YELLOW VESTS, PESSIMISTIC BELIEFS, AND CARBON TAX AVERSION*

Thomas Douenne and Adrien Fabre†

Abstract

Using a representative survey, we find that after the Yellow Vests movement, French people would largely reject a tax & dividend policy, i.e., a carbon tax whose revenues are redistributed uniformly to each adult. They overestimate their net monetary losses, wrongly think that the policy is regressive, and do not perceive it as environmentally effective. We show that changing people's beliefs can substantially increase support. Although significant, the effects of our informational treatments on beliefs are small. Indeed, the respondents that oppose the tax tend to discard positive information about it, which is consistent with distrust, uncertainty, or motivated reasoning.

JEL classification: D72; D91; H23; H31; Q58

Keywords: Climate policy; Carbon tax; Beliefs; Preferences; Tax aversion

*Acknowledgments: We are grateful to Mouez Fodha, Fanny Henriet and Katheline Schubert for their comments and their help with securing funding. We also thank Stefano Carattini, Linus Mattauch, Joseph Stiglitz, and Thierry Verdier; the seminar participants at the Paris School of Economics, Columbia University, the University of Pennsylvania, the University of Amsterdam, the University of Quebec at Montréal, KU Leuven, Erasmus University Rotterdam, Environmental Defense Fund, EIEE-CMCC (Milan), OECD, and OFCE; and conference participants at FSR Climate Annual Conference (Florence), ADRES (Lyon), and FAERE (Rennes). We are thankful to Christina Hobbs, Paul-Hervé Tamokoué Kamga and Adrien Montalbo for the proofreading. We are grateful to the editor Erzo Luttmer and three anonymous referees for their insightful comments and suggestions. We acknowledge financial support from the Cepremap, EUR PGSE (ANR-17-EURE-0001) ANR (ANR16-CE03-0011), and Université Paris 1 Panthéon-Sorbonne economics doctoral school (ED 465).

†Douenne: Paris School of Economics, Université Paris 1 Panthéon-Sorbonne, 48 Boulevard Jourdan, 75014, Paris, France (email: thomas.douenne@psemail.eu); Fabre: Paris School of Economics, Université Paris 1 Panthéon-Sorbonne, 48 Boulevard Jourdan, 75014, Paris, France (email: adrien.fabre@psemail.eu)

1 Introduction

The French government initially committed to an ambitious trajectory for the price of carbon.[1] Initiated in 2014 at 7€/tCO_2, the French carbon tax reached 44.6€/tCO_2 in 2018 and was supposed to continue growing to reach 86.2€/tCO_2 by 2022. However, at the end of 2018, the same government that had accelerated the price trajectory decided to abandon it and froze the tax at its current level for an undetermined period of time. This turnaround in French climate policy is the direct consequence of the popular protest by the "Yellow Vests", which started in opposition to the carbon tax.[2] Among several factors, the negative impact of the tax on households' purchasing power has certainly been a key driver of public discontent. The increasing revenues from the carbon tax were mostly used to fund the budget rather than redistributed to households, raising concerns over the distributive effects of the policy. To tackle the negative impact of carbon taxation on households' purchasing power, economists have proposed a scheme known as "tax & dividend", i.e., a carbon tax whose revenue is redistributed uniformly to each adult. 3,354 American economists recently supported this strategy in The Wall Street Journal "To maximize the fairness and political viability of a rising carbon tax". Therefore, an implicit assumption is that with a design that ensures that the properties of the tax are aligned with people's *preferences*, one should be able to generate support for it. However, is this truly sufficient? In this paper, we show that to understand the link between the properties of a policy and its support, one has to account for a critical ingredient: *beliefs*.

The objective of this paper is to understand how the beliefs regarding a policy form and then determine the attitudes towards it. Recent events undoubtedly make the French carbon tax an interesting case study. To explain French attitudes towards carbon taxation, we surveyed a representative sample of 3,002 French households. We focus on a carbon "tax & dividend" policy, i.e., an increase in the carbon tax by 50€/tCO_2, the revenue of which is

[1] Specifically, the "Contribution Climat-Énergie" is a *sectoral* carbon tax specific to fossil fuels.

[2] Following a massive petition against rising gasoline prices in November 2018, hundreds of thousands of people started protesting. They wore recognizable fluorescent clothing, gathered at roundabouts and toll booths every day and demonstrated in Paris each Saturday. The Yellow Vests expressed a general concern over their purchasing power and discontent with French elites and institutions.

redistributed uniformly to each adult. Our reform allows one to clearly specify the distributive effects of the policy, in contrast to the one abandoned by the government. Only 10% of respondents approve of the reform, and 70% disapprove of it (the rest do not know or do not want to answer). We analyze the perceptions of three well-known determinants of the acceptance of a carbon tax: the impact on one's purchasing power, the progressivity of the scheme, and the scheme's environmental effectiveness. We compare the subjective beliefs regarding the impacts on one's purchasing power to the objective distribution computed using official household survey data. This comparison shows that people largely overestimate the incidence of the tax on their household. For instance, while 70% of households are expected to benefit from this reform, only 14% think that they would. Similarly, while the scheme proposed in our survey is progressive, a large majority of individuals perceive it as regressive. In addition, a majority of respondents do not believe that such a policy would reduce pollution and combat climate change. Using information reported on their energy equipment and usage, we are able to compute a respondent-specific estimate of the tax incidence on their purchasing power. This estimation enables us to examine the heterogeneity in what we call *biases* about the perceived tax incidence. We find that the people most opposed to the policy, and in particular those that support the Yellow Vests, are the most biased, i.e., the most inclined to overestimate their losses. Thus, one may wonder whether pessimistic beliefs lead to policy rejection or the causality runs in the opposite direction.

To disentangle the effect of initial beliefs on attitudes towards the policy from the reverse effect of attitudes on perceptions, we investigate the effect of providing new information to respondents through random treatments. Respondents randomly receive (or do not receive) a piece of information about the progressivity and/or the effectiveness of the policy as well as customized information—derived from our respondent-specific estimation—on whether their household is expected to win or lose from the policy. We also specify that this latter information is correct in five cases out of six, a probability that we carefully estimated out-of-sample. Our first observation is that our treatments generally fail to change pessimistic beliefs. For example, among those who would benefit from the reform despite pessimistically believing that they would lose, only 12% are convinced that they would gain when we

disclose our estimation to them. Even worse, respondents revise their beliefs in an asymmetric way, giving more weight to new information when it shows that they would lose from the reform, i.e., when it provides them with arguments against the tax. We also find evidence consistent with motivated reasoning[3] in the formation of beliefs since those who already approved of the reform are more likely to correctly revise their beliefs, while those most opposed to it such as supporters of the Yellow Vests tend to discard new information unless it goes against the tax. Moreover, we find that this phenomenon is accentuated among highly educated people, suggesting that it stems from an adaptive advantage rather than a cognitive deficiency.

We use the randomly displayed information as instruments to estimate the causal effect of holding certain beliefs (measured as binary variables) on policy support. In the case of self-interest (taken as one's beliefs about winning or losing purchasing power from the policy), we supplement these treatments by testing the support for a different policy, a tax & *targeted* dividend; people become eligible for this dividend when their incomes are below a certain threshold, and we vary the threshold across respondents to create exogenous variations in eligibility. The method we use in this case is noteworthy since it creates random variation in the beliefs of winning around the eligibility thresholds and enables us to estimate the causal effect of this belief using a fuzzy regression discontinuity design (RDD). Our results indicate that convincing people of the actual incidence and effectiveness of the policy could lead to majority support. Indeed, we find that self-interest has a large effect on support for the policy: the belief that one does not lose from it increases the acceptance rate by more than 50 p.p. Similarly, believing that the tax is environmentally effective increases the approval rate of the reform by more than 40 p.p. We also provide non-causal evidence that believing in the progressivity of the scheme has a large effect on support. Overall, these results suggest that the rejection of carbon taxation does not typically result from clashing principles, such as a disinterest in the climate or a dislike of price instruments but rather from overly pessimistic beliefs about the properties of the reform. Provided that people's opposition to the policy

[3]Motivated reasoning is the "tendency to find arguments in favor of conclusions we want to believe to be stronger than arguments for conclusions we do not want to believe" (Kunda, 1990). This psychological mechanism can be a driver of the well-known "confirmation bias"

reinforces their pessimism—which our results suggest—their biases are reinforced such that new information might only furtherpush their attitude in one direction.[4]

The contribution of this paper is twofold. First, it contributes to a recent literature that has emerged to understand the political economy of climate policies since this issue is becoming critical in the public debate. For a thorough review of this literature, we refer the reader to Carattini et al. (2018) and suggest the more synthetic Klenert et al. (2018), as well as Millner & Ollivier (2016) for a review of the political obstacles to environmental policies. Stern et al. (1993) is an early work proposing and testing a model of attitudes on environmental quality intended to disentangle egoistic from altruistic motives on the one hand and beliefs from values on the other hand. Among all possible attitudes, they show that beliefs about consequences for self-interest are the only predictor of the willingness to pay Pigouvian taxes. Using a postelectoral survey in Switzerland, Thalmann (2004) also finds a correlation between carbon tax acceptance and self-interest, proxied by the number of cars owned. In surveys on British, Swedish, and Swiss respondents, Bristow et al. (2010), Brannlund & Persson (2012), and Carattini et al. (2017), respectively, document a higher approval rate when the reform addresses distributional issues. Baranzini & Carattini (2017) report that a majority of the people they interviewed in Geneva do not believe that the tax would be effective, which confirms what Dresner et al. (2006b) find with focus groups in the UK. Surveying Norwegian people, Kallbekken & Sælen (2011) show that self-interest matters for acceptance but less than concerns for environmental effectiveness or distributional effects. Using US data, Anderson et al. (2019) argue that ideology explains most of the support for carbon taxation and suggest that this effect would dominate that of self-interest.

