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 August 2023

We are thankful for excellent comments from Raj Chetty, Matthew Weinzierl, Austan Goolsbee, Steve Levitt, James Poterba, Dario Tortarolo, Sutirtha Bagchi and seminar participants at the NBER-Public Economics, University of Michigan, University of Chicago, University of Chicago-Advances in Field Experiments, 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. Xinmei Yang 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 August 2023
JEL No. C93,H26,I22,Z13

ABSTRACT

Do perceptions about how the government spends tax dollars affect the willingness to pay taxes? We designed a field experiment to test this hypothesis in a natural, high-stakes context and via revealed preferences. We measure how taxpayers perceive the government is using their tax dollars, such as the percentage of their property taxes that funds public schools. We find that even though accurate information is available, taxpayers still hold substantial misperceptions. We use an information-provision experiment to induce exogenous shocks to these perceptions. Using administrative data on property tax appeals, we measure the causal effect of perceived government spending on the willingness to pay taxes. We find that perceptions about government spending have a significant effect on the probability of filing a tax appeal and in a manner that is consistent with reciprocal motivation: individuals are more willing to pay taxes if they believe that the government services funded by those taxes will be of greater personal benefit to them. We discuss implications for the study of tax morale.

Matias Giaccobasso
Anderson School of Management - UCLA
110 Westwood Plaza, C 3.10
Los Angeles, CA 90077
mgiaccobasso@ad.ucla.edu

Brad C. Nathan
Naveen Jindal School of Management
The University of Texas at Dallas
800 W Campbell Rd.
Richardson, TX 75080
brad.nathan@utdallas.edu

Ricardo Perez-Truglia
Haas School of Business
University of California, Berkeley
545 Student Services Building #1900
Berkeley, CA 94720-1900
and NBER
ricardotruglia@berkeley.edu

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

Why is tax compliance higher in some countries than in others? Why are some individuals more willing to pay their taxes than others? There are two schools of thought that offer potential explanations: institutions and tax morale. Abundant research shows that institutions have a large effect 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). On the contrary, little causal evidence shows that tax morale actually matters (Luttmer and Singhal, 2014). In this paper, we attempt to advance our understanding of tax morale by means of a natural field experiment in a high-stakes context and via revealed preferences.

Tax morale encompasses various potential mechanisms. We focus on one specific mechanism: our hypothesis is that individuals are more willing to pay taxes if they believe that the government services funded by those taxes will be of greater personal benefit to them. Our hypothesis is related 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." Our hypothesis also relates 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). To test our hypothesis, which for the sake of brevity we from hereon call "reciprocal motivation," we conducted an experiment to measure taxpayers' *perceptions* about where their tax dollars go and determine how these *perceptions* affect their willingness to pay taxes.

Our experiment leverages the context of property taxes, which represents an important source of revenue for governments in the United States and around the world. 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.[1] 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 context offers two key advantages to test our hypothesis of reciprocal motivation. 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. For instance, households with children enrolled in local public schools benefit directly from publicly funded education, whereas households with no children enrolled in local public schools do not. We refer to households with children enrolled in public schools as "households *with* children" and

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

those without as "households *without* children." The second advantage of this setting is that we can study the willingness to pay taxes 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.[2] In a nutshell, households can use the subjective nature of the property appraisal process in their favor. If they feel like their taxes are too high, they can file a tax appeal to reduce their tax burden.[3]

We conducted a field experiment in Dallas County, Texas. We focus on this county primarily because, from a logistical perspective, it is more practical to implement a field experiment in a single location. Dallas County is the second-largest county in Texas, with an estimated population of about 2.6 million in 2020 (U.S. Census Bureau, 2021) – Dallas County alone 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.[4]

We sent a letter to a sample of households inviting them to participate in an online survey. Our main subject pool comprises 2,110 respondents who completed the survey between April and May of 2021, when subjects could file a protest of their property taxes with the county. Our survey elicited whether the household has children enrolled in public schools to identify which subjects benefit directly from public school spending and which do not. We conducted an information-provision experiment a few weeks before households faced the opportunity to file a tax appeal. We then matched survey responses to administrative records from the county assessor's office. The rich administrative data allowed 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 public school spending. 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 measure the respondents' misperceptions about where their tax dollars go by comparing their guesses about the school share to the true estimates from administrative records. To study the causal effect of beliefs about government spending, the survey embeds an information-provision experiment. After eliciting

[2] 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.

[3] For more details about how tax protests work, see the discussion in Section 2.3 and also Nathan et al. (2020) and Jones (2019).

[4] 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).

respondents' prior beliefs, we inform a random half of them about the true value of their respective school shares. By doing so, we can assess how that information affects their posterior beliefs, as measured by our survey, and their decisions to file a tax appeal, as measured by 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 the reciprocal motivation 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 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 government services than they originally thought.

The principle of reciprocal motivation could have implications for tax redistribution. When taxpayers learn that their tax dollars are being spent in communities other than their own, they may be less willing to pay taxes because they do not receive benefits from the taxes they pay. We explore this additional hypothesis using a second treatment arm. Specifically, we take advantage of the significant redistribution of property taxes across school districts that occurs in some states. In Texas, this redistribution is dictated by legislation often referred to by the media as "Recapture Plan" or "Robin Hood Plan."[5] Thus, in the second treatment arm, we measure households' perceptions about the share of their school funding that is redistributed away from or toward their own school district. For the sake of brevity, in the remainder of the paper, we refer to this as the "recapture share." For example, a recapture share of 10% would mean that 10% of the district's school tax revenue is not spent in that district and instead is transferred to disadvantaged school districts.

We can measure the causal effects of the perceived recapture share using the information-provision experiment. Take the example of a district that is a net contributor to the recapture plan. According to the reciprocal motivation mechanism, the belief about the recapture share should not affect the decision to file a tax appeal for households *without* children because the diverted funds are being used for a service that does not benefit them directly anyways. By contrast, households *with* children should be more likely to protest upon learning that some of their tax payments are being diverted to other districts because they were benefiting directly from the diverted funds.

[5] For the full history of property tax recapture in Texas, see for example Villanueva (2018).

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%. We also find significant variation in how the recapture system affects school districts in our sample, with some school districts diverting as much as 57% of their school districts' property taxes and others receiving as much as 23% additional funds from other districts. Owners can protest "directly" on their own, which is 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 on school taxes nor on recapture) protested directly in 2021. These tax protests are consequential. For instance, 65.4% of protests led to a decrease in assessed home value, resulting in average tax savings of $579 in the first year alone.

The results of the first treatment arm indicate that even though the information is publicly available and easily accessible, most households have misperceptions about their respective school shares. When provided with factual information, we observe that households strongly update their beliefs. We use these information shocks to estimate the causal effects of these beliefs. The estimates are consistent with the predictions from the model of reciprocal motivation. Upon learning that their school shares are higher, households *with* children become *less* likely to protest, whereas households *without* children become *more* likely to protest. The effects of the perceptions about government spending are statistically and economically significant. Our baseline estimates imply that increasing the (perceived) school share by 10 percentage points (pp) would cause a drop of 3.67 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. The effects amount to 11% and 10% of the corresponding baseline protest rates, respectively. These results are robust to a host of alternative specifications and falsification tests.

To assess whether the results were surprising or predictable, we conducted a forecast survey using a sample of 56 experts, most of whom are professors researching related topics. After receiving a brief explanation of the experiment, experts are asked to forecast the experimental findings. Only a few of them were able to accurately predict the experimental findings. Most experts predicted that beliefs on school share would have no effect on the likelihood of filing a tax appeal, perhaps forming their predictions based on the results from the existing tax morale literature.

The results of the second treatment arm, about the share of funds being recaptured, are unfortunately imprecisely estimated and thus largely inconclusive. We find that respondents have significant misperceptions about the recapture share and that they update their beliefs significantly when provided with information in the experiment. However, both the levels of misperception and updating are smaller relative to the corresponding findings for the school share. As a result, the information shocks for the recapture share are not nearly as strong as those for the school share, and the causal effects of the beliefs about the recapture share are very imprecisely estimated. It is important to note that the level of misperceptions and belief updating is difficult to anticipate prior to conducting the experiment. So, while ex-ante we expected that both treatment arms would be adequately powered to detect effects, ex-post we found out that we were under-powered for the second treatment arm. In an effort to mitigate publication bias (DellaVigna and Linos, 2022), we still report the analysis for the second treatment arm. Consistent with the hypothesis of reciprocal motivation, the belief about recapture share does not have significant effects on the decision to file a tax appeal among households *without* children – although this finding must be taken with a grain of salt due to the lack of sufficient statistical power. We do not find evidence of significant positive effects for households *with* children – however, the coefficient is so imprecisely estimated that we cannot rule out large positive effects.

Property taxes work almost identically across counties in Texas and similarly throughout the country (Dobay et al., 2019; World Bank, 2019; Nathan et al., 2020).[6] These similarities imply that our results from Dallas County can be reasonably generalizable to other U.S. counties. We discuss the external validity of our results more thoroughly in the Conclusion section of the paper. Moreover, replicating our field experiment in other U.S. counties would be straightforward. Indeed, we propose the use of property tax protests as a novel context to study taxpayers' preferences and tax payments. Tax compliance is affected by state capacity and also varies across taxes and across taxpayers. In developed countries with high state capacity, tax evasion for some taxes and some taxpayers is more difficult due, for instance, to tax withholding and third-party reporting. For this reason, a large share of the literature on tax compliance has been conducted in developing countries with lower state capacity. In contrast, in our context, we can observe if taxpayers want to pay more taxes when they can legally pay less taxes through protesting. In this way, our context allows us to study new questions affecting taxpayers' decisions to pay taxes.[7] We provide detailed accounts of the implementation and data sources that other researchers can follow, and we are happy to share data, code, tips, and additional resources.

[6] For instance, property taxes provide a significant source of school funding in most of the U.S. (Chen, 2021), and other states also redistribute property taxes across school districts, similar to the Texas recapture system (Youngman, 2016).

[7] Another notable advantage of our setting is that it uses publicly available data, which facilitates replication efforts and avoids potential conflict of interest in partnerships with government organizations.

Our study relates and contributes to the literature on the role of tax morale in tax compliance decisions. Unlike the vast amount of causal evidence showing that institutions matter, there is little causal evidence showing that tax morale matters (Luttmer and Singhal, 2014; Slemrod, 2019). We contribute to this literature by providing experimental evidence showing that tax morale can be a significant factor in practice. Moreover, we make methodological contributions that other researchers can follow to better explore the role of tax morale.

A series of papers use correspondence experiments to study tax morale (Slemrod, 2019). They usually 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 (e.g., for a seminal contribution, see Blumenthal et al., 2001). These messages of moral suasion sometimes include information related to government services (see e.g., Castro and Scartascini, 2015; Bott et al., 2020; Bowers et al., 2020; De Neve et al., 2021; Bergolo et al., 2021; Carrillo et al., 2021). This literature shows that while messages with information related to deterrence (e.g., audit probabilities or penalties) are highly effective, messages of moral suasion are largely ineffective. For example, 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."[8] Based on this evidence, one natural interpretation is that institutions are an important driver of tax compliance, but tax morale plays a negligible role. Indeed, the findings from this literature constitute our preferred explanation for why, when asked to forecast the results of our intervention, the majority of experts guessed that information on school taxes would have no effect on tax compliance. However, the results of our study challenge this view. We argue that tax morale matters, but existing correspondence experiments cannot uncover the effects of tax morale due to methodological limitations.

Two innovations in our methodology allow us to shed light on tax morale, both of which are possible due to the novel research design linking data from a survey experiment to administrative tax data at the individual level. This approach is new to this stream of literature and rare even in broader economic research (Bergolo et al., 2020). First, the messages in previous research have typically sought to affect individuals' tax morale by influencing individuals' preferences. However, such preferences are based on historical life experiences and may be too difficult to change with a simple message (e.g., "it is important to contribute your part"). Instead of trying to influence preferences, we propose to study tax morale by inducing changes in beliefs. This is related to a growing literature showing that simple information-provision

[8] Furthermore, they arrive at the same conclusion when they focus on the sub-group of moral suasion messages, including specific information on how public goods and services are funded.

experiments can have significant and long-lasting effects on perceptions and expectations on a range of topics, such as macroeconomic expectations (Cavallo et al., 2017) and salary perceptions (Cullen and Perez-Truglia, 2022).

The second innovation is our ability to measure heterogeneous effects by linking survey data and administrative data. In the context of tax morale, there is scope for highly heterogeneous effects of information. As illustrated by our results on the school share, the same piece of information can have effects in opposite directions for different groups of subjects (i.e., households *with* children vs. households *without* children). It is possible that these large effects across different groups cancel each other out, on average, which would lead to the erroneous conclusion that tax morale is irrelevant for tax compliance. Using survey data to identify which households have children enrolled in public schools and which do not, we can measure the effects of the information separately for each group.[9] Another reason to expect heterogeneous effects relates to how subjects update their beliefs in response to new information. Households that underestimate their school share may adjust their prior beliefs upward when given accurate information, whereas households that overestimate their school share may adjust their beliefs downward when provided with the same information. Again, it is possible that these large effects across different groups cancel each other out, on average, which would lead to the erroneous conclusion that tax morale is irrelevant for tax compliance. Our survey allows us to measure prior and posterior beliefs, thus allowing us to fully elucidate the effects of information on perceptions.

Some studies that do not use correspondence experiments find some suggestive evidence that, consistent with our results, rewarding taxpayers with public services has a positive effect on their subsequent tax compliance. For example, 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 nonexperimental evidence from Manaus, Brazil, showing that households with access to the city sewer system are more likely to pay property taxes.

In comparison to the existing literature, our research design allows us to disentangle the underlying causal mechanisms. In fact, many interventions combine multiple features, making it impossible to identify the precise mechanisms at play. For example, the intervention

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

of 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 the social recognition, the construction of the sidewalk, or both. Similarly, the message in Bergolo et al. (2021) includes normative language that describes information about government spending and information about tax evasion, making it impossible to disentangle whether the effects are driven by the information on spending or by the information on tax evasion.

Our study is unique in two additional dimensions. Our field experiment is the first to measure taxpayers' perceptions about the destination of their tax dollars. Second, the vast majority of related experiments were conducted in developing countries such as Argentina, Brazil, Haiti and Malawi. In those contexts, tax enforcement is low and, as a result, tax compliance is low. Thanks to the use of a novel margin of tax compliance, the decision to file a tax appeal, we are able to study tax morale in a high enforcement context of a developed country.

Our findings are also related to a few other studies in political economy, such as Cullen et al. (2020), which provides quasi-experimental evidence that tax evasion decreases when the political party of the taxpayer is in control of the presidency, and Huet-Vaughn et al. (2019), which provides related laboratory evidence showing that the ideological match between the taxpayer and specific tax expenditures affects the willingness to pay taxes. Beyond tax compliance, recent quasi-experimental evidence demonstrates how the salience of government spending can affect electoral outcomes (Huet-Vaughn, 2019; Ajzenman and Durante, 2022).

Finally, our study also relates to a small but growing literature on the interplay between tax policy and normative considerations. The normative principle of benefit-based taxation was a prominent and at times leading approach among tax theorists in the early twentieth century (Seligman, 1908; Musgrave, 1959). However, the modern optimal taxation literature has largely ignored normative considerations, instead focusing solely on efficiency aspects of taxation (Weinzierl, 2018; Scherf and Weinzierl, 2020). Moreover, a growing body of work seeks to incorporate other normative considerations into the design of tax policy (Mankiw and Weinzierl, 2010; Weinzierl, 2014; Saez and Stantcheva, 2016).[10] This literature is new and mostly theoretical. There is some empirical evidence, but 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 presents the conceptual framework. Section 4 discusses the experimental

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

design and implementation. Sections 5 and 6 present the results. The last section concludes.

2 Institutional Context

2.1 Property Taxes and Public Schools

In Dallas County, property taxes fund various public services, such as schools, parks, roads, and police and fire departments.[11] In 2021, the average home in Dallas County was worth \$327,690, and the average estimated property tax bill was \$6,370, implying an effective tax rate of 1.94%.[12] 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 make up the largest share of property taxes, accounting on average for nearly half (49.78%) of the total property tax bill. There is some variation in the share of school taxes between households. For example, in our subject pool the 10th percentile of the school share is 41.57% and the 90th percentile is 57.01%.[13] The second largest component is the city tax (accounting for approximately 28% of property taxes), followed by hospital (10%), county (8%), college (4%), and special district (<1%) taxes.

Dallas County has 16 main 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 the K-12 public school(s) in their ISD. All households must pay school taxes, regardless of whether they have children enrolled in public schools. The public schools in Dallas County are generally of good quality, although there are some differences.[14] 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 funding and do not receive funding from district property taxes. However, sending children to private schools can be expensive. The average tuition cost for private schools in Dallas County is \$12,374 per student as of 2022.[15] According to data from the 2020 U.S. Census, about 90% of K–12

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

[12] There is substantial 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).

[13] 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 – for more details, see (Nathan et al., 2020).

[14] 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).

[15] Data accessed from `https://www.privateschoolreview.com/exas/dallas-county` on January 5, 2022.

students in Dallas County attend public schools.

2.2 Property Tax Recapture

To make public school funding more equitable across school districts, Texas enacted a redistribution system in 1993 called the "Recapture Plan" or "Robin Hood Plan." This plan diverts school tax funds from "property-wealthy" districts to "property-poor" districts. Due to the significant amounts of taxes involved, the recapture system has been a topic of heated debate among politicians and the general public (Dallas Morning News, 2018). The recapture system has been amended several times since its inception, including a change in 2019 that slowed the growth in the recaptured amounts. However, the degree of redistribution remains significant under the current recapture formula (Texas Education Agency, 2021c).

In this paper, we focus on the *net* redistribution, which is the difference between the taxes recaptured by the state from the district (if any) and the amount distributed from that state pool to the districts. In Appendix B.2, we provide more details about the recapture system, including the recapture formula. Unlike the school share, the recapture share does not vary at the household-level and only varies at the ISD-level (see, e.g., Appendix B.3). A wide variation in the recapture share occurs across the 14 ISDs that are included in our subject pool. Four ISDs are net contributors: the highest giver is Highland Park ISD, which has 57.3% of its school taxes diverted towards more disadvantaged school districts. The remaining ten districts are net receivers: the highest receiver is Mesquite ISD, which in addition to the school taxes it raises, can spend an additional 23.3% thanks to the funds recaptured from other districts.

2.3 Tax Protests

Each year, the Dallas Central Appraisal District (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/or by mail.[16] 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 taxes, city taxes). After the notifications are sent, households have a month from the notification date to file a protest if they disagree 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.

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

Homeowners can file a protest directly or hire an agent to help them with their protests. In exchange for representation, 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 4.5 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.[17] After reviewing the argument, the DCAD can (and often does) make an offer by mail or phone to reduce the assessed value of the home. 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 either immediately or at the billing date if it is later (i.e., on October 1st in 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. To avoid costly in-person appraisals, the DCAD uses statistical models and large datasets (e.g., recent home sales) to formulate an estimated market value for each property. 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 the interpretation of our results because it implies that households are not trying to objectively "correct" estimates from the DCAD. Instead, they are presenting 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 more details, see Nathan et al., 2020) and not necessarily to complain about the county's estimate of their home value.

3 Conceptual Framework

To formalize the logic of the reciprocal motivation mechanism, we introduce a simple model of how the provision of government services and recapture affect the decision to file a protest.

[17] To protest online, homeowners need to look up their account (e.g., searching for their own 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).

Let subscript $j \in \{C, NC\}$ represent the two types of households: those *with* children enrolled in public schools ($j = C$) and those *without* ($j = NC$). The probability that a household of type j protests its taxes is given by:

$$Prob(j\ protests) = \Phi(P_j), \tag{1}$$

where $\Phi()$ is the cumulative distribution function from a standard normal and $P_j \in (-\infty, +\infty)$ is a latent variable representing the tax morale of the household. Note that by construction, for any variable x, the sign of $\frac{\partial \Phi(P_j)}{\partial x}$ will be equal to the sign of $\frac{\partial P_j}{\partial x}$. For this reason, and for the sake of brevity, the following analysis focuses on the latent variable P_j. Let B_j be how much households in group j benefit from each dollar spent on government services. Consider the following relationship:

$$P_j = \gamma_0 + \gamma \cdot B_j \tag{2}$$

where γ_0 is a constant and $\gamma < 0$ represents the reciprocal motivation: that is, when households benefit directly from government expenditures, they are less likely to protest their taxes.

3.1 The Effects of the School Share

Let S be the government expenditures in the local public school district and NS be the government expenditures in other local government services (e.g., police, parks, roads). The two types of households benefit from the two types of government expenditure in the following manner:

$$B_C = \alpha^S \cdot S + \alpha^{NS} \cdot NS \tag{3}$$

$$B_{NC} = \alpha^{NS} \cdot NS \tag{4}$$

The parameters α^S and α^{NS} capture how households benefit from different types of expenditure. The parameter α^S denotes how much a household *with* children enrolled in public school benefits per dollar spent in public schools. α^{NS} denotes how much households (regardless of whether they have children) benefit per dollar spent on non-school government expenditures.

Next, we conduct a simple normalization. Let $G = S + NS$ denote total expenditures and $s = \frac{S}{G}$ denote school expenditures as a fraction of total expenditures, which we previously defined as school share. It is important to note that while we do not incorporate misperceptions into this simple framework, in practice, the "s" that matters is the one perceived by the taxpayer when deciding whether to protest. We thus can rewrite equations (3) and (4) as follows:

$$B_C = G \cdot (\alpha^S \cdot s + \alpha^{NS} \cdot (1 - s)) \tag{5}$$

$$B_{NC} = G \cdot \alpha^{NS} \cdot (1 - s) \tag{6}$$

Combining equations (2), (5), and (6), we obtain the following:

$$P_C = \gamma_0 + \gamma \cdot G \cdot (\alpha^S \cdot s + \alpha^{NS} \cdot (1 - s)) \tag{7}$$

$$P_{NC} = \gamma_0 + \gamma \cdot G \cdot \alpha^{NS} \cdot (1 - s) \tag{8}$$

Using equations (7) and (8), we can see what happens to protest rates if the school share increases. Let us start with households *without* children:

$$\frac{\partial P_{NC}}{\partial s} = -\gamma \cdot G \cdot \alpha^{NS} > 0 \tag{9}$$

Intuitively, when the school share is increased, that unambiguously means that households *without* children benefit less from government services, and thus are more likely to protest. For the households *with* children, the effect could go either way:

$$\frac{\partial P_C}{\partial s} = \gamma \cdot G \cdot (\alpha^S - \alpha^{NS}) \tag{10}$$

Intuitively, whether households *with* children are more or less likely to protest will depend on whether they benefit more from the school expenditures or the non-school expenditures. If they prefer school expenditures ($\alpha^S > \alpha^{NS}$) then they will be less likely to protest when the school share increases. If they prefer the non-school expenditures ($\alpha^{NS} > \alpha^S$) then they will be more likely to protest when the school share goes up. In either case, if we subtract equation (9) from (10), we obtain the following:

$$\frac{\partial P_C}{\partial s} - \frac{\partial P_{NC}}{\partial s} = \gamma \cdot G \cdot \alpha^S < 0 \tag{11}$$

In other words, when the school share goes up, while the effect on households *with* children may be negative or positive, it has to be smaller than the corresponding effect for household *without* children. The intuition is straightforward. When the school share goes up, both households *with* children and households *without* children lose in the non-school expenditures. However, for households *with* children, at least they gain in school expenditures. For that reason, even if the (latent) probability of protesting goes up for a household *with* children, it should go up less than for households *without* children because households *with* children at least have something to gain. This can be summarized in the following prediction:

Prediction 1: *When increasing the school share, the effect on the (latent) probability of protesting should be lower for households* with *children in public schools than for households* without *children in public schools.*

One special case worth mentioning is when households *with* children in public schools benefit more from school expenditures than from non-school expenditures. Intuitively, unlike the benefits from non-school expenditures (e.g., police, roads), which are spread over the entire community, the benefits from school expenditures are concentrated on a subset of the population (households *with* children enrolled in public schools). For that reason, it is plausible that the households *with* children prefer school expenditures over non-school expenditures:

Corollary 1: *If $\alpha^S > \alpha^{NS}$, an increase in the school share should negatively affect the (latent) protest probability of households* with *children in public schools and positively affect the (latent) protest probability of households* without *children in public schools.*

These predictions are based on some assumptions. First, this setup assumes that benefits from non-school services are the same for households *with* children as for households *without* children. However, the main predictions will still hold as long as the parameters are close enough between the two types of households. Second, our model assumes that households are entirely selfish and thus households *without* children do not benefit at all from school spending. In practice, these taxpayers may feel happy to help other parents in the community, they may benefit from schools in the future, or value public schools because they had children in schools in the past. Alternatively, they may benefit from school spending for selfish reasons if, for instance, school spending reduces crime in the neighborhood. Nevertheless, in Appendix A.1 we show that the main prediction still holds under more general assumptions.

3.2 The Effects of Recapture of School Taxes

It is straightforward to extend this simple model to include redistribution of school taxes. Non-school expenditures are still NS. School expenditures are now $S \cdot (1 - r)$, where $r \in [-\inf, 1]$ is what we previously defined as the recapture share and represents the direction and intensity of the effects of recapture on the funding available for the local school district.[18] If there is no recapture, or if there is recapture but the local school district does not lose or gain in net terms, $r = 0$ and we are back to the original model. A positive value of r means

[18] The value of r can be below -1 because, in theory, a school district could receive through recapture more than 100% of the amount it raised in school taxes.

that the school district is a net contributor to the recapture system. More precisely, r is the fraction of school taxes raised in the district that are transferred to disadvantaged school districts and therefore cannot be spent in the local school district. For example, $r = 0.1$ would indicate that 10% of local school taxes are redistributed to other school districts. On the other hand, a negative value of r means that the school district is a net beneficiary of the recapture system and thus can spend more in schools than what the district raised in school taxes. More precisely, for each dollar raised locally in school taxes, the school district can spend an additional $-r$ dollars thanks to net transfers from wealthier school districts. For example, $r = -0.1$ would indicate that the local school district can spend the school taxes it collects plus an additional 10% from the amount recaptured.

We can extend equations (3) and (4) to incorporate recapture into the model:

$$B_C = \alpha^S \cdot S \cdot (1 - r) + \alpha^{NS} \cdot NS \tag{12}$$

$$B_{NC} = \alpha^{NS} \cdot NS \tag{13}$$

We combine equations (12) and (13) with equation (2), and then rearrange them as follows:

$$P_C = \gamma_0 + \gamma \cdot \alpha^S \cdot S \cdot (1 - r) + \gamma \cdot \alpha^{NS} \cdot NS \tag{14}$$

$$P_{NC} = \gamma_0 + \gamma \cdot \alpha^{NS} \cdot NS \tag{15}$$

We can see what would happen if we increased the recapture share:

$$\frac{\partial P_{NC}}{\partial r} = 0 \tag{16}$$

Households *without* children in the school district do not benefit from school taxes, regardless of whether their school district gives or receives funding from the recapture system, so their willingness to pay taxes is not affected by recapture.

$$\frac{\partial P_C}{\partial r} = -\gamma \cdot \alpha^S \cdot S \cdot > 0 \tag{17}$$

For households *with* children, in turn, more recapture means fewer benefits for their local school district, and they are thus less willing to pay taxes.

We can also subtract (16) from (17) to show the following:

$$\frac{\partial P_C}{\partial r} - \frac{\partial P_{NC}}{\partial r} = -\gamma \cdot \alpha^S \cdot S \cdot > 0 \tag{18}$$

Again, there is an unambiguous prediction about the differential effect between households *with* children and *without* children. These results are summarized in the following prediction:

Prediction 2: *An increase in the recapture share should increase the (latent) protest probability for households* with *children in public schools, but should not affect the (latent) protest probability for households* without *children in public schools.*

However, we must take this prediction with a grain of salt. Our setup assumes that households are totally selfish. However, this assumption may be misleading: as the survey data show, and contrary to the prediction of the selfish model, there is quite a bit of support for the recapture system. A more realistic model would include altruism. For example, when funds are transferred from advantaged to disadvantaged districts due to recapture, households may appreciate that their tax dollars are helping the most disadvantaged households, even if that means that their own children will have fewer resources. In Appendix A.2, we provide an extension of this framework that incorporates altruism and show that Prediction 2 may no longer hold.

4 Data, Experimental Design, and Implementation

4.1 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).[19] 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. Our starting point to define the experimental sample is 400,193 non-commercial, owner-occupied, residential properties, which we will refer to as the "Universe" sample.[20] When necessary, we

[19] The latest version of the data is available in https://www.dallascad.org/DataProducts.aspx. We downloaded most of the baseline information on 04/16/2021, the day the DCAD notified the proposed values for 2021.

[20] We arrived at this subsample 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, among others. See Appendix B.1 for a comprehensive description of the selection criteria.

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

Out 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." The sample criteria, explained thoroughly in Appendix B.1, ensures a wide representation of beneficiaries and contributors to the recapture system. More specifically, we over-sample households from ISDs within Dallas County that contribute the most to the recapture system (Carrolton, Coppell, and Highland Park) in order to increase variation in the recapture share. All homeowners in these three districts were selected for the letter sample. We also over-sample 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).[21]

Panel (a) of Table 1 presents descriptive statistics for some key variables based on the information available in the administrative records. Column (1) corresponds to the universe, while column (2) corresponds to the letter sample.[22] 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%). Additionally, because we over-sampled properties in school districts that are property richer, the average recapture share is positive (1.23%), which means that the average property included in the letter sample is located in a school district that transfers part of their school taxes revenues to other property poorer school districts. In terms of protest history, the homeowners selected to receive the letter seem slightly more likely to fill a protest directly (e.g., in 2020, the direct protest rate was 8.84% for this sample vs. 7.99% in the universe sample).

4.2 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 an URL to access the survey. We mailed our letters so that they would be delivered close to the time that homeowners in Dallas County could start filing tax appeals. Appendix C shows a sample envelope, and Appendix D 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-

[21] 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.

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

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 itself included a physical address for the researcher and a link to the study's website (see Appendix E 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 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. As previously mentioned, the letter also mentioned the proposed value of the recipient's home and the estimated amount of property tax for 2021.

Most 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.[23] As a second incentive, some subjects were informed that they would enter a raffle for 20 prizes worth $100 each.[24]

4.3 Survey Design

In this section, we summarize the main features of the survey.[25] We start by asking a critical question, that is, whether the respondent's household has children enrolled in grades K–12 at their local public school district and, if so, how many. This critical information is not included in the tax agency's administrative records. Thus, the analysis would be impossible without this question, particularly the heterogeneity analysis concerning the framework of reciprocal motivation, which is the main form of heterogeneity that we anticipate in the RCT pre-registration.

The module on school taxes can be summarized as follows:

[23] This walkthrough included hyperlinks to relevant websites and screenshots of a sample protest using information for a fictitious household for added clarity. To access these instructions, subjects received a URL and a code on the final screen of the survey. A copy of the Web instructions is included in Appendix F. Nathan et al. (2020) show that these instructions have a significant positive effect on the probability of protesting.

[24] 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 B.5. 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).

[25] A sample of the full survey instrument is attached as Appendix H. We included methodological notes as pop-up windows, which are reported in Appendix B.4.

- **Step 1 (Elicit Prior Belief):** We begin by providing the estimated total property tax amount of 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 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 or not. 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, as opposed to providing the information directly in the letter, to measure beliefs prior to information provision. To take advantage of the effect of information on prior beliefs, 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).

The following module is about the recapture share. Some subjects may not know about or understand recapture. Thus, we start with a couple of short paragraphs summarizing the recapture system. The rest of the module follows the same structure as described above for Step 1 through Step 3. We elicit beliefs about the recapture share in two stages. First, we ask respondents to guess if their school district will receive more, the same, or less taxes than what households in their district paid in school taxes. The following stage is quantitative in nature. If the respondent selects "More" (or "Less") in the first question, we ask them to guess how much *more* (or *less*) funding their school district will receive as a share of the district's school tax revenues due to recapture, using any amount between 0% and 100%. We then conduct Step 2 (information-provision experiment) and Step 3 (elicitation of posterior beliefs).

We cross-randomize subjects to receive the two pieces of information about school taxes and recapture, respectively, with a 50% probability for each. Thus, roughly 25% of the sample receives both pieces of information, 25% receives the first piece of information only, 25% receives the second piece of 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 one that serves as a secondary outcome in the analysis of the effects of beliefs. We ask respondents if they plan to file a protest this year on a 1-4 likelihood scale. 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 descriptive purposes, we include questions asking respondents' gender, age, ethnicity, education, and political party. To provide complementary evidence, towards the end of the survey, we include additional questions that are described in more detail in the following sections.

4.4 Implementation

We timed the intervention so that our letters would arrive early enough before the protest deadline 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 began to receive survey responses and visits to the study's website on April 22, 2021.[26] 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,

[26] Appendix B.3 contains more descriptive information about the sample of homeowners who answered the survey, and Appendix B.6 contains more details about the timing of survey responses and discusses in detail attrition rates and balance tests.

school share, and recapture 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).[27] Among respondents, the median time to complete the survey was 11.3 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 92.1% 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 drop responses that, as explained in the RCT pre-registration, could not be excluded *ex ante* due to data availability. We drop 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 B.1).

When studying perceptions through survey data, it is important to properly deal with outlier beliefs. Some individuals may provide guesses that are wildly inaccurate, not because they truly hold such extreme beliefs, but because they misunderstand the question, make a typo, or do not pay enough attention to the question. The "information shocks" for these individuals can be large but meaningless, which can induce substantial attenuation bias to the causal estimates. 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.[28] After applying these filters, 2,110 respondents remain, constituting our main subject pool. 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 more lax definitions of outliers. Finally, we

[27] 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.

[28] For more details on the distribution of outlier observations, see Appendix B.7.

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 show the average pre-treatment characteristics according to the administrative records (e.g., home value, number of bedrooms, tax rate). 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 (thanks to the large sample sizes) but typically small in magnitude.[29] The comparison between columns (2) and (3) indicate 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 one 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.[30] This is largely by design, as we crafted the letter to attract the attention of households who were 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.[31] 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, the average subject owns a home with an assessed market value of $349,988 and property taxes of $7,738 (corresponding to an average tax rate of 2.21%). Panel (b) of Table 1 reports some descriptive statistics based on information collected in the survey. The average respondent is 49.6 years old, 42.9% are women, 44.3% are White, and 38.3% have a college degree. And 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

[29] 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 over-sampled certain types of households.

[30] For a more detailed discussion of the differences between survey respondents and non-respondents, see Appendix B.3.

[31] Additionally, our instructions likely make it easier for survey respondents to file an appeal, as documented in Nathan et al. (2020).

and thus should not be affected by the treatment assignment.[32] 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.[33] Appendix B.6 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 they already made the decision to start the survey.

4.5 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.[34] To get a sense of the baseline protest rate, we consider subjects in the control group (i.e., those who do not receive any information on school taxes or recapture). Approximately 30.1% of these owners file a tax appeal in 2021. These tax protests are consequential: 65.4% of the protests lead to a decrease in the assessed value of the home, which, in turn, translated into \$579 in average tax savings in the first year alone.[35]

Owners can file their own protests, which is the main focus of this paper. For the sake of 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% use an agent.

As stated in the RCT pre-registration, there are several features of our experimental setting and the administrative protesting process that lead us to expect that the experimental effects are concentrated on whether owners file tax protests directly and not on protests

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

[33] 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.

[34] Information on whether property owners protested their property taxes was downloaded from the DCAD website for the last time on June 22, 2021.

[35] These calculations are based on data downloaded from the DCAD website on December 8, 2021, which compared to June 22, 2021 data, contain additional information on the resolution of the protests. The remaining protests are not resolved by December 2021 (12.2%) or resolved without a change in the assessed home value (22.4%).

through an agent. Indeed, that is precisely the reason why, when forming the subject pool to be invited to the survey, we filtered out households whose owners had protested through an agent in previous years. Additionally, while we acknowledge that the information included in the experimental treatment could affect the decision to hire or fire an agent, we think this is unlikely. According to conversations with households, tax agents, and representatives from assessor's offices, households typically sign contracts with agents well in advance of the date when the proposed values are announced. Some households even sign long-term contracts to file protests on behalf of the owner over many years. If these contracts are based on flat fees, agents would have an incentive to protest mechanically, particularly if the cost of protesting is low. Additionally, homeowners in all counties in Texas are required to complete and submit a form to their Appraisal District to terminate an agent, creating a stickiness in the relationship between agents and households. This implies that agents frequently protest on behalf of owners every year. 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. For the sake of completeness, we report the effects on protests through agents' tax protests, but we expect our intervention to have no effects on this margin.

4.6 Expert Prediction Survey

To assess whether the experimental results are surprising, we conduct a forecast survey with a sample of experts. A sample of the full survey instrument is attached as Appendix I. In this survey, which follows best practices (DellaVigna et al., 2019), we describe the experiment and ask experts to forecast key results in a way comparable to experimental estimates. More precisely, we elicit their prediction of the effect of a 10 pp shock to the belief about the school share, separately for households *with* and *without* children. We then conduct the corresponding elicitations for beliefs about the recapture share.

We invited experts to participate in our survey in two ways. First, we posted the survey on the Social Science Prediction Platform from July 13, 2021, to December 31, 2021. Second, on November 2021 we invited by email a sample of 238 professors with published research on related topics. The final sample includes 56 experts' responses. Of these, 21.4% responded to the survey through the Social Science Prediction Platform, and the remaining 78.6% responded through our email invitation.[36] The final sample is made up of professors (82.1%), Ph.D. students (12.5%), postdocs (3.6%) and other researchers (1.8%). Most of the respondents (78.6%) are economists, 66.1% report having done research on taxation, and

[36] Among the responses from the Social Science Prediction Platform, we require that they either are academics, already have a Ph.D. or are currently pursuing one.

25% have done research on preferences for redistribution.

5 Perceptions about School Spending

5.1 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.[37] 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 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 1(a) illustrates the distribution of misperceptions about the school share for the 2,110 observations in the subject pool before the experimental treatment.[38] The x-axis corresponds to the difference between the actual school share (i.e., potential feedback) and that perceived by respondents. For the sake of 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 1(a), where more observations fall in the right half of the histogram (corresponding to an under-estimation) than in the left half (corresponding to an over-estimation). 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 B.7.

[37] See Appendix G 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 at the time of deciding whether to protest.

[38] Appendix B.7 contains additional information for the full survey sample without excluding any outliers.

5.2 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 as of 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. Specifically, 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 (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.[39] This Bayesian updating model can be summarized by the following linear relationship:

$$s_i^{post} - s_i^{prior} = \alpha \cdot \left(s_i^{feed} - s_i^{prior}\right) \tag{19}$$

Intuitively, Bayesian learning predicts that, when shown feedback, respondents who overestimate the school share would revise their beliefs downward, whereas respondents who underestimate the school share would revise their beliefs upward. Figure 1(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 update of the belief ($s_i^{post} - s_i^{prior}$). Intuitively, the x-axis shows the maximum revision we would expect if the respondent were to fully react 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 1(b) correspond to the subjects who are shown feedback about the school share. Consistent with significant updating, there is a strong relationship between the updated beliefs and prior

[39] In the typical model in the literature, the results assume 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).

gaps: an additional percentage point (pp) in perception gap is associated with an actual revision that is 0.809 pp higher.

The gray squares in Figure 1(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 that is 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 that is 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:

$$s_i^{post} - s_i^{prior} = \tau + \alpha \cdot \left(s_i^{feed} - s_i^{prior}\right) \cdot T_i^S + \beta \cdot \left(s_i^{feed} - s_i^{prior}\right) + \epsilon_i \tag{20}$$

This regression forms the basis for the first-stage of the 2SLS model. In this model, parameter α represents true learning arising from the information provision (not spurious learning), while parameter β captures spurious learning. The parameter α can be calculated from the estimates in Figure 1(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.

Appendix B.7 provides some additional results and robustness checks. For example, we show that belief update is not different between households *with* and *without* children and 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).

5.3 Econometric Model

Let P_i^{2021} denote the main outcome of interest: an indicator variable equal to 100 for individuals filing a protest in 2021 (i.e., post-treatment) and 0 otherwise. As discussed in the conceptual model in Section 3, and as noted in the RCT pre-registration, the effects of the school share information treatment on protests are expected to be different depending on whether the household has children enrolled in public schools. 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. Therefore, we can use the following econometric specification to estimate our parameters of interest:

$$P_i^{2021} = \beta_0 + \beta_C^S \cdot C_i \cdot s_i^{post} + \beta_{NC}^S \cdot (1 - C_i) \cdot s_i^{post} + \beta_1 \cdot C_i + \epsilon_i \tag{21}$$

where ϵ_i is the usual error term. The two parameters of interest are β_C^S and β_{NC}^S. According to Prediction 1, we expect $\beta_C^S - \beta_{NC}^S < 0$. Moreover, according to Corollary 1, we expect $\beta_C^S < 0$ and $\beta_{NC}^S > 0$. Posterior beliefs (s_i^{post}) could be correlated to a host of omitted variable biases. Therefore, we estimate equation (21) using a Two-stage Least-Squares (2SLS) model that leverages the exogenous variation in posterior beliefs induced by the information-provision experiment. More precisely, we estimate the following model:

$$\begin{aligned} P_i^{2021} = & \beta_0 + \beta_C^S \cdot C_i \cdot s_i^{post} + \beta_{NC}^S \cdot (1 - C_i) \cdot s_i^{post} + \beta_1 \cdot C_i + \\ & + \beta_2 \cdot C_i \cdot (s_i^{feed} - s_i^{prior}) + \beta_3 \cdot (1 - C_i) \cdot (s_i^{feed} - s_i^{prior}) + X_i \beta_X + \epsilon_i \end{aligned} \tag{22}$$

The endogenous variables are $C_i \cdot s_i^{post}$ and $(1 - C_i) \cdot s_i^{post}$, for which we use the excluded instruments $C_i \cdot T_i^S \cdot (s_i^{feed} - s_i^{prior})$ and $(1 - C_i) \cdot T_i^S \cdot (s_i^{feed} - s_i^{prior})$.[40]

We can illustrate the intuition behind the model using a simple example. Consider a pair of subjects with children enrolled in public schools that share the same bias about the school share: both underestimate the actual school share by 20 pp. Suppose that 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

[40] Note that equation (22) 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).

+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.[41]

The term X_i in equation (22) corresponds to a set of additional control variables. In principle, the 2SLS model leverages experimental variation, so additional control variables are not needed for causal identification. However, the inclusion of additional control variables can be helpful, for example, in reducing the variance of the error term and thus improving the statistical precision (McKenzie, 2012). The vector of control variables includes basic pre-treatment information, such as the household's prior history of tax appeals.[42]

Following the regression specification that we use to study the effects of the school share (equation (21)), it is straightforward to define the regression specification to study the effects of recapture. Indeed, since these two information treatments are cross-randomized for the same sample, we estimate all effects simultaneously in a single 2SLS regression.

5.4 2SLS Estimates

The 2SLS estimates for the school share are presented in the top half of Table 2. In column (1) of Table 2, the dependent variable is the main outcome of interest: an indicator variable that equals 100 if the subject protests directly in 2021 and 0 otherwise. According to Prediction 1, the difference in the effects of school share between households *with* children and *without* children should be negative. Consistent with that prediction, the difference between the coefficients with and without children is negative (-0.644), large in magnitude and statistically significant (p-value=0.012). Under an additional assumption, Corollary 1 predicts

[41] Typically in 2SLS models, if treatment effects are heterogeneous, the estimates identify the local average treatment effects of beliefs (Imbens and Angrist, 1994). More precisely, in our study, our estimates would give a higher weight to subjects whose beliefs are more affected by the information-provision experiment. By construction, this weight will be higher for subjects with greater prior misperceptions and, holding the misperceptions constant, those who react more strongly to feedback.

[42] The full set of additional control variables includes the log 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 homestead binding 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). As reported in Section 5.5, as falsification tests, in some cases, we report estimates that use pre-treatment outcomes (i.e., measured before 2021). In these cases, the variables that control the history of the protest correspond to the period prior to the measurement of the outcome variable. For example, if the dependent variable is the decision to protest in 2020, we control for a set of indicator variables corresponding to the history of protests in 2016-2019.

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. The results presented in column (1) of Table 2 are also consistent with those predictions: the coefficient for households *with* children is negative (-0.367) and borderline statistically significant (p-value=0.096), while the coefficient for households *without* children is positive (0.277) and statistically significant (p-value=0.032).

These coefficients 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 from column (1) of Table 2 indicate that this change would cause a decrease of 3.67 pp ($= 0.367 \cdot 10$) in the probability of filing a protest for households *with* children and an increase of 2.77 pp ($= 0.277 \cdot 10$) in the probability of protesting for households *without* children. These effects would be roughly equivalent to 11% and 10% of the baseline protest rates (33.86 pp and 28.83 pp, reported in the lower rows of Table 2).

To further illustrate the magnitude of these effects, we can convert them into a money metric. For that, we take advantage of the estimates from Nathan et al. (2020), which is another study on tax protests in Dallas County. 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 3.67 pp (for households *with* children) and 2.77 pp (for households *without* children) against that benchmark. For households *with* children, the effect of 3.67 pp would be equivalent to an effect of -\$172 on the expected tax savings ($= \frac{3.67 \cdot 100}{2.14}$). For households *without* children, the 2.77 pp effect on the protest probability would be equivalent to an effect of \$129 on the expected tax savings ($= \frac{2.77 \cdot 100}{2.14}$).

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 to protest 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. For reference, at the time of the survey, 45.4% report that they are very likely to protest (this corresponds to the baseline rate, combining subjects *with* and *without* children who do not receive any feedback), which is higher than the actual protest rate in the administrative data, 30.06%. For example, a respondent may report a high probability of protesting in the survey, but then do not protest due to filing frictions (Nathan et al., 2020). It is important to note that the stated intention to protest is significantly correlated with whether the individual actually files a protest, but that correlation is far from perfect: the correlation coefficient is 0.410 for the no-feedback group (p-value<0.001).[43] Due to this imperfect correlation, the effects on the intention to protest at the time of answering the survey should not be expected to be "mechanically" the same as the effects on actual protests.

The results from column (2) of Table 2 are consistent with the results from column (1). In column (2), the coefficient for households *with* children 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). Again, most importantly, 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).

A common concern when using 2SLS estimation is the potential for weak instruments (Stock et al., 2002). Given the strong belief updating documented in Section 5.2, weak instruments should not be a concern in our setting. Nevertheless, for a more rigorous assessment, Table 2 reports the Cragg-Donald F-statistic, which is commonly used to diagnose weak instruments. The value of this statistic in each regression is well above the rule of thumb of $F > 10$ proposed by Stock et al. (2002): it equals 30.10 and 30.22, respectively, in columns (1)–(2) of Table 2. It is important to note that our preferred specification corresponds to the 2SLS estimates because they can be interpreted more easily. Nevertheless, due to the strong first stage (i.e., strong belief updating), the reduced-form estimates are also statistically significant and qualitatively similar. We report these results, together with the first stage estimates, in Appendix B.8.

5.5 Robustness Checks

As explained in Section 4.5, it is highly unlikely that the information provided in our survey would affect protests through an agent. However, for completeness, we report these results in column (3) of Table 2. 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 (-0.028 and -0.033) for both households *with* and *without* children, precisely estimated with standard errors smaller than in column (1), and statistically insignificant (p-values of 0.816 and 0.518). The difference between the coefficients for households *with* and *without* children is close to zero (0.006), precisely estimated and statistically insignificant (p-value=0.966).

[43] Among respondents who report being very likely to protest, 56.8% end up protesting directly or through an agent. On the other hand, among respondents who do not report being very likely to protest, 16.8% end up protesting.

To investigate the robustness of the results, column (4) of Table 2 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 as the dependent variable the protest decision in a pre-treatment year (2020) rather than in the post-treatment year (2021). Intuitively, since the information was provided in 2021, it could not possibly have an effect on 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); most importantly, 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 2(a). The x-axis denotes the year of the dependent variable (i.e., whether the owner protests directly in 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 2(a), which takes the value 0.175, corresponds to the coefficient from column (4) of Table 2. 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 2SLS model used for the results in equation (22) 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. To probe this linearity assumption, Figure 2(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 2(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. Additionally, this figure shows that outliers do not drive the regression results discussed above. In a similar spirit, as discussed in Section 5.4, Appendix B.8 shows that the 2SLS estimates are consistent with the reduced-form estimates.

Table 3 presents additional robustness checks. Columns (1) and (2) of Table 3 reproduce the baseline specification given by columns (1) and (2) of Table 2 for reference. Columns (3) through (12) of Table 3 present the results under alternative specifications. The specification from columns (3) and (4) is identical to the specification from columns (1) and (2), except that we include some additional control variables: the 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). Note that these variables are measured at the end of the survey, but some respondents did not finish the entire survey. Therefore, the inclusion of these additional controls reduces the number of observations, which is the main reason why we exclude these variables from the set of baseline controls. The results from columns (3) and (4) are similar to the baseline results from columns (1) and (2). If anything, the inclusion of the additional controls yields effects that are slightly stronger (-0.714 vs. -0.644 and -0.744 vs. -0.678). Columns (5) and (6) report the results of an alternative specification that does not include any additional control variables at all. The results are again similar in direction, size, and statistical significance.

In columns (7) through (10) of Table 3, we consider alternative definitions of outliers in prior misperceptions. The baseline specification is already conservative in that it excludes the extreme top and bottom 5% of the distributions. In columns (7) and (8), we use a less stringent definition of outliers based on the top and bottom 2.5% instead of the top and bottom 5%. The results are similar to those of the baseline specification of columns (1) and (2), although the coefficients are slightly smaller in magnitude. In columns (9) and (10), we consider an even more lax definition of outliers, excluding only the top and bottom 1% of misperceptions. These coefficients remain consistent with those from the baseline specification of columns (1) and (2), although again the magnitudes are somewhat smaller. These results are consistent with the arguments in Section 4.4 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, columns (11) and (12) are identical to the baseline specification from columns (1) and (2), except that they exclude respondents who do not pass the attention check included at the end of the survey. Consistent with the attention argument, when we focus on subjects who pass the attention check, the coefficients increase somewhat.

Appendix B.8 discusses additional robustness tests and additional results that include average treatment effects of the school feedback treatment, heterogeneous effect by individual characteristics such as age, gender, education, and self-identification with a political party, heterogeneous effects by direction of the prior gap (i.e., overestimate vs. underestimate), additional survey outcomes, and an alternative specification that estimates the 2SLS baseline specification in separate regressions for the school share feedback and the recapture feedback.

5.6 Comparison to Expert Predictions

Next, we compare our experimental results with expert predictions, as shown in Figure 3. Panel (a) presents the predictions of experts for households *with* children, and panel (b) presents the predictions for households *without* children. The histograms correspond to the distribution of expert predictions for the effect of a 1 pp increase in the school share.[44] The solid vertical red line in each panel represents the corresponding estimate from the baseline 2SLS model (column (1) of Table 2), and the red shading denotes the corresponding confidence intervals.

Figure 3 shows that our experimental findings are not obvious to the sample of experts. Our experimental results are consistent with a minority of experts who predicted that the school share belief would have a negative effect on the protest rate for households *with* children (panel (a)) and a positive effect for households *without* children (panel (b)). They are also consistent with the mean of the experimental estimates in these two panels. However, the forecasts of the majority of experts are inconsistent with the experimental results: most experts predict either zero effect or an effect of the opposite sign compared to the experimental findings. In addition, only a few expert predictions are close to the experimental estimates, even if we account for the sampling variation in the experimental estimates. More precisely, for households *with* children, only 41.1% of predictions are within the 90% confidence interval of the experimental estimate. For households *without* children, only 17. 9% of the predictions are within the 90% confidence interval of the experimental estimate. That the majority of experts' predictions do not coincide with the experimental findings may not be surprising, since their predictions are consistent with the general takeaway from the extant literature on how messages of moral suasion affect tax compliance, which suggests that deterrence nudges are effective, whereas tax morale messages are less effective or have no effects whatsoever (see Antinyan and Asatryan (2019)).

At the end of the survey, we ask the experts to express how confident they feel about their

[44] To make the elicitation easier, in the prediction survey, we ask subjects to predict the effects of a 10 pp increase in the school share. In Figure 3, we divide those predictions by 10 to obtain the effect per 1 pp, so it can be compared directly with the 2SLS estimates.

forecasts. One notable finding is that experts do not feel confident about their predictions: on a scale of 1 to 5, where 1 is "not confident at all" and 5 is "extremely confident," the average confidence is 2.07.[45] In any case, we find that the comparison between the forecasts and experimental estimates is similar if we weight the forecasts by the confidence of experts (results reported in Appendix B.9).

5.7 Non-Experimental Evidence

In this section, we present some non-experimental evidence that complements the experimental evidence presented above. Our survey included a question asking respondents to choose between hypothetical policies, in the spirit of Weinzierl (2014) and Saez and Stantcheva (2016). More precisely, we include a question about public school taxes. 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 will want the household *with* children to pay more taxes than the household *without* children, because the former benefits more from this government service. We find that most (58.8%) of the respondents behave according to the reciprocal mechanism, that is, they assign a higher tax burden to the household *with* children even though both homes are worth the same.[46] This evidence suggests that the logic of reciprocal motivation resonates with most taxpayers.

As a last piece of anecdotal evidence, there is a feature of the property tax policy in the state of Texas that suggests reciprocal motives are at play. It is highly unlikely that people 65 years or older have children of school age. In Texas, homeowners who fall into this age group are eligible for a special exemption with respect to their school taxes. This exemption ensures that their school tax payments remain fixed at the amount they paid in the year they turned 65, regardless of future increases in the value of their property (Texas Comptroller, 2021). This policy, which benefits households unlikely to have school-aged children, reflects a benefit-based approach to taxation.

[45] More precisely, 25.0% of experts selected "not confident at all," 51.8% selected "slightly confident," 19.6% selected "somewhat confident," 3.57% selected "very confident," and 0% selected "extremely confident."

[46] For more details, see Appendix B.10.

6 Perceptions about Recapture

6.1 Accuracy of Prior Beliefs

Unlike the information on the school share, the information on recapture is not readily available in the Notice of Appraised Value from the DCAD. However, households can be informed about the recapture system through its media coverage. Also, it is probably widely known that the recapture system redistributes from more to less advantaged districts. As a result, if a homeowner knows whether he or she lives in a more or less advantaged district, that information alone may be enough to form a decent guess about the recapture share.

Figure 4(a) shows a histogram of the degree of misperceptions about the recapture share. The x-axis corresponds to the difference between the actual recapture share versus that perceived by the respondents. A minority of subjects have accurate perceptions: around 20% of subjects guess the recapture share to be within $\pm$ 5 pp of the actual share. Misperceptions are significant in magnitude: the mean absolute error is 11.36 pp. However, the mean absolute error for the recapture share (11.36 pp) is substantially less pronounced than that of the school share (16.57). The fact that misperceptions for the recapture share are smaller than those for the school share implies that there is less scope for the information provision experiment to update beliefs and thus less statistical power for the 2SLS estimates.

Unlike misperceptions about the school share, misperceptions about the recapture share have no systematic bias: on average, subjects overestimate the recapture share by just 0.28 pp. This can be seen directly from Figure 4(a), which shows that households are roughly equally likely to be in the left half of the histogram (corresponding to over-estimation) as in the right half (corresponding to an under-estimation). Appendix B.7 contains a host of additional information on the distribution of prior beliefs (e.g., the comparison of gaps between households *with* and *without* children).

6.2 Belief Updating

Next, we summarize how subjects update their beliefs in reaction to the information provision about the recapture share. Figure 4(b) shows the results as a binned scatterplot. The x-axis corresponds to the gaps in prior beliefs, and the y-axis denotes the belief updating. The x-axis of Figure 4(b) shows the theoretical revision that we would expect if the respondents fully responded to the information provided, while the y-axis shows the revision observed in practice. The red circles in Figure 4(b) correspond to the subjects who receive feedback on the recapture share. Consistent with significant learning, there is a strong relationship between the belief revisions and prior gaps: an additional percentage point (pp) in perception gap is associated with a revision that is 0.632 pp higher. The gray squares from Figure 4(b) correspond to subjects who do not receive information about the school share. In turn, gray squares indicate a statistically significant (p-value<0.001) but economically small (0.099) degree of spurious revision. Most importantly, the degree of true learning corresponds to the difference in slopes between subjects who are shown the feedback and subjects who are not shown the feedback. This difference is large ($0.533 = 0.632 - 0.099$) and highly statistically significant (p-value<0.001). This difference suggests that a 1 pp information shock induces a 0.533 effect in posterior beliefs. Although large, this rate of information pass-through (0.533) is quite smaller than the corresponding rate for the school share (0.757).

Many reasons help explain the weakly updated beliefs about recapture. For example, respondents may feel more confident in their prior beliefs about recapture, or they may have lower trust in the recapture feedback. Indeed, the recapture estimates that we use for the feedback are based on a number of assumptions, so subjects may naturally find the recapture feedback less persuasive. Last, subjects may pay less attention to recapture feedback due to survey fatigue, as this information appears later in the survey. The most important implication of the weaker belief updating for recapture share (relative to school share) is that it will result in less variation in posterior beliefs and, thus, less precisely estimated 2SLS coefficients. Appendix B.7 contains additional robustness checks, for example, showing that the degree of belief updating does not differ between households *with* and *without* children, and that households did not use the information on the school share to update beliefs about the recapture share.

6.3 2SLS Estimates

Let r_i^{post} be the posterior belief about the funds recaptured from individual i's school district, in percentage points, as defined in the conceptual framework from Section 3 above. Positive values indicate that individual i's district is a net contributor to the recapture system; in other words, $r_i^{post} = 10$ means that 10% of school taxes from household i's district are redistributed to disadvantaged school districts. Negative values, on the contrary, represent situations where individual i's school district benefits from recapture: $r_i^{post} = -10$ means that the school district can spend the school taxes it raises plus an additional 10% from the amount recaptured. We use the following econometric specification:

$$P_i^{2021} = \beta_0 + \beta_C^R \cdot C_i \cdot r_i^{post} + \beta_{NC}^R \cdot (1 - C_i) \cdot r_i^{post} + \beta_1 \cdot C_i + \epsilon_i \quad (23)$$

The two parameters of interest are β_C^R and β_{NC}^R for households *with* and *without* children, respectively. Prediction 2 states that $\beta_C^R > 0$ and $\beta_{NC}^S = 0$ (and, as a result, $\beta_C^R - \beta_{NC}^R > 0$).

As mentioned in Section 5.3, we estimate the effects of school share and recapture share jointly in the same 2SLS regression. Thus, we identify the effects of the recapture share using 2SLS to exploit the variation in posterior beliefs induced exogenously by the information provision experiment.

The 2SLS estimates for the recapture share are presented in panel (b) of Table 2. In column (1) of Table 2, the dependent variable indicates if the subject protests directly in 2021. The causal effects of the beliefs about the recapture share are very imprecisely estimated, so the results for this treatment arm are largely inconclusive. Consistent with the hypothesis of reciprocal motivation, the belief about recapture share does not have significant effects on the decision to file a tax appeal among households *without* children: the coefficient is positive (0.498) and borderline statistically insignificant (p-value=0.101). This finding must be taken with a grain of salt, however: since the coefficient is imprecisely estimated, we cannot rule out large effects, positive or negative.

To illustrate how imprecisely estimated this coefficient is, note that the standard error for recapture share is 135% larger than the corresponding standard error for school share (0.303 vs. 0.129). In other words, the effects for recapture share should be more than twice as high as the effects of school share to have enough power to detect statistically significant effects. The less precise estimation for the coefficients for recapture share occurs for two reasons, both of which are difficult to anticipate ex-ante in the experimental design. First, as explained in Section 6.1, the misperceptions about recapture share were smaller (mean absolute difference of 11.36 pp) than those about school share (mean absolute difference of 16.57 pp). Second, as documented in Section 6.2, conditional on a level of misperceptions, subjects updated their beliefs more strongly in response to the feedback about school share than in response to the feedback about recapture share.

We do not find evidence of significant positive effects for households *with* children. The coefficient for households *with* children is positive (0.076) but statistically insignificant (p-value=0.875). Again, this coefficient is so imprecisely estimated that it does not really constitute evidence against the hypothesis of reciprocal motivation, because we cannot rule out very large positive effects. More precisely, the 95% confidence interval cannot rule out a positive coefficient of up to 1.02, which is several times the magnitude of the effects documented for the first treatment arm. Likewise, the difference between the coefficients for households *with* versus *without* children is statistically insignificant (p-value=0.454), but it is very imprecisely estimated so we cannot rule out large differences.

The coefficients from column (2) of Table 2 show that the results for recapture share

are similar if we look at the intention to protest instead of the actual protest decision.[47] In addition, column (3) shows that effects on protests through agents are also null.[48] The coefficients from column (4) show the event-study falsification exercise: i.e., the dependent variable is whether the household protested in 2020. As expected, the estimates are close to zero and statistically insignificant.[49] Finally, Table 3 shows that the null results for recapture share hold under alternative specifications. In the same spirit as the additional tests discussed for the school share analysis, Appendix B.8 reports some additional results that show that the (lack of) effects for the recapture shre are not due to non-linearities or outliers, and compares the 2SLS estimates to the expert predictions.

7 Conclusions

Compared to abundant causal evidence on the importance of institutions for tax compliance, little causal evidence shows that tax morale is important. In this paper, we attempt to fill this gap by providing evidence from a natural field experiment. Our novel research design studies tax morale by linking data from a survey experiment to administrative tax records at the individual level. Our subjects are homeowners who pay property taxes and have the opportunity to appeal their property tax assessment. We find that even though accurate information is publicly available and easily accessible, households have large misperceptions about how tax dollars are spent. Through an information-provision experiment, we corrected misperceptions about where their tax dollars go. The effects of the information provision experiment are consistent with our hypothesis of reciprocal motivation. After learning that a higher share of property taxes funds public schools, households *with* children enrolled in public schools become less likely to appeal their property taxes and households *without* children become more likely to appeal their property taxes.

A common consideration in any empirical study is related to the external validity of the

[47] More precisely, column (2) of Table 2 is identical to column (1), except that the dependent variable is the intention to protest instead of whether the household actually files a protest. As in column (1), the estimates from column (2) are all statistically insignificant. The only coefficient from column (1) that is borderline significant, for households *without* children (p-value=0.101), is not even close to being statistically significant in column (2), and furthermore, it has the opposite sign.

[48] More precisely, column (3) of Table 2 uses the dependent variable that indicates whether the household ever protested through an agent. As expected, the coefficients are statistically insignificant (p-values of 0.249 and 0.359 for households *with* and *without* children, respectively) and the difference between the two coefficients (-0.207) is also statistically insignificant (p-value=0.486).

[49] Column (4) of Table 2 uses the protest decision in a pre-treatment year (2020) as a dependent variable. As expected, the coefficients from column (4) (0.164 and -0.039 for households *with* and *without* children, respectively) are statistically insignificant (p-values of 0.694 and 0.867); the difference between the two (0.203) is also statistically insignificant (p-value=0.664). We find similar results if we expand this falsification test to other pre-treatment years (results presented in Appendix B.8).

estimates.[50] In this regard, the sample of survey respondents who paid attention to our letter are households that might have been particularly inclined to file a property tax protest, and it is possible that these households have a higher propensity to react to our treatments than households that did not respond to our survey. Furthermore, we conducted our experiment in a context with relatively high quality public schools. However, it is possible that responses to treatment would be weaker or null in contexts where the public goods are of low quality or in contexts of high corruption. That is, if households believe that tax revenues do not turn into good services or are stolen by politicians, they may not care whether the tax dollars flow to public schools or some other service. We also focus on a single tax and, although it is the second most important tax in terms of revenue in the United States, each tax has its idiosyncrasies.[51] Indeed, this paper focuses on specific beliefs such as the share of school taxes. However, we seek to make a more general methodological contribution: our research design can be used to study other mechanisms under the umbrella of tax morale. For example, this approach could be used to assess the willingness to pay taxes in response to changes in the perceived quality of government spending or perceived corruption.

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 that those taxes 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 the pockets of taxpayers 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 some hints for policy makers looking to boost 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), but it can reduce compliance

[50] In the language of List (2020), we view our results as a wave-1 insight that establishes initial causality and produces first tests of theory.

[51] For example, based on the tax and the country or sub-national government some taxes are withheld and some are not.

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' perception that their tax dollars are well spent, they also may increase the average tax compliance.

References

Ajzenman, N. and R. Durante (2022). Salience and Accountability: School Infrastructure And Last-Minute Electoral Punishment. *Economic Journal, forthcoming*.

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.

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., M. Leites, R. Perez-Truglia, and M. Strehl (2020). What Makes a Tax Evader? *NBER Working Paper No. 28235*.

Bergolo, M. L., R. Ceni, G. Cruces, M. Giaccobasso, and R. Perez-Truglia (2021). Tax Audits as Scarecrows: Evidence from a Large-Scale Field Experiment. *American Economic Journal: Economic Policy, forthcoming*.

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. Ã. Sørensen, and B. Tungodden (2020). You've 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. *NBER Working Paper No. 27412*.

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)*.

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: Macroeconomics 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 (2020). Political alignment, attitudes toward government and tax evasion. *American Economic Journal: Economic Policy, forthcoming*.

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.

Dallas Morning News (2018). Do your schools get your property tax dollars? *July 4, 2018*.

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, forthcoming*.

DellaVigna, S. and E. Linos (2022). Rcts to scale: Comprehensive evidence from two nudge units. *Econometrica 90*(1), 81–116.

DellaVigna, S., D. Pope, and E. Vivalt (2019). Predict science to improve science. *Science 366*(6464), 428–429.

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.

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.

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*.

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. Technical report, National Bureau of Economic Research.

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.

Musgrave, R. (1959). The Theory of Public Finance. *McGraw-Hill*.

Nathan, B., R. Perez-Truglia, and A. Zentner (2020). My Taxes are Too Darn High: Why Do Households Protest their Taxes? *NBER Working Paper No. 27816*.

Nathan, B., R. Perez-Truglia, and A. Zentner (2023). Paying Your Fair Share: Perceived Fairness and Tax Compliance. *Working Paper*.

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.

Scherf, R. and M. Weinzierl (2020). Understanding Different Approaches to Benefit-Based Taxation. *Fiscal Studies 41*(2), 385–410.

Seligman, E. R. A. (1908). Progressive Taxation in Theory and Practice. Technical Report 4.

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.

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 Comptroller (2021). Frequently Asked Questions. *https://comptroller.texas.gov/taxes/property-tax/exemptions/age65older-disabled-faq.php*.

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*.

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.

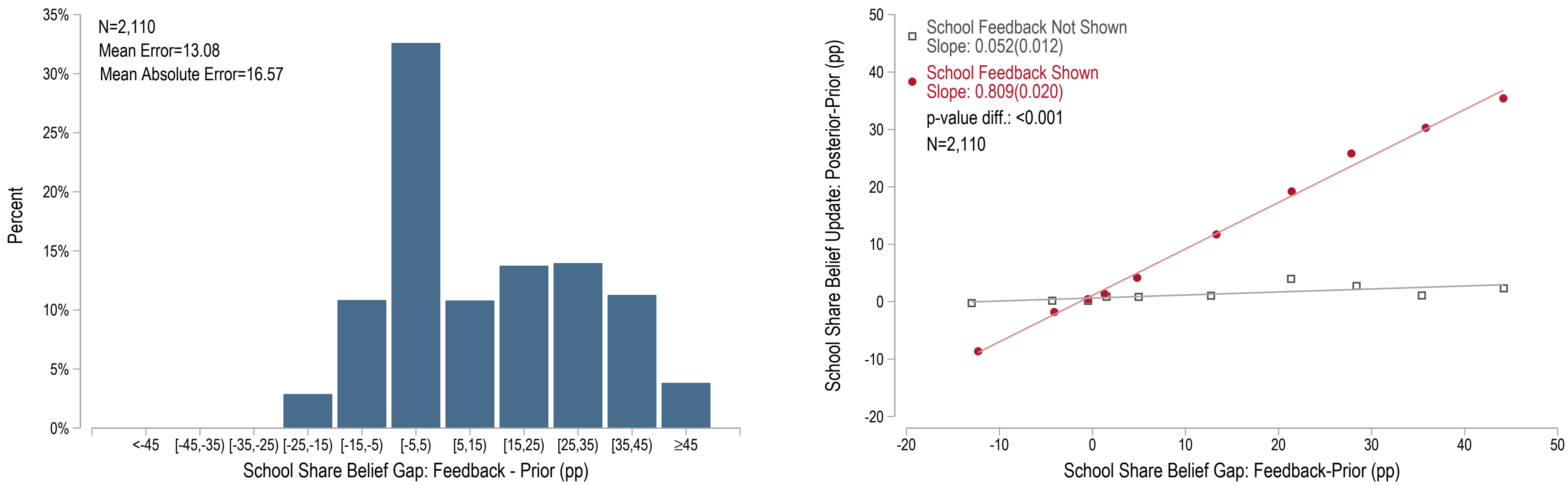
Weinzierl, M. (2018). Revisiting the Classical View of Benefit-based Taxation. *The Economic Journal 128*(612), F37–F64.

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

Youngman, J. (2016). *A Good Tax.* New York: Columbia University Press.

Figure 1: Perceptions about the Share of Property Taxes Going to Public Schools

(a) Gap in Prior Beliefs **(b)** Belief Updating

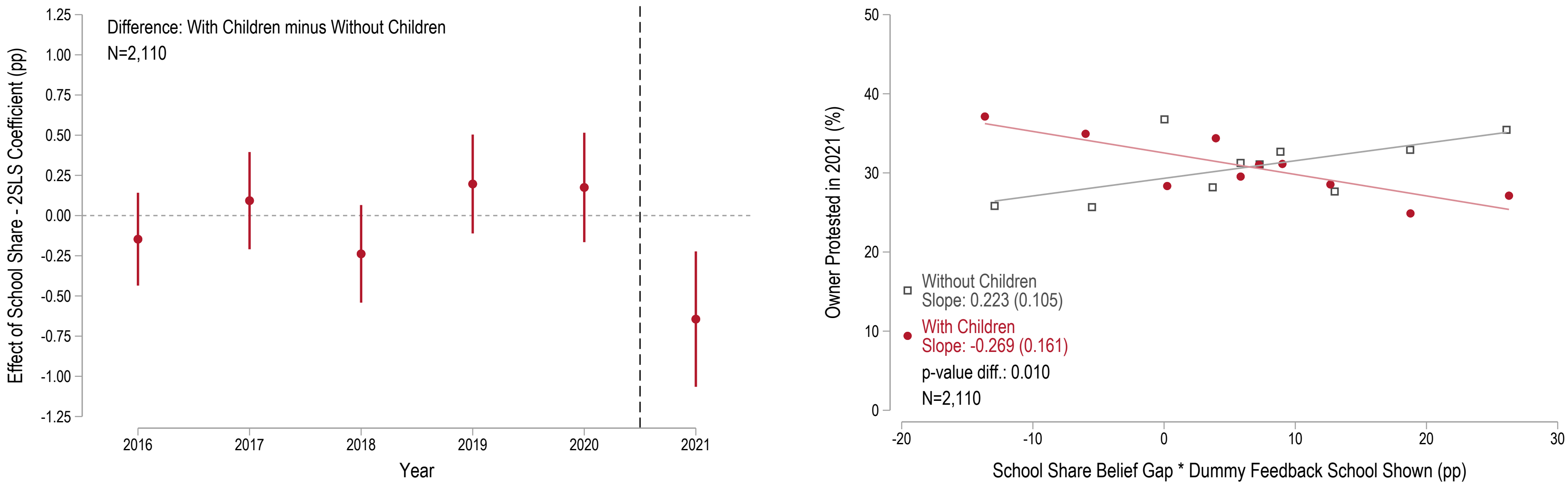


Notes: Panel (a) shows the gap in prior beliefs about the school share. The x-axis reports the difference between the actual school share and respondents' prior beliefs about the school share in 10 pp-width bins. The y-axis reports the percentage of survey respondents in each bin. The upper left corner 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 2: The Effects of School Share Perceptions on Protests: Additional Robustness Checks

(a) Event-Study Analysis

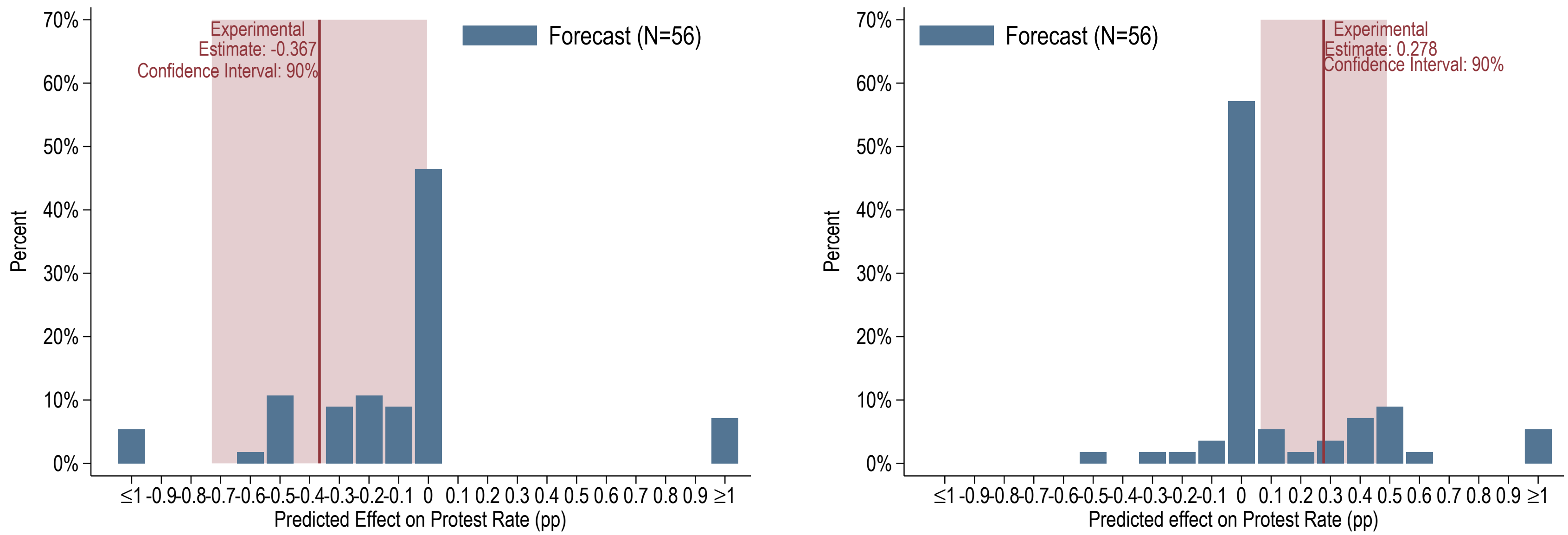
(b) Binned Scatterplot (Reduced Form)



Notes: Panel (a) reports an event-study analysis of the differential effect of school share belief on the protest probability for households *with* children versus *without* children. The estimates plotted in this figure correspond with the 2SLS point estimate based on equation (22), 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 2. 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 20 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 B.6). 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.

Figure 3: The Effects of School Share on Protests: Comparison to Expert Predictions

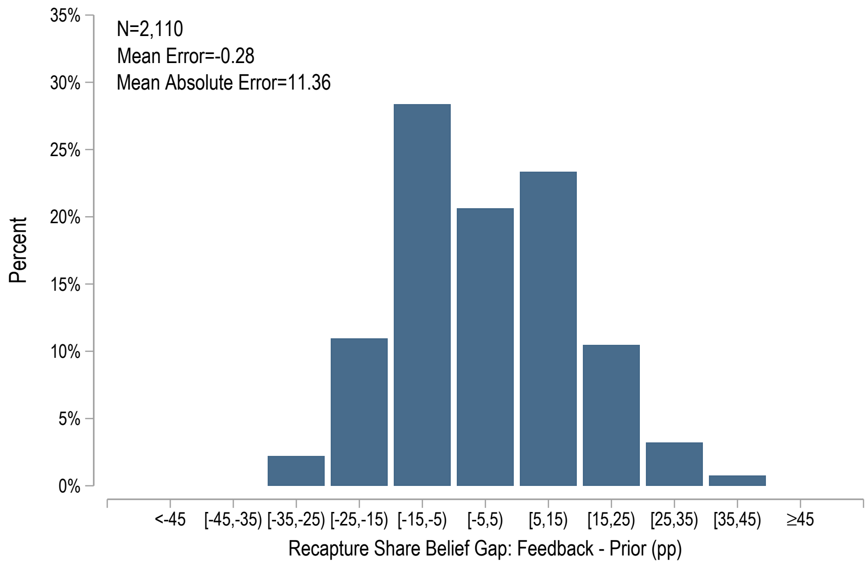
(a) With Children in Public School **(b)** Without Children in Public School



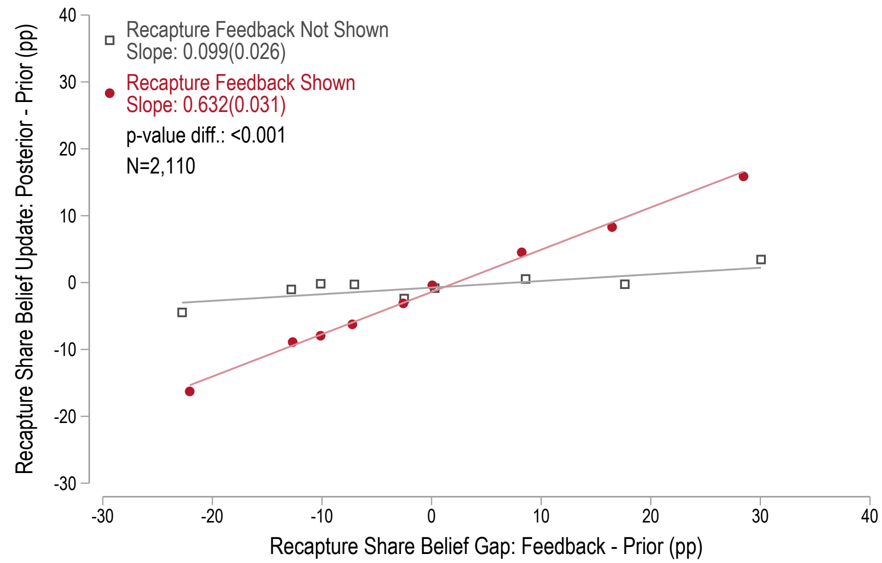
Notes: This figure shows the distribution of expert predictions about the effects of a 1 pp increase in school share beliefs on the probability that a homeowner files a protest directly for households *with* children enrolled in the public school district (panel (a)) and households *without* children enrolled in the public school district (panel (b)), based on the data collected in the forecast survey. To make the elicitation easier, in the prediction survey we asked subjects to predict the effects of a 10 pp increase in beliefs about school share. For this figure, we divide those predictions by 10 and we obtained the effect per 1 pp so these coefficients can be compared directly to the 2SLS estimates. In both panels, we pooled responses that were greater than 1 in absolute value into the corresponding extreme bins. The vertical red solid line corresponds to the experimental estimate based on the 2SLS specification reported in Table 2. The shaded area (in pink) corresponds to the 90% confidence interval. The full questionnaire for the prediction survey can be found in Appendix I.

Figure 4: Perceptions about the Share of School Taxes Affected by Recapture

(a) Gap in Prior Beliefs



(b) Belief Updating



Notes: Panel (a) shows the gap in prior beliefs about the recapture share. The x-axis reports the difference between the actual recapture share and respondents' prior beliefs about the recapture 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 recapture share and respondents' prior beliefs about the recapture 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 recapture 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 recapture 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.

Table 1: Balance of Households' Characteristics across Treatment Groups

				Treatment Arm				
	Universe (1)	Letter Sample (2)	Subject Pool (3)	No Feedback (4)	Recapture Feedback (5)	School Feedback (6)	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 (%)	7.986	8.841	18.104	23.121	15.000	19.883	14.684	0.001
	(0.043)	(0.102)	(0.838)	(1.852)	(1.538)	(1.764)	(1.527)	
2020 Agent Protested (%)	8.042	6.313	1.611	1.156	2.222	1.754	1.301	0.505
	(0.043)	(0.087)	(0.274)	(0.470)	(0.635)	(0.580)	(0.489)	
2019 Owner Protested (%)	6.085	6.625	13.507	15.029	11.111	14.230	13.755	0.268
	(0.038)	(0.089)	(0.744)	(1.570)	(1.354)	(1.544)	(1.486)	
2018 Owner Protested (%)	5.801	6.452	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 (%)	5.599	5.687	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 (%)	4.423	4.630	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 for different pre-treatment characteristics of the homeowner 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 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: 2SLS Estimates: Main Results

	P_D^{2021} (1)	I^{2021} (2)	P_A^{2021} (3)	P_D^{2020} (4)
Panel (a): Effects of School Share:				
With Children	-0.367*	-0.408*	-0.028	0.110
	(0.221)	(0.234)	(0.118)	(0.181)
Without Children	0.277**	0.269*	-0.033	-0.065
	(0.129)	(0.144)	(0.051)	(0.097)
Difference (Children - No Children)	-0.644**	-0.678**	0.006	0.175
	(0.256)	(0.275)	(0.129)	(0.207)
Panel (b): Effects of Recapture Share:				
With Children	0.076	-0.313	-0.321	0.164
	(0.485)	(0.541)	(0.278)	(0.417)
Without Children	0.498	-0.101	-0.114	-0.039
	(0.303)	(0.325)	(0.124)	(0.234)
Difference (Children - No Children)	-0.422	-0.212	-0.207	0.203
	(0.563)	(0.620)	(0.297)	(0.468)
Cragg-Donald F-Statistic	30.10	30.22	30.10	30.02
Mean Outcome (Baseline):				
With Children	33.86	47.20	7.09	25.98
Without Children	28.83	44.87	4.08	22.19
Observations	2,110	2,090	2,110	2,110

Notes: Significant at *10%, **5%, ***1%. Robust standard errors in parentheses. This table reports 2SLS estimates of equation (22) discussed in Section 5.3. Panel (a) reports the estimates 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. Panel (b) reports analogous results but for the recapture share treatment effects. 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 subject's protest likelihood in 2021 ("Do you intend to protest this year?") The dependent variable in column (3) is an indicator variable that takes the value 100 if the subject used an agent to protest in 2021. The dependent variable in column (4) is an indicator variable that takes the value 100 if the subject protested directly in 2020. Mean outcomes at baseline correspond with the mean of the dependent variables computed using the group of subjects who did not receive feedback about the school share nor recapture share (the control group).

Table 3: 2SLS Estimates: Robustness Checks

	P_D^{2021} (1)	I^{2021} (2)	P_D^{2021} (3)	I^{2021} (4)	P_D^{2021} (5)	I^{2021} (6)	P_D^{2021} (7)	I^{2021} (8)	P_D^{2021} (9)	I^{2021} (10)	P_D^{2021} (11)	I^{2021} (12)
Panel (a): Effects of School Share												
With Children	-0.367*	-0.408*	-0.429*	-0.457*	-0.299	-0.322	-0.330*	-0.250	-0.226	-0.088	-0.369	-0.418*
	(0.221)	(0.234)	(0.225)	(0.235)	(0.232)	(0.246)	(0.190)	(0.205)	(0.168)	(0.191)	(0.237)	(0.247)
Without Children	0.277**	0.269*	0.285**	0.286**	0.292**	0.299**	0.196*	0.321**	0.197*	0.256**	0.301**	0.324**
	(0.129)	(0.144)	(0.133)	(0.146)	(0.136)	(0.147)	(0.119)	(0.132)	(0.116)	(0.130)	(0.139)	(0.153)
Difference (Children - No Children)	-0.644**	-0.678**	-0.714***	-0.744***	-0.591**	-0.622**	-0.525**	-0.571**	-0.423**	-0.344	-0.671**	-0.743**
	(0.256)	(0.275)	(0.262)	(0.278)	(0.269)	(0.286)	(0.224)	(0.244)	(0.203)	(0.231)	(0.274)	(0.290)
Panel (b): Effects of Recapture Share												
With Children	0.076	-0.313	0.141	-0.222	-0.018	-0.382	0.166	0.135	0.065	0.013	0.231	-0.059
	(0.485)	(0.541)	(0.478)	(0.536)	(0.541)	(0.593)	(0.417)	(0.451)	(0.330)	(0.373)	(0.442)	(0.492)
Without Children	0.498	-0.101	0.436	-0.125	0.275	-0.249	0.414	-0.129	0.247	-0.051	0.473	-0.051
	(0.303)	(0.325)	(0.307)	(0.325)	(0.287)	(0.303)	(0.273)	(0.291)	(0.243)	(0.265)	(0.318)	(0.338)
Difference (Children - No Children)	-0.422	-0.212	-0.295	-0.098	-0.293	-0.133	-0.248	0.264	-0.182	0.063	-0.242	-0.009
	(0.563)	(0.620)	(0.559)	(0.616)	(0.613)	(0.666)	(0.500)	(0.536)	(0.394)	(0.438)	(0.527)	(0.579)
Cragg-Donald F-Statistic	30.10	30.22	29.68	29.68	21.50	21.44	35.26	35.34	47.35	47.55	34.35	34.35
Mean Outcome (Baseline):												
With Children	33.86	47.20	34.68	47.58	33.86	47.20	35.00	47.10	33.11	47.95	36.27	50.00
Without Children	28.83	44.87	29.12	45.10	28.83	44.87	29.77	45.64	29.53	46.09	29.33	44.28
Observations	2,110	2,090	2,070	2,070	2,110	2,090	2,335	2,309	2,482	2,454	1,807	1,807
Baseline Controls	√	√	√	√			√	√	√	√	√	√
Additional Controls			√	√								
5% Outliers	√	√	√	√	√	√					√	√
2.5% Outliers							√	√				
1% Outliers									√	√		
Attention Check											√	√

Notes: Significant at *10%, **5%, ***1%. Robust standard errors in parentheses. This table reports 2SLS estimates of equation (22) discussed in Section 5.5. Panel (a) reports the 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. Panel (b) reports analogous results but for the recapture share treatment effects. Columns (1) and (2) correspond to our preferred specification reported in columns (1) and (2) in Table 2 (for reference). The rest of the columns in this table use the same dependent variables from columns (1) and (2). Columns (3) and (4) add additional control variables collected in the survey: age, gender, college degree, and political party. Columns (5) and (6) report estimates where no control variables are included at all. Columns (7) and (8) drop 2.5% of the outliers at each tail of the distribution (instead of the 5% used in the baseline specification). Columns (9) and (10) drop 1% of the outliers at each tail. Columns (11) and (12) restrict the samples to subjects who passed the attention check included in the questionnaire (see Appendix H for the survey). Each mean outcome corresponds with the mean of the dependent variable for subjects who did not receive feedback about the school share nor the recapture share (the control group).