NBER WORKING PAPER SERIES

WHERE DO MY TAX DOLLARS GO? TAX MORALE EFFECTS OF PERCEIVED GOVERNMENT SPENDING

Matias Giaccobasso
Brad C. Nathan
Ricardo Perez-Truglia
Alejandro Zentner

Working Paper 29789
http://www.nber.org/papers/w29789

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2022, revised September 2024

We are thankful for excellent comments from Raj Chetty, Matthew Weinzierl, Austan Goolsbee, Steve Levitt, James Poterba, Joel Slemrod, Dario Tortarolo, Sutirtha Bagchi and seminar participants at the NBER-Public Economics, UC Santa Barbara, UC Berkeley, University of Michigan, University of Chicago, University of Chicago-Advances in Field Experiments, CESifo, EU Tax Observatory, Federal Reserve Bank of Chicago, RIDGE, IIPF, Journees LAGV, and NOVAFRICA. This project was reviewed and approved in advance by the Institutional Review Board at The University of Texas at Dallas. The field experiment was pre-registered in the AEA RCT Registry (#0007483). To prevent contamination of the subject pool (e.g., that subjects could read about the hypotheses being tested), we posted the RCT pre-registration immediately after the deadline to file a protest had passed, but before conducting any analysis of the data. After the study is accepted for publication, we will share all the code and data through a public repository. Matias Giaccobasso is grateful for funding from the Research Council of Finland, grant # 346253. Xinmei Yang and Miriam Malament provided superb research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Where Do My Tax Dollars Go? Tax Morale Effects of Perceived Government Spending
Matias Giaccobasso, Brad C. Nathan, Ricardo Perez-Truglia, and Alejandro Zentner
NBER Working Paper No. 29789
February 2022, revised September 2024
JEL No. C93,H26,I22,Z13

ABSTRACT

Do perceptions about government spending affect willingness to pay taxes? We test this hypothesis with a natural field experiment that focuses on the allocation of property taxes to public schools. Our results show that taxpayers often misperceive the destination of their tax dollars. By introducing shocks to households' perceptions via an information-provision experiment, we find that perceptions of how tax dollars are used significantly affect the probability of filing a tax appeal. Moreover, the effects are consistent with reciprocal motivations: individuals are more willing to pay taxes if they believe that the government services funded by those taxes will provide greater personal benefit.

Matias Giaccobasso
Arkadiankatu 7
00100, Helsinki, Finland
Helsinki 00100
Finland
matias.giaccobasso@vatt.fi

Brad C. Nathan
brad.nathan@rutgers.edu

Ricardo Perez-Truglia
University of California, Los Angeles
110 Westwood Plaza
Los Angeles, CA 90095
and NBER
ricardotruglia@gmail.com

Alejandro Zentner
Naveen Jindal School of Management
University of Texas at Dallas
800 W Campbell Rd.
Richardson, TX 75080
azentner@utdallas.edu

A data appendix is available at http://www.nber.org/data-appendix/w29789

1 Introduction

Do individuals' perceptions of how the government spends tax dollars influence their willingness to pay taxes? From an individual's point of view, the taxes that a single person pays have a negligible impact on the quantity or quality of government services that the individual receives. In other words, one individual can free-ride on the taxes paid by all the other taxpayers. However, we hypothesize that perceptions about government spending may still affect tax compliance through tax morale. More precisely, if individuals perceive that their taxes are being used to fund government services that are of personal benefit to them, they may feel a heightened *moral obligation* to pay their taxes. From hereon, we refer to this as the "reciprocal motivation" hypothesis.[1] The reciprocal motivation channel may help explain differences in tax morale, such as why some individuals are more willing to pay taxes than others. This channel could also be relevant for policymakers, as government communications about public finances might influence tax morale. In this paper, we test the reciprocal motivation hypothesis using a natural field experiment in a high-stakes context and through revealed preferences.

Our experiment leverages the context of property taxes, which represent an important source of revenue for governments in the United States and around the world.[2] For instance, U.S. property tax revenues in 2019 were estimated at $577 billion (Tax Policy Center, 2021a), nearly three times higher than the corporate income tax.[3] In the United States, virtually all counties rely heavily on property taxes to fund key government services such as schools, parks, and roads. School funding typically makes up the largest component of property taxes.

This setting offers two key advantages to test our hypothesis. First, our research design leverages the straightforward path between property taxes and the government services they fund, allowing us to identify who benefits from what. In particular, households with children enrolled in local public schools benefit directly from education funded by property taxes, whereas households without children enrolled in local public schools do not. For brevity, we hereinafter refer to households with children enrolled in local public schools as "households *with* children" and those without as "households *without* children."

The second advantage of this setting is that we can study the willingness to pay taxes

[1] This channel is similar to what Luttmer and Singhal (2014) call *reciprocal motivation*: "the willingness to pay taxes in exchange for benefits that the state provides to them (...) even though their pecuniary payoff would be higher if they didn't pay taxes." This channel is also related to a normative principle known as *benefit-based taxation*, which can be briefly described as the "idea of basing tax liabilities on how much an individual benefits from the activities of the state" (Weinzierl, 2018).

[2] For other studies on property taxes, see for example Cabral and Hoxby (2012), Jones (2019), Avenancio-Leon and Howard (2022), Nathan et al. (2020) and Dzansi et al. (2022).

[3] For reference, the 2019 federal income tax generated $1.717 trillion in revenue and corporate income tax generated $230 billion (Tax Policy Center, 2021b).

via revealed preferences using households' decisions to file property tax appeals, also known as tax protests (Nathan et al., 2020). Filing an appeal is a consequential, high-stakes action that households can take to reduce the amount they have to pay in property taxes.[4] In a nutshell, households can use the subjective nature of the property appraisal process in their favor. If they feel their taxes are too high, they can file a tax appeal to reduce their tax burden. The majority of these tax protests are successful, and they typically save households hundreds of dollars per year (Nathan et al., 2020).

We conducted the field experiment in Dallas County, Texas. We focus on one county because, from a logistical perspective, it is more practical to implement a field experiment in a single location. With an estimated population of about 2.6 million in 2020 (U.S. Census Bureau, 2021), Dallas County is the second-largest county in Texas and has a larger population than 15 of the 50 U.S. states. The county is diverse along many dimensions, such as ethnicity, and has a relatively even distribution of Democrat and Republican supporters.[5]

We sent an invitation letter to a sample of 78,128 households, encouraging them to participate in an online survey. The response rate was 2.7%, which aligns with the expected rate based on the recruitment method (Sinclair et al., 2012; Nathan et al., 2020). Our main subject pool comprises 2,110 respondents who completed the survey a few weeks before the deadline to file a property tax appeal between April and May of 2021. Our survey elicited key characteristics of the household, such as whether it has children enrolled in public schools. We match survey responses to administrative records from the county assessor's office. The rich administrative data allows us to determine, among other things, if the survey respondent subsequently filed a tax appeal.

Our experimental design can be summarized as follows. First, we measure respondents' perceptions about the share of their own property taxes that corresponds to school taxes and thus funds local public schools. For brevity, in the remainder of the paper, we refer to this percentage as the household's "school share." The school share for the average household in Dallas County is about 49.78%. We can assess respondents' misperceptions about the allocation of their tax dollars by comparing their estimates of the school share to the actual figures from administrative records. To study the causal effect of beliefs about how the government spends its tax dollars, the survey embeds an information-provision experiment. After eliciting respondents' prior beliefs, we provide a random half of them with information about their respective true school shares. We can assess whether the provision of information

[4] When studying attitudes towards taxation, social scientists rely primarily on survey data. However, survey data have some well-known limitations, such as social desirability bias. For example, some individuals may *say* that they are willing to pay more in taxes but would *choose* otherwise when facing real stakes.

[5] For example, in the 2012 presidential election, Barack Obama received 57% of the votes in Dallas County, whereas Mitt Romney received 42% (the remaining 1% of votes went to third-party candidates).

influences their posterior beliefs, as indicated by survey responses, and their decisions to file a tax appeal, as reflected in administrative data.

The information-provision experiment creates exogenous variation in respondents' posterior beliefs about the fraction of their property taxes that funds local schools. To illustrate, a subject who perceives her or his school share amount to be 30% may be informed that the actual share is 50%. According to our hypothesis and as noted in the randomized control trial (RCT) pre-registration, the expected effects of the information shock depend on whether the household has children enrolled in public schools. Upon learning that the school share is higher than originally thought, households *with* children should become less likely to file a tax appeal because they learn that they benefit more from the allocation of government services than they originally believed. Conversely, households *without* children enrolled in public schools should become more likely to file a tax appeal because they learn that they benefit less from the allocation of government services than they originally thought.

Our experiment also included a second treatment arm providing information on how the county redistributed part of the tax revenues from richer to poorer school districts (also known as "recapture system"). Before conducting the experiment, we expected that both treatment arms would be adequately powered to detect effects. However, we found that we were underpowered in the second treatment arm. For transparency, we still report the full analysis for the second treatment arm in the appendix.

Before any adjustment resulting from tax appeals, the average subject in our sample owns a home worth $349,988 and pays $7,738 in annual property taxes. There is significant variation in the degree to which households benefit from public education, which is important for our research design: households *with* children account for 25.5% of the sample, and households *without* children account for the remaining 74.5%. Owners can protest "directly" on their own (the main focus of this paper) or they can hire an agent to protest on their behalf. For reference, 30.1% of homeowners in the control group (i.e., those who did not receive any information treatment) protested directly in 2021.

The results indicate that even though the information is publicly available and easily accessible, most households have misperceptions about their respective school shares. When provided with accurate information, we observe that households strongly update their beliefs. Moreover, consistent with the reciprocal motivation hypothesis, the information affects the probability of tax appeals differently for households *with* and *without* children. We start by measuring the average treatment effects of the information. On average, both households *with* and *without* children increase their perception about the share of taxes going to schools. Households *with* children become *less* likely to protest, while households *without* children become *more* likely to file an appeal.

To better assess the magnitude of these effects, we use a Two-Stage Least Squares (2SLS) model. This model estimates the causal effect of beliefs about school share on the probability of protesting by leveraging the exogenous shocks to beliefs induced by the information-provision experiment. We find that the effects are not only statistically significant but economically significant too. Our baseline estimates imply that increasing the (perceived) school share by 10 percentage points (pp) would cause a drop of 4.09 pp in the probability of filing a protest among households *with* children and an increase of 2.78 pp in the probability of protesting among households *without* children. These effects amount to 12.1% and 9.6% of the corresponding baseline protest rates, respectively. Furthermore, we show that the effects on protests were consequential in that they subsequently affected the assessed home values. These results are robust to a host of alternative specifications and falsification tests.

Property taxes function in much the same way across counties in Texas and similarly throughout the country (Dobay et al., 2019; World Bank, 2019; Nathan et al., 2020).[6] On one hand, these similarities suggest that our results from Dallas County may be reasonably generalizable to other U.S. counties. On the other hand, there are some unique features of our sample that should be considered when extrapolating to other contexts. First, there are some significant differences between our survey respondents and the broader population of homeowners in Dallas County – for example, survey respondents are more likely to file a tax protest. As a result, the average effects of the information in the subject pool may differ from those among the general population. Second, we conducted our experiment in an area with relatively high-quality public schools. The effects of information about school spending might be much weaker, or even null, in areas where public goods are of lower quality or where local governments are perceived as corrupt. Third, while we examined the effects on tax appeals, the effects may be different for other margins of tax compliance, such as tax delinquency or tax evasion. More broadly, in the language of List (2020), our results can be seen as a wave-1 insight that establishes initial causality and provides the first tests of theory.

Our study relates and contributes to the literature on tax compliance.[7] To explain why tax compliance varies among taxpayers and countries, there are two schools of thought: institutions and tax morale. Abundant research shows that institutions have large effects on tax compliance (Slemrod, 2019). For example, the introduction of withholding and third-party reporting caused a massive increase in tax compliance (Bagchi and Dušek, 2021). In comparison, compelling evidence supporting the role of tax morale is more limited (Luttmer and Singhal, 2014; Slemrod, 2019). Our study provides novel evidence showing that tax

[6] For instance, property taxes provide a significant source of school funding in most of the U.S. (Chen, 2021).
[7] The role of tax morale in shaping individuals' attitudes toward taxation, and in particular the reciprocity mechanism, has also been studied in the lab or using "stated" preferences in the tax aversion literature (e.g., Sussman and Olivola, 2011; Kessler and Norton, 2016; Douenne and Fabre, 2022).

morale *can* be a significant driver of tax compliance and provides evidence on a specific mechanism: reciprocal motivation.

There are a few related studies providing suggestive evidence that are consistent with our results which find that rewarding taxpayers with public services has a positive effect on their subsequent tax compliance.[8] Cullen et al. (2021) provides quasi-experimental evidence that tax evasion decreases when the political party of the taxpayer is in control of the presidency.[9] Carrillo et al. (2021) conducted an experiment in which 400 taxpayers from an Argentine municipality were randomly selected to be publicly recognized for their tax compliance and were awarded the construction of a sidewalk near their homes. They found that their intervention had a positive effect on subsequent tax compliance. Krause (2020) found that tax payments increased 27% as a consequence of an intervention that increased municipal garbage removal in some randomly selected census blocks in Carrefour, Haiti. Lastly, Kresch et al. (2023) provides non-experimental evidence from Manaus, Brazil, showing that households with access to the city sewer system are more likely to pay property taxes. We contribute to the literature in two ways. First, we provide experimental evidence to identify causal effects. Second, we identify and quantify more precisely the causal mechanisms at play. For instance, ours is the first study to measure taxpayers' perceptions about the destination of their tax dollars. Moreover, interventions from previous studies combine multiple features, making it challenging to identify the precise mechanisms at play. For example, the intervention in Carrillo et al. (2021) jointly awards taxpayers with social recognition and the construction of a sidewalk near their homes. The bundled nature of this experimental intervention makes it impossible to identify whether the effects on tax compliance are due to social recognition, the construction of the sidewalk, or a combination of both.