In the present paper, we also study how acceptance depends on these three motives (i.e., self-interest, perceived environmental effectiveness and progressivity). We contribute to the literature by providing robust evidence for causal effects where past studies essentially show

[4]The "campaign effect" documented by Anderson et al. (2019) (in the case of referenda in Washington state in the US) is an example of how support for a carbon tax can decrease substantially after it enters the public debate. This may explain why the acceptance of an increase in the carbon tax plummeted with the Yellow Vests movement, down from a level of 48% (ADEME, 2018) in the middle of the range of that in other countries (Brechin, 2010). This effect confirms that the French carbon tax may be an insightful case study to understand what could happen in other countries when a controversial policy is publicly debated.

correlations, often relying on proxies such as fuel consumption to proxy for self-interest (e.g., Thalmann, 2004; Kallbekken & Sælen, 2011; Anderson et al., 2019). In contrast, we neither assume that people are fully rational nor that they have perfect information. Thus, our methodology offers a novel examination of the political economy of climate policies since it allows one to disentangle erroneous *beliefs* from the pure effects of *preferences*.[5] The paper also quantifies biases regarding the costs of the carbon tax. To the best of our knowledge, this is the first study that compares subjective beliefs and objective data about the private costs that arise from carbon taxation. Given the intense public debate over the incidence of such a policy, identifying and measuring the discrepancy between actual impacts and their subjective perception is critical.

Beyond the case of carbon pricing, our paper contributes to the literature on the formation of political beliefs. Recent research has shown how beliefs on inequality and social mobility affect people's attitudes regarding distributive policies (e.g., Cruces et al., 2013; Kuziemko et al., 2015; Alesina et al., 2018). Our paper expands this literature by investigating the relationship between beliefs and attitudes on climate policies. Using a representative survey, our paper shows how beliefs towards a policy causally impact attitudes towards it. It also shows that people's attitudes towards a policy can be linked to how they process new infomation about it. Among other possible mechanisms, our results are consistent with theories of motivated reasoning (Kunda (1990); see Bénabou & Tirole (2016) for a recent review) that have thus far been mostly tested in a lab (e.g., Redlawsk, 2002; Thaler, 2019). In our context, motivated reasoning could be a manifestation of *tax aversion*, which we can define as a "gut" rejection of a tax (or taxation in general) that influences beliefs about the properties of a tax such as its effectiveness, fairness, or equivalence to a measure labeled differently. This would explain how popular protests against the carbon tax affected people's beliefs on related policies like the tax & dividend studied in this paper.

The remainder of the paper is organized as follows. In Section 2, we describe our survey and other data sources. In Section 3, we compare subjective perceptions to objective data and

[5]We take preferences over policies as the mapping from beliefs (on facts) to attitudes (on policies), i.e., how attitudes are determined as a function of beliefs. Conversely, motivated reasoning represents the feedback loop from attitudes to beliefs.

measure the bias regarding the impacts of carbon taxation. In Section 4, we study the formation of beliefs and propose several mechanisms to rationalize people's pessimism. In Section 5, we estimate the effects of changing people's beliefs about tax incidence and effectiveness on acceptance. Section 6 concludes. Further results and methodological complements are reported in the Appendix and in an online Appendix.

2 Context, survey, and data

2.1 Context of the study

The Yellow Vests constitute a singular protest movement. Although they are overrepresented within the far left and right, they are supported by a large fraction of the French from across the political spectrum.[6] Thousands of small-scale protests were organized autonomously on social networks, and the movement was remarkably independent from political parties and unions. Before the emergence of the movement, none of the major political parties campaigned against the carbon tax, and this policy did not trigger specific opposition until the increase in oil prices brought it to the forefront of the debate.[7] The opposition then quickly gained ground, notably through Facebook, where a petition against the tax and a call to protest on roundabouts were widely spread. These protests initially occurred every day and did not start decreasing until December 2018, when the government responded with a set of measures including supplements to low wages and modest pensions and the abandonment of the carbon tax increases initially scheduled. The fading movement came to an almost complete halt at the end of April 2019 when the government agreed to some of the demands for greater purchasing power and direct democracy (Boyer et al., 2020).

A simple interpretation of these protests could be that French people are far more con-

[6] Table H.1 in the online Appendix H provides data on our respondents' attitude towards the Yellow Vests depending on their sociodemographic characteristics and left-right leaning. This shows that the support for the movement is widespread. People at the center of the political spectrum are the least supportive with 46% having a positive attitude towards the Yellow Vests vs. 66% for the whole population.

[7] Fuel prices peaked in October 2018. The movement gained momentum at that time, leading to the first massive protest on November 17.

cerned with their purchasing power than climate change. However, our companion paper documents that a large majority of French people are aware of and concerned about climate change and supportive of various climate policies, such as a tax on air travel, green investments or stricter pollution norms, (Douenne & Fabre, 2020),[8] and our survey suggests that French people's willingness-to-pay for a carbon tax is similar to that of other countries (see online Appendix L). Instead, French people may simply not regard a carbon tax as the appropriate policy to address climate change. Thus, the present paper sheds light on people's beliefs about the carbon tax, how these beliefs form and how they affect policy support.

2.2 Our survey

2.2.1 Survey data collection

We conduced the survey in February and March 2019, three months after the government decided to abandon the planned increase in the carbon tax. The 3,002 responses were collected through the survey company Bilendi. This company maintains a panel of French respondents whom they can email with survey links. The respondents received 3€ if they fully completed the survey. The respondents who choose to respond are first channeled through screening questions that ensure that the final sample is representative in terms of six sociodemographic characteristics: gender, age (5 brackets), education (4), socioprofessional category (8), town size (5) and region (9). We relaxed the quotas by 5% to 10% relative to the actual proportions to facilitate the sampling process. Table A.1 in the Appendix A shows that our sample is still extremely representative. Nonetheless, observations are weighted to correct for small differences between the sample and population frequencies (e.g., in education). The median time to complete the survey was 19 minutes. We ensured that all questions requiring some concentration were in the first half of the survey. We took several steps to ensure the best possible data quality. Our representative sample was obtained after excluding the inattentive and quickest respondents. We confirm in the online Appendix M that this

[8]The levels of awareness and concern are similar to those of other countries (Stokes et al., 2015). For instance, 72% know that climate change is anthropogenic compared to 66% in the US (Gallup, 2019).

sampling restriction does not affect the main results.

2.2.2 The survey

The full survey in French can be found online,[9], and we provide the translated questionnaire in the online Appendix G. It contains several random branches and treatments that are independent of one another. Figure 2.1 shows the sequence of information or treatments (represented by ellipses) and questions (boxes). This section presents each part of the survey in turn.

Information intervention on environmental issues The survey opens with a brief presentation: three short sentences to welcome the participant, introduce ourselves as "two social science researchers", and explain that the survey will last 15 to 20 minutes. We then randomly display two blocks of information: one on climate change and another on particulate matter (i.e., air pollution). This informational treatment divides the sample into four groups who receive either one block of information, the other block of information, no block of information, or both blocks of information. The objective of these treatments is to see whether providing salient information on the consequences of climate change or air pollution affects respondents' answers later in the survey. The climate change information includes temperature trends for the long-run future, concerning facts on current and expected impacts, and a claim that keeping global warming below 2 °C is technically feasible. The particulate information consists of the estimated impact on French mortality (48,000 deaths per year) and life expectancy (reduced by 9 months on average in France) and the assertion that reducing fuel consumption would improve health. We save the time spent on each block and display links to scientific references to support the information.

Household characteristics In addition to the six quota strata, the sociodemographic characteristics include zip code, household structure, and the income of the respondent and of the household. A block on energy characteristics contains questions that allow us to estimate the

[9] preferences-pol.fr/doc_q.php#_e

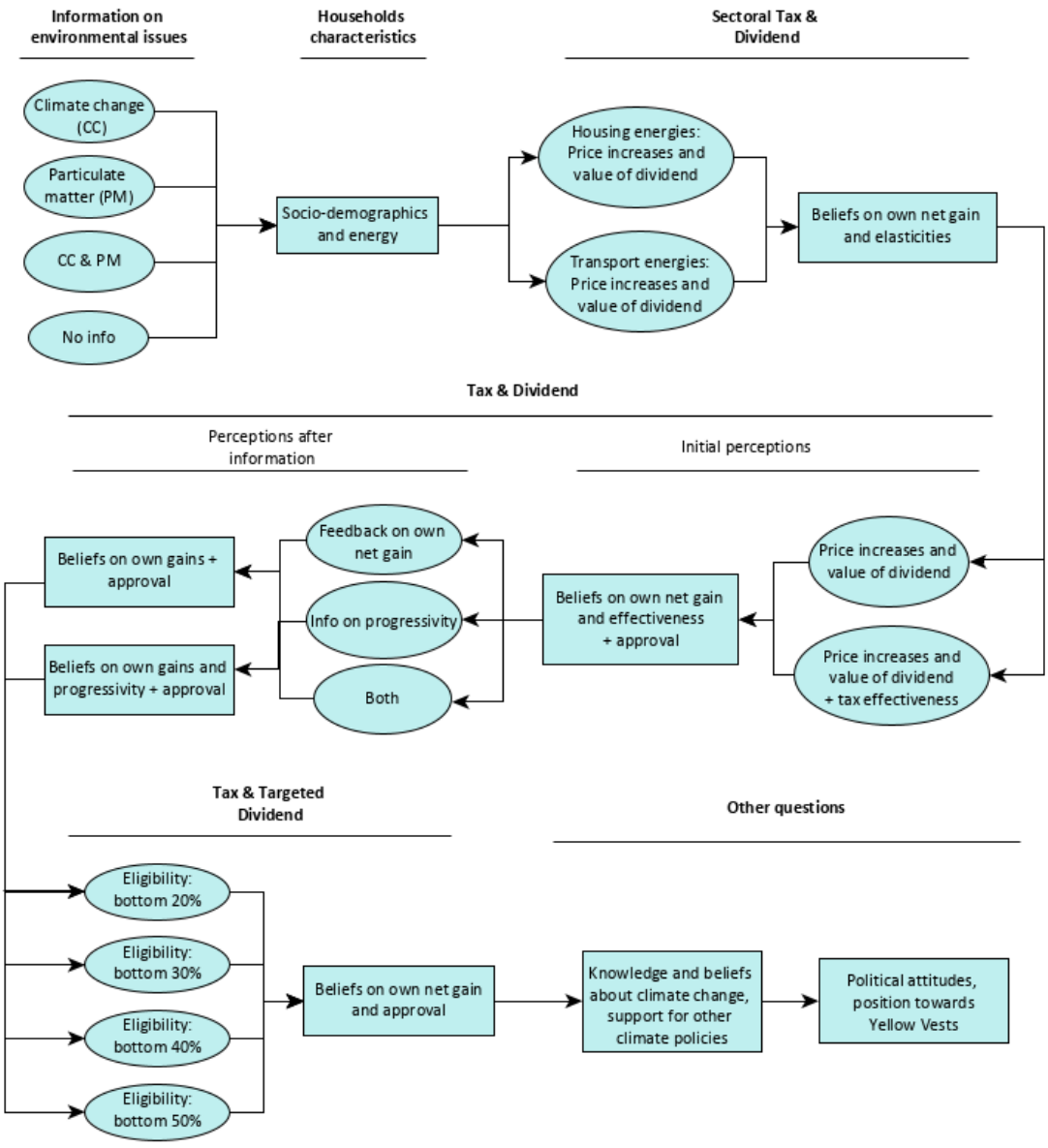


Figure 2.1: Sequence of information or treatments (ellipses) and questions (boxes).

NOTE: The sequence of informative treatments and of questions on beliefs and support for different tax & dividend policies is informative of how beliefs are revised in view of new information and allows us to estimate the causal effects of these beliefs on policy support.

impact of a carbon tax increase on housing expenditures (energy source and home size) and on transport expenditures (number of vehicles, type(s) of fuel, distance traveled last year, and average fuel economy). The distributions of the answers are in line with official statistics, as

shown in Table A.2 in the Appendix A.

Sectoral tax & dividend We first randomly allocate the respondent to one of the two sectors to which the French carbon tax applies: housing *or* transport. We present a specific policy to them: a sectoral tax & dividend, i.e., an increase in housing *or* transport energy taxes that would finance a lump-sum transfer to all adults.[10] We detail the price increases that would follow and the value of the dividend they would receive: for the housing energy tax, $+13\%$ for gas and $+15\%$ for heating oil together with a yearly transfer of 50€ per adult; and for the transport energy tax, $+0.11$€ per liter of gasoline and $+0.13$€/L for diesel with a yearly transfer of 60€ per adult. These figures are equivalent to an increase in the carbon price on these sources of energy by 50€/tCO_2, but we do not mention the name "carbon tax" at this stage since we do not want people to think that it also falls on the other sector. The value of the dividends was obtained such that the policy is budget neutral and assuming typical price elasticities (see Section 2.3.1). We present the policy starting with "The government studies..." to capture the effect of distrust in the government that could arise in the actual political process.

Then, we ask the respondent whether their household would win, lose, or be unaffected by the reform in terms of purchasing power (*win/lose category* thereafter). Depending on their answer, we further ask respondents to estimate their expected gain (or loss) among 5 (or 6) intervals. The interval thresholds are tailored to each respondent since they are computed in proportion to the number of consumption units (c.u.) of the household (as defined by Eurostat).[11] Similarly, households' gains and losses are always expressed per consumption unit in the analysis. To obtain accurate answers, we did not incentivize the questions by using monetary rewards. Indeed, Sapienza & Zingales (2013) show that people think that economic experts are too optimistic regarding the carbon tax, and so incentivizing the answers could have led respondents to misreport their true beliefs and shift them towards what they think

[10] We chose to redistribute the money per adult instead of per consumption unit to make the scheme more understandable. We limited the number of beneficiaries to two per household to better align this policy with current welfare benefits that depend on the number of consumption units.

[11] For instance, for a single-member household (c.u.=1), the intervals of the expected gain (in €/year) are $(0, 10)$, $(10, 20)$, etc.; and for a childless couple (c.u.=1.5), these intervals are $(0, 15)$, $(15, 30)$, etc.

the researchers expect. Finally, to see whether people think that the tax achieves its purpose of incentiving people, we ask respondents to estimate their own elasticity and that of French people in general. To this end, we borrow the phrasing of Baranzini & Carattini (2017) and ask for the expected decrease in consumption (in 5 brackets) that would follow a 30% increase in the price of heating (or equivalently, an increase of 0.50€/L in fuel prices).

Tax & dividend

Initial perceptions Our main reform of interest is an increase of 50€/tCO_2 in the French carbon tax, which concerns *both* housing and transport.[12] The revenues generated are again redistributed equally so that each adult receives a yearly lump-sum compensation of 110€. We now explicitly present the reform as an increase in the carbon tax, although, as before, we do not give the implicit carbon price but rather the effect on energy prices (the same as before but on both sectors) and the value of the dividend.[13] We randomly display a new informational treatment stating that "scientists agree that a carbon tax would be effective in reducing pollution", thereby splitting the sample in two. After describing the reform and providing (or not providing) this information, a first block of questions elicits the respondent's perceptions. We ask about their win/lose category and their subjective net gain in purchasing power in the same manner as for the sectoral tax with adapted intervals. We then ask them whether the reform would be effective in reducing pollution and combating climate change. Finally, we ask, "Would you approve of this reform?" and let the respondent choose between "Yes", "No" and "PNR (I don't know, I don't want to answer)".[14] In the following, we say that respondents *approve* of a reform if they respond "Yes" and *accept* the reform if they do *not* respond "No". Table I.1 in the online Appendix I describes the rates of support for the tax & dividend policies at different stages of the survey.

[12] Electricity and industries are exempt from the French carbon tax because they are already covered by the EU-ETS.

[13] For the exact phrasing, see question 35 in the online Appendix G.

[14] In English, "PNR" stands for "Prefer Not to Respond".

Perceptions *after information* To assess how people form their beliefs and measure the importance of self-interest and fairness motives in the acceptance of the reform, we then provide some information on the effect of the reform. To half of the sample selected at random, we provide customized information explaining the following: "In five cases out of six, a household with your characteristics would [win/lose] through the reform. (The characteristics taken into account are: heating using [energy source] for an accommodation of [surface] m^2; [distance] km travelled with an average consumption of [fuel economy] L for 100 km.)". In Section 2.3.2, which details how we compute each respondent's net gain, we show that our prediction that a household wins or loses is correct in 83% of cases, which is our "five cases out of six". To a third of the sample selected at random, we explain that "this reform would increase the purchasing power of the poorest households and decrease that of the richest, who consume more energy". To the remaining sixth of the sample, we provide both the customized feedback on their net gain and the information on progressivity.[15]

Then, we again ask about the win/lose category (i.e., if the respondent's household would win, lose or be unaffected by the reform) and for the approval of the reform. We also ask respondents about the perceived advantages and disadvantages of the policy, including the effect on the poorest households. To the latter half of the sample, immediately after the treatment on progressivity, we explicitly ask whether they think that the reform would benefit the poorest since most respondents appeared not to believe our information.

Tax & targeted dividend To disentangle the effect of self-interest from other motives for acceptance in Section 5, we then propose to respondents an alternative reform where only some people are eligible for the dividend. Specifically, we propose one of four alternative reforms where the payments, which are still equal among recipients, are targeted to adults whose income is below some threshold. The four possible thresholds correspond to the 20th, 30th, 40th, and 50th percentiles of the income distribution. They are computed using inflated deciles of individual income from the *Enquête sur les Revenus Socio-Fiscaux* (ERFS

[15] In total, one-half of the respondents are treated with the information on progressivity. For the customized feedback, we choose to treat two-thirds of the sample to increase our statistical power on the identification strategy that exploits a subsample of those who received the feedback (see Section 5.1).

2014) produced by Insee (the French national statistics bureau).[16] We randomly allocate respondents whose income lies between two thresholds to the reform defined by one of these thresholds. For example, a person at the 25th percentile of the income distribution has a one-in-two chance to face the reform targeted to the bottom 30%, where they are eligible for the dividend; and a one-in-two chance to face the reform targeted to the bottom 20%, where they are not. When the income is close to only one threshold (i.e., when its percentile in the distribution is below 20 or within $[50;70]$), the allocated reform corresponds to that one. When the respondent's income is distant from all thresholds, i.e., when it is in the *top* 30% (above 2220€/month), the reform they face is determined by the income of the household's second adult. Finally, when both adults (or the only adult) in the household are in the top 30%, their reform is allocated randomly from the four variants. Table 2.1 details the income thresholds and dividends of the four variants and the proportion of respondents allocated to each of them, along with the proportion one would expect from the official data. The two sets of figures match almost perfectly, indicating that our sample is representative according to income.

Table 2.1: Characteristic of the targeted reform by the payment target.

Targeted percentiles	≤ 20	≤ 30	≤ 40	≤ 50
Income threshold (€/month)	780	1140	1430	1670
Payment to recipients (€/year)	550	360	270	220
Proportion of respondents	.356	.152	.163	.329
Expected proportion of respondents	*.349*	*.156*	*.156*	*.339*

NOTE: This table reads as follows: when targeted people are those below the 20th percentile (≤ 20), all adults with an income below 780€/month receive a dividend of 550€/year; 0.356 of our respondents are assigned to this policy (for which they may or may not be eligible depending on their income) compared to 0.349 if our survey were *exactly* representative of the true income distribution of the French population.

We describe to each respondent the variant they face: the price increases, the income threshold and the value of the dividend. We also specify how many persons in their household would be eligible for the payment. Finally, we ask respondents again for their an-

[16] Income sources that are entitlements for the household rather than its members, such as certain welfare benefits, are divided equally among the two oldest adults in the household.

ticipated win/lose category and their approval. The random variation in eligibility creates exogenous variation in the win/lose belief, which we use to estimate its causal effect on acceptance in a fuzzy RDD.

Other questions We do not detail the other questions of the survey because we devote a companion paper to their analysis, Douenne & Fabre (2020). In these questions, we examine the opinions on environmental policies, including other ways to use the revenues from a carbon tax. We measure the knowledge and perceptions of climate change and ask specific questions on the influence of climate change on the choice to give birth and one's willingness to change one's lifestyle. We study the use, availability of, and satisfaction with public transportation and active mobility. We also ask questions regarding political preferences, including attitude toward the Yellow Vests. Finally, we let the respondent express any comments in a text box.

2.3 Official household surveys

In addition to our survey, we use three official household surveys produced by Insee[17] for two purposes. First, we use the consumer and the transport surveys to compute the objective distribution of the increases in fuel expenditures and the revenue from the tax (and hence the value of the dividend). Second, we use the housing survey to compute a respondent-specific estimate of their objective net gain.

2.3.1 Eliciting objective aggregates and distributions

We use the database constructed by Douenne (2020) that matches the consumer and transport surveys. It includes over 10,000 households for whom it provides information on all their revenues and expenditures—including their energy bills—together with many sociodemo-

[17]The additional datasets are the consumer survey *Budget de Famille* (BdF 2011), the transport survey *Enquête Nationale Transports et Déplacements* (ENTD 2008) and the housing survey *Enquête Logement* (EL 2013). We use data from National Accounts to homogeneously inflate households' sectoral expenditures from each dataset we use to make them representative of the most recent trend and comparable across datasets. For more information about these surveys, see the Appendix C.1.

graphic characteristics. From this combined dataset, we are able to determine the increase in expenditures that households would face following changes in tax rates and compute the total tax revenue to be redistributed in a lump-sum (see the Appendix C.2 for more details). We thereby obtain the distribution of households' *objective* net gains in purchasing power implied by the policies proposed. Formally, the net gain γ_h of household h (the notations are explained in the Appendix B) can be expressed as follows:

$$\gamma_h = N_h^a \cdot D - \Delta E_h^{transport} - \Delta E_h^{housing} \tag{1}$$

where D denotes the value of the dividend, N_h^a is the number of adults receiving it in this household, and ΔE is the increases in their energy expenditures. Our computations use the typical elasticities found in the literature on French households: -0.4 for transport and -0.2 for housing.[18] We assume that consumers bear 80% of the incidence of energy taxes.

2.3.2 Computing households' expected net gains

Simulating expected net gains To measure each respondent's bias and to provide customized feedback on their win/lose category, we estimate their net gain by using equation (1). We directly compute the increase in transport fuel expenditures $\Delta E^{transport}$ from the information reported on their yearly distance traveled and the average fuel consumption of their vehicles. From the housing survey—which provides information on households' housing energy bills—we compute $\Delta E^{housing}$ and regress it on household characteristics. We then use the coefficients obtained to compute $\widehat{\Delta E}^{housing}$ for the households of our own survey respondents. The specification we chose is presented in the Appendix C.3 and compared with alternative specifications and prediction methods.

Assessing feedback's accuracy To assess the accuracy of our prediction, we test it out-of-sample on the consumer survey. Indeed, for households in this survey, we both obtain the true net gain γ and use our prediction to estimate $\widehat{\gamma}$ and then assess the likelihood of

[18] These values correspond to the short-run uncompensated price elasticities estimated by Douenne (2020) and are in line with previous findings on French households (e.g., Clerc & Marcus, 2009; Bureau, 2011).

correctly predicting their win/lose category. Because we made the prediction by using a different survey than the one on which we tested it, we avoided the risk of overfitting. For five households out of six, we correctly predict whether their purchasing power would increase or decrease due to the policy. We make this ratio symmetrical to balance the shares of overly optimistic and overly pessimistic feedback. Among the households in BdF predicted to win, 83.4% were actual winners, while among those predicted to lose, 83.4% were actual losers (see Figure C.1 in the Appendix D.2). Assuming that the characteristics reported by our respondents are correct, there is no reason to believe that the probability of error is higher or lower when simulations are applied to our survey respondents.[19]

3 Pessimistic beliefs

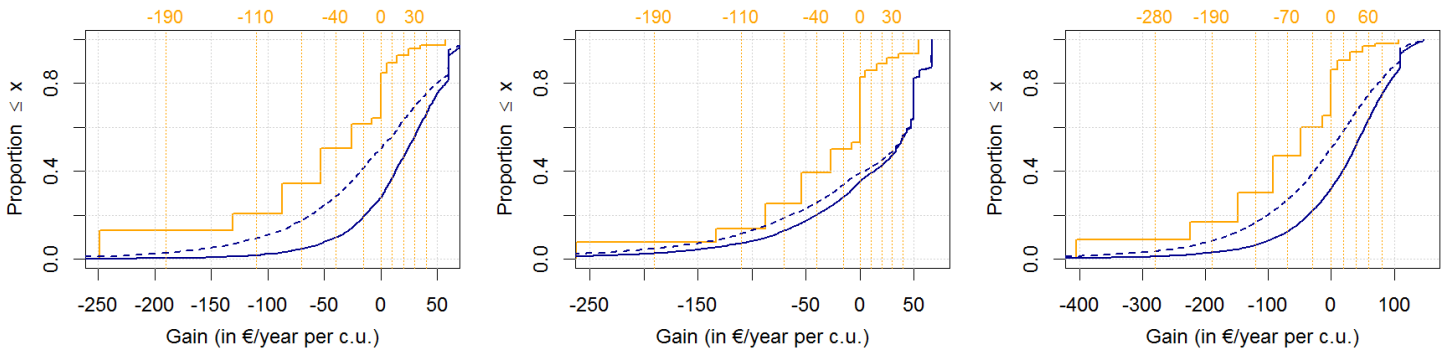
3.1 Self-interest

Overestimation of policy costs The analysis of respondents' beliefs about the impact of both sectoral taxes and the carbon tax & dividend policy shows a clear overestimation of the incidence of these taxes on their household. Figure 3.1 compares the CDF of the objective net gains estimated using official data with the subjective beliefs from our survey for both sectoral taxes and the tax & dividend policy. It is evident from these figures that on average, respondents overestimate the cost of these policies, even in the extreme case of perfectly inelastic expenditures. This result holds both for sectoral carbon taxes on transport and housing energy and the carbon tax & dividend policy. The average objective net gains from the taxes on transport, housing, and both (i.e. the tax & dividend) are 18€ per consumption unit (c.u.), 6€ per c.u., and 24€ per c.u., respectively. Extrapolating from our survey, we instead find average subjective net gains of −61€, −43€, and −89€, respectively.[20] For the tax & div-

[19] In particular, a critical assumption is that people correctly reported their distance traveled and the average fuel economy of their vehicles so that the computation of $\Delta E^{transport}$ is correct. As shown in Table A.1 in the Appendix A, the values reported by respondents follow a distribution very similar to that found in the official statistics.

[20] The subjective intervals are translated into numerical values, assuming that the distribution within each interval is the same as that in the Insee data. Within each bin, we simply take the actual average for the CDF. Among the several methods that we considered to assign numerical values, all realistic ones yield identical

idend, while 70% of households should benefit (in monetary terms) from it, only 14% think that they would (and 22% see themselves as unaffected).[21] The median gap of 116€ between objective and subjective gains for this policy indicates a substantial bias towards loss from typical respondents. This bias is widespread, as we find that 89% of respondents underestimate their purchasing power gains relative to our household-specific estimation. (The full distribution of respondents' bias is provided in Figure C.2 in the Appendix C.3.) This proportion remains as high as 77% when assuming inelastic expenditures, which provides a lower bound on the share who underestimate their net utility gain.



(a) Transport fuels (b) Housing energy (c) Both

NOTE: The dashed blue lines represent the distributions of the objective gains in the extreme case of totally inelastic expenditures. The vertical dotted orange lines show the limits of the intervals of the answers to the subjective gains.

Figure 3.1: Cumulative Density Function of objective (dark blue) vs. subjective (orange) net gains from our tax & dividend.

Heterogeneity in bias To characterize profiles of individuals more likely to misperceive their gains, we regress what we call a *large bias* over many respondent characteristics. A large bias is defined as a gap between objectively estimated and subjective net gains greater than 110€ per c.u. for our tax & dividend policy.[22] The results given in Table 3.1 show that

results, and we find that the policy costs are overestimated, even in the most conservative approach (taking the maximal bounds of intervals).

[21] For the transport and housing energy taxes, the objective proportions of winners are very similar at 74% and 67%, respectively, while the subjective shares are 16% and 17%, respectively (with 22% and 30% unaffected, respectively).

[22] Indeed, our estimation differs from the true objective gain by more than 110€ in only 5% of cases. This definition ensures that for the 55% of respondents with a large bias, there is a clear gap between their subjective perceptions and their actual net gains. Other definitions yield very similar results.

Table 3.1: Determinants of bias in subjective gains.

	Large bias ($\lvert\widehat{\gamma}-g\rvert > 110$)		
	OLS	*logistic*	*OLS*
Initial tax: PNR (I don't know)			−0.179
			(0.023)
Initial tax: Approves			−0.284
			(0.031)
Yellow Vests: PNR	0.039	0.035	0.024
	(0.036)	(0.035)	(0.036)
Yellow Vests: understands	0.081	0.062	0.041
	(0.025)	(0.024)	(0.025)
Yellow Vests: supports	0.108	0.103	0.051
	(0.026)	(0.025)	(0.026)
Yellow Vests: is part	0.202	0.193	0.147
	(0.048)	(0.040)	(0.047)
Environmentalist	−0.064	−0.061	−0.025
	(0.026)	(0.026)	(0.026)
Left-right: Left	−0.066	−0.044	−0.045
	(0.063)	(0.065)	(0.061)
Left-right: Center	−0.062	−0.048	−0.046
	(0.065)	(0.068)	(0.064)
Left-right: Right	−0.024	−0.010	−0.026
	(0.064)	(0.066)	(0.063)
Left-right: Extreme-right	−0.076	−0.057	−0.088
	(0.066)	(0.069)	(0.065)
Left-right: Indeterminate	−0.009	0.017	−0.007
	(0.061)	(0.063)	(0.060)
Controls: Sociodemo, political leaning	✓	✓	✓
Observations	3,002	3,002	3,002
R^2	0.061		0.098

NOTE: Standard errors are reported in parentheses. For the logit, the average marginal effects are reported and not the coefficients. The omitted variables are *Yellow Vests: opposes* and *Left-right: Extreme-left*. The list of controls can be found in the Appendix F. A large bias is defined as a difference between the subjective (g) and objectively estimated ($\widehat{\gamma}$) net gain larger than 110€/year per c.u.

having a large bias is largely idiosyncratic: when controlling for a large set of variables[23] (column 1), R^2 remains small (0.06). Nevertheless, we identify several variables having a significant effect on *large bias*, even when controlling the false discovery rate at 5%.[24] While we find that environmentalists are less likely to display a large bias and that the standard left/right political leaning has no significant effect, the degree of support for the Yellow Vests movement is positively associated with a large bias. This effect increases with the degree of adhesion, up to 20 p.p. for individuals who reported being part of the movement. Column (3) additionally includes one's attitude towards the policy when it is first presented as a covariate. We see that people who approve of the policy at the initial stage are 28 p.p. less likely to have a large bias than those who do not accept it and 10 p.p. less likely than those who do not know.

Overall, the biases are large and closely related to one's convictions. However, the direction(s) of the causality between beliefs and attitude towards the policy is not resolved at this stage. The results of Section 4 suggest that people form their beliefs differently depending on their political views and identity, while Section 5 shows that perceived outcomes causally influence support.

3.2 Environmental effectiveness

A well-established result in the literature on the acceptability of climate policies is the perceived ineffectiveness of Pigouvian instruments (e.g., Dresner et al., 2006a; Kallbekken et al., 2011; Baranzini & Carattini, 2017). In particular, people do not see carbon taxes as effective in combating climate change. Our findings confirm this result. Among the respondents who did not receive the information on environmental effectiveness, only 15% answered "Yes" when asked whether our tax & dividend would be effective in reducing pollution and fighting climate change, 68% answered "No" and 18% answered that they did not

[23] The control variables used throughout the paper are described in the Appendix F.

[24] To conduct the multiple testing procedure (following Benjamini & Hochberg, 1995), instead of associating each dummy with a different null hypothesis, we used F-tests of joint nullity for the dummies of each categorical variable and for two additional triplets of variables: those related to household composition and those related to income.

know.

An explanation sometimes encountered to explain perceptions of ineffectiveness is that most people believe that energy consumption is quite inelastic (Kallbekken & Sælen, 2011; Carattini et al., 2018). In our survey, we randomly elicited respondents' perceived price elasticity of either housing or transport energy for French people. We find that higher perceived elasticities are associated with a higher likelihood to believe that the policy is effective, even when controlling for many respondents' characteristics. However, these effects are too modest to explain the perceived ineffectiveness (see Appendix D.3). Indeed, among respondents who perceive the policy to be environmentally ineffective, almost half anticipate responses to price changes larger than those in the literature.[25]

A more plausible explanation for perceived ineffectiveness is that people do not believe that the policy would be sufficient to *substantially* affect pollution and climate change. Taking respondents' average anticipated elasticities for transport and housing energy (that are fairly accurate[25]), our tax & dividend policy should reduce French greenhouse gas (GHG) emissions by 5.7 Mt of CO_2 equivalents (CO_2e) each year. This reduction corresponds to 0.8% of French annual emissions and 0.01% of global emissions.[26] Thus, although respondents do anticipate responses to price incentives, our results suggest that they do not perceive a 50€/tCO_2 national carbon tax to be a proportionate reaction to climate change.

3.3 Progressivity

It is often argued that a critical barrier to the acceptance of carbon taxation is its perceived distributional impact, in particular the higher burden imposed on lower income households (Bristow et al., 2010; Brannlund & Persson, 2012; Gevrek & Uyduranoglu, 2015). However, the literature has shown that redistributing the revenue of a carbon tax through uni-

[25]Overall, the average subjective elasticities are close to the estimates of the literature for transport (at -0.45) and somewhat overestimated for housing (-0.43). Among those who declared that the policy was not effective, 45% (resp. 43%) anticipated an aggregate elasticity at or below -0.5 for housing (resp. for transport), while elasticities obtained from the literature are approximately -0.2 for housing and -0.4 for transport.

[26]The computations are based on household carbon emissions simulated from official data. In 2014, French GHG consumption-based emissions were equal to 712 $MtCO_2e$ (CGDD, 2019). In 2017, global emissions were 53.5 $GtCO_2e$ (UNEP, 2018).

form lump-sum transfers—i.e., a tax & dividend—can make the policy progressive (West & Williams, 2004; Bento et al., 2009; Williams et al., 2015), including for France (Bureau, 2011; Douenne, 2020). Figure C.3 in Appendix C.4 displays the average net gain by income decile for our tax & dividend. This figure shows that lower income households would gain more than richer households, both in relative and in absolute terms. However, among respondents who did not receive the information on the progressivity, only 19% of respondents think the policy would benefit the poorest households, compared to 60% who declare that it would not and 20% who do not know.

4 Determinants of beliefs

The previous section has shown that people's low acceptance of our tax & dividend correlates with pessimistic beliefs about the properties of the scheme. As knowledge about these properties has been shown to be decisive for acceptance (Carattini et al., 2018), it is important to assess how people form their beliefs. In the following, we test respondents' reactions to information about their gains, environmental effectiveness, and progressivity. If overly pessimistic views simply reflected a lack of knowledge, we would expect them to revise their beliefs after they receive new information, which is what we refer to as "updating".

4.1 Self-interest

4.1.1 Pessimism in the revision of beliefs

Our respondent-specific estimation of net gains (see Section 2.3) enables us to tell respondents that given their characteristics, they have a 5-in-6 chance to "win" or "lose" from the policy. We can then examine how they update their beliefs about their win/lose category after receiving this information. The full transition matrices of people's beliefs are given in Tables D.2 and D.3 in the Appendix D.2. More concisely, Table 4.1 reports the share of respondents whose beliefs after being informed are aligned with our feedback with the corresponding 95% binomial confidence intervals. It shows a highly asymmetric response depending on

the feedback received. On the one hand, for the 24% of individuals who receive "lose" feedback ($\widehat{\Gamma} = 0$), the *ex post* belief is on average consistent with the fact that 83% of them do in reality lose from the tax & dividend policy, with 86% endorsing our "lose" feedback. If anything, these people would rather tend to *agree too much* with our noisy signal, especially when excluding people who initially consider themselves to be unaffected (i.e., focusing on $g^0 \neq 0$). On the other hand, the 76% who received "win" feedback ($\widehat{\Gamma} = 1$) appear to be much more conservative in their revision since only 25% of them endorse the "win" feedback. Among the respondents who initially thought that they would lose in this group, a mere 12% switch their answer from "lose" to "win". This is in sharp contrast to the respondents who initially thought they would win and receive "lose" feedback since 82% of them endorse our prediction. Thus, pessimistic beliefs are persistent to our treatment, but optimistic beliefs are not.

Table D.4 in the Appendix D.2 conducts the same analysis for the 28% of respondents whose gain is very positive or very negative, i.e., above 110€ per c.u. in absolute terms. For these respondents, our out-of-sample prediction of the win/lose category is correct in 99% of the cases, as seen in Figure C.1 in the Appendix D.2. The alignments with our feedback are similar between the whole sample and these respondents for whom we are certain to make a correct prediction. The similarity of alignments for different prediction accuracies rules out the possibility that a large fraction of respondents do not update their beliefs because their private information would be *truly* more accurate than our prediction.

4.1.2 Causal effect of feedback on beliefs

To be relevant, our feedback on respondents' win/lose category had to be customized and as accurate as possible. Since our information intervention was not a randomized treatment, we have yet to identify the causal effect of receiving *win* vs. *lose* feedback on respondents' beliefs. Despite this non-random assignment of the treatment, there is still a way to estimate the causal effect of the feedback on people's beliefs. The binary win/lose feedback is a variable $\widehat{\Gamma}$ that jumps from 0 to 1 when our continuous estimation of respondents' net gains $\widehat{\gamma}$ exceeds the zero threshold. The following equation enables us to estimate the threshold effect

Table 4.1: Share of respondents with new beliefs aligned with feedback.

	Aligned with feedback: $G^F = \widehat{\Gamma}$	
	Feedback:	
	win ($\widehat{\Gamma} = 1$) (75.8%)	lose ($\widehat{\Gamma} = 0$) (24.2%)
Initial belief winner ($g^0 > 0$) (14.0%)	78.8% [73.2%;83.4%]	81.5% [65.0%;91.3%]
Initial belief unaffected ($g^0 = 0$) (21.7%)	21.6% [17.6%;26.2%]	44.9% [33.5%;56.8%]
Initial belief loser ($g^0 < 0$) (64.3%)	12.2% [10.3%;14.5%]	93.9% [90.9%;96.0%]
Initial belief affected ($g^0 \neq 0$) (78.3%)	26.1% [23.7%;28.7%]	92.9% [89.8%;95.1%]
All (100%)	25.1% [23.0%;27.3%]	85.7% [82.2%;88.7%]

NOTE: The 95% confidence intervals for the binomial probabilities are given in brackets. The Table reads as follows: among those who initially think that they would win ($g^0 > 0$) but are told that they are expected to lose ($\widehat{\Gamma} = 0$), 81.5% agree that they would lose ($G^F = 0$). The feedback $\widehat{\Gamma}$ is not a random draw but a deterministic outcome of the characteristics reported by the respondents in the survey. The subsample used is the 1,968 respondents who received feedback.

around zero net gain in an RDD, and thus to obtain the effect of receiving *win* compared to *lose* feedback:

$$G_i^F = \alpha_0 + \alpha_1 \widehat{\Gamma}_i + \alpha_{\gamma,1} \widehat{\gamma}_i + \alpha_{\gamma,2} \widehat{\gamma}_i^2 + \alpha_{\mathbf{C}} \mathbf{C_i} + \alpha_{\mathbf{I}} \mathbf{I_i} + \eta_i \tag{2}$$

The dependent variable G^F represents respondents' belief about their gain after the feedback. It is equal to 0 if they believe they lose, and 1 otherwise. $\mathbf{C}$ is a set of respondents' characteristics, and $\mathbf{I}$ a set of variables describing their household's income. The two terms $\widehat{\gamma}$ and $\widehat{\gamma}^2$ allow for a quadratic specification to control for estimated net gains to properly isolate the effect of receiving positive rather than negative feedback. To better identify the threshold effect in the RDD, we focus on the subsample of respondents whose estimated

net gain $\widehat{\gamma}$ is close to zero (less than 50€ per annum in absolute value). Table 4.2 displays the results, including our main specification (column 1), the same estimation performed on the full sample (2), a modified version of (1) where interaction terms have been added to study heterogeneous treatment effects (3), and a specification where the dependent variable corresponds to the belief to *win* instead of to *not lose* (4).

Table 4.2: Effect feedback on belief of winning.

	Believes does not lose			Believes wins
	(1)	(2)	(3)	(4)
Predicted winner ($\widehat{\Gamma}$)	0.269	0.208	0.246	0.153
	(0.058)	(0.035)	(0.064)	(0.045)
Initial tax Acceptance (A^0)	0.306	0.331	0.015	0.310
	(0.066)	(0.038)	(0.087)	(0.051)
Yellow Vests supporter			0.019	
			(0.058)	
$\widehat{\Gamma} \times A^0$			0.166	
			(0.079)	
$\widehat{\Gamma} \times$ Yellow Vests supporter			−0.151	
			(0.069)	
$\widehat{\Gamma} \times G$			0.080	
			(0.079)	
Controls: Incomes (piecewise continuous) estimated gains, sociodemo, other motives	✓	✓	✓	✓
Controls: initial win/lose category (G)			✓	
Subsample	$\lvert\widehat{\gamma}\rvert < 50$		$\lvert\widehat{\gamma}\rvert < 50$	$\lvert\widehat{\gamma}\rvert < 50$
Observations	757	1,968	757	757
R^2	0.301	0.320	0.419	0.253

NOTE: Standard errors are reported in parentheses. The list of controls can be found in the Appendix F.

Everything else being equal, after receiving the feedback, respondents predicted winners are 27 p.p. more likely to believe they do not lose than respondents predicted losers (column 1). They are also 15 p.p. more likely to believe that they win (column 4). Interestingly,

column 3 shows that the effect of the informational treatment is heterogeneous. From the two interaction terms, we see that among the respondents who initially accepted the tax & dividend policy—i.e., when it was first presented—the effect of the *win* feedback on the belief not to lose goes from 27 p.p. to 41 p.p., while for those supportive of the Yellow Vests—who declared to be part of or to support the movement—it goes down to less than 10 p.p. Thus, the information provided to respondents is processed differently depending on their attitude towards the policy. While the treatment is rather effective at convincing those most favorable to the policy about its impact on their household, those most opposed to it do not appear to be receptive to this information.

4.1.3 Mechanisms

There are several ways to rationalize pessimistic and heterogeneous beliefs against the tax & dividend. We propose the following four mechanisms: distrust, uncertainty, motivated reasoning, and intentional misreporting.

Distrust The first mechanism is that respondents distrust what we present to them. They may perceive our information to be biased, think that we wrongly estimated their likelihood of winning and that we are overly optimistic.[27] As a result, they may discount our new information relative to their prior belief or assign relatively more weight to our information when it is pessimistic. This distrust may stem from an impression that experts understate the costs of a carbon tax or that the government will break its promise to pay the dividend. For instance, Sapienza & Zingales (2013) report that 51% of Americans are skeptical that their governments would deliver on using the proceeds of a carbon tax to reduce other taxes (see also Dresner et al., 2006a; Hsu et al., 2008). A similar level of skepticism regarding the dividend could explain much of the pessimism about net gains.

[27] Another possibility is that respondents give too much value to their private information relative to the base rate information. That is, pessimistic winners might be overconfident in seeing themselves as specific so that they partly discard the new information, e.g., by thinking that they are part of the one-sixth for whom our prediction is erroneous, perhaps because they believe that they always lose more than others from new policies.

Uncertainty The second mechanism stems from people's uncertainty regarding their gain. That uncertainty would make them see their possible gain as a distribution (see Stiglitz, 2019). Then, instead of reporting the average of this distribution, loss-averse people would conservatively estimate their gains. Also, the effect of uncertainty on updating is ambiguous. On the one hand, uncertain people could be more likely to rely on our base rate information; but on the other hand, their subjective probability of losing could remain high despite our information.

Motivated reasoning The third mechanism to explain the observed asymmetry in belief revision is that some people have a strong skeptical attitude towards the carbon tax, which affects the formation of their beliefs. They would engage in motivated reasoning, i.e., update their beliefs in a way that is consistent with their initial views (Druckman & McGrath, 2019; Little, 2019) rather than integrate information in a way that leads to accurate conclusions. Although motivated reasoning is linked to distrust since it also involves neglecting information, in the case of distrust, people discard information because they do not trust its source; for motivated reasoning, they dismiss the information when its content contradicts preexisting views. Motivated reasoning entails a deviation from Bayesian updating—contrary to the first two mechanisms—but it can still be *rationalized* as a psychological adaptation to preserve one's sense of identity (Kahan, 2013). By analyzing the determinants of a correct update, the online Appendix J.1 shows that those who initially disapprove of the tax & dividend and those who are part of the Yellow Vests movement are significantly less likely to correctly update their win/lose belief after feedback (by 18 and 14 p.p., respectively), even controlling for their prior belief. This finding is consistent with motivated reasoning since political views and identity are correlated with the way people form their beliefs (see the online Appendix J.2 for a discussion). However, we did not demonstrate the causality, and this correlation could also stem from higher distrust (or higher uncertainty) among these groups.

Intentional misreporting A fourth possibility is that some respondents intentionally report overly pessimistic beliefs compared to what they actually think. This could stem from a

rejection of the tax and could follow from strategic thinking if they believe that their survey answers might influence policy-makers. Such respondents could be aware that they would gain but still reject the tax for other motives, even more so if they are still uncertain about their gain. Their misreporting could also be due to a type of motivated reasoning that would not directly affect their beliefs but rather induce them to misreport what they think. This could help them justify their rejection of the policy, even more so in that it could be costly for their ego to admit that they were wrong to reject the policy.

4.2 Environmental effectiveness

Table D.5 in the Appendix D.3 reports the effect of displaying relevant information on the belief that our tax & dividend is environmentally effective. The effect of reporting a scientific consensus on environmental effectiveness (E) is positive and statistically significant, but its magnitude—approximately 5 p.p.—seems modest given that the question immediately follows the information intervention. The effects of information on climate change (CC) or particulates (PM) are smaller, and only CC is significant, which is understandable since we displayed the information at the very beginning of the survey and it does not mention any environmental policy. As suggested by Millner & Ollivier (2016), given the complexity of the mechanisms at play, drawing a causal link between the causes and consequences of environmental problems requires considerable cognitive effort, making it difficult to convince one of the effectiveness of policies that decentralize efforts to address pollution. Finally, we observe that our information interventions have no significant effect on beliefs about the causes and consequences of climate change. Overall, these treatments appear to be insufficient to change most people's minds about climate change and carbon tax effectiveness.

4.3 Progressivity

Table D.6 in the Appendix D.4 reveals the absence of an effect of explaining that our tax & dividend is progressive on perceived progressivity. The correlation between the two is close to 0 (at -0.006) and even has an unexpected negative sign. Column (2) of the same table

clarifies why our treatment does not change the overall share of people who think that the policy is regressive. Those who have a large bias in their perception of gains are in fact *more* prone to perceive *regressivity* once provided the information by 13 p.p. This result may be a manifestation of the boomerang effect with people inclined to motivated reasoning, which has already been documented for Republican attitudes on climate change in the US (Zhou, 2016). Indeed, Hovland et al. (1953) showed that when someone is pressured to make a certain choice, psychological reactance (theorized by Brehm, 1966) can cause them to resist this pressure by adopting an opposite alternative. Although the effect on those without a large bias is not significant, providing them with information is associated with a lower perceived regressivity by 5 p.p. To conclude, without a deep explanation of the underlying mechanisms, the progressivity of the policy remains unintuitive for most people, and we cannot convince them easily.

5 How beliefs determine attitudes

Our results clearly indicate that, at present, a carbon tax is unlikely to be accepted in France. However, we have also shown that people display overly pessimistic perceptions about the true effects of the policy. Most of them overestimate the negative impact on their purchasing power, think that the policy is regressive, and do not see it as environmentally effective. In this section, we examine to what extent the low acceptance rate reflects intrinsic preferences or incorrect perceptions. The question we address is whether convincing people about the actual incidence of the policy and its effectiveness would be sufficient to generate public support.

5.1 Self-interest

Identification challenge Among the three-quarters of the respondents expected to win from our tax & dividend, 62% both consider that they would not win and disapprove of the policy. We want to estimate to what extent knowing that they would win would lead them

to approve of the reform. Because respondents thinking that they would win might differ in many respects from those thinking they would not, we need a specific identification strategy to estimate the causal effect of the perception of winning on approval.

Main identification strategy To identify the effect of self-interest on acceptance *ceteris paribus*, we exploit exogenous variations in gains and losses. To do so, we consider a tax & targeted dividend, where respondents are randomly assigned to a compensation scheme for which they are eligible or not depending on their income (see Section 2.2.2). Formally, we denote $I_{i,1}$ as the income percentile of respondent i and $I_{i,2}$ as that of the second adult in their household if there is one. We define the eligibility of adult $j \in \{1;2\}$ as follows:[28]

$$T_{i,j} = \begin{cases} 0, & \text{if } I_{i,j} > t_i \\ 1, & \text{otherwise} \end{cases} \tag{3}$$

where $t_i \in \mathscr{T} = \{20;30;40;50\}$ is the eligibility threshold randomly allocated to household i (see Section 2.2.2). Since eligibility increases the likelihood of believing that one wins from the policy—without necessarily implying it—our method leads to a fuzzy RDD, where eligibility corresponds to the intention to treat and the respondents who believe that they will win from the tax and targeted dividend correspond to the treated. Formally, we denote G_i^T as a dummy variable equal to 0 if respondent i thinks that they would lose from the tax & targeted dividend and 1 otherwise. Similarly, A_i^T is a dummy variable equal to 0 if respondent i disapproves of this policy and 1 otherwise. We can then write the model as a two-stage least squares with the following first-stage equation:

$$G_i^T = \alpha_0 + \alpha_{T,1} T_{i,1} + \alpha_{T,2} T_{i,2} + \alpha_{T,3} (T_{i,1} \times T_{i,2}) + \sum_{k \in \mathscr{T}} \alpha_k \mathbb{1}_{t_i = k} + \alpha_S S_i + \alpha_{\mathbf{C}} \mathbf{C_i} + \alpha_{\mathbf{I}} \mathbf{I_i} + \eta_i \tag{4}$$

where $\mathbf{C_i}$ is a vector of respondent characteristics, S_i is a dummy variable equal to 1 when there is a single adult in the household, and $\mathbf{I_i}$ is a vector of income variables defined as

[28] As explained in Section 2.2.2, we explicitly limit the number of beneficiaries to two per household.

$\left(I_{i,j},\ \left(\min(I_{i,j}-k,\ 0)\right)_{k=20,70}\right)'_{j=1,2}$. $\mathbf{I_i}$ allows for a continuous piecewise linear relationship in incomes with slope changes at the 20th and 70th percentiles. We also introduce fixed effects for the policy assigned $\mathbb{1}_{t_i=k}$ ($k \in \mathscr{T}$) to control for preferences regarding the specificities of the policy, i.e., the share of the population targeted by the policy and the value of the dividend. To obtain more precise estimates, we control for initial acceptance of our tax & dividend since this should explain much of the variation in the dependent variable. Finally, the second stage is written as follows:

$$A_i^T = \beta_0 + \beta_1 \widehat{G}_i^T + \sum_{k \in \mathscr{T}} \beta_k \mathbb{1}_{t_i=k} + \beta_S S_i + \beta_{\mathbf{C}} \mathbf{C_i} + \beta_{\mathbf{I}} \mathbf{I_i} + \varepsilon_i \tag{5}$$

where $\widehat{G_i}^T$ denotes the fitted value of G_i^T from the first-stage regression. As seen from the first-stage results shown on Table 5.1 below, the eligibility of both respondents and households' second adults are positively correlated with beliefs about winning, and so both instruments are relevant. The exclusion restriction states that conditional on income, being eligible affects approval solely through beliefs on winning. The RDD procedure employed in the first stage ensures that this is the case. Conditional on income, eligibility is random, and by controlling for the specific policy assigned ($\mathbb{1}_{t_i=k}$), it should affect acceptance only through self-interest.

Alternative IV identification As an alternative identification strategy, we exploit a methodology similar to the main specification—i.e., a fuzzy RDD—but applied it to the customized feedback. Indeed, we use our estimation of respondents' net gains $\widehat{\gamma}$ as the assignment variable and the binary win/lose feedback $\widehat{\Gamma}$ as the intention to treat. Since our feedback $\widehat{\Gamma}$ (which jumps from 0 to 1 at the threshold of zero net gain) is predictive of the belief about the win/lose category after feedback (see Section 4.1.2), G^F, we can determine the effect of this belief on acceptance after feedback, A^F. This alternative fuzzy RDD leads to the following two-stage least squares model:

$$G_i^F = \alpha_0 + \alpha_1 \widehat{\Gamma}_i + \alpha_{\gamma,1} \widehat{\gamma}_i + \alpha_{\gamma,2} \widehat{\gamma}_i^2 + \alpha_{\mathbf{C}} \mathbf{C_i} + \alpha_{\mathbf{I}} \mathbf{I_i} + \eta_i \tag{6}$$

$$A_i^F = \beta_0 + \beta_1 \widehat{G}_i^F + \beta_{\gamma,1}\widehat{\gamma}_i + \beta_{\gamma,2}\widehat{\gamma}_i^2 + \beta_{\mathbf{C}}\mathbf{C_i} + \beta_{\mathbf{I}}\mathbf{I_i} + \varepsilon_i \tag{7}$$

where $\widehat{G}_i^F$ denotes the fitted value of G_i^F from the first-stage regression. The first stage equation thus corresponds to our main specification of the RDD estimated in Section 4.1.2 (column (1) in Table 4.2). The identification assumption of this second IV states that conditional on estimated net gains ($\widehat{\gamma}$)—which we control for with a quadratic specification—receiving a win feedback ($\widehat{\Gamma} = 1$) affects approval solely through self-interest. As before, we restrict our analysis to respondents close enough to the threshold by retaining only those with net gains below 50€ per annum in absolute value ($|\widehat{\gamma}| < 50$).

Results The first-stage regression results are given in Table 5.1. The effective F-statistics (Olea & Pflueger, 2013) range from 15.6 to 21.3, indicating that both targeted transfers and feedback are strong instruments. Table 5.2 provides the second-stage results. Overall, the estimated effect of self-interest indicate that believing that one would not lose increases acceptance by more than 50 p.p. (53 p.p. in our main specification). Both IV strategies yield consistent results, although they apply to different policies since revenue recycling is not designed in the same manner and are estimated on different subsamples since compliers are not the same in these two specifications.

Alternative specifications for robustness In our main specification (column 1 in Table 5.1), we exclude households where none of the adults have an income from the 10th to 60th percentiles to keep only those close enough to the thresholds. In an alternative specification (2), we replicate the same estimation using the full sample. In (3), we also compare our results with a simple OLS regression on the full sample. Finally, we investigate alternative versions of the previous models in Appendix E. We estimate the effect of winning instead of not losing, and on approval instead of acceptance (Table E.1). We estimate our main specification with the slope of incomes changing at additional thresholds (30th, 40th, 50th or 60th percentile). Finally, we allow for heterogeneous effects along the income dimension

Table 5.1: First-stage regressions results for self-interest

	Believes does not lose		
	Targeted Dividend (G^T)		After feedback (G^F)
	(1)	(2)	(4)
Transfer to respondent (T_1)	0.199	0.224	
	(0.034)	(0.030)	
Transfer to spouse (T_2)	0.172	0.156	
	(0.042)	(0.039)	
$T_1 \times T_2$	−0.145	−0.158	
	(0.045)	(0.037)	
Predicted winner ($\widehat{\Gamma}$)			0.269
			(0.058)
Initial policy Acceptance (A^0)	0.123	0.154	0.306
	(0.041)	(0.033)	(0.066)
Controls: Incomes (piecewise continuous) estimated gains, sociodemo, other motives	✓	✓	✓
Controls: Policy assigned	✓	✓	
Subsample	[p10; p60]		$\lvert\widehat{\gamma}\rvert < 50$
Effective F-statistic	15.6	23.8	21.3
Observations	1,969	3,002	757
R^2	0.221	0.196	0.301

NOTE: In (1,2), we use the random eligibility for the dividend (conditional on income) as a source of the exogenous variation in the belief. In (4), we use the discontinuity in the win/lose feedback when the net gain switches from negative to positive. The column numbers correspond to second-stage results, as given in Table 5.2.

(Table E.2).

Overall, our main point estimate is robust to these alternative specifications. For example, compared to our main point estimate of 53 p.p., the full sample and simple OLS point estimates are close at respectively 48 p.p. and 44 p.p., and we find no significant heterogeneity in the effect along the income dimension.

Table 5.2: Effect of self-interest on acceptance

	Acceptance ("Yes" or "Don't know" to policy support)			
	Targeted Dividend (A^T)			After Feedback (A^F)
	IV: random target/eligibility		*OLS*	*IV: discontinuity in feedback*
	(1)	(2)	(3)	(4)
Believes does not lose (G)	0.534	0.476	0.438	0.644
	(0.132)	(0.106)	(0.014)	(0.170)
Initial policy Acceptance (A^0)	0.356	0.354	0.361	0.420
	(0.041)	(0.034)	(0.026)	(0.074)
Controls: Incomes (piecewise continuous) estimated gains, sociodemo, other motives	✓	✓	✓	✓
Controls: Policy assigned	✓	✓	✓	
Subsample	[p10; p60]			$\lvert\widehat{\gamma}\rvert < 50$
Effective F-statistic	15.6	23.8		21.3
Observations	1,969	3,002	3,002	757
R^2	0.320	0.308	0.472	0.541

NOTE: Standard errors are reported in parentheses. The list of controls can be found in the Appendix F. The source of the exogenous variation in the belief used in the first-stage estimations for the targeted dividend is the random assignment of the income threshold, which determines eligibility for the dividend. The first stage for the non-targeted dividend instead exploits the discontinuity in the win/lose feedback when the net gain switches from negative to positive.

5.2 Environmental effectiveness

Main identification strategy One of the strongest barriers to carbon tax implementation is a widespread perception of its environmental ineffectiveness. Our objective is therefore to assess to what extent learning about the environmental benefits of the tax could increase support. To identify this effect, we estimate a two-stage least squares (2SLS) model where the first stage uses random information to predict beliefs about environmental effectiveness, while the second stage regresses acceptance on the fitted exogenous variations in these beliefs. Because information on particulate matter (Z_{PM}) is poorly correlated with beliefs regarding effectiveness, we restrict the set of instruments to our informational treatments on

the scientific consensus (Z_E) and climate change (Z_{CC}). Although these treatments do not have a very large effect on people's beliefs (as discussed in Section 4.2), these instruments are significantly related to our endogenous variable. Denoting $\dot{A^0}$ as the dummy for initial approval of the tax & dividend and $\dot{E}$ as the dummy for the belief that the policy is environmentally effective, we can write a 2SLS model as follows:

$$\dot{E}_i = \alpha_0 + \alpha_1 Z_{E,i} + \alpha_2 Z_{CC,i} + \alpha_{\mathbf{C}} \mathbf{C_i} + \eta_i \tag{8}$$

$$\dot{A}_i^0 = \beta_0 + \beta_1 \widehat{\dot{E}_i} + \beta_{\mathbf{C}} \mathbf{C_i} + \varepsilon_i \tag{9}$$

where $\widehat{\dot{E}_i}$ denotes the fitted value of $\dot{E}_i$ from the first-stage regression and $\mathbf{C}$ is a vector of characteristics. Acknowledging that our information intervention could affect acceptance motives other than effectiveness alone, we include other motives in our list of control variables to avoid potential bias.

Results The first-stage regression results can be found in Table 5.3. To avoid the problem of weak instruments in our main specification, we adopt strict definitions for our variables (i.e., the answer *"Yes"*, denoted by a dot, to the belief in effectiveness and approval) since they yield a higher effective F-statistic: 11 instead of 6 for broad definitions (*not "No"*). Table 5.4 reports the results of the second stages. They all consistently indicate a strong positive and significant effect of beliefs about environmental effectiveness on support for the policy. All else equal, believing that the tax is effective increases the likelihood of approving it by 42 p.p. (1).

Alternative specifications for robustness checks In addition to the 2SLS (specification 1), we estimate an OLS (2) model with strict definitions for our variables. We also estimate other specifications with different definitions for our variables. The 2SLS in column (3) employs acceptance instead of approval as the dependent variable. In the Appendix E, we estimate a 2SLS with broad definitions only, as well as two OLS regressions ("Yes" on acceptance and *not "No"* on acceptance). As a robustness check, we also estimate our

Table 5.3: First-stage regressions results for environmental effectiveness

	Environmental effectiveness	
	"Yes"	not "No"
	(1; 3)	(A4)
Info on Environmental Effectiveness (Z_E)	0.059	0.062
	(0.014)	(0.017)
Info on Climate Change (Z_{CC})	0.028	0.030
	(0.013)	(0.017)
Controls: Sociodemo, other motives, incomes, estimated gains	✓	✓
Effective F-statistic	11.2	6.0
Observations	3,002	3,002
R^2	0.123	0.121

NOTE: Regarding the column names, (A4) refers to columns with alternative second stages in Table E.3. We use the information randomly displayed about climate change (Z_{CC}) and the effectiveness of carbon taxation (Z_E) as sources of exogenous variation in the belief. We chose the set of instruments that maximizes the effective F-statistics. The Sargan test does not reject the validity of our overidentification restrictions (p-value of 0.93).

main specification using a limited information maximum likelihood (LIML). As a robustness check, we also estimate our main specification using a limited information maximum likelihood (LIML).

Overall, our main result is robust to these alternative specifications. The effect estimated with the OLS (2) is only slightly lower than the LATE estimated in our main specification—38 vs. 42 p.p. Given the exogeneity of our instruments, the only concern of the relative weakness of our instruments is a potential bias towards the OLS, which—as suggested by the results of column (2)—would entail estimates that are too conservative in our case. Finally, we obtain identical results when running a 2SLS or an LIML for our main specification (1). Regarding the broad definition of acceptance, the LIML estimate is broadly consistent with the 2SLS result (3), though it is somewhat higher: 64 p.p. vs. 51 p.p. (column (A2) of Table E.3 in the Appendix E).

Table 5.4: Effect of believing in environmental effectiveness on approval

	Initial Tax & Dividend		
	Approval ($\dot{A}^0$)		Acceptance (A^0)
	IV	*OLS*	*IV*
	(1)	(2)	(3)
Believes in effectiveness ($\dot{E}$)	0.416	0.374	0.505
	(0.168)	(0.013)	(0.242)
Instruments: info E.E. & C.C.	✓		✓
Controls: Sociodemo, other motives, incomes, estimated gains	✓	✓	✓
Effective F-statistic	11.2		11.2
Observations	3,002	3,002	3,002
R^2	0.161	0.342	0.218

NOTE: Standard errors are reported in parentheses. The list of controls can be found in the Appendix F. The dependent variable corresponds to either initial approval (answer "Yes" to support of the policy) or acceptance (answer not "No"). The first stage exploits the information randomly displayed about climate change (C.C.) and the effectiveness of carbon taxation (E.E.) as exogenous instruments.

5.3 Progressivity

Since informing respondents does not convince them that our tax & dividend policy is progressive (see Section 4.3), we cannot perform an IV estimation to identify the causal effect of understanding the progressivity on support for the policy. In our online Appendix K, we estimate how one's belief in progressivity—interacted with other motives—correlates with acceptance using simple OLS and logit regressions. Controlling for many respondent characteristics and other motives for support, the effect of progressivity remains statistically significant and as high as 27 p.p. in our preferred specification. Of course, this result should be interpreted with caution since we can still suspect that the results are affected by unobserved confounders and reverse causality.

To conclude, these results show that convincing citizens of the true incidence and effectiveness of a tax & dividend could substantially increase support for such a policy. Extrap-

olating from the causal effects that we find for self-interest and environmental effectiveness, we estimate that if everyone could be convinced that the reform is environmentally effective, acceptance would reach 61% (and approval would reach 45%) and that if everyone could learn their win/lose category, acceptance would reach 47%. Thus, if people could be convinced on more than one motive, the tax & dividend would presumably be supported by a large majority, although we cannot causally identify the interaction effect of several motives combined (see the online Appendix K for non causal estimates). Our results also qualify the findings of Anderson et al. (2019), who suggest that ideology better predicts carbon tax acceptance than self-interest. By distinguishing beliefs from preferences, our results suggest that ideology plays an indirect role by shaping beliefs about one's self-interest and that beliefs directly affect acceptance.

6 Conclusion

In this paper, we study how beliefs about a policy form and then determine attitudes towards it. We investigate this question through the study of carbon taxation in France during the Yellow Vests movement that started to protest fuel price increases. Our analysis is based on a new survey and official household survey data, enabling us to compare subjective beliefs with objective impacts on French households. We find that 70% disapprove of a carbon tax & dividend policy, which can be explained by pessimistic beliefs about its properties. Of our survey respondents, 89% overestimate its negative impact on their purchasing power, and most of them do not perceive it as environmentally effective or progressive. Pessimistic beliefs appear correlated with people's support for the scheme: the more they oppose the mechanism, the more pessimistic they are. The causality between beliefs and attitudes towards the policy could actually go in both directions. People more opposed to the tax are more (pessimistically) biased in their treatment of *new* information with respect to it, suggesting that political views and identity shape beliefs about the impacts of taxation. We simultaneously find that beliefs causally determine acceptance and that if people could be convinced about the incidence and effectiveness of a tax & dividend, this policy would likely

be accepted by a majority, given the large effects of these motives (approximately 50 p.p. each).

However, our treatments that provide accurate arguments in favor of the scheme mostly fail to convince people. The pessimism could be related to a strong distrust of the government, as documented, e.g., in Alesina et al. (2018) and Algan et al. (2019), echoing recent findings that the ambition of climate policies increases with the level of trust (Rafaty, 2018). These results leave us with three main challenges. First, since it is unlikely that the issue of trust can be resolved in the short run, it seems necessary to find climate policies that would be accepted by a majority. We address this question in a companion paper (Douenne & Fabre, 2020) in which we assess both knowledge and beliefs about climate change and the preferred policies of French people. Second, since trust in the government needs to be restored in the longer run, it is crucial to analyze what causes distrust and how it can be overcome. Third, it is important to assess to what extent the mechanisms of belief formation and their effects on political attitudes we document can be generalized to other policies and other contexts. Although rejection of the tax may be lower in a different country, biases in perceptions and political polarization may occur everywhere. Thus, a lesson must be learned for policy design and implementation to avoid another carbon tax debacle *à la Française*.

References

ADEME. Représentations sociales de l'effet de serre. Technical report, 2018. 5

A. Alberini, A. Bigano, M. Ščasný, & I. Zvěřinová. Preferences for Energy Efficiency vs. Renewables: What Is the Willingness to Pay to Reduce CO2 Emissions? *Ecological Economics*, 2018. Link. 81

A. Alesina, S. Stantcheva, & E. Teso. Intergenerational Mobility and Preferences for Redistribution. *American Economic Review*, 2018. Link. 6, 39

Y. Algan, E. Beasley, D. Cohen, M. Foucault, & M. Péron. Qui sont les gilets jaunes et leurs soutiens ? Technical report, CEPREMAP et CEVIPOF, 2019. 39, 75

S. Anderson, I. E. Marinescu, & B. Shor. Can Pigou at the Polls Stop US Melting the Poles? Working Paper 26146, National Bureau of Economic Research, 2019. 5, 6, 38

A. Baranzini & S. Carattini. Effectiveness, earmarking and labeling: testing the acceptability of carbon taxes with survey data. *Environmental Economics and Policy Studies*, 2017. Link. 5, 12, 20

R. Bénabou & J. Tirole. Mindful Economics: The Production, Consumption, and Value of Beliefs. *Journal of Economic Perspectives*, 2016. Link. 6

Y. Benjamini & Y. Hochberg. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 1995. Link. 20

A. M. Bento, L. H. Goulder, M. R. Jacobsen, & R. H. von Haefen. Distributional and Efficiency Impacts of Increased US Gasoline Taxes. *American Economic Review*, 2009. Link. 22

R. P. Berrens, A. K. Bohara, H. C. Jenkins-Smith, C. L. Silva, & D. L. Weimer. Information and effort in contingent valuation surveys: application to global climate change using

national internet samples. *Journal of Environmental Economics and Management*, 2004. Link. 81

P. C. Boyer, T. Delemotte, G. Gauthier, V. Rollet, & B. Schmutz. Les déterminants de la mobilisation des Gilets jaunes. *Revue économique*, 2020. Link. 7

R. Brannlund & L. Persson. To tax, or not to tax: preferences for climate policy attributes. *Climate Policy*, 2012. Link. 5, 21

S. Brechin. Public opinion: a cross-national view. In *Handbook of Climate Change and Society*. Routledge, lever-tracy, constance edition, 2010. ISBN 978-1-135-99850-9. 5

J. W. Brehm. *A theory of psychological reactance*. Academic Press, Oxford, England, 1966. 29

A. L. Bristow, M. Wardman, A. M. Zanni, & P. K. