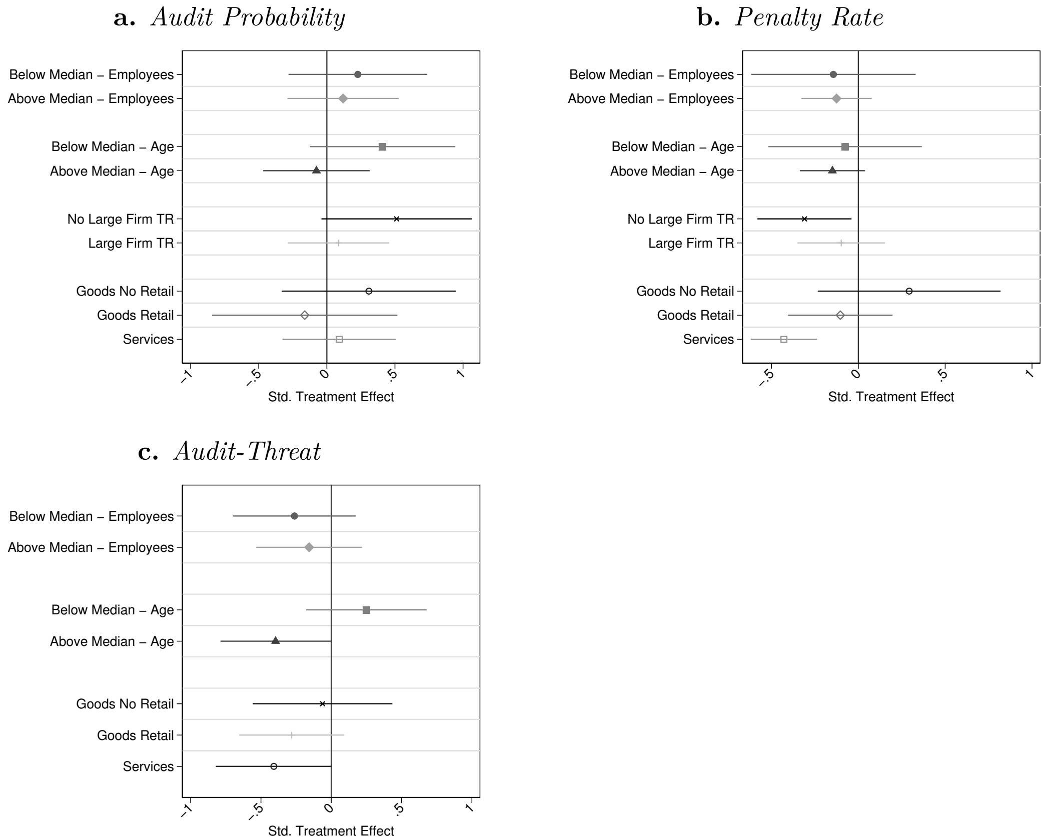


c. *Audit-Threat*



Notes: * significant at the 10% level, ** at the 5% level, *** at the 1% level. Standard errors are clustered at the firm level. Treatment effects are estimated for the first year post-treatment (October 2015–September 2016 vs. October 2014–September 2015) and are based on an augmented version of the difference-in-differences specification reported in equation (2) which compares treated firms that received different signals on p and θ. In all cases, we include an additional set of dummies for quintiles of the pre-treatment VAT payments, which are the groups from which we drew the sample of "similar firms" to calculate p and θ, and the corresponding interactions with the time variable. The results are based on Poisson regressions with variables expressed in percentage terms, so the coefficients can be interpreted directly as elasticities. To capture the differential effects by group, we include a dummy variable for the interaction between the group of interest and the coefficient that captures the treatment effect. We also include all additional corresponding interactions with the quintile, time, and treatment variables. More details about the baseline specification can be found in Table 3 and Section 5.2.1. More details about the augmented model can be found in Section B.8. Panel (a) presents the effect of providing different information regarding p in the *audit-statistics* message and panel (b) does the same for θ. Panel (c) compares the two *audit-threat* messages, i.e., the 50% threat of audit vs. the 25% threat of audit. Estimates should be interpreted as the effect of an additional percentage point of p and θ respectively in the information included in the letters on post-treatment VAT payments. The heterogeneity analysis is based on four characteristics: number of employees, age of the firm (i.e., the number of years registered with the IRS), type of tax return filed by the company, and sector of operation. The median number of employees in the experimental sample is 2.5, while the median age of the firms is 12.7 years. Sixty-eight percent of firms filed a comprehensive VAT tax return in 2013 while the remaining 32% filed a simplified VAT tax return. Regarding the distribution of sector of operation, 29% of firms are non-retail sellers of goods, 22% are retail sellers, and 49% are service providers. In all cases, analysis is restricted to letters that were not returned by the postal service.

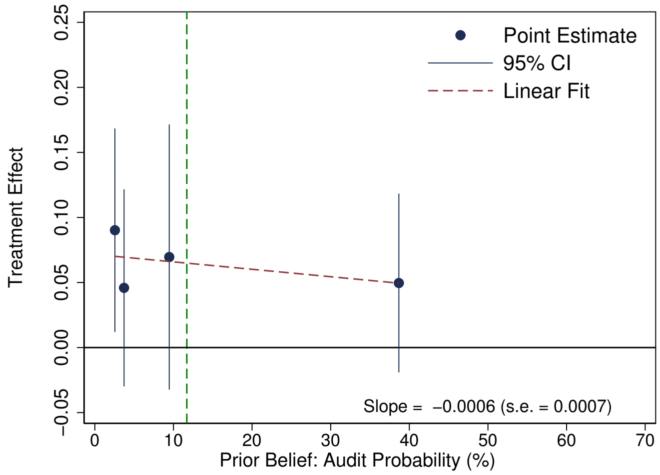
Figure B.14: Elasticities of Tax Payments with Respect to Audit Probability and Penalty Rate in *Audit-Statistics* and *Audit-Threat* Sub-Treatments by Type of Firm, Pre-Treatment Falsification Test