Our setting has some differences with respect to the existing literature on tax morale that are worth discussing. We study taxes in a high enforcement context. The existing institutions, such as tax payments via escrow (Cabral and Hoxby, 2012), tax lien sales, foreclosure of delinquent properties, and wage garnishments, lead to extremely high rates of property tax compliance. In our subject pool, for example, only 0.42% failed to pay their property taxes.[10] In comparison, the related studies were conducted in developing countries such as Argentina, Brazil, Haiti, and Malawi, where more than 50% of households failed to pay their property taxes. In low enforcement contexts, there is scope to study tax morale by looking at outcomes such as tax evasion or tax delinquency. However, in high enforcement

[8] Beyond taxation, recent quasi-experimental evidence demonstrates how the salience of government spending can affect electoral outcomes (Huet-Vaughn, 2019; Ajzenman and Durante, 2022).

[9] In the context of a laboratory study, Huet-Vaughn et al. (2019) provides related laboratory evidence showing that partisanship can affect attitudes towards taxation.

[10] For the property taxes from the 2021 fiscal year, which were due on January, 31st of 2022, 9 out of the 2,110 households (0.42%) of our subject pool had not paid their property taxes by October 22nd, 2022.

contexts, households end up paying their taxes whether they like it or not. By looking at a different margin of tax compliance (legal tax avoidance via appeals), we can study tax morale even in a context of high enforcement. However, it is important to note that these differences may limit the direct comparability of our findings to those of studies focusing on tax evasion. For instance, illegal tax evasion, legal tax avoidance, and legal property tax protests are likely viewed through different moral frameworks by taxpayers: tax evasion is probably viewed in a much more negative light than tax protests.

Our findings can also speak to a literature that uses correspondence experiments to study tax morale (Slemrod, 2019). These studies typically randomize a message of moral suasion, such as a reminder that paying taxes is the right thing to do, and then measure the effects of that message on subsequent tax compliance.[11] Antinyan and Asatryan (2019) conducted a meta-analysis of about 1,000 treatment effects from 45 randomized control trials and concluded that "interventions pointing to elements of individual tax morale (...) are on average ineffective in curbing tax evasion, while deterrence nudges (...) are potent catalysts of compliance."[12] In a review, Slemrod (2019) reaches a similar conclusion.[13] We want to highlight two advantages of our experimental design that enabled us to detect tax morale effects in ways that would have been difficult with traditional correspondence studies. First, because we randomized the information provision within the survey, we can confirm that subjects actually saw the information. In traditional correspondence studies, the vast majority of subjects may not read the correspondence carefully or might discard it without even opening it (Bottan and Perez-Truglia, 2020; Nathan et al., 2020). Second, it is crucial to consider the possibility of highly heterogeneous effects of information. The same piece of information can have opposite effects on different groups, such as households *with* or *without* children or those with prior beliefs that under- or over-estimate the truth.[14] These opposing effects across different groups may, on average, cancel each other out, leading to the erroneous conclusion that tax morale is irrelevant to tax compliance decisions. Our methodological approach allows us to disentangle these sources of heterogeneity. We hope these methodological insights aid researchers to better explore tax morale in other settings.

[11] For a seminal contribution, see Blumenthal et al. (2001).

[12] The messages of moral suasion used in these studies sometimes, but not always, include information related to government services (see e.g., Castro and Scartascini, 2015; Hallsworth et al., 2017; Bowers et al., 2020; De Neve et al., 2021; Carrillo et al., 2021; Bergolo et al., 2023). When looking specifically at the group of interventions including information on government services, Antinyan and Asatryan (2019) arrive at the same conclusion. As for the broad group of moral suasion messages, interventions including information on government services are also ineffective. However, despite the average findings suggesting null effects, there are a few exceptions (e.g., Del Carpio, 2014; Bott et al., 2020).

[13] Slemrod (2019) concludes, "In sum, a plethora of studies have failed to find evidence that appeals to tax morale, defined broadly, affect taxpayer behavior in the short run when delivered via a one-time mailing."

[14] For related evidence on the importance of treatment heterogeneity in the context of tax morale, see Castro and Scartascini (2015).

Lastly, this study is related to a small but growing body of work that seeks to incorporate normative considerations into tax policy design (Mankiw and Weinzierl, 2010; Weinzierl, 2014; Saez and Stantcheva, 2016).[15] This literature is new and mostly theoretical. There is some empirical evidence, but it is limited to survey data, such as asking individuals to choose between hypothetical tax policies (Weinzierl, 2014; Saez and Stantcheva, 2016; Weinzierl, 2017). We fill this gap in the literature by providing evidence based on real-world behavior and in a natural, high-stakes context.

The remainder of the paper proceeds as follows. Section 2 describes the institutional context. Section 3 discusses the data, experimental design and implementation. Section 4 presents the results. The last section concludes.

2 Institutional Context and Conceptual Framework

2.1 Property Taxes and Public Schools

In Dallas County, property taxes fund various public services, such as schools, roads, parks, and police and fire departments.[16] In 2021, the average home in Dallas County was worth $327,690. The average estimated property tax bill was $6,370, implying an effective tax rate of 1.94%.[17] Texas does not have a state income tax. To compensate, revenues from property taxes fund a greater share of local government services in Texas than in many states. School taxes comprise the largest share of property taxes, accounting on average for nearly half (49.78%) of the total property tax bill. There is variation in the share of school taxes between households. For example, in our subject pool, the school share is 41.57% for the 10th percentile and 57.01% for the 90th percentile.[18] The second largest component is the city tax (accounting for approximately 28% of property taxes), followed by the hospital (10%), county (8%), college (4%), and special district (<1%) taxes.

Dallas County has 16 major Independent School Districts (ISDs). Homeowners who live within the geographical boundaries of a given ISD jurisdiction are subject to the tax rate for that ISD. Households also have the right to send their children to their district's K–12 public schools. All households must pay school taxes, regardless of whether they have

[15] For instance, the normative considerations related to equality of opportunity or poverty alleviation.

[16] In this sub-section, we present the most important features of the institutional context. More details on the definition of the samples of interest and additional information on the property tax system in Texas are reported in Appendices A.1 and A.2.

[17] There is heterogeneity in the effective tax rate that households pay, with some households paying a rate that is as much as 1 pp below or above the average rate – for more details, see Nathan et al. (2023).

[18] These differences are due to a host of factors such as differences in jurisdictional tax rates across districts and household-specific exemptions such as the homestead cap – more details in Nathan et al. (2020).

children enrolled in public schools. The public schools in Dallas County are generally of great quality.[19] Alternatively, homeowners can send their children to private schools, opt for homeschooling, or enter a lottery for the chance to send their children to charter schools, which are tuition-free public schools that receive state and federal funding and do not receive funding from the district's property taxes.[20]

Our definition of "households *with* children" comprises households with children enrolled in local public schools administered by the ISDs since these are funded with property taxes. The category "households *without* children" includes households without school-aged children as well as those with school-aged children but attending private schools, charter schools, or home-schooling. Note that while charter schools are public institutions, we do not include them in the definition of "households *with* children" because charter schools are not funded with property taxes. In the subject pool, the category "households *without* children" is comprised of 83.1% of households without school-aged children and a small minority of 16.9% of households with school-aged children but attending private schools, charter schools, or home-schooled. Unfortunately, we did not include a survey question to differentiate between the three subcategories. However, based on publicly available statistics, our best guess is that the category "households *without* children" includes approximately 4.1% of households with children in private schools, 6.4% with children in charter schools, and 6.4% with home-schooled children.[21]

2.2 Tax Protests

Each year, the DCAD performs market value appraisals for all homes in the county. Each appraisal results in a "proposed value" for the home, which is an estimate of the home's market value as of January 1st. The DCAD makes this information available to all homeowners through its website and by mail.[22] The notice includes additional information, such as the estimated taxes due based on the property's proposed values and how property taxes are allocated across jurisdiction types (e.g., school and city taxes). After the notifications are sent, households have a month from the notification date to file a protest if they disagree

[19] For example, according to `www.GreatSchools.org`, 100% of the schools in the Highland Park ISD have above-average ratings in Texas, whereas 43% of schools in the Mesquite ISD have below-average ratings (data accessed on November 4, 2021).

[20] ISDs in Texas can contract with charter schools (Senate Bill 188). No such contract exists in Dallas County.

[21] These calculations are based on estimates from the Texas Public Policy Fundation (2022) indicating that in the Dallas-Fort Worth-Arlington metropolitan area, 79.3% of students were enrolled in local public schools, 7.9% in charter schools, 5.0% in private schools and 7.8% were home-schooled in the 2020-2021 school year.

[22] A sample notification, called the "Notice of Appraised Value," is shown in Appendix H. This notification is available online for every household and is also sent by mail to some households (e.g., households with proposed values that increased from the previous year).

with the proposed value. In 2021, the DCAD notified the proposed values on April 16; as a result, the deadline to protest was May 17.

Homeowners can file a protest directly or hire an agent to help them. Agents typically charge a combination of a flat fee and a percentage of the tax savings, which can be as high as 50% of the savings. We explain in Section 3.5 that our main focus is on direct protests. Homeowners can file a direct protest by mail using a form included with their mailed notice, or they can file a protest online using a simple tool called uFile.[23] After reviewing the argument, the DCAD can (and often does) make an offer by mail or phone to reduce the home's assessed value. If the homeowner refuses to pay this settlement value or the DCAD does not offer a settlement, the appeal proceeds to a formal hearing with the Appraisal Review Board. Once the protests are resolved, the new tax amount becomes payable immediately or at the billing date if it is later (i.e., on October 1st, 2021). Unpaid taxes eventually become delinquent (e.g., unpaid 2021 property taxes became delinquent on January 31, 2022).

A key feature of this setting is the difficulty in estimating home market values for homes that have not been sold recently, a process that involves significant ambiguity and subjectivity. The DCAD uses statistical models and large datasets (e.g., recent home sales) to estimate each property's market value. However, even multibillion-dollar companies like Zillow and Redfin have a hard time estimating market values using statistical models (Parker and Friedman, 2021). This ambiguity in home value is important for interpreting our results because it implies that households are not trying to objectively "correct" estimates from the DCAD. Instead, they present a data point (e.g., the sale price of a neighboring home) to support their protest. This distinction is consistent with what was expressed in our conversations with officials from some of the county appraisal districts in Texas. Their prevailing view is that households use the subjective nature of the appraisal process as an excuse to complain about their taxes being too high (for details, see Nathan et al., 2020) and not necessarily to complain about the county's estimate of their home value.

Tax protests in Dallas County operate similarly to how they operate across all 241 counties in Texas. Indeed, Appraisal District Boards (e.g., DCAD in Dallas) are political subdivisions of the Texas Comptroller of Public Accounts, and homeowners in every county in Texas can protest using the same form provided by the state. For this reason, the prevalence of appeals is likely similar across counties in Texas. Although some form of property value protests exists in other states in the United States, we are unaware of any compiled statistics at

[23] To protest online, homeowners need to look up their account (e.g., searching for their names or addresses) and then follow some straightforward steps in the uFile system. To protest by mail, households who received a notification from the DCAD can use the protest form included with the notification, and households that did not receive a notification can file by mailing a printed form that can be obtained online on either the DCAD's or the Texas Comptroller's website. In 2020, about 75% of direct protests were filed online, while the remaining 25% were filed by mail (Nathan et al., 2020).

the national level. However, we know there are similarities in how tax appeals work, and therefore, appeal rates in other states may be similar to those of Texas. For instance, 16.0% of all Dallas County households filed a protest in 2020.[24] In comparison, the percentage of households who filed a tax appeal in Cook County, Illinois, ranged from 10% to 21% annually between 2002 and 2015, with a mean of 14.6% (Avenancio-Leon and Howard, 2022). Nevertheless, tax appeals may be less common in other states. For instance, some counties charge a processing fee for filing an appeal.[25] In some states, such as California, appraised values are updated only when properties are sold, which provides less incentive to file a tax appeal if the homeowner purchased the home several years ago (Nathan et al., 2020). As a result, we estimate that the rates of tax appeal in San Francisco were below 4% in recent years.[26]

2.3 Conceptual Framework

In this section, we outline the key predictions of the reciprocal motivation channel. For brevity, we discuss the main intuitions here and provide the formal model in Appendix C.

Consider a government that allocates a household's tax dollars into two spending categories. A portion of the tax dollars funds local public schools referred to as *school* expenditures, while the remainder is allocated to a residual category, *non-school* expenditures, which includes services such as police and roads. There are two types of households: those *with* children, who benefit from both school and non-school expenditures and those *without* children, who do not benefit from school spending but do benefit from non-school expenditures in the same way as households *with* children. The reciprocal motivation channel assumes that households are more willing to pay taxes (and thus less likely to file a tax appeal) when they perceive that they benefit more from the government services funded by their tax dollars.

Assume there is an increase in a household's perceived school share.[27] We are interested in how this change in perception affects the likelihood of the household filing a tax appeal. The reciprocal motivation channel makes a key prediction: there should be a differential effect between households *with* children and those *without* children. The intuition is straightforward.

[24] See column (1) in Table 1.

[25] For example, San Francisco County charges a fee of $60 and Los Angeles County charges a $46 fee.

[26] According to San Francisco Chronicle (2023), there were 1,878 residential and commercial appeals in San Francisco during the 2022-2023 cycle and 6,836 during the 2023-2024 cycle. Based on the number of residential and commercial properties in San Francisco County (source: `https://www.sfassessor.org/news-information/property-data-0`) we estimate appeal rates of 0.8% and 3.2%, respectively.

[27] This simple model assumes that we can change the belief about the school share while holding other beliefs constant. In practice, however, households might become more pessimistic about how efficiently school funds are used in response to the school share information. In that case, households *with* children might become more likely to protest.

On the one hand, when the school share increases, households *without* children are more likely to protest since they do not benefit from school expenditures and thus perceive more of their tax dollars as being used for services they do not consider valuable. On the other hand, for households *with* children, the direction of the effect depends on their preference between school and non-school expenditures. For example, if they prefer school expenditures, households *with* children should be less likely to protest when their perceived school share increases.[28] Regardless of this preference, the effect on the probability of protesting should always be lower for households *with* children than for those *without* children. Intuitively, when the perceived school share increases, both types of households lose in non-school expenditures; however, households *with* children at least gain in school expenditures. In the remainder of this paper, we leverage a field experiment to test these predictions.[29]

3 Data, Experimental Design, and Implementation

3.1 Data and Sample Selection

To carry out our experiment, we use publicly available administrative data on property taxes and property tax protests from the Dallas County Appraisal District (DCAD).[30] This information includes details about ownership, address, and property characteristics, like square footage and the number of bedrooms and bathrooms, for the different taxing jurisdictions (community college, hospital, 31 cities, 16 ISDs, 12 Special Districts and the county itself). Additionally, the data include historical yearly records of proposed and certified market values, exemption amounts, taxable values, and tax rates. Furthermore, detailed information is available on property tax protest records, separating protests conducted directly by the owner and protests conducted with the help of an agent. The raw data available on the DCAD website contains information on more than 800,000 residential and commercial properties. We begin with a sample of 400,193 owner-occupied residential properties, which we will refer to as the "Universe" sample.[31] When necessary, we supplement the administrative records

[28] Or, if they prefer non-school expenditures, they should become more likely to protest.

[29] Appendix C presents several extensions. For instance, it considers cases where households *without* children derive value from expenditures in public schools because they are not selfish, have benefited from public schools in the past, or expect to benefit from public school expenditures in the future.

[30] The latest version of the data is available in https://www.dallascad.org/DataProducts.aspx. We downloaded most of the baseline information on April 16, 2021, the day the DCAD notified the proposed values for 2021. Data corresponding to the post-treatment period was downloaded for the current version of the paper on April 25, 2024.

[31] We arrived at this sample by applying several filters, such as excluding commercial properties, non-owner-occupied residences, and properties in two ISDs − Ferris and Grapevine-Colleyville− from which only a marginal area belongs to Dallas County. See Appendix A.2 for details of the selection criteria.

with data from other sources, such as the National Change of Address (NCOA) records.

Of the 400,193 properties, we selected a sub-sample of 78,128 households to receive a letter inviting them to participate in an experimental survey. We will refer to this sample as the "Letter Sample." We designed the sample criteria to ensure a wide representation of richer and poorer school districts. More specifically, we oversample households from the richer ISDs in Dallas County (Carrolton, Coppell, and Highland Park). All homeowners in these three districts were selected for the letter sample. We also oversample households who experienced increases in their estimated taxes because they are more likely to consider filing a tax protest (Jones, 2019; Nathan et al., 2020).[32] The sample criteria are explained in more detail in Appendix A.2.

Panel (a) of Table 1 presents descriptive statistics for some key variables based on the information available in the administrative records at baseline. Column (1) corresponds to the universe, while column (2) corresponds to the letter sample.[33] By construction, properties in the letter sample are more expensive and consequently pay more in property taxes, although the share of property taxes that correspond to school taxes is similar for the letter sample (50.60%) and the universe (49.77%). Regarding protest history, the homeowners selected to receive the letter seem slightly more likely to file a protest directly (e.g., in 2020, the direct protest rate was 8.83% for this sample vs. 7.96% in the universe sample).

3.2 Subject Recruitment

We sent a letter to the 78,128 households in the letter sample, inviting them to participate in an online survey. The letter included a URL to access the survey. We mailed our letters so they would be delivered close to the time that homeowners in Dallas County could start filing tax appeals. Appendix D shows a sample envelope, and Appendix E shows a sample letter. We include several features to indicate the legitimacy of the letters. For example, the letters were sent on behalf of researchers at The University of Texas at Dallas, a well-known institution in Dallas County. The envelope featured the university logo, the name of a professor from that university, and non-profit organization postage. The letter included a physical address for the researcher and a link to the study's website (see Appendix F for a screenshot of the website). It also provided contact information for the researchers and the Institutional Review Board. The letter salutation included each recipient's name, and recipients' names and addresses were printed at the bottom of the second page so that

[32] More precisely, for the 11 remaining districts, we sorted the data by the percentage increase in the estimated property tax bill (relative to 2020) and a randomly generated number. We then selected the first 5,200 properties within each school district to be invited to the survey.

[33] Appendix A.3 contains a more detailed description of each subgroup and a more thorough discussion of the property characteristics.

they appeared through the envelope window. In cases where properties were jointly owned by multiple individuals (typically, husband and wife), we sent one letter to the address but listed all owners on the letter. The letter also mentioned the proposed value of the recipient's home and the estimated amount of property tax for 2021.

Importantly, we can link the survey responses to the administrative records. In addition to the opportunity to contribute to a research study, we included two additional incentives for survey participation. First, the letters indicated that detailed, step-by-step instructions on how to file a protest online or by mail would be provided at the end of the survey.[34] As a second incentive, some subjects were informed that they would enter a raffle for 20 prizes worth $100 each.[35]

3.3 Survey Design

In this section, we summarize the main features of the survey.[36] We start by asking a critical question: whether the respondent's household has children enrolled in grades K–12 at their local public school district and, if so, how many. This important information is not included in the tax agency's administrative records. The survey can be summarized as follows:

- **Step 1 (Elicit Prior Belief):** We begin by providing the estimated total property tax amount for the respondent's home in 2020 (based on administrative records). We then explain that this total amount is the sum of different components, such as school, city, and hospital taxes. We asked respondents to guess their school share in 2020, using any amount between 0% and 100%.

- **Step 2 (Information-Provision Experiment):** For every subject, we calculate the "correct" answer to the previous question based on administrative records. We then randomize whether the subject sees the correct answer. Each subject faces a 50% probability of being shown this information. To avoid respondents making inferences from the act of receiving information, we make the randomization explicit. On the first screen, we inform respondents that some participants will be randomly chosen to receive the information and

[34] This walk-through included hyperlinks to relevant websites and screenshots of a sample protest using information from a fictitious household for added clarity. To access these instructions, subjects received a URL and a code on the final screen of the survey. Appendix G includes a copy of the Web instructions. Nathan et al. (2020) show that these instructions have a significant positive effect on the protest probability.

[35] All respondents were entered into the same raffle, but only a random half of respondents were informed about the raffle in the letter (i.e., before deciding whether to participate in the survey). This randomization aimed to assess the effectiveness of raffle prizes in increasing response rates, which can be useful information for future researchers conducting similar field experiments. The results are presented in Appendix A.4. Overall, the raffle message slightly increased the participation rate by 0.2 pp, an effect that is statistically significant (p-value = 0.047) but economically small (5.4% of the baseline rate).

[36] A sample of the full survey instrument is attached as Appendix I.

that they will find out on the next screen if they are selected. On the next screen, we inform respondents whether they are chosen to receive feedback.

- **Step 3 (Elicit Posterior Belief):** We re-elicit the guess they provided in Step 1, which we do for all subjects, regardless of whether they received information. To avoid asking the exact same question twice, we asked about their 2021 taxes (i.e., the most recent year) instead of their 2020 taxes (i.e., the year prior to our intervention). To avoid subjects making inferences based on the opportunity to re-elicit their guesses (e.g., subjects inferring that we ask again only if their answer in Step 1 is incorrect), we explicitly inform them that all survey participants have this opportunity, regardless of their initial guesses.

To learn about the causal effects of beliefs, it is critical to leverage information on prior beliefs. When provided feedback during the information-provision experiment, individuals who underestimate may update their beliefs upward, and those who overestimate may adjust their beliefs downward. Some individuals may have accurate priors and thus may not make any updates. Whether an individual's probability of protesting increases, decreases, or remains the same should depend on the individual's beliefs before receiving the information. For this reason, we conducted the information-provision experiment within the survey instead of providing the information in the letter to measure beliefs prior to information provision.

The survey also included a module with a second treatment arm where we provided information about the "recapture share:" i.e., the share of school taxes from one's school district that are redistributed towards poor or from rich school districts. Because we found out that we were underpowered to detect effects, we discuss the design and results related to this second treatment arm in Appendix B. We cross-randomized subjects to receive up to two pieces of information, with a 50% probability for each. Thus, roughly 25% of the sample receives both pieces of information, 25% receives the school share information only, 25% receives the recapture share information only, and 25% receives no information at all.

These questions form the core of the survey. We also include a series of additional questions, including the intention to appeal, which serves as a secondary outcome. For descriptive purposes, we include questions about the respondent's gender, age, ethnicity, and education. To provide complementary evidence, we include some additional questions that are described in more detail below.

3.4 Implementation

We timed the intervention so that our letters would arrive early enough before the deadline to appeal as required to influence the recipient's decision. We created the letters on April 16th, 2021, as soon as the administrative data, including 2021 proposed values, became available.

To accelerate delivery, we used a mailing company in Dallas County (i.e., the same county as all recipients). The mailing company dropped off the letters at the local post office on April 20, 2021, and estimated that most would be delivered in the next couple of days. Consistent with this projection, we received survey responses and visits to the study's website starting April 22nd, 2021.[37] Survey responses were linked to each homeowner's information from the administrative records, including whether the subject protested directly or with the help of an agent in any year from 2016 to 2020, property characteristics, home value, tax amount, and school share.

Of the 78,128 households invited to the survey, 2,966 started the survey (i.e., completed at least the first couple of questions), and 2,821 completed the two key modules (i.e., up to the posterior belief on recapture). The implied response rate of 3.6% ($=\frac{2{,}821}{78{,}128}$) is comparable to the response rate of 3.7% from a previous study in the same context and using a similar recruitment method (Nathan et al., 2023). Moreover, the response rate of 3.6% is on the same order of magnitude as the response rate of surveys that use this recruitment method (4.7%, as reported in Sinclair et al., 2012).[38] Among respondents, the median time to complete the survey was 11.2 minutes. At the end of the survey, we included an attention check similar to that used in other studies (Bottan and Perez-Truglia, 2020), which 92.7% of respondents successfully passed. This passing rate is relatively high for a survey study, especially given that the attention check was located at the very end of the survey when fatigue was likely at its highest.

Of the 2,821 survey responses, we dropped responses that, as explained in the RCT pre-registration, could not be excluded *ex-ante* due to data availability. We dropped 36 responses from subjects who, according to the DCAD's records, had already filed a protest before starting our survey and 23 additional subjects who responded to the survey after the deadline to file a protest, as the survey information could not have affected their decisions to protest. We similarly dropped 185 subjects who, according to the DCAD's records, had already hired a tax agent before starting our survey (for more details, see Appendix A.2).

When studying perceptions through survey data, it is important to deal with outlier beliefs properly. Some individuals may provide wildly inaccurate guesses, not because they genuinely hold such extreme beliefs, but because they misunderstand the question, make a typo, or do not pay enough attention. The "information shocks" for these individuals can be large but meaningless, which can induce substantial attenuation bias to the causal estimates.

[37] Appendix A.3 contains more descriptive information about the sample of homeowners who answered the survey, and Appendix A.5 contains more details about the timing of survey responses and discusses in detail attrition rates and balance tests.

[38] The 4.7% response rate corresponds to a mailing of a personally-addressed postcard inviting a household to complete a web-based survey using a unique alphanumeric code.

To reduce sensitivity to outliers, we follow standard practice in information-provision experiments and drop respondents with the most extreme misperceptions in their prior beliefs (see e.g., Fuster et al., 2022; Cullen and Perez-Truglia, 2022; Bottan and Perez-Truglia, 2020). For the baseline specification, we use a conservative definition of outliers that drops 467 subjects from the bottom 5% and top 5% of the distribution of prior misperceptions.[39] After applying these filters, 2,110 respondents remain, constituting our main subject pool (an implied 2.7% net response rate). Since these exclusions are based on pre-treatment variables (e.g., prior beliefs), they should not compromise the validity of the experimental variation. As a robustness check, we reproduce the analysis with less strict definitions of outliers. Finally, we provide several sharp falsification tests to address any potential concerns about the internal validity of the results, such as event-study analyses.

Panel (a) of Table 1 shows the average pre-treatment characteristics according to the administrative records (e.g., home value, number of bedrooms). A comparison between columns (1) and (2) shows that the households invited to the survey are largely similar to the universe of households: for most characteristics, the differences are statistically significant (due to the large sample sizes) but typically small in magnitude.[40] The comparison between columns (2) and (3) indicates that the households who responded to the survey are largely similar to the sample of households who were invited to participate in the survey. There is a key difference, however: relative to survey non-respondents, survey respondents are more likely to have filed a protest in the recent past and also more likely to protest in 2021.[41] This is mainly by design, as we crafted the letter to attract the attention of households interested in tax protests. As a result, subjects who were at least considering filing a protest in 2021 are more likely to pay attention to the letter and, thus, more likely to notice the survey link included in the letter. Moreover, our letter promises instructions on how to file a protest as a reward for participation, so it is natural that households who are considering filing a protest would be more likely to participate in the survey.[42] Indeed, this higher propensity to protest among survey respondents is consistent with the results from Nathan et al. (2020), who use a similar recruiting method to collect survey responses in this same context.

Prior to any adjustment resulting from protests, column (3) of Table 1 shows that the average subject owns a home with an assessed market value of $349,988 and property taxes of $7,738 (an average tax rate of 2.21%). Panel (b) of Table 1 reports descriptive statistics based on information collected in the survey. The average respondent is 49.6 years old, 42.9%

[39] For more details on the distribution of outlier observations, see Appendix A.6.

[40] The households invited to the survey are not exactly representative of the universe of households because, as explained above, we applied some filters and intentionally oversampled certain types of households.

[41] Appendix A.3 presents more details of the differences between survey respondents and non-respondents.

[42] Our instructions likely make it easier for survey respondents to file an appeal (Nathan et al., 2020).

are women, 44.3% are White, and 38.3% have a college degree. Moreover, the proportion of households *with* and *without* children who answered our survey, 25.5% and 74.5%, respectively, approximately matches the proportion of families who have or do not have children in Dallas County: 32.3% and 67.4%, respectively (Statistical Atlas, 2023).

Columns (4) through (7) of Table 1 break down the average characteristics in each of the four treatment groups. All characteristics shown in Table 1 are determined pre-treatment and thus should not be affected by the treatment assignment.[43] Column (8) reports p-values for the null hypothesis that the average characteristics are equal across the four treatment groups. Table 1 shows that, consistent with successful random assignment, the observable characteristics are balanced across treatment groups.[44] Appendix A.5 presents alternative versions of the randomization balance tests, such as breaking the sample down by households *with* and *without* children. We also show that response rates to the survey and attrition among participants are orthogonal to treatment assignment, which is expected given that subjects can receive information treatments only after starting the survey.

3.5 Outcomes of Interest

As stated in the RCT pre-registration, the main outcome of interest is a dummy variable indicating whether the household protested directly in 2021.[45] To get a sense of the baseline protest rate, we consider subjects in the control group (i.e., those not selected to receive any feedback). Approximately 30.1% of these owners file a tax appeal in 2021. These tax protests are consequential in that they often reduce the home's assessed value. Indeed, we can use the administrative records to measure whether the protests were consequential. For example, in our sample, 65.4% of the protests lead to a decrease in the home's assessed value – these lower assessments can translate into tax savings in the current or future years.

For brevity, in the rest of the paper, we use the term "protest" as a shorthand for direct protests by the homeowner, unless explicitly stated otherwise. Households also have the option to hire an agent to file a protest on their behalf. In addition to 30.1% of owners who protest directly, 4.8% protest through an agent. Although households can hire an agent to protest on their behalf, we designed our experiment focusing on direct protests. Indeed, when

[43] Some questions, such as the respondent's gender, are asked after the information-provision stage. However, treatment assignments should not affect these responses. For example, we do not expect information on school spending to change responses regarding gender or educational level.

[44] The difference is statistically significant for one of the variables (owner protest in 2020). Given the large number of tests conducted, some differences may be statistically significant just by chance. To be safe and follow best practices in field experiments (Athey and Imbens, 2017), we include this variable in the set of control variables in all regressions.

[45] Information on whether property owners protested their property taxes was downloaded from the DCAD website for the last time on April 25, 2024.

forming the subject pool to invite to the survey, we filtered out households whose owners had protested through an agent in previous years. For completeness, we report the effects on protests through agents, but we expect our intervention to have no effects on this margin.[46]

We included one question after the information-provision stage to serve as a secondary outcome: We asked respondents if they plan to file a protest this year on a 4-point likelihood scale. We construct an indicator variable that equals 100 if, at the end of the survey, the subject responds "very likely" to the question on the likelihood to protest in 2021 and 0 otherwise. This outcome allows us to detect short-term effects on the *intention* to protest, even if those effects do not materialize into actual protests. For reference, in the control group, 45.4% report that they are very likely to protest. Most importantly, the stated intention to protest is significantly correlated with whether the individual actually files a protest. However, that correlation is far from perfect: the correlation coefficient is 0.410 in the control group (p-value<0.001).[47]

4 Results

4.1 Average Treatment Effects

As an initial examination of the data, we report the average treatment effects of the information separately for households *with* and *without* children. This approach should be taken cautiously, as it does not account for heterogeneous belief updating, making it a less efficient use of the data. However, its simplicity offers a natural starting point.

The results are presented in Figure 1, which reports both the raw mean difference between treatment and control groups, as well as conditional on a set of baseline control variables.[48] Panel (a) shows the average effects on the posterior belief about school share. This panel

[46] We did not provide the information to the agents. And while the information provided to the homeowners could potentially affect their decisions to hire or fire an agent, we think this is unlikely because of the characteristics of the contracts between homeowners and agents in this setting (e.g., agents protests mechanically every year, homeowners need to submit a form to the county to remove an agent which creates stickiness). Lastly, consistent with these institutional considerations, Nathan et al. (2020) show that their mail intervention had large effects on direct protests but negligible effects on protests through agents.

[47] Among respondents who report being very likely to protest, 56.8% end up protesting directly or through an agent. Among respondents who do not report being very likely to protest, 16.8% end up protesting.

[48] All the control variables are pre-treatment characteristics: the logarithm of total market value in 2021, the growth in total market value between 2020 and 2021, an indicator for positive growth, an indicator of whether the property value was re-evaluated in 2021, the 2021 estimated property taxes (in logs), a dummy for homestead exemption in 2021, an indicator for a binding homestead cap in 2021, the household's effective tax rate, a dummy variable for multiple owners, a dummy variable for condos, the total living area, the number of bedrooms, the number of full baths, the building age, a set of dummies for school districts, the survey start date, and indicator variables for whether the household protested in each pre-treatment period since 2016 (one set for direct protests and another set for protests through agents).

shows that providing information has a positive and significant effect on the posterior belief about school share by 10.9 pp (p-value<0.001) for households *with* children and by 10.6 pp for households *without* children (p-value<0.001). Indeed, the difference between households *with* and *without* children is statistically insignificant (p-value=0.823). This evidence suggests that consistent with the results that we report in the next section, households typically underestimate the share of school taxes.

In turn, Figure 1(b) shows the average effects on the probability of filing a tax appeal. More precisely, the dependent variable is an indicator that equals 100 if the subject protested directly in 2021 and 0 otherwise. This figure shows that the information on school share beliefs *decreases* the probability of filing a tax protest for households *with* children (by about -4.77 pp), while it *increases* for households *without* children (by about 4.79 pp). For households *with* children, this average effect is statistically insignificant when estimated without including control variables (p-value=0.230) but becomes statistically significant when including the set of baseline controls (p-value=0.074). For households *without* children, this effect is statistically significant, excluding and including controls (p-values of 0.039 and 0.011, respectively). Moreover, the difference in the effects between households *with* and *without* children is statistically significant in both specifications (p-values of 0.038 without controls and 0.005 with controls). These results provide preliminary evidence consistent with the reciprocal motivation mechanism.

In the following section, we explore the experiment's results in greater detail. First, we document the initial misperceptions and further examine how individuals updated their beliefs in response to the information. Second, we employ a 2SLS model to assess the causal impact of these beliefs on the decision to file a tax appeal.

4.2 Accuracy of Prior Beliefs

Transparency and accountability efforts have made information about property taxes publicly available. Each year, the Dallas Central Appraisal District (DCAD) provides homeowners in Dallas County with a Notice of Appraised Value, which contains a detailed breakdown of the household's property taxes by tax jurisdiction, including the share of their property taxes that funds public schools.[49] But the ease of access to this information does not mean that everyone searches for it or uses it. Many other contexts show that people often misperceive

[49] See Appendix H for a sample of this notice, with the breakdown by tax jurisdiction shown on the second page. The county uses the prior year's jurisdictional tax rates to estimate taxes due in the Notice of Appraised Value because the tax rates for the current year are set later in the year. In practice, tax rate changes are uncommon, so approximation errors are typically negligible. In our study, we use the same definition of estimated taxes because these are the relevant object of study and represent the subjects' best approximation when deciding whether to protest.

easily accessible information, such as the official inflation rate (Cavallo et al., 2017) or recent trends in national home prices (Bottan and Perez-Truglia, 2020).

Figure 2(a) illustrates the distribution of misperceptions about the school share for the 2,110 observations in the subject pool before the experimental treatment.[50] The x-axis corresponds to the difference between the actual school share (i.e., potential feedback) and respondents' perceptions. For brevity, we use the term feedback to refer to potential feedback. A minority of subjects have accurate perceptions: more precisely, only 32.6% of subjects guess the school share to be within ± 5 pp of the actual school share. Misperceptions are quite large on average: the mean absolute error is 16.57 pp. The large degree of misperceptions implies sufficient scope for the information provision experiment to shock beliefs. Another interesting feature of prior beliefs is that the misperceptions show a systematic bias: on average, subjects underestimate the school share by 13.08 pp, as indicated by the mean error. This systematic bias is quite noticeable in Figure 2(a), where more observations fall in the right half of the histogram (corresponding to an underestimation) than in the left half (corresponding to an overestimation). It is important to note that households *with* children do not have more accurate perceptions about the school share than households *without* children. We discuss this in detail in Appendix A.6.

4.3 Belief Updating

We find that taxpayers update their inaccurate beliefs when provided with accurate feedback. To model belief updating, we use a simple Bayesian model that has been shown to accurately represent belief formation in other information-provision experiments on a wide range of topics, such as inflation expectations (Cavallo et al., 2017), salary expectations (Cullen and Perez-Truglia, 2022), and home price expectations (Fuster et al., 2022).

We use the subscript i to index the subjects. We use the variable s_i^{prior} to represent subject i's belief about the school share right before the information-provision stage. We use the variable s_i^{feed} to represent the value of the feedback about the school share that the subject can potentially receive in the experiment. We define the variable T_i^S as a binary variable that equals 1 if subject i is selected to receive that information about the school share and 0 if not. We define the variable s_i^{post} as the posterior belief about the school share: s_i^{post} represents the perceived school share after the taxpayer sees or does not see the feedback.

An individual shown feedback will form her posterior belief (s_i^{post}) as the average of the prior belief (s_i^{prior}) and the feedback (s_i^{feed}), weighted by a parameter α that captures the degree of learning. This parameter can range from 0 (individuals ignore the feedback) to 1

[50] Appendix A.6 contains additional information for the entire survey sample without excluding any outliers.

(individuals fully adjust to the feedback) and is a function of the relative precision of the prior belief with respect to the precision of the feedback.[51] This Bayesian updating model can be summarized by the following linear relationship:

$$s_i^{post} - s_i^{prior} = \alpha \cdot \left(s_i^{feed} - s_i^{prior} \right) \tag{1}$$

Intuitively, Bayesian learning predicts that, when shown feedback, respondents who overestimate the school share would revise their beliefs downward. In contrast, respondents who underestimate the school share would revise their beliefs upward. Figure 2(b) estimates this Bayesian learning model using a binned scatterplot. The x-axis corresponds to the gaps in prior beliefs ($s_i^{feed} - s_i^{prior}$), and the y-axis corresponds to the belief updating ($s_i^{post} - s_i^{prior}$). The x-axis shows the maximum revision we would expect if the respondent fully reacted to the information, and the y-axis shows the actual revision. In the case of no updating, the observations should form a horizontal line; in the other extreme, under full updating, the observations should form a 45-degree line. The red circles in Figure 2(b) correspond to the subjects shown feedback about the school share. Consistent with significant updating, there is a strong relationship between the updated beliefs and prior gaps: an additional percentage point (pp) in the perception gap is associated with an actual revision of 0.809 pp higher.

The gray squares in Figure 2(b) correspond to subjects who do not receive information about the school share. In the absence of feedback, these subjects should not update their beliefs. However, in practice, individuals may revise their beliefs in the direction of the feedback for spurious reasons even when they do not receive feedback. For example, respondents may reassess their answers or correct typos when asked a question a second time, leading to an answer closer to the truth. The gray squares indicate a weak relationship between belief updating and prior gaps in the group that was not shown the feedback: an additional 1 pp in the prior gap is associated with an actual revision of 0.052 pp higher. This effect is statistically significant (p-value<0.001) but economically very small. This result is consistent with other information-provision experiments that show evidence of spurious revisions (e.g., Fuster et al., 2022; Cullen and Perez-Truglia, 2022).

We can exploit the random assignment from the information-provision experiment to control for spurious learning:

$$s_i^{post} - s_i^{prior} = \tau + \alpha \cdot \left(s_i^{feed} - s_i^{prior} \right) \cdot T_i^S + \beta \cdot \left(s_i^{feed} - s_i^{prior} \right) + \epsilon_i \tag{2}$$

This regression forms the basis for the first stage of the 2SLS model. In this model,

[51] In the typical model in the literature, the results assume a normal distribution of priors and feedback and assume that the variance of the prior and the variance of the feedback are independent of the mean of the prior. For more details, see Hoff (2009).

parameter α represents true learning arising from the information provision, while parameter β captures spurious learning. The parameter α can be calculated from the estimates in Figure 2(b). Specifically, the parameter α corresponds to the difference in the regression slopes between subjects who receive feedback and those who do not. The estimated α is large ($0.757 = 0.809 - 0.052$) and highly statistically significant (p-value<0.001). This difference suggests that a 1 pp information shock causes a change of 0.757 pp in the subject's posterior belief. This shows that, although subjects did not fully update to the feedback, they were close to updating fully. This finding of imperfect updating is consistent with other information-provision experiments, and it is likely due to some subjects mistrusting the source of the feedback or simply not paying enough attention to the survey.[52]

4.4 Econometric Model

To measure the causal effect of posterior beliefs on the probability of filing an appeal, we use the same econometric models used in other information-provision experiments (see e.g., Cullen and Perez-Truglia, 2022; Bottan and Perez-Truglia, 2022).[53] Let P_i^{2021} be an indicator variable equal to 100 for individuals filing a protest in 2021 (i.e., post-treatment) and 0 otherwise. Since the effects of school share are expected to be different for households *with* and *without* children, let $C_i \in \{0, 1\}$ be an indicator variable that equals 1 if the household has a child enrolled in a local public school and 0 otherwise. Consider the following equation:

$$P_i^{2021} = \beta_0 + \beta_C^S \cdot C_i \cdot s_i^{post} + \beta_{NC}^S \cdot (1 - C_i) \cdot s_i^{post} + \beta_1 \cdot C_i + \epsilon_i \tag{3}$$

where ϵ_i is the usual error term. The two parameters of interest are β_C^S and β_{NC}^S. We expect $\beta_C^S < 0$ and $\beta_{NC}^S > 0$, and $\beta_C^S - \beta_{NC}^S < 0$. Posterior beliefs ($s_i^{post}$) could be correlated with a host of omitted variables. Therefore, we estimate equation (3) using a 2SLS model that leverages the exogenous variation in posterior beliefs induced by the information-provision experiment. More precisely, we estimate the following model:

$$\begin{aligned} P_i^{2021} = &\beta_0 + \beta_C^S \cdot C_i \cdot s_i^{post} + \beta_{NC}^S \cdot (1 - C_i) \cdot s_i^{post} + \beta_1 \cdot C_i + \\ &+ \beta_2 \cdot C_i \cdot (s_i^{feed} - s_i^{prior}) + \beta_3 \cdot (1 - C_i) \cdot (s_i^{feed} - s_i^{prior}) + X_i \beta_X + \epsilon_i \end{aligned} \tag{4}$$

The endogenous variables are $C_i \cdot s_i^{post}$ and $(1 - C_i) \cdot s_i^{post}$, for which we use the excluded

[52] Some additional results are presented in the Appendix. Appendix A.6 shows that the belief updating is not different between households *with* and *without* children. Appendices A and B show that learning from feedback is compartmentalized (i.e., subjects do not use the information about the school share to update beliefs about the recapture share, or vice versa).

[53] Since the two information treatments were cross-randomized within the same sample, we estimate all effects simultaneously in a single 2SLS regression. For the sake of brevity, however, the econometric model presented below only includes the terms corresponding to the school share arm.

instruments $C_i \cdot T_i^S \cdot (s_i^{feed} - s_i^{prior})$ and $(1 - C_i) \cdot T_i^S \cdot (s_i^{feed} - s_i^{prior})$.[54]

We can illustrate the intuition behind the model using a simple example. Consider a pair of households *with* children that have the same bias about the school share: both underestimate the actual school share by 20 pp. Suppose we randomly assign information about the true school share to one of them. We expect that relative to the subject who does not get the information, the subject who receives the information adjusts his or her perceived school share upwards. For the sake of argument, assume that the subject who does not receive the information continues to underestimate the actual school share by 20 pp and that the subject who does receive the information reacts to it by underestimating the school share by just 10 pp. Therefore, the information provision is equivalent to a +10 pp shock to the perceived school share. We can then check the behavior of this pair of households in the weeks after they receive the information. For example, the +10 pp shock to the perceived school share could translate to a lower probability of filing a protest. Assume that the +10 pp shock to the belief causes a 2 pp drop in the probability of protesting. Combining these two results, we obtain an estimate $\beta_C^S = -0.2$. That is, each 1 pp increase in the perceived school share reduces the probability of protesting by 0.2 pp.[55]

The term X_i in equation (4) corresponds to the same set of controls described in Section 4.1. In principle, the 2SLS model leverages experimental variation, so additional control variables are not needed for causal identification. However, including additional control variables can be helpful, for example, in reducing the variance of the error term and thus improving the statistical precision (McKenzie, 2012).

4.5 2SLS Results

Table 2 shows the results from the 2SLS model and the corresponding reduced form and first-stage models. Unlike Figure 1, which reports the average effects of being assigned to the information treatment, estimates in Table 2 account for heterogeneous belief updating. In panel (a) of Table 2, the dependent variable in column (1) is our main outcome variable: protesting taxes directly in 2021. According to the conceptual framework, the difference in the effects of school share between households *with* children and *without* children should be negative. Consistent with this prediction, the difference between the coefficients with

[54] Note that equation (4) controls for the prior gaps in beliefs ($C_i \cdot (s_i^{feed} - s_i^{prior})$) and $(1 - C_i) \cdot (s_i^{feed} - s_i^{prior})$). The inclusion of these control variables ensures that the excluded instruments isolate the information shocks that are driven purely by the random assignment of the feedback (T_i^S).

[55] As is typical in 2SLS models, if treatment effects are heterogeneous, the model estimates the local average treatment effects of beliefs (Imbens and Angrist, 1994). In our context, it would mean that the treatment effects would put more weight on the individuals who updated their beliefs the most in response to the information.

and without children is negative (-0.687), large in magnitude, and statistically significant (p-value=0.007). Similarly, we expect that an increase in the perceived school share should decrease the probability of protesting for households *with* children but should have the opposite effect for households *without* children.[56] The results are also consistent with these predictions: the coefficient for households *with* children is negative (-0.409) and statistically significant (p-value=0.062), while the coefficient for households *without* children is positive (0.278) and statistically significant (p-value=0.031).

These coefficients in column (1) of Table 2(a) are not only statistically significant but also economically large. As a thought experiment, consider what would happen if the perceived school share increases by 10 pp – for reference, this is roughly the magnitude of the average update in beliefs due to the information shock. The estimates indicate that this change would cause a decrease of 4.09 pp ($= 0.409 \cdot 10$) in the probability of filing a protest for households *with* children and an increase of 2.78 pp ($= 0.278 \cdot 10$) in the probability of protesting for households *without* children. These effects would be roughly equivalent to 12.1% and 9.6% of the baseline protest rates (33.86 pp and 28.83 pp, reported in the lower rows of Table 3).

To further illustrate the magnitude of these effects, we can convert them into a money metric. We take advantage of the estimates from Nathan et al. (2020). Using a regression kink design, they estimate that an increase of \$100 in the expected tax savings causes an increase of 2.14 pp in the probability of protesting. We can compare the effects of 4.09 pp (for households *with* children) and 2.78 pp (for households *without* children) against that benchmark. For households *with* children, the effect of 4.09 pp would be equivalent to an effect of -\$191 on the expected tax savings ($= \frac{4.09 \cdot 100}{2.14}$). For households *without* children, the 2.78 pp effect on the protest probability would be equivalent to an effect of \$130 on the expected tax savings ($= \frac{2.78 \cdot 100}{2.14}$).

Column (1) of Table 2(b) presents the reduced form specification, i.e., an OLS regression of the outcome of interest on the instrumental variable. Due to the strong first stage (i.e., strong belief updating), the reduced-form estimates are statistically significant and qualitatively similar to the baseline 2SLS estimates. The coefficients are a bit smaller in the reduced form specification than in the 2SLS specification (e.g., -0.296 versus -0.409 for households *with* children, or 0.224 versus 0.278 for households *without* children), which is expected since the first stage results, presented in panel (c) of Table 2, show that homeowners did not fully update their beliefs incorporating the information provided in the treatment. Besides the expected difference in magnitude, the reduced-form coefficients are consistent with the 2SLS results in both direction and statistical significance.

[56] The negative effect for households *with* children is based on the additional assumption that these households prefer school expenditure over non-school expenditure – for more details, see Appendix C.1.

The 2SLS model used in our preferred specification, reported in panel (a) of Table 2, assumes a linear relationship between school share and the probability of protesting. This means that a 1 pp increase in the perceived school share should have the same effect on the probability of protesting regardless of whether we start at a low or a high value of the prior belief. This is a natural starting point because of its simplicity and because it is a common specification in the literature on information-provision experiments. Reduced form estimates are also useful to probe this linearity assumption. Figure 3(b) presents a binned scatterplot representation of the reduced-form effects of the information provision experiment. The x-axis corresponds to the interaction between the information disclosure and the prior gap (i.e., the excluded instrument). The y-axis corresponds to the probability of protesting in 2021. This binned scatterplot includes all the same control variables used in the 2SLS model. Figure 3(b) tries to assess whether the relationship between the interaction term on the horizontal axis and the protest probability on the vertical axis is linear, and the figure shows that a linear fit is a reasonable functional form assumption for this context. In other words, an additional percentage point in the school feedback treatment seems to have the same incremental effect on the probability of protesting, regardless of whether we start from a prior belief that is somewhat below or somewhat above the accurate feedback.[57] Further, this figure shows that outliers do not drive the regression results discussed above.

Column (2) of Table 2 is identical to column (1), except that it uses a different dependent variable: an indicator variable that equals 100 if, at the end of the survey, the subject responds "very likely" to the question on the likelihood of protesting in 2021 and 0 otherwise. This outcome measures the intention to protest and allows us to measure whether the effects of the information lead to an intention to protest immediately after the information is provided. As discussed above, the correlation between the intention to protest and the actual protests is far from perfect, so the effects should not be expected to be "mechanically" the same across these two outcome variables. The results from column (2) of Table 2 for the stated intention to protest are consistent with the results from actual protests in column (1). In column (2), the coefficient for households *with* children in panel (a) is negative (-0.408) and similar in magnitude to the corresponding coefficient from column (1) and statistically significant (p-value=0.080). The coefficient for households *without* children is positive (0.269), on the same order of magnitude as the coefficient from column (1), and statistically significant (p-value=0.062). The difference between the coefficients for households *with* children versus those *without* children (-0.408 and 0.269) is statistically significant (p-value=0.014). The estimated effects in panels (b) and (c) of column (2) are also consistent with the corresponding

[57] To test the linearity of this relationship, we estimated a model including a quadratic term, which turned out to be statistically insignificant.

coefficients in column (1).

In Table 3, we present some additional results and robustness checks. For reference, columns (1) and (2) present the results from the baseline specification, corresponding to columns (1) and (2) from panel (a) of Table 2. A concern when using 2SLS estimation is the potential for weak instruments (Stock et al., 2002). Given the strong belief updating documented in Section 4.3, weak instruments should not be a concern in our setting. Nevertheless, we computed the Cragg-Donald F-statistic, which is commonly used to diagnose weak instruments. The F-statistics reported in columns (1) and (2) of Table 2, of 30.10 and 30.22, are substantially above the rule of thumb of F > 10 proposed by Stock et al. (2002).

As explained in Section 3.5, it is highly unlikely that the information provided in our survey would affect protests through an agent. Nevertheless, in column (3) of Table 3, we report the results of protests through agents for completeness. In this column, we report estimates from the same regression from column (1) but using protests conducted by agents as the dependent variable. As expected, the coefficients from column (3) are close to zero (-0.015 and -0.030) for both households *with* and *without* children, precisely estimated with standard errors smaller than in column (1), and statistically insignificant (p-values of 0.906 and 0.566). The difference between the coefficients for households *with* and *without* children is close to zero (0.044), precisely estimated, and statistically insignificant (p-value=0.743).

Column (4) of Table 3 provides a falsification test. In this column, we exploit the timing of the information intervention in an event-study fashion. Specifically, we estimate the same baseline regression from column (1), except that we use the protest decision in a pre-treatment year (2020) rather than in the post-treatment year (2021) as the dependent variable. Intuitively, since the information was provided in 2021, it could not possibly affect the decision to protest a year earlier (2020). We, therefore, expect the coefficients from this falsification exercise to be close to zero and statistically insignificant. The results reported in column (4) confirm our expectations. The estimated effects are close to zero (0.110 and -0.065, for households *with* and *without* children, respectively), precisely estimated with standard errors smaller than in column (1), and statistically insignificant (p-values of 0.545 and 0.504); the difference between households *with* children and *without* children is also close to zero (0.175) and statistically insignificant (p-value=0.398). Indeed, we can extend this same falsification test to other pre-treatment years for which we have readily available data. For ease of exposition, the results are presented in a graphical form in Figure 3(a). The x-axis denotes the year of the dependent variable (i.e., whether the owner protests directly in the years 2016 through 2021). This figure focuses on the main result, which corresponds to the difference in coefficients between households *with* children versus *without* children. For example, the 2020 coefficient from Figure 3(a), which takes the value 0.175, corresponds to the coefficient from column (4) of Table 3. As expected, for each pre-treatment year (2016–2020), the coefficients are close to zero and statistically insignificant; by contrast, the coefficient is negative and statistically significant in the post-treatment year (2021).

The main focus of our paper is whether the beliefs about government spending affect the household's decision to file a tax appeal insofar as that choice reveals the willingness to pay taxes. However, in some contexts, it may also be helpful to know whether the protests induced or dissuaded by our treatment were consequential, for example, by changing the home's assessed value. To explore that additional hypothesis, columns (5) through (7) focus on downstream outcomes. In column (5) of Table 3, we measure the effects on the probability of a *successful* protest. More precisely, the dependent variable takes the value of 100 if the subject protested directly in 2021 *and* the protest resulted in a reduction of the assessed home value, and 0 otherwise (i.e., either if the owner did not protest or if the owner protest was not successful). The coefficients reported in column (5) are similar to (and statistically indistinguishable from) those in column (1). This similarity suggests that the marginal protests induced by the treatment were largely successful in changing the assessed market value and that the marginal protests dissuaded by the treatment would have been largely successful too. Column (6) illustrates the same result but uses an alternative outcome equal to the percent change in the assessed home value due to the direct protest.[58] The coefficients reported in column (6) are not directly comparable in magnitude to those reported in column (1) because the scale of the outcome is substantially different: e.g., the mean outcome of the percent change in the assessed market value for the reference group is 1.13 for household *with* children whereas for the direct protest variable is 33.86. However, in terms of the direction of the effects and their statistical significance, the conclusions are the same. Column (7) uses a third measure of success, equal to the estimated percent change in the tax liability due to the direct protest. The effects on tax savings in the first year go in the expected direction but are more muted and statistically insignificant. The tax savings during the first year were small because the homestead cap was binding for a significant fraction of the households in our sample.[59] However, the reduction in assessed home values should eventually be reflected in tax savings during subsequent years since it caps future increases in appraised values.

Finally, columns (8) and (9) of Table 3 present estimates on outcomes that incorporate data from the two years after the intervention: 2022 and 2023. Estimating effects on an extended post-treatment period aims to test whether the treatment induces new protests or simply changes their timing and whether treatment effects carry over a more extended period. In column (8) of Table 3, the dependent variable takes the value 100 (200, 300) if

[58] Consistent with the definition used in column (5), this variable takes the value of 0 either when the subject did not protest or in cases where the protest was unsuccessful.

[59] See Nathan et al. (2020) for a detailed explanation of how the homestead cap works.

the subject protested directly once (twice, three times) in the years 2021, 2022, and 2023, and 0 if the subject did not protest at all. In column (9), we use an indicator variable that takes the value 100 if the individual protested at least once during that period and 0 otherwise. If our treatment in 2021 only changed the timing of protests, but not the decision to protest or not, we should not observe any effects on these two longer-term outcomes. Estimates reported in columns (8) and (9) show statistically significant effects that align with our baseline specification and thus rule out that effects are explained only by changes in the timing of protests. Furthermore, when considering the total number of protests for households *with* children (column (8)), the effect is twice as large (-0.856 vs. -0.409), and the same is observed for the difference between the two types of households. This suggests that not only did the information change the protests in 2021, but the effects may have spilled over to subsequent years too.[60]

4.6 Robustness Checks and Additional Results

Table 4 presents additional robustness checks. In panel (a), we report estimates using the direct protest in 2021 as the outcome variable, whereas, in panel (b), we use the intention to protest variable. Estimates reported in column (1) replicate our baseline specification for reference. Column (2) is identical to the baseline specification but includes additional control variables measured in the survey.[61] On the other hand, column (3) presents results from a specification that does not include any control variables. In both cases, the results are similar in direction, size, and statistical significance compared to the preferred specification.

Columns (4) and (5) of Table 4 analyze the robustness of our results to less stringent definitions of outliers for prior beliefs. While in the baseline specification we exclude the top and bottom 5%, in column (4) we exclude the top and bottom 2.5% and in column (5) we exclude the top and bottom 1%. The results are similar to those of the baseline specification in column (1), although slightly smaller in magnitude. These results are consistent with the arguments in Section 3.4 that we should be cautious when including extreme misperceptions because they probably reflect a lack of attention or mistakes rather than legitimate misperceptions. To explore this further, column (6) excludes respondents who do not pass the attention check included at the end of the survey. Consistent with the attention argument, the coefficients are slightly larger when we focus on subjects who pass the attention check.

[60] Additional results using the extended post-treatment period data are reported in Appendix A.7.

[61] The additional control variables included are: respondent's age, a dummy for individuals that self-identify as White, a dummy for gender, a dummy for college degree, and a dummy for political party (which equals 1 for individuals who self-identify as Democrat). These variables are measured at the end of the survey, but some respondents did not finish the entire survey. Including these additional controls reduces the number of observations, which is the main reason why we exclude these variables from the set of baseline controls.

To partially address potential sample selection, in columns (7) and (8) of Table 4, we use inverse probability weights to re-weight the observations in the subject pool to match the letter sample and universe sample, respectively. The results are similar to the baseline results. For example, in the baseline estimates, the treatment effects for the difference between households *with* children and *without* children are -0.687 and -0.678 for the protest probability and the intention to protest, respectively. The estimates when re-weighting to match the letter sample (universe sample) are -0.614 and -0.799 (-0.610 and -0.738), respectively.[62]

Some additional results and robustness checks are presented in the Appendix. In the baseline specification, the category households *without* children combines a strong majority of households that do not have children with a smaller fraction that has children but sends them to private school, charter school or home-school. The rationale for this specification is that none of these households benefit from school expenditures financed by their property taxes. However, in theory, there could be meaningful differences between these sub-groups. For instance, households who send their children to charter schools may react differently because even though they do not benefit from the local school taxes, they do benefit from public education financed by state-level and federal-level taxes. For that reason, in Appendix A.7, we report results for a more detailed breakdown of households (i.e., children in public schools vs. no children vs. children not in public schools), and we show that the main findings are consistent. Appendix A.7 also reports estimates from alternative specifications, excluding all variables related to the recapture treatment, event studies for downstream outcomes, and additional year-by-year estimates for the extended post-treatment period (2021-2023). Appendix A.8 uses as dependent variables other questions included after the information provision stage.

Lastly, Appendix A.9 shows results from a forecast prediction survey that we conducted to assess whether the experimental results are surprising.[63] Specifically, we elicited predictions from 56 experts with publications on related topics about the effect of a 10 pp shock to the belief about the school share, for households *with* and *without* children. The results from this exercise suggest that our experimental findings are not obvious to the sample of experts, and only a minority of the experts predict effects close to the experimental estimates.

4.7 Non-Experimental Evidence

We present non-experimental evidence that complements the experimental evidence presented above. Our survey asked respondents to choose between hypothetical policies in the spirit of

[62] The separate effects for households *with* children and *without* children have larger standard errors but are consistent in sign and magnitude with the baseline results.

[63] The survey instrument used for the forecasting prediction survey is replicated in Appendix J.

Weinzierl (2014) and Saez and Stantcheva (2016). More precisely, we present the respondent with a hypothetical situation in which two households (A and B) own homes worth \$200,000 each. Both households are identical, except that household A has two children enrolled in the public school district, and household B has no children enrolled in the public school district. The respondent must levy a total tax of \$8,000, which can be spread between the two households in any way (e.g., assign all the burden to household A, all the burden to household B, or anything in the middle). According to the hypothesis of reciprocal motivation, the respondents should want the household *with* children to pay more taxes than the household *without* children because the former benefits more from this government service. Consistent with this prediction, most (58.8%) respondents assign a higher tax burden to the household *with* children even though both homes are worth the same.[64] This evidence suggests that the logic of reciprocal motivation resonates with most taxpayers.

5 Conclusions

We conducted a natural field experiment to examine the reciprocal motivation hypothesis: Are households more willing to pay taxes if they believe they benefit from government spending? We leverage the fact that households have significant misperceptions about how their tax dollars are allocated. We designed a field experiment to induce exogenous changes in households' perceptions by providing households with information on the destination of their tax dollars. The results reveal that perceptions about where tax dollars go significantly influence the likelihood of filing a tax appeal in the direction predicted by the reciprocal motivation channel. Specifically, when households with children enrolled in public schools learned that a larger portion of their taxes supports local schools, they were less inclined to file tax appeals, indicating an increased willingness to pay due to greater perceived personal benefit. In contrast, households *without* children were more likely to protest their taxes after learning the same information, as they perceived less personal benefit. These effects are both statistically significant and economically meaningful, highlighting the impact of perceptions of government spending on tax compliance.

Our evidence contributes to understanding why some individuals are more willing to pay taxes than others and why tax compliance varies across countries. It has been documented that nations with the most effective government services, such as the Scandinavian countries, also exhibit the highest levels of tax morale (Kleven, 2014). The concept of reciprocal motivation offers a natural explanation for this correlation: better government services cause taxpayers to be more willing to pay their taxes. Moreover, this reciprocal mechanism can lead

[64] For more details, see Appendix A.8.

to self-reinforcing cycles. For example, if taxpayers develop a negative bias about how much they benefit from their tax dollars, they may become less inclined to comply with their tax obligations. This reduced compliance can lead to lower government revenue and diminished services, turning the initial bias into a self-fulfilling prophecy. Investigating these implications, both theoretically and empirically, remains an important avenue for future research.

Our results stress the challenges of public communication policies. First, we document large misperceptions about government spending, even when such information is publicly available. For governments interested in educating their citizens on how tax dollars are spent, they should do more than post information on a website. Additionally, governments may want to simplify the connection between the taxes they collect and the government services they support. In fact, local governments in the United States are already doing this by breaking down property taxes into specific components, such as the school tax and the hospital tax. Even in the simple context of property taxes, however, we still find that taxpayers have large misperceptions about how their tax dollars are spent. In the case of state and federal governments, tax dollars follow a complicated path from taxpayers' pockets to the provision of public services. As a result, there is probably much more room for improvement in how the state and federal governments communicate with their taxpayers.

Our experimental intervention was designed to disentangle causal mechanisms, not to increase average tax compliance. Nevertheless, our findings provide hints for policymakers looking to improve tax compliance. Our results underscore the challenges and limitations of transparency policies and information campaigns. For example, a message highlighting a government service (e.g., public schools) can boost tax compliance among individuals who benefit most from that service (e.g., households *with* children). However, it can reduce compliance from taxpayers who do not benefit from that service (e.g., households *without* children). As a result, these effects may cancel each other out, resulting in a null average effect on tax compliance. In some cases, this approach may even backfire. Our findings suggest that governments may be able to use reciprocal motives to boost average tax compliance, but only if they are willing to target information (e.g., informing households *with* children about public school spending). Also, governments could try to persuade taxpayers that their tax dollars are spent efficiently or that their tax payments are not captured by corrupt politicians or wasted by bureaucrats. To the extent that these messages raise the average taxpayers' perceptions that their tax dollars are well-spent, they also may lead to higher overall tax compliance.

References

Ajzenman, N. and R. Durante (2022, 07). Salience and Accountability: School Infrastructure And Last-Minute Electoral Punishment. *The Economic Journal 133*(649), 460–476.

Antinyan, A. and Z. Asatryan (2019). Nudging for tax compliance: A meta-analysis. *ZEW-Centre for European Economic Research Discussion Paper* (19-055).

Athey, S. and G. W. Imbens (2017). The econometrics of randomized experiments. In *Handbook of economic field experiments, Vol. 1*, pp. 73–140.

Avenancio-Leon, C. F. and T. Howard (2022, 02). The Assessment Gap: Racial Inequalities in Property Taxation. *The Quarterly Journal of Economics 137*(3), 1383–1434.

Bagchi, S. and L. Dušek (2021). The effects of introducing withholding and third-party reporting on tax collections: Evidence from the U.S. state personal income tax. *Journal of Public Economics 204*, 104537.

Bergolo, M. L., R. Ceni, G. Cruces, M. Giaccobasso, and R. Perez-Truglia (2023, February). Tax Audits as Scarecrows: Evidence from a Large-Scale Field Experiment. *American Economic Journal: Economic Policy 15*(1), 110–53.

Blumenthal, M., C. Christian, and J. Slemrod (2001). Do Normative Appeals Affect Tax Compliance? Evidence From a Controlled Experiment in Minnesota. *National Tax Journal 54*(1), 125–138.

Bott, K. M., A. W. Cappelen, E. Ø. Sørensen, and B. Tungodden (2020). Youâve got mail: A randomized field experiment on tax evasion. *Management science 66*(7), 2801–2819.

Bottan, N. and R. Perez-Truglia (2020). Betting on the House: Subjective Expectations and Market Choices. *American Economic Journal: Applied Economics (forthcoming).*

Bottan, N. L. and R. Perez-Truglia (2022). Choosing Your Pond: Location Choices and Relative Income. *The Review of Economics and Statistics 104*(5), 1010–1027.

Bowers, J., N. Chen, C. Grady, and M. Winters (2020). Can information about taxation and improved public services increase tax compliance? lessons from malawi. *Evidence in Governance and Politics (EGAP).*

Cabral, M. and C. Hoxby (2012). The hated property tax: salience, tax rates, and tax revolts. Technical report, National Bureau of Economic Research.

Carrillo, P. E., E. Castro, and C. Scartascini (2021). Public good provision and property tax compliance: Evidence from a natural experiment. *Journal of Public Economics 198*, 104422.

Castro, L. and C. Scartascini (2015). Tax Compliance and Enforcement in the Pampas Evidence From a Field Experiment. *Journal of Economic Behavior & Organization 116*, 65–82.

Cavallo, A., G. Cruces, and R. Perez-Truglia (2017). Inflation expectations, learning, and supermarket prices: Evidence from survey experiments. *American Economic Journal: Macroeco-*

nomics 9(3), 1–35.

Chen, G. (2021). An Overview of the Funding of Public Schools. *https://www.publicschoolreview.com/blog/an-overview-of-the-funding-of-public-schools*.

Cullen, J., N. Turner, and E. Washington (2021, August). Political alignment, attitudes toward government and tax evasion. *American Economic Journal: Economic Policy 13*(3), 135–66.

Cullen, Z. and R. Perez-Truglia (2022). How Much Does Your Boss Make? The Effects of Salary Comparisons. *Journal of Political Economy 130*(3), 766–822.

De Neve, J.-E., C. Imbert, T. Tsankova, and M. Luts (2021). How to Improve Tax Compliance? Evidence from Population-wide Experiments in Belgium. *Journal of Political Economy 129*(5), 1425–1463.

Del Carpio, L. (2014). Are the neighbors cheating? evidence from a social norm experiment on property taxes in peru. *Unpublished Manuscript, Princeton University*.

DellaVigna, S., N. Otis, and E. Vivalt (2020). Forecasting the Results of Experiments: Piloting an Elicitation Strategy. *AEA Papers and Proceedings 110*, 75–79.

Dobay, N., F. Nicely, A. Sanderson, and P. Sanderson (2019). The best (and worst) of international property tax administration. Technical report, Council On State Taxation.

Douenne, T. and A. Fabre (2022). Yellow vests, pessimistic beliefs, and carbon tax aversion. *American Economic Journal: Economic Policy 14(1)*, 81–110.

Dzansi, J., A. Jensen, D. Lagakos, and H. Telli (2022, April). Technology and tax capacity: Evidence from local governments in ghana. Working Paper 29923, National Bureau of Economic Research.

Fuster, A., R. Perez-Truglia, M. Wiederholt, and B. Zafar (2022). Expectations with Endogenous Information Acquisition: An Experimental Investigation. *Review of Economics and Statistics 104*(5), 1059–1078.

Hallsworth, M., J. A. List, R. D. Metcalfe, and I. Vlaev (2017). The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of public economics 148*, 14–31.

Hoff, P. D. (2009). *A first course in Bayesian statistical methods*. Springer Science & Business Media.

Huet-Vaughn, E. (2019, 2). Stimulating the Vote: ARRA Road Spending and Vote Share. *American Economic Journal: Economic Policy 11*(1), 292–316.

Huet-Vaughn, E., A. Robbett, and M. Spitzer (2019). A taste for taxes: Minimizing distortions using political preferences. *Journal of Public Economics 180*, 104055.

Imbens, G. W. and J. D. Angrist (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica 62*(2), 467–475.

Jones, P. (2019). Loss Aversion and Property Tax Avoidance. *Working Paper*.

Kessler, J. B. and M. I. Norton (2016). Tax aversion in labor supply. *Journal of Economic Behavior Organization 124*, 15–28.

Kleven, H. J. (2014, November). How can scandinavians tax so much? *Journal of Economic Perspectives 28*(4), 77–98.

Krause, B. (2020). Balancing purse and peace: tax collection, public goods and protests. *Berkeley, CA: Agricultural and Resource Economics, University of California, Berkeley*.

Kresch, E. P., M. Walker, M. C. Best, F. Gerard, and J. Naritomi (2023). Sanitation and property tax compliance: Analyzing the social contract in Brazil. *Journal of Development Economics 160*, 102954.

List, J. A. (2020). Non est Disputandum de Generalizability? A Glimpse into The External Validity Trial. *NBER Working Paper No. 27535*.

Luttmer, E. F. P. and M. Singhal (2014). Tax Morale. *Journal of Economic Perspectives 28*(4), 149–168.

Mankiw, N. G. and M. Weinzierl (2010). The optimal taxation of height: A case study of utilitarian income redistribution. *American Economic Journal: Economic Policy 2*(1), 155–176.

McKenzie, D. (2012). Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics 99*(2), 210–221.

Nathan, B., R. Perez-Truglia, and A. Zentner (2020). My Taxes are Too Darn High: Why Do Households Protest their Taxes? *American Economic Journal: Economic Policy (forthcoming).*

Nathan, B., R. Perez-Truglia, and A. Zentner (2023). Paying Your Fair Share: Perceived Fairness and Tax Compliance. *NBER Working Paper No 32588*.

Parker, W. and N. Friedman (2021). Zillow Quits Home-Flipping Business, Cities Inability to Forecast Prices. *The Wall Street Journal, November 2 2021*.

Saez, E. and S. Stantcheva (2016). Generalized social marginal welfare weights for optimal tax theory. *American Economic Review 106*(1), 24–45.

San Francisco Chronicle (2023). S.F. hit with avalanche of requests to lower property taxes. Hereâs what happens now. *https://www.sfchronicle.com/realestate/article/sf-property-tax-18550413.php*.

Sinclair, M., J. O'Toole, M. Malawaraarachchi, and K. Leder (2012). Comparison of response rates and cost-effectiveness for a community-based survey: postal, internet and telephone modes with generic or personalised recruitment approaches. *BMC Medical Research Methodology 12*(1), 132.

Slemrod, J. (2019). Tax Compliance and Enforcement. *Journal of Economic Literature 57*(4), 904–954.

Statistical Atlas (2023). The Demographic Statistical Atlas of the United States: Household Types in Dallas County, Texas. `https://statisticalatlas.com/county/Texas/Dallas-County/Household-Types`. Accessed: 2023-03-27.

Stock, J. H., J. H. Wright, and M. Yogo (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business and Economic Statistics 20*(4), 518–529.

Sussman, A. B. and C. Y. Olivola (2011). Axe the tax: Taxes are disliked more than equivalent costs. *Journal of Marketing Research 48(SPL)*, S91–S101.

Tax Policy Center (2021a). Amount of Revenue by Source. *https://www.taxpolicycenter.org/statistics/property-tax-revenue*.

Tax Policy Center (2021b). Amount of Revenue by Source. *https://www.taxpolicycenter.org/statistics/amount-revenue-source*.

Texas Education Agency (2021a). Excess Local Revenue. *https://tea.texas.gov/finance-and-grants/state-funding/excess-local-revenue*.

Texas Education Agency (2021b). Texas Public School Finance Overview: Biennium 2020-2021.

Texas Education Agency (2021c). What is House Bill 3? *https://tea.texas.gov/about-tea/government-relations-and-legal/government-relations/house-bill-3*.

Texas Public Policy Fundation (2022). Metro area student enrollment 2012-2021. *Accessed from: https://www.texaspolicy.com/where-are-they-enrollment-trends-in-k-12-education/*. Accessed: 2024-08-27.

U.S. Census Bureau (2021). Population, Dallas County, Texas. *https://www.census.gov/quickfacts/fact/table/dallascountytexas/POP010220*.

Villanueva, C. (2018). What is Recapture? *Center for Public Policy Priorities Report, August 30, 2018*.

Weinzierl, M. (2014). The promise of positive optimal taxation: normative diversity and a role for equal sacrifice. *Journal of Public Economics 118*, 128–142.

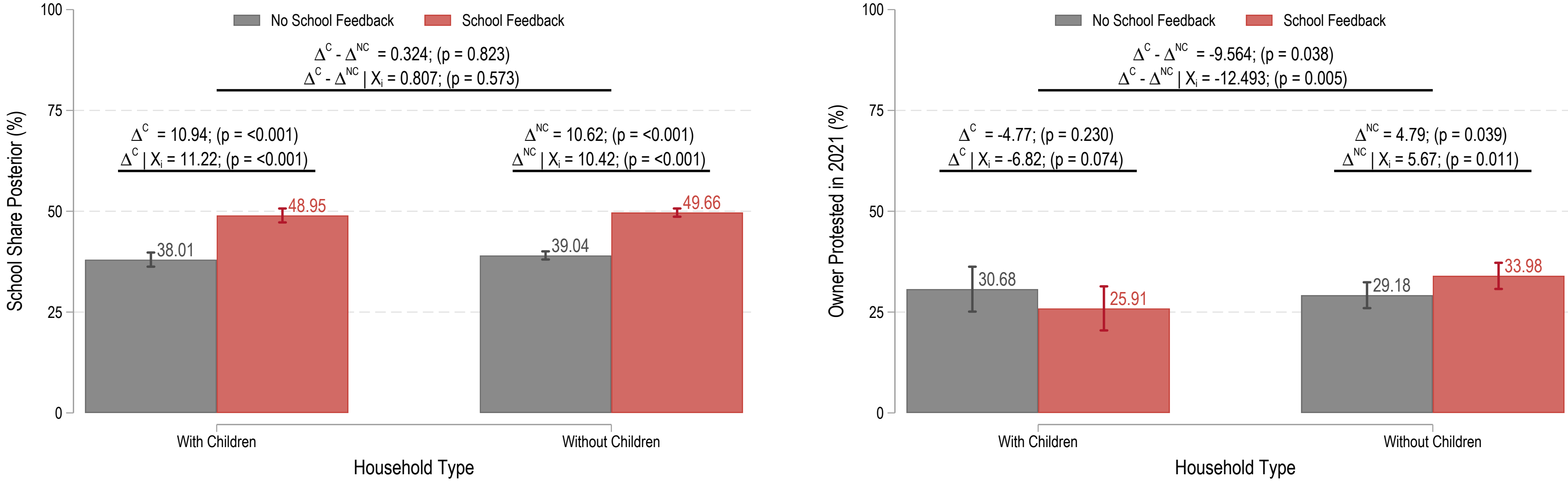
Weinzierl, M. (2017). Popular acceptance of inequality due to innate brute luck and support for classical benefit-based taxation. *Journal of Public Economics 155*, 54–63.

Weinzierl, M. (2018). Revisiting the Classical View of Benefit-based Taxation. *The Economic Journal 128*(612), F37–F64.

World Bank (2019). The Administrative Review Process for Tax Disputes: Tax Objections and Appeals in Latin America and the Caribbean.

Figure 1: Average Treatment Effects of School Feedback

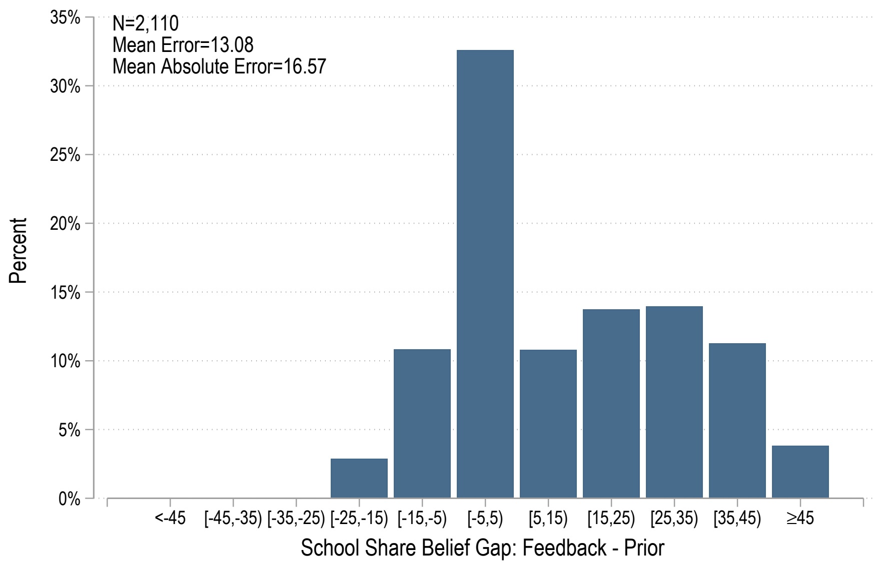
(a) School Share Posterior **(b)** 2021 Owner Protest



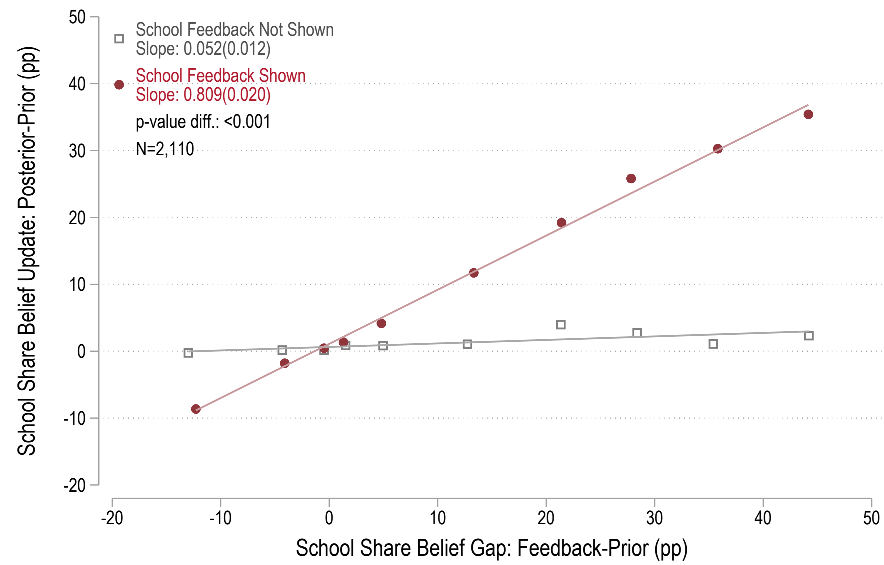
Notes: This figure illustrates the average treatment effects of the school feedback treatment. Results are based on 538 households *with* children and 1,572 households *without* children. Panel (a) reports the mean of the school share posterior by treatment group, separately for households *with* and *without* children. Grey bars represent the average school share posterior for the group that did not receive the school feedback treatment, while red bars represent the average school share posterior for the group that did receive it. In addition, we report: (1) the difference in raw means between the treatment and the control group (Δ^C, and Δ^{NC}), (2) the difference in means conditional on the covariates included in the baseline specification ($\Delta^C|X$ and $\Delta^{NC}|X$), (3) the corresponding p-value of the equality of means test for treated *versus* control groups within each household type, and (4) the p-value corresponding to the comparison of these differences between household types (i.e., a double-difference test). Panel (b) replicates the analysis but uses the probability of protesting as the dependent variable.

Figure 2: Perceptions about the Share of Property Taxes Going to Public Schools

(a) Gap in Prior Beliefs

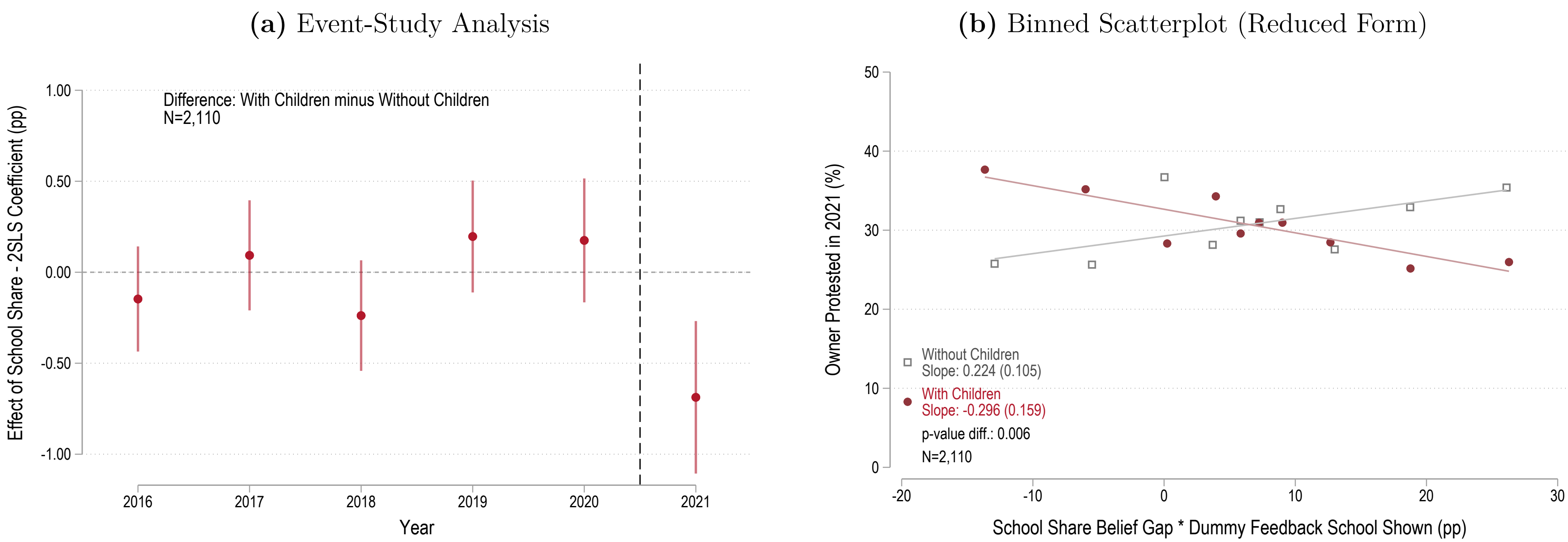


(b) Belief Updating



Notes: Panel (a) shows the gap in prior beliefs about the school share. The x-axis reports the difference between the actual school share and respondents' prior beliefs about the school share in 10 pp-width bins. The y-axis reports the percentage of survey respondents in each bin. The upper left corner reports the total number of observations, the mean error, and the mean absolute error. Panel (b) shows how respondents update their beliefs using a binned scatterplot (using ten bins corresponding to each decile of the School Share Belief Gap). The x-axis reports the difference between the actual school share and respondents' prior beliefs about the school share. The y-axis reports the difference between posterior and prior beliefs (i.e., belief updating). Red circles (gray squares) represent the average update within each bin for the group of homeowners that were selected (were not selected) into the school share treatment. Each line corresponds to the fitted values from separate OLS regressions where the dependent variable is the actual update and the independent variable is the school share belief gap. The coefficients associated with the gap variable are reported in the upper left corner, as well as their robust standard errors (in parentheses), the p-value of the difference in the slopes, and the number of observations included in the analysis.

Figure 3: The Effects of School Share Perceptions on Protests: Additional Robustness Checks



Notes: Panel (a) reports an event-study analysis of the differential effects of school share beliefs on the probability of protesting for households *with* children versus *without* children. The estimates plotted in this figure correspond with the 2SLS point estimate based on equation (4), with 90% confidence intervals based on robust standard errors. The coefficient plotted for 2021 is the coefficient reported in the "difference" row of panel (a), column (1) of Table 3. The remaining coefficients come from similar regressions but using the outcomes in pre-treatment years as falsification tests and restricting the pre-treatment controls to the corresponding years. The vertical dashed line separates the post-treatment year (2021) from the pre-treatment years (2016-2020). Panel (b) depicts a scatterplot representation of the reduced-form effect for households *with* and *without* children separately, using red circles and gray squares respectively and 10 equally-sized bins. The x-axis corresponds to the interaction between the prior school share belief gap (defined as the difference between the actual school share and the prior belief about the school share) and a dummy variable that indicates if the homeowner was selected into the school share treatment group. The y-axis corresponds to the probability of a direct protest in 2021. Each line corresponds to a separate OLS binned scatterplot regression, including the same control variables used in the 2SLS specification. Control variables for the protest history depend on the year in which the outcome is measured. For instance, if the outcome corresponds to the protest in 2019, the protest history controls include protests in 2016, 2017, and 2018; and so on. The coefficients reported in the lower left corner and their (robust) standard errors are based on a unique regression that interacts the key variables with a dummy for having children at school (for the results in table form, see Table 2). In addition we report the p-value of the difference in the effect for the two groups and the number of observations used in the estimation.

Table 1: Balance of Households' Characteristics across Treatment Groups

	Universe (1)	Letter Sample (2)	Subject Pool (3)	Treatment Arm: No Feedback (4)	Treatment Arm: Recapture Feedback (5)	Treatment Arm: School Feedback (6)	Treatment Arm: Both Feedback (7)	p-value test (8)
Panel (a): Admin. Records Variables								
2021 Home Value ($1,000)	327.688	359.145	349.988	365.355	330.631	365.198	340.088	0.163
	(0.651)	(1.632)	(6.774)	(14.907)	(10.302)	(16.461)	(12.037)	
2021 Property Tax Amount ($1,000s)	6.372	7.645	7.738	8.018	7.448	7.960	7.546	0.292
	(0.013)	(0.028)	(0.129)	(0.296)	(0.218)	(0.287)	(0.228)	
School Share (%)	49.777	50.600	50.726	50.603	50.566	50.701	51.029	0.140
	(0.017)	(0.016)	(0.079)	(0.155)	(0.160)	(0.155)	(0.158)	
Recapture Share (%)	-1.957	1.227	1.622	1.852	1.054	2.505	1.130	0.351
	(0.021)	(0.068)	(0.325)	(0.678)	(0.633)	(0.672)	(0.622)	
2020 Owner Protested (%)	7.969	8.832	18.057	23.121	14.815	19.883	14.684	0.000
	(0.043)	(0.102)	(0.838)	(1.852)	(1.530)	(1.764)	(1.527)	
2020 Agent Protested (%)	8.059	6.322	1.659	1.156	2.407	1.754	1.301	0.375
	(0.043)	(0.087)	(0.278)	(0.470)	(0.660)	(0.580)	(0.489)	
2019 Owner Protested (%)	6.055	6.587	13.365	15.029	10.926	14.035	13.569	0.238
	(0.038)	(0.089)	(0.741)	(1.570)	(1.344)	(1.535)	(1.478)	
2018 Owner Protested (%)	5.788	6.443	13.460	13.680	12.407	14.815	13.011	0.697
	(0.037)	(0.088)	(0.743)	(1.510)	(1.420)	(1.570)	(1.452)	
2017 Owner Protested (%)	5.589	5.674	10.853	11.561	11.111	11.891	8.922	0.400
	(0.036)	(0.083)	(0.677)	(1.405)	(1.354)	(1.430)	(1.230)	
2016 Owner Protested (%)	4.412	4.623	7.773	8.478	6.667	8.187	7.807	0.705
	(0.032)	(0.075)	(0.583)	(1.224)	(1.074)	(1.212)	(1.158)	
Multiple Owners (%)	22.173	23.886	24.645	22.929	24.444	25.146	26.022	0.693
	(0.066)	(0.153)	(0.938)	(1.847)	(1.851)	(1.917)	(1.893)	
Living Area (1,000s Sq. Feet)	2.048	2.182	2.313	2.317	2.302	2.331	2.302	0.959
	(0.002)	(0.004)	(0.022)	(0.046)	(0.042)	(0.046)	(0.040)	
Number of Bedrooms	3.119	3.345	3.428	3.432	3.398	3.423	3.459	0.609
	(0.001)	(0.003)	(0.016)	(0.032)	(0.033)	(0.034)	(0.031)	
Number of Baths	2.062	2.171	2.273	2.274	2.272	2.292	2.253	0.883
	(0.001)	(0.003)	(0.017)	(0.034)	(0.033)	(0.039)	(0.032)	
Panel (b): Survey Variables								
With Children (%)			25.498	24.470	25.370	26.316	25.836	0.918
			(0.949)	(1.889)	(1.874)	(1.946)	(1.889)	
Female (%)			42.898	44.922	43.774	40.990	41.887	0.574
			(1.086)	(2.200)	(2.157)	(2.191)	(2.145)	
Age			49.608	49.711	49.381	50.438	48.945	0.146
			(0.234)	(0.470)	(0.481)	(0.461)	(0.460)	
Race: White (%)			44.300	44.727	47.818	44.422	40.265	0.103
			(1.092)	(2.200)	(2.178)	(2.220)	(2.134)	
Education: Grad. Degree (%)			38.309	39.844	37.761	38.446	37.240	0.841
			(1.069)	(2.166)	(2.114)	(2.173)	(2.104)	
Prior Belief: School Share (%)			37.642	37.741	37.186	37.935	37.726	0.918
			(0.394)	(0.804)	(0.760)	(0.790)	(0.800)	
Prior Belief: Recapture Share (%)			1.910	1.799	1.372	2.945	1.570	0.216
			(0.287)	(0.632)	(0.505)	(0.593)	(0.564)	
Observations	400,193	78,128	2,110	519	540	513	538	

Notes: Average pre-treatment characteristics of homeowners' properties disaggregated by sample. Column (1) corresponds to the universe of non-commercial, owner-occupied residences that pay property taxes. Column (2) corresponds to homeowners that were selected to receive a letter with the invitation to answer the survey. Column (3) corresponds to homeowners that answered the survey and belong to the subject pool used in our preferred specifications for the main analysis. Column (4) is based on homeowners who were not selected to receive any information (control group). Column (5) is based on homeowners selected to receive information on the recapture share only. Column (6) is based on homeowners selected to receive information on the school share only. Column (7) is based on homeowners selected to receive information on both the school share and the recapture share. Column (8) reports the p-value of a test of equal means across the four treatment groups. Standard errors are reported in parentheses. The statistics in panel (a) are based on administrative records available on the DCAD's website. The statistics in panel (b) are based on survey responses.

Table 2: Main Results: 2SLS, Reduced-Form, and First-Stage

	(1)	(2)
	P_D^{2021}	I^{2021}
a. Effect of School Share (2SLS)		
With Children	-0.409*	-0.408*
	(0.219)	(0.234)
Without Children	0.278**	0.269*
	(0.129)	(0.144)
(Difference Children - No Children)	-0.687***	-0.678**
	(0.255)	(0.275)
	P_D^{2021}	I^{2021}
b. Effect of School Share Belief Gap * Dummy Feedback School Shown (Reduced Form)		
With Children	-0.296*	-0.288*
	(0.159)	(0.172)
Without Children	0.224**	0.217*
	(0.105)	(0.117)
(Difference Children - No Children)	-0.520***	-0.506**
	(0.191)	(0.208)
	s^{post}	s^{post}
c. Effect of School Share Belief Gap * Dummy Feedback School Shown (First Stage)		
With Children	0.736***	0.741***
	(0.044)	(0.044)
Without Children	0.808***	0.805***
	(0.026)	(0.026)
(Difference Children - No Children)	-0.072	-0.064
	(0.051)	(0.051)
Mean Outcome (Baseline)		
With Children	33.86	47.20
Without Children	28.83	44.87
Observation	2,110	2,090

Notes: Significant at *10%, **5%, ***1%. Robust standard errors in parentheses. This table shows the 2SLS reduced form, and first-stage estimates corresponding to the school share treatment effect. We present the coefficients for households *with* children and households *without children* separately, as well as the difference between these two types of households. In panel (a), the dependent variable in column (1) is an indicator variable that takes the value 100 if the subject protested directly in 2021. The dependent variable in column (2) is an indicator variable that takes the value 100 if the subject answered "very likely" to the question on the likelihood of protesting in 2021. Reported estimates are based on the 2SLS econometric model given by model 4 and discussed in Section 4.4. Panel (b) reports the reduced form effects, i.e., it represents the effect of the information included in the feedback message with respect to the prior beliefs without considering how much subjects actually learn from the experiment. In this case, the dependent variables are the same as in panel (a). Finally, the first-stage estimates are reported in panel (c). Estimates in this panel use the school share posterior as the dependent variable and the reported coefficients can be interpreted as the percentage point update in school share posteriors per percentage point of prior school share misperception. Estimates reported in panels (b) and (c) are based on OLS regressions and use the same set of control variables discussed in Section 4.4. Mean outcomes at baseline correspond to subjects who did not receive feedback about the school share nor the recapture share.

Table 3: 2SLS Estimates: Additional Results

	P_D^{2021} (1)	I^{2021} (2)	P_A^{2021} (3)	P_D^{2020} (4)	$P_{D,won}^{2021}$ (5)	ΔMV^{2021} (6)	ΔT^{2021} (7)	$\#P_D^{2021-2023}$ (8)	$\text{Any}_D^{2021,2023}$ (9)
Effects of School Share:									
With Children	-0.409*	-0.408*	0.015	0.110	-0.450**	-0.047**	-0.011	-0.856**	-0.439*
	(0.219)	(0.234)	(0.123)	(0.181)	(0.190)	(0.021)	(0.012)	(0.430)	(0.242)
Without Children	0.278**	0.269*	-0.030	-0.065	0.136	0.005	0.003	0.294	0.144
	(0.129)	(0.144)	(0.051)	(0.097)	(0.111)	(0.009)	(0.007)	(0.234)	(0.140)
Difference (Children - No Children)	-0.687***	-0.678**	0.044	0.175	-0.586***	-0.052**	-0.014	-1.150**	-0.583**
	(0.255)	(0.275)	(0.134)	(0.207)	(0.220)	(0.022)	(0.014)	(0.493)	(0.280)
Cragg-Donald F-Statistic	30.10	30.22	30.10	30.02	30.10	30.10	30.10	30.10	30.10
Mean Outcome (Baseline):									
With Children	33.86	47.20	7.09	25.98	20.47	1.13	0.84	74.80	47.24
Without Children	28.83	44.87	4.08	22.19	19.39	1.66	1.04	66.33	45.41
Observations	2,110	2,090	2,110	2,110	2,110	2,110	2,110	2,110	2,110

Notes: Significant at *10%, **5%, ***1%. Robust standard errors in parentheses. This table reports 2SLS estimates of equation (4) discussed in Section 4.4. corresponding to the school share treatment effect. We present the coefficients for households *with* children and households *without* children, as well as the difference between these two types of households. The dependent variable in column (1) is an indicator variable that takes the value 100 if the subject protested directly in 2021, and 0 otherwise. The dependent variable in column (2) is an indicator variable that takes the value 100 if the subject answered "very likely" to the question on the subject's protest likelihood in 2021 ("Do you intend to protest this year?"). The dependent variable in column (3) corresponds to an indicator variable that takes the value 100 if the subject used an agent to protest in 2021, whereas in column (4) corresponds to an indicator variable that takes the value 100 if the subject protested directly in 2020. In Column (5), the dependent variable is an indicator that takes the value of 100 if the protest by the owner was successful, and 0 otherwise. Column (6) reports the effects on the market value *savings* from protesting. Negative estimates indicate that post-protest market values are higher. Column (7) uses savings in post-protest estimated tax liabilities as the dependent variable. Again, negative values indicate that post-protest taxes are higher. The dependent variables in columns (8) and (9) consider the protesting behavior in the 2021-2023 period. In column (8), the dependent variable is the total number of protests in 2021-2023 (multiplied by 100 to make it comparable to estimates in column (1). In column (9), the dependent variable is an indicator that takes the value 100 if the subject protested directly in 2021, 2022, or 2023. Mean outcomes at baseline are computed using the group of subjects who did not receive feedback about the school share nor recapture share (the control group).

Table 4: 2SLS Estimates: Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent Variable: P_D^{2021}							
a. Effects of School Share on P_D^{2021}								
With Children	-0.409*	-0.473**	-0.336	-0.364*	-0.252	-0.413*	-0.384	-0.556*
	(0.219)	(0.224)	(0.231)	(0.189)	(0.167)	(0.236)	(0.248)	(0.325)
Without Children	0.278**	0.286**	0.292**	0.196*	0.198*	0.303**	0.230	0.054
	(0.129)	(0.133)	(0.136)	(0.119)	(0.116)	(0.139)	(0.143)	(0.153)
Difference (Children - No Children)	-0.687***	-0.759***	-0.628**	-0.561**	-0.450**	-0.716***	-0.614**	-0.610*
	(0.255)	(0.261)	(0.268)	(0.223)	(0.203)	(0.273)	(0.285)	(0.357)
Observations	2,110	2,070	2,110	2,335	2,482	1,807	2,110	2,091
	Dependent Variable: I_D^{2021}							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
b. Effects of School Share on I_D^{2021}								
With Children	-0.408*	-0.457*	-0.322	-0.250	-0.088	-0.418*	-0.420	-0.461
	(0.234)	(0.235)	(0.246)	(0.205)	(0.191)	(0.247)	(0.262)	(0.360)
Without Children	0.269*	0.286**	0.299**	0.321**	0.256**	0.324**	0.379**	0.277
	(0.144)	(0.146)	(0.147)	(0.132)	(0.130)	(0.153)	(0.163)	(0.176)
Difference (Children - No Children)	-0.678**	-0.744***	-0.622**	-0.571**	-0.344	-0.743**	-0.799***	-0.738*
	(0.275)	(0.278)	(0.286)	(0.244)	(0.231)	(0.290)	(0.308)	(0.397)
Observations	2,090	2,070	2,090	2,309	2,454	1,807	2,090	2,071
Baseline Controls	√	√		√	√	√	√	√
Additional Controls		√						
5% Outliers	√	√	√			√	√	√
2.5% Outliers				√				
1% Outliers					√			
Attention Check						√		
Re-weighted (Exp.)							√	
Re-weighted (Univ.)								√

Notes: Significant at *10%, **5%, ***1%. Robust standard errors in parentheses. This table reports 2SLS estimates of equation (4) discussed in Section 4.6 corresponding to the school share treatment effect. We present the coefficients for households *with* children and households *without* children separately, as well as the difference between these two types of households. In the upper panel of the table, we report coefficients corresponding to regressions where the outcome variable is our main outcome of interest, i.e., direct protests, while column (2) reports estimates on the intention to protest. For reference, column (1) corresponds to our preferred specification reported in columns (1) and (2) in Table 3. Column (2) includes additional control variables collected in the survey: age, gender, college degree, and political party. Column (3) reports estimates where no control variables are included at all. In column (4) we drop 2.5% of the outliers at each tail of the distribution (instead of the 5% used in the baseline specification), while in column (5) we drop 1% of the outliers at each tail. Column (6) restricts the sample to subjects who passed the attention check included in the questionnaire (see Appendix I for the survey). Column (7) reports re-weighted estimates where inverse probability weights are used to match the letter sample. Column (8) does the same but for the universe of non-commercial properties. Estimates are based on inverse probability weighted 2SLS regressions. Weights are obtained from a logit model that regresses a response dummy over the variables included in the administrative records reported in panel (a) of Table 1.