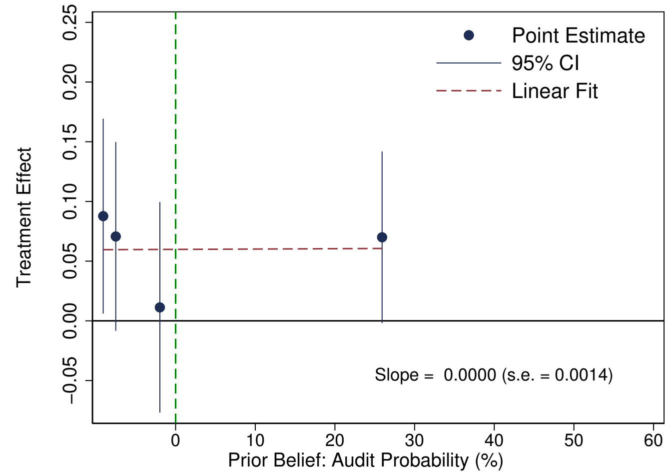
Notes: * significant at the 10% level, ** at the 5% level, *** at the 1% level. Standard errors are clustered at the firm level. Estimates presented in the figure correspond to falsification tests that compare two pre-treatment periods (October 2014–September 2015 vs. October 2013–September 2014) and are based on an augmented version of the difference-in-differences specification reported in equation (2), which compares treated firms that received different signals on p and θ. In all cases, we include an additional set of dummies for quintiles of the pre-treatment VAT payments, which are the groups from which we drew the sample of "similar firms" to calculate p and θ and the corresponding interactions with the time variable. The results are based on Poisson regressions with variables expressed in percentage terms, so the coefficients can be interpreted directly as elasticities. To capture the differential effects by group, we include a dummy variable for the interaction between the group of interest and the coefficient that captures the treatment effect. We include as well all additional corresponding interactions with the quintile, time, and treatment variables. More details about the baseline specification can be found in Table 3 and Section 5.2.1. More details about the augmented model can be found in Section B.8. Panel (a) presents the effect of providing different information regarding p in the *audit-statistics* message and panel (b) regarding θ. Panel (c) compares the two *audit-threat* messages, i.e., the 50% threat of audit vs. the 25% threat of audit. Estimates should be interpreted as the effect on post-treatment VAT payments of an additional percentage point of p and θ respectively in the information in the letters. The heterogeneity analysis is based on four characteristics: number of employees, age of firm (i.e., number of years registered with the IRS), type of tax return filed by the company, and sector of operation. The median number of employees in the experimental sample is 2.5, while the median age of the firms is 12.7 years. Sixty-eight percent of firms filed a comprehensive VAT tax return in 2013 while the remaining 32% filed a simplified VAT tax return. Regarding the distribution of sector of operation, 29% of firms are non-retail sellers of goods, 22% are retail sellers, and 49% are service providers. In all cases, analysis is restricted to letters that were not returned by the postal service.

Figure B.15: Effect of *Audit-Statistics* vs. *Baseline* by Prior Beliefs

a. Audit probability - *prior (calibrated to 11.7%)*



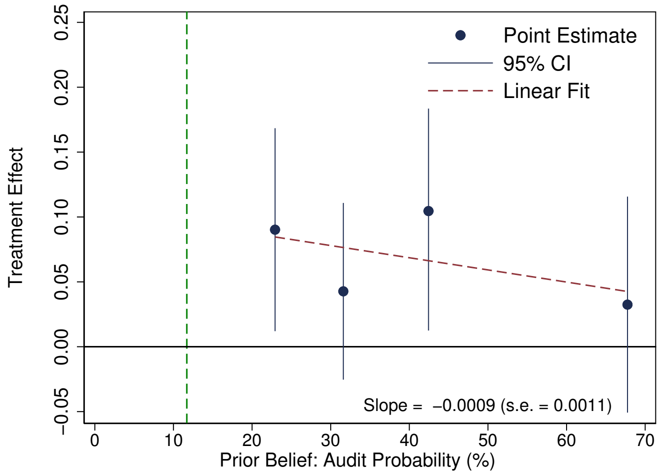
b. Audit probability - *prior* - *signal (calibrated to 11.7%)*



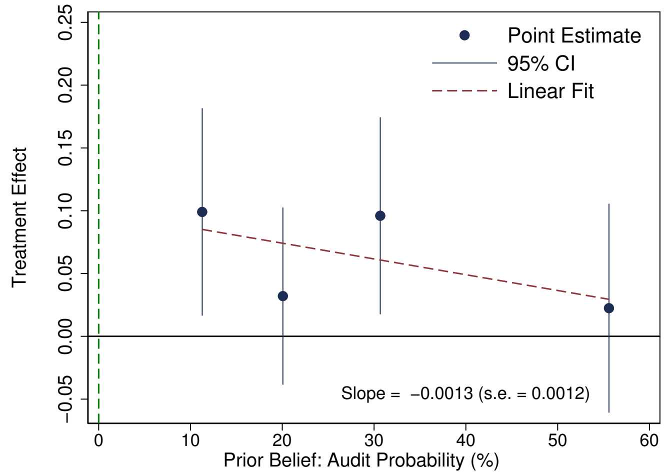
Notes: Panel (a) plots the first-year effect (October 2015–September 2016 vs October 2014–September 2015) of the *audit-statistics* letter on total VAT payments by quartiles of prior beliefs, while panel (b) reports the same results by quartiles of the difference between the prior belief and the signal sent in the *audit-statistics* message (N = 11,989). The prior belief, that is, belief before the experiment, is computed as $\hat{p}_i = 1 - \left(1 - \frac{\alpha_0 + N_i}{\alpha_0 + \beta_0 + T_i}\right)^3$ with parameters $\alpha_0 = 0.13$ and $\beta_0 = 1$ such that the mean prior belief about the probability of being audited at least once in the following three years matches the actual average probability observed in our sample. In panel (b), the signal for the placebo group was randomly assigned using the same strategy as for the *audit-statistics* group. The red dashed line represents the linear fit corresponding to the four estimates. In panel (a), the dashed vertical line represents the average actual probability of being audited as provided in the *audit-statistics* letters (11.7%). In panel (b), it represents the point at which the prior belief and the signal provided are equal. In both panels, each dot represents the estimated treatment effect for each quartile of the variable considered. Regressions are estimated using the baseline difference-in-differences specification reported in equation (1), but additional interaction terms are included for each quartile. More details are reported in the notes to Table 2 and Section 4. All effects are depicted with 95% confidence intervals. The results are based on Poisson regressions, which means the coefficients can be interpreted directly as semi-elasticities. Confidence intervals are computed with standard errors clustered at the firm level.

Figure B.16: Effect of *Audit-Statistics* vs. *baseline* by Prior Beliefs

a. Effect of audit probability signal
by quartile of prior beliefs (calibrated to 40.5%)



b. Effect of audit probability signal
by quartile of prior minus signal (calibrated to 40.5%)



Notes: Panel (a) plots the first-year effect (October 2015–September 2016 vs October 2014–September 2015) of the *audit-statistics* letter on total VAT payments by quartiles of prior beliefs, while panel (b) reports the same results by quartiles of the difference between the prior belief and the signal sent in the *audit-statistics* message (N=11,989). The prior belief is computed as $\hat{p}_i = 1 - \left(1 - \frac{\alpha_0+N_i}{\alpha_0+\beta_0+T_i}\right)^3$ where $\alpha_0 = 1.36$ and $\beta_0 = 1$ such that the mean prior belief about the probability of being audited at least once in the following three years matches the average probability perceived by firms in the *baseline* and *public-goods* groups according to survey answers (40.5%). In panel (b), the signal for the placebo group was randomly assigned using the same strategy as for the *audit-statistics* group. The red dashed line represents the linear fit corresponding to the estimates for the four quartiles. In panel (a), the green dashed line represents the average perceived probability (40.5%) for the comparison group, and in panel (b) it represents the point at which the prior belief and the signal provided are equal. In both panels, each dot represents the estimated treatment effect for each quartile of the variable considered. Regressions are estimated using the baseline difference-in-differences specification reported in equation (1), but additional interactions terms are included for each quartile. More details are reported in the notes to Table 2 and Section 4. All effects are depicted with 95% confidence intervals. The results are based on Poisson regressions, which means the coefficients can be interpreted directly as semi-elasticities. Confidence intervals are computed with standard errors clustered at the firm level.

Table B.1: Comparison of Firm Characteristics for Different Groups

		Experimental Sample			
	All firms (1)	Main (2)	Secondary (3)	Invited to the survey (4)	p-value test Survey. vs Non-Survey (5)
Share paid VAT (3 months pre-mailing)	0.778	0.927	0.894	0.926	0.875
	(0.001)	(0.002)	(0.005)	(0.004)	
Amount of VAT paid (3 months pre-mailing)	3.717	1.894	1.744	1.894	0.993
	(0.033)	(0.022)	(0.067)	(0.048)	
Years registered in tax agency	14.208	15.233	19.437	14.445	0.001
	(0.039)	(0.132)	(0.202)	(0.159)	
Share audited between 2013-2015	0.065	0.102	0.141	0.079	<0.001
	(0.001)	(0.003)	(0.007)	(0.006)	
Number of employees	12.653	4.838	5.600	6.431	0.000
	(0.657)	(0.208)	(0.093)	(0.867)	
Share file comprehensive tax return in 2013	0.447	0.684	0.999	0.569	<0.001
	(0.001)	(0.004)	(0.000)	(0.008)	
Share no retail goods sector	0.246	0.288	0.432	0.263	<0.001
	(0.001)	(0.003)	(0.008)	(0.007)	
Share retail goods sector	0.132	0.216	0.328	0.150	<0.001
	(0.001)	(0.003)	(0.007)	(0.006)	
Share services sector	0.622	0.496	0.240	0.588	<0.001
	(0.001)	(0.004)	(0.007)	(0.008)	
N	120,142	16,392	4,048	3,845	

Notes: This table presents the average characteristics for different subsamples of the universe of firms registered with the tax agency (standard deviations in parentheses). Column (1) includes all firms that submitted at least one VAT payment in 2014 or 2015. Column (2) includes the subset of firms selected for the experimental sample according to the criteria described in section 3.2. Column (3) represents a group of high-risk firms selected by the IRS and that received the *audit-threat* letter. Column (4) corresponds to firms in the main sample with valid email addresses on file with the IRS. These are the firms that were selected to participate in the online survey conducted after the experiment. Column (5) reports the p-value of the means test between firms in the experimental sample that were invited to the survey and firms that were not. All data are based on administrative tax records (monthly payments, annual tax returns, and auditing registers), and Columns (2) to (4) are restricted to firms with successful letter delivery, which is the group of firms used in the empirical analysis. Robust standard errors are provided in parentheses.

Table B.2: Delivery Status by Treatment Arm

	Main Sample					Secondary Sample		
	Audit Statistics (1)	Public Goods (2)	Audit Endogeneity (3)	Baseline (4)	p-value test (5)	Audit Threat (25%) (6)	Audit Threat (50%) (7)	p-value test (8)
Delivered (%)	72.215	72.094	73.004	72.869	0.781	77.360	78.086	0.555
	(0.396)	(0.888)	(0.876)	(0.879)		(0.875)	(0.861)	
In Process (%)	0.219	0.117	0.234	0.274	0.660	0.131	0.303	0.211
	(0.041)	(0.068)	(0.095)	(0.103)		(0.076)	(0.114)	
Returned (%)	19.694	21.057	20.569	19.312	0.290	11.932	11.953	0.982
	(0.352)	(0.807)	(0.798)	(0.781)		(0.678)	(0.675)	
Missing (%)	7.873	6.732	6.194	7.545	0.011	10.577	9.658	0.302
	(0.238)	(0.496)	(0.476)	(0.522)		(0.643)	(0.615)	
N	12,791	2,555	2,567	2,558		2,288	2,309	

Notes: This table reports the delivery status of the letters by treatment arm. Columns (1) through (4) report the distribution of delivery statuses by treatment arm for firms in the main sample, while column (5) reports the p-value of the joint equality test. Columns (6) through (8) provide an analogous breakdown for the secondary sample. Information on delivery status is provided by the Uruguayan Post Office using the certified delivery service. "Delivered" letters are letters that were delivered by the post office and signed for by the recipient. "In Process" letters are letters that were in the process of being delivered, but not yet delivered, when the post office provided us with the delivery report. "Returned" letters are letters returned to the sender as the post office was not able to deliver them. Finally, there are some letters for which the post office did not provide any information. These are coded as "Missing."

Table B.3: Tax Payments: Summary Statistics

	Mean (1)	SD (2)	10th (3)	25th (4)	50th (5)	75th (6)	90th (7)
VAT Amounts							
Post-treatment	6.47	7.77	0.44	1.30	3.74	8.48	16.55
Pre-Treatment	7.77	8.21	0.95	1.97	4.83	10.93	19.49
Retroactive VAT Amounts							
Post-treatment	0.30	1.40	0.00	0.00	0.00	0.00	0.62
Pre-Treatment	0.40	1.85	0.00	0.00	0.00	0.00	0.77
Concurrent VAT Amounts							
Post-treatment	6.16	7.51	0.33	1.16	3.52	8.07	15.84
Pre-Treatment	7.37	7.85	0.86	1.83	4.47	10.28	18.72
Other Taxes Amounts							
Post-treatment	3.30	5.43	0.00	0.95	1.81	3.52	7.42
Pre-Treatment	4.05	8.54	0.04	1.44	2.13	4.36	8.58
Total Taxes Amounts							
Post-treatment	9.77	11.31	1.04	2.68	6.14	12.39	23.33
Pre-Treatment	11.82	13.61	1.84	3.64	7.45	15.83	27.20
Some Payment							
Post-treatment	0.96	0.20	1.00	1.00	1.00	1.00	1.00
Pre-Treatment	1.00	0.06	1.00	1.00	1.00	1.00	1.00
More Than 6 payments							
Post-treatment	0.71	0.45	0.00	0.00	1.00	1.00	1.00
Pre-Treatment	0.89	0.32	0.00	1.00	1.00	1.00	1.00
More Than 3 payments							
Post-treatment	0.91	0.29	1.00	1.00	1.00	1.00	1.00
Pre-Treatment	0.98	0.15	1.00	1.00	1.00	1.00	1.00
VAT in Tax Return							
Post-treatment	7.19	9.93	0.00	0.49	3.84	9.69	18.97
Pre-Treatment	7.79	7.91	0.14	1.88	5.20	11.57	18.91

Notes: The statistics in this table correspond to firms that received the *baseline* letter (N = 2,064). The pre-treatment period spans from October 1, 2014 to September 30, 2015, and the post-treatment period from October 1, 2015 to September 30, 2016. "VAT Amounts" correspond to total VAT payments and withholdings (the sum of retroactive and concurrent VAT amounts). "Retroactive VAT Amounts" correspond to VAT payments and withholdings submitted for liabilities incurred two or more months prior to payment (e.g., VAT payments made in March 2016 for September 2015). "Concurrent VAT Amounts" include VAT payments made for the current or previous month. "Other Taxes Amounts" include payments of corporate income tax, wealth tax, and other taxes specific to a given business activity. "Total Taxes Amounts" is the sum of "VAT Amounts" and "Other Taxes Amounts." "VAT in Tax Return" is the final tax liability calculated in the yearly tax return that firms submit to the Uruguayan IRS.

Table B.4: Response Rates by Treatment Arm

	Audit Statistics (1)	Public Goods (2)	Audit Endogeneity (3)	Baseline (4)	Audit Threat (5)
a. Total firms by treatment arm					
Total	10,272	2,017	2,039	2,064	4,048
b. Invited to the survey					
Total	2,408	490	480	467	22
%	23.44	24.29	23.54	22.63	0.54
c. Started the survey					
Total	584	117	135	107	5
%	24.25	23.88	28.13	22.91	22.73
d. Owners that responded the survey					
Total	449	87	105	82	2
%	76.88	74.36	77.78	76.64	40.00
e. Non-Skip Rate - audit probability					
Total	351	64	82	65	1
%	92.61	92.75	95.35	95.59	100.00
f. Non-Skip Rate - penalty size					
Total	320	62	76	64	1
%	88.89	95.38	91.57	98.46	100.00

Notes: This table describes the composition of the online survey participants by treatment arm. Panel (a) reports the total number of firms by treatment arm. Panel (b) reports the raw number of invitations sent to firms in the experimental sample. The denominator for the percentage reported in this panel is the number of firms in each treatment arm reported in panel (a). Panel (c) reports the number of firms that started the survey (i.e., answered at least the first two questions). The denominator for the percentage reported in this panel is calculated using the totals reported in panel (b)as. Panel (d) reports the number of individuals who identified themselves as owners or who did not answer the question regarding their occupation in the firm, conditional on having started the survey (i.e., using the totals reported in panel (c) as the denominator). Finally, panels (e) and (f) report the non-skip rates–i.e., the response rate conditional on having answered the previous question–by treatment arm for the two key questions in the survey on perceptions of audit probability and penalty rate.

Table B.5: Effect of Treatment Status on Probability of Answering the Relevant Questions

	Answer: Perceived p		Answer: Perceived θ	
	Unconditional (1)	Conditional (2)	Unconditional (3)	Unconditional (4)
Treated	0.019	0.009	-0.010	-0.052
	(0.033)	(0.030)	(0.033)	(0.040)
Observations	1,105	267	1,105	267

Notes: * significant at the 10% level, ** at the 5% level, *** at the 1% level. Robust standard errors reported in parentheses. This table reports the effect of receiving the *audit-statistics* letter on probability of answering the two key questions in our survey analysis. The treatment variable is defined as one if the firm received the *audit-statistics* letter and zero if the firm received the *baseline* or the *public-goods* letter (our pooled control group). Column (1) reports the effect of receiving the *audit-statistics* letter on the probability of answering the audit probability question, as column (3) does for the penalty rate question. Because the treatment could have induced differential drop-out rates, Columns (2) and (4) report the same results but conditioning the sample to individuals who answered every single question up to the audit probability and penalty rates questions. We use OLS regressions to estimate the treatment effects. In all cases, the analysis is restricted to individuals who are owners of the firm or who did not answer the question regarding their occupation at it.

Table B.6: Balance Test for Survey Responses Conditional on Having Answered Perceptions Questions

	Main Sample				
	Audit Statistics (1)	Public Goods (2)	Audit Endogeneity (3)	Baseline (4)	p-value test (5)
Age	27.752	24.441	25.892	29.765	0.012
	(0.561)	(1.260)	(1.052)	(1.269)	
Female (%)	48.864	48.485	50.617	52.308	0.955
	(2.668)	(6.199)	(5.590)	(6.243)	
Capital city (%)	66.860	68.254	62.821	56.923	0.419
	(2.542)	(5.912)	(5.508)	(6.190)	
Very small firms (%)	11.966	9.231	14.634	15.152	0.685
	(1.735)	(3.618)	(3.927)	(4.447)	
Professional services (%)	50.829	44.118	44.578	52.941	0.540
	(2.631)	(6.066)	(5.489)	(6.098)	
Years registered with tax agency	16.366	13.385	15.025	17.894	0.036
	(0.539)	(0.925)	(1.033)	(1.221)	
Operates in one location (%)	61.582	62.687	64.634	67.164	0.829
	(2.589)	(5.953)	(5.312)	(5.781)	
One employee (%)	66.376	45.098	57.407	54.762	0.027
	(3.129)	(7.037)	(6.792)	(7.773)	
Audited in the last 3 years (%)	6.849	13.235	12.048	7.246	0.190
	(1.324)	(4.140)	(3.595)	(3.144)	
N	365	68	83	69	

Notes: This table reports the averages for different self-reported characteristics in the survey data by treatment group. The last column of each sample reports the p-value of a test in which the null hypothesis is that the mean is equal for all the treatment groups. "Age," "Female," and "Capital city" refer to individual characteristics of the respondent. "Very small firms" refers to firms that participate in a special VAT regime with a fixed VAT annual payment. "Professional services" include individuals that pay VAT because they provide professional services. Observations included in the analysis are restricted to individuals who are owners of the firm or who did not answer the question regarding their occupation at it and who answered the two questions relevant to our analysis, that is, those regarding perceived audit probability and penalty rate.

Table B.7: Average Effects of *Audit-Statistics*, *Audit-Endogeneity,* and *Public-Goods* Messages: Alternative Specifications and Data Sources

	Made VAT Payments			VAT Payments			VAT Tax Returns
	≥ 1	≥ 3	≥ 6				Balance
	Probit (1)	Probit (2)	Probit (3)	Poisson (4)	OLS (5)	Tobit (6)	Poisson (7)
a. *Audit-Statistics* (10,272 firms [6,088]) vs *Baseline* (2,064 firms [1,270])							
Post-Treatment	0.095	-0.019	-0.045	0.070***	0.459***	0.451***	0.056*
	(0.132)	(0.064)	(0.040)	(0.021)	(0.147)	(0.153)	(0.029)
Pre-Treatment				0.009	0.069	0.052	0.007
				(0.020)	(0.154)	(0.155)	(0.032)
b. *Audit-Endogeneity* (2,039 firms [1,233]) vs *Baseline* (2,064 firms [1,270])							
Post-Treatment	0.109	-0.021	-0.039	0.071***	0.444**	0.452**	0.059
	(0.161)	(0.084)	(0.052)	(0.028)	(0.199)	(0.206)	(0.041)
Pre-Treatment				-0.005	-0.034	-0.048	0.011
				(0.028)	(0.219)	(0.220)	(0.040)
c. *Public-Goods* (2,017 firms [1,240]) vs *Baseline* (2,064 firms [1,270])							
Post-Treatment	-0.099	-0.140	-0.116**	0.051**	0.301*	0.332*	-0.021
	(0.200)	(0.091)	(0.053)	(0.025)	(0.177)	(0.184)	(0.038)
Pre-Treatment				-0.003	-0.019	-0.015	0.015
				(0.024)	(0.186)	(0.187)	(0.039)

Notes: * significant at the 10% level, ** at the 5% level, *** at the 1% level. Standard errors reported in parentheses are clustered at the firm level. Treatment effects are estimated using the difference-in-differences specification reported in equation (1), which compares treated firms to control firms and pre-treatment to post-treatment periods using yearly aggregated variables. In the first row of each panel ("Post-Treatment"), the coefficient corresponds to a comparison between a post-treatment period and a pre-treatment period (October 2015–September 2016 vs October 2014–September 2015). The second row ("Pre-Treatment") presents a falsification test where two pre-treatment periods are compared (October 2014–September 2015 vs October 2013–September 2014). Additional details about the model are reported in the notes to Table 2 and Section 4. Panel (a) compares the *audit-statistics* message with the *baseline* letter, while panels (b) and (c) replicate the comparison for the *audit-endogeneity* and *public-goods* messages respectively. Column (1) presents the treatment effect on the probability of making at least one VAT payment in the post-treatment period using a Probit model. Columns (2) and (3) are identical to column (1), except that the dependent variable is defined as making at least three or six payments in the post-treatment period respectively. Columns (4), (5), and (6) present different estimation strategies for the intensive margin, i.e., the total amount of VAT paid. Column (4) corresponds to the baseline specification (a Poisson model), while column (5) uses an OLS regression and column (6) a Tobit regression. Column (7) reports the result of estimating a Poisson model on the final VAT liability calculated from annual tax returns. The number of observations corresponding to specifications reported in column (7) is reported between brackets.

Table B.8: Effects of *Audit-Statistics* and *Audit-Threat* Sub-Treatments: Alternative Specifications

	Made VAT Payments			VAT Payments			VAT Tax Returns
	≥ 1	≥ 3	≥ 6				Balance
	Probit (1)	Probit (2)	Probit (3)	Poisson (4)	OLS (5)	Tobit (6)	Poisson (7)
a. *Audit-Statistics* (N=10,272 firms [6,088])							
Audit Probability (%)							
Post-Treatment	0.205	0.481	0.121	-0.063	-0.517	-0.128	-0.218
	(1.333)	(0.669)	(0.380)	(0.242)	(1.783)	(1.851)	(0.257)
Pre-Treatment				0.141	1.141	1.190	-0.038
				(0.164)	(1.336)	(1.343)	(0.230)
Penalty Size (%)							
Post-Treatment	0.717	-0.353	-0.005	-0.033	-0.233	-0.152	-0.223
	(0.575)	(0.333)	(0.202)	(0.118)	(0.976)	(0.995)	(0.138)
Pre-Treatment				-0.128	-1.130	-1.148	-0.092
				(0.108)	(0.964)	(0.970)	(0.114)
b. *Audit-Threat* (N=4,048 firms [3,236])							
Post-Treatment	0.044	0.229	0.007	0.217	1.375	1.298	0.354**
	(0.435)	(0.240)	(0.170)	(0.142)	(0.901)	(0.997)	(0.178)
Pre-Treatment				-0.185	-1.215	-1.246	-0.249
				(0.157)	(1.072)	(1.117)	(0.167)

Notes: * significant at the 10% level, ** at the 5% level, *** at the 1% level. Standard errors reported in parentheses are clustered at the firm level. Treatment effects are estimated using the difference-in-differences specification reported in equation (2), which compares treated firms that received different signals on p and θ. In all cases, we include an additional set of dummies for quintiles of the pre-treatment VAT payments, which are the groups from which we drew the sample of "similar firms" to calculate p and θ and the corresponding interactions with the time variable. The results are based on Poisson regressions with variables expressed in percentage terms, which means the coefficients can be interpreted directly as elasticities. Panel (a) presents the effect of providing different information regarding p and θ in the *audit-statistics* message. Panel (b) compares the two *audit-threat* messages, i.e., the 50% threat of audit vs. the 25% threat of audit. In the "Post-Treatment" rows, the coefficient reported corresponds to a comparison between a post-treatment period and a pre-treatment period (October 2015–September 2016 vs October 2014–September 2015). In the "Pre-Treatment" rows, we present a falsification test where two pre-treatment periods are compared (October 2014–September 2015 vs October 2013–September 2014). Columns (1) and (2) show the effect on the extensive margin using two alternative strategies. Column (1) presents the treatment effect on the probability of making at least one VAT payment in the post-treatment period using a Probit model. Columns (2) and (3) are identical to column (1), except that the dependent variable is defined as making at least three or six payments in the post-treatment period respectively. Columns (4), (5), and (6) present different estimation strategies for the intensive margin, i.e., the total amount of VAT paid. Column (4) corresponds to the baseline specification (a Poisson model), while column (5) uses an OLS regression and column (6) uses a Tobit regression. Column (7) reports the result of estimating a Poisson model on the final VAT liability calculated from annual tax returns. The number of observations corresponding to specifications reported in column (7) is reported in square brackets.

Table B.9: Effects of *Audit-Statistics* Sub-Treatments: Alternative Specification of $p \cdot \theta$

	By Time Horizon		By Payment Timing		By Tax Type	
	First Year (1)	Second Year (2)	Retroactive (3)	Concurrent (4)	Non-VAT (5)	VAT + Non-VAT (6)
$p \cdot \theta$ (%)						
Post-Treatment	-0.292	-0.516	2.503	-0.406	0.398	0.017
	(0.529)	(0.560)	(3.095)	(0.532)	(0.531)	(0.439)
Pre-Treatment	-0.011	-0.011	-3.193	0.132	-0.173	-0.094
	(0.415)	(0.415)	(2.701)	(0.412)	(0.528)	(0.355)

Notes: * significant at the 10% level, ** at the 5% level, *** at the 1% level. Standard errors reported in parentheses are clustered at the firm level. Treatment effects are estimated using the difference-in-differences specification reported in equation (B.1), which compares treated firms that received different signals on $p \times \theta$. In all cases, we include an additional set of dummies for quintiles of the pre-treatment VAT payments, which are the groups from which we drew the sample of "similar firms" to calculate p and θ and the corresponding interactions with the time variable. This table presents the effect of providing different information regarding $p \times \theta$ in the *audit-statistics* message. In the "Post-Treatment" row, the coefficient reported corresponds to a comparison of a post-treatment period and a pre-treatment period. In the "Pre-Treatment" row, we present a falsification test where two pre-treatment periods are compared. Columns (1) and (2) report the effect of treatment by time horizon. The post-treatment effect reported in column (1) corresponds to the difference-in-differences estimate that compares October 2015–September 2016 vs. October 2014–September 2015. The post-treatment effect reported in column (2) is analogous but uses the second year after the treatment as the post-treatment period (i.e., October 2016 – September 2017). For the falsification tests, column (1) is based on a comparison of October 2014 – September 2015 and October 2013 – September 2014 while column (2) compares October 2014–September 2015 and October 2012–September 2013. Columns (3) and (4) present the first-year effect of treatment on retroactive (3) and concurrent (4) VAT payments. Columns (5) and (6) report the first-year results by type of tax. Column (5) presents the effect of the treatment on other (non-VAT) tax payments, while column (6) reports its effect on the total amount of taxes paid by the firms during the same period. In all cases, we restrict analysis to firms that effectively received the letter as reported by the postal service.

Table B.10: Average Effects of *Audit-Statistics*, *Audit-Endogeneity,* and *Public-Goods* Messages: Decomposition of Taxes

	Total Taxes (1)	VAT (2)	Other (3)	Other Taxes: Corporate (4)	Other Taxes: Wealth (5)	Other Taxes: PIT Withholdings (6)
a. *Audit-Statitstics* (10,272 firms) vs *Baseline* (2,064 firms)						
Post-Treatment	0.073***	0.070***	0.086**	0.058*	0.198	0.078
	(0.020)	(0.021)	(0.037)	(0.030)	(0.129)	(0.048)
Pre-Treatment	0.014	0.009	0.008	0.024	-0.057	0.030
	(0.021)	(0.020)	(0.043)	(0.030)	(0.166)	(0.046)
b. *Audit-Endogeneity* (2,039 firms) vs *Baseline* (2,064 firms)						
Post-Treatment	0.078***	0.071***	0.090*	0.069	0.204	-0.003
	(0.028)	(0.028)	(0.054)	(0.054)	(0.133)	(0.077)
Pre-Treatment	0.017	-0.005	0.056	0.063	0.019	0.072
	(0.028)	(0.028)	(0.055)	(0.049)	(0.169)	(0.080)
c. *Public-Goods* (2,017 firms) vs *Baseline* (2,064 firms)						
Post-Treatment	0.056**	0.051**	0.067	0.058	0.135	0.005
	(0.024)	(0.025)	(0.043)	(0.043)	(0.129)	(0.059)
Pre-Treatment	-0.015	-0.003	-0.038	-0.045	-0.005	-0.008
	(0.026)	(0.024)	(0.054)	(0.053)	(0.162)	(0.065)

Notes: * significant at the 10% level, ** at the 5% level, *** at the 1% level. Standard errors reported in parentheses are clustered at the firm level. Treatment effects are estimated using the difference-in-differences specification reported in equation (1), which compares treated firms to control firms and pre-treatment to post-treatment periods, using yearly aggregated variables. The results are based on Poisson regressions, which means the coefficients can be interpreted directly as semi-elasticities. Panel (a) compares the *audit-statistics* message to the *baseline* message, while panels (b) and (c) replicate the comparison for the *audit-endogeneity* and *public-goods* messages respectively. In the first row of each panel ("Post-Treatment"), the coefficient reported corresponds to a comparison of a post-treatment period (October 2015–September 2016) and a pre-treatment period (October 2014–September 2015). The second row ("Pre-Treatment") presents a falsification test where two pre-treatment periods are compared (October 2014–September 2015 and October 2013–September 2014). Column (1) reports the treatment effects on total tax payments. Columns (2) and (3) break down total tax payments into VAT and other tax payments. Columns (4), (5), and (6) break down other taxes into corporate income tax, property taxes, and personal income tax withholding respectively. In all cases, we restrict the analysis to firms that effectively received the letter as reported by the postal service.

Table B.11: Elasticities of Tax Payments with Respect to Audit Probability and Penalty Rate, *Audit-Statistics* and *Audit-Threat* Sub-Treatments: Breakdown of Taxes

				Other Taxes		
	Total Taxes (1)	VAT (2)	Other (3)	Corporate (4)	Wealth (5)	PIT Withholdings (6)
a. *Audit-Statitstics* (10,272 firms)						
Audit Probability (%)						
Post-Treatment	0.038	-0.063	0.109	0.316	-0.125	-0.288
	(0.208)	(0.242)	(0.240)	(0.293)	(0.358)	(0.573)
Pre-Treatment	0.063	0.141	-0.035	0.002	-0.609	0.517
	(0.147)	(0.164)	(0.230)	(0.257)	(0.652)	(0.509)
Penalty Size (%)						
Post-Treatment	-0.001	-0.033	0.061	0.001	0.021	0.608**
	(0.092)	(0.118)	(0.103)	(0.123)	(0.184)	(0.299)
Pre-Treatment	-0.078	-0.128	0.018	-0.066	0.450	-0.129
	(0.087)	(0.108)	(0.119)	(0.132)	(0.283)	(0.281)
b. *Audit-Threat* (4,048 firms)						
Audit Probability (%)						
Post-Treatment	0.233**	0.217	0.002	-0.060	0.146	2.567***
	(0.111)	(0.142)	(0.176)	(0.210)	(0.146)	(0.989)
Pre-Treatment	-0.257	-0.185	-0.067	-0.047	-0.100	-1.480
	(0.164)	(0.157)	(0.148)	(0.181)	(0.194)	(1.013)

Notes: * significant at the 10% level, ** at the 5% level, *** at the 1% level. Standard errors reported in parentheses are clustered at the firm level. Treatment effects are estimated using the difference-in-differences specification reported in equation (2), which compares treated firms that received different signals on p and θ. In all cases, we include an additional set of dummies for quintiles of the pre-treatment VAT payments, which are the groups from which we drew the sample of "similar firms" to calculate p and θ and the corresponding interactions with the time variable. The results are based on Poisson regressions with variables expressed in percentage terms, which means the coefficients can be interpreted directly as elasticities. Panel (a) presents the effect of providing different information regarding p and θ in the *audit-statistics* message. Panel (b) compares the two *audit-threat* messages, i.e., the 50% threat of audit vs. the 25% threat of audit. In the first row of each panel ("Post-Treatment"), the coefficient reported corresponds to a comparison of a post-treatment period (October 2015–September 2016) and a pre-treatment period (October 2014–September 2015). The second row ("Pre-Treatment") presents a falsification test where two pre-treatment periods are compared (October 2014–September 2015 and October 2013–September 2014). Column (1) reports the treatment effects on total tax payments. Columns (2) and (3) break down total tax payments into VAT and other tax payments. Columns (4), (5), and (6) break down other taxes into corporate income tax, property taxes, and personal income tax withholding respectively. In all cases, we restrict the analysis to firms that effectively received the letter as reported by the postal service.

Table B.12: Effects of *Audit-Statistics* by Prior Beliefs ($\hat{p}$) and Information Treatment's Audit Probability (p), and by Previous Audit Experience

		By $\hat{p}$		Audited in 2001-2015	
	All (1)	$\hat{p} < p$ (2)	$\hat{p} \geq p$ (3)	Yes (4)	No (5)
a. *Audit-Statitstics* (10,272 firms) vs *Baseline* (2,064 firms)					
Post-Treatment	0.070***	0.061***	0.086***	0.082**	0.066***
	(0.021)	(0.023)	(0.028)	(0.039)	(0.025)
Pre-Treatment	0.009	0.011	0.004	-0.028	0.023
	(0.020)	(0.020)	(0.024)	(0.043)	(0.022)

Notes: * significant at the 10% level, ** at the 5% level, *** at the 1% level. Standard errors reported in parentheses are clustered at the firm level. Treatment effects are estimated using the difference-in-differences specification reported in equation (1), which compares treated firms to control firms and pre-treatment to post-treatment periods, using yearly aggregated variables. In the first row ("Post-Treatment"), the coefficient reported corresponds to a comparison between a post-treatment period and a pre-treatment period (October 2015–September 2016 vs October 2014–September 2015). The second row ("Pre-Treatment") presents a falsification test where two pre-treatment periods are compared (October 2014–September 2015 vs October 2013–September 2014). Additional details about the model are reported in the notes to Table 2 and Section 4. Column (1) shows results from the main specification, i.e., the first year post-treatment effect estimates reported in Table 2. Column (2) reports the effect of the *audit-statistics* message on firms whose prior beliefs about probability of being audited were below the signal (p) reported in the letter they received. Column (3) reports the results for firms whose prior beliefs were above the reported p. In both cases, the reference group is the *baseline* group. Column (4) reports estimates for firms that were audited at least once between 2001 and 2015, and column (5) for the group of firms not audited during that period.

C Solution to the Model

The optimal evasion is given by maximizing the expected utility:

$$\max_{E\in[0,Y]} \frac{1-p\left(\frac{E}{Y}\right)-\epsilon}{1-\sigma}\Big(Y-\alpha\tau(Y-E)\Big)^{1-\sigma}+\frac{p\left(\frac{E}{Y}\right)+\epsilon}{1-\sigma}\Big(Y-\alpha\tau(Y-E)-(1+\theta)\tau E\Big)^{1-\sigma}$$

The FOC for the interior solution is:

$$\begin{aligned}\left(1-p\left(\frac{E}{Y}\right)-\epsilon\right)\Big(Y-\alpha\tau(Y-E)\Big)^{-\sigma}\alpha\tau-\frac{\frac{\partial p\left(\frac{E}{Y}\right)}{\partial E}}{Y(1-\sigma)}\Big(Y-\alpha\tau(Y-E)\Big)^{1-\sigma}-\\ -\left(p\left(\frac{E}{Y}\right)+\epsilon\right)\Big(Y-\alpha\tau(Y-E)-(1+\theta)\tau E\Big)^{-\sigma}(1+\theta-\alpha)\tau+\\ +\frac{\frac{\partial p\left(\frac{E}{Y}\right)}{\partial E}}{Y(1-\sigma)}\Big(Y-\alpha\tau(Y-E)-(1+\theta)\tau E\Big)^{1-\sigma}=0\end{aligned}$$

$$\begin{aligned}\Big(Y-\alpha\tau(Y-E)\Big)^{-\sigma}\left((1-p\left(\frac{E}{Y}\right)-\epsilon)\alpha\tau-\frac{\frac{\partial p\left(\frac{E}{Y}\right)}{\partial E}}{Y(1-\sigma)}(Y-\alpha\tau(Y-E))\right)=\\ \Big(Y-\alpha\tau(Y-E)-(1+\theta)\tau E\Big)^{-\sigma}\left(\left(p\left(\frac{E}{Y}\right)+\epsilon\right)(1+\theta-\alpha)\tau-\frac{\frac{\partial p\left(\frac{E}{Y}\right)}{\partial E}}{Y(1-\sigma)}(Y-\alpha\tau(Y-E)-(1+\theta)\tau E)\right)\end{aligned}$$

In the traditional *A&S* specification, with $\alpha=1$, $p_1=0$, and $\epsilon=0$, we can obtain a closed analytical form for the elasticities between the VAT payments and p and θ:

$$\frac{\partial log\left(\tau(Y-E)\right)}{\partial p}=-(1-\tau)\frac{-\frac{1}{\sigma}\left(\frac{p\theta}{1-p}\right)^{-\frac{\sigma+1}{\sigma}}\frac{\theta(1+\theta)}{(1-p)^2}}{\left(\theta\left(\frac{p\theta}{1-p}\right)^{-\frac{1}{\sigma}}+1\right)^2\left(\tau-(1-\tau)\left(\frac{\left(\frac{p\theta}{1-p}\right)^{-\frac{1}{\sigma}}-1}{1+\theta\left(\frac{p\theta}{1-p}\right)^{-\frac{1}{\sigma}}}\right)\right)}$$

$$\frac{\partial log\left(\tau(Y-E)\right)}{\partial \theta}=-(1-\tau)\frac{-(1+\theta)\frac{1}{\sigma}\left(\frac{p\theta}{1-p}\right)^{-\frac{\sigma+1}{\sigma}}\frac{p}{1-p}-\left(\frac{p\theta}{1-p}\right)^{-\frac{1}{\sigma}}\left(\left(\frac{p\theta}{1-p}\right)^{-\frac{1}{\sigma}}-1\right)}{\left(\theta\left(\frac{p\theta}{1-p}\right)^{-\frac{1}{\sigma}}+1\right)^2\left(\tau-(1-\tau)\left(\frac{\left(\frac{p\theta}{1-p}\right)^{-\frac{1}{\sigma}}-1}{1+\theta\left(\frac{p\theta}{1-p}\right)^{-\frac{1}{\sigma}}}\right)\right)}$$

The closed-form expressions are still available for the extensions with $\alpha<1$ and $\epsilon>0$, but not for $p_1>0$ (we thus use standard numerical methods to compute these elasticities).

D Additional Material

Figure D.1: Mass Advertising Campaign Sample: Billboard Poster, United Kingdom's Tax Authority, 2012

Notes: Advertising campaign by the United Kingdom's tax authority, HMRC (Her Majesty's Revenue and Customs), 2012. Previously hosted at http://www.gov.uk/sortmytax (no longer available, accessed through http://web.archive.org/)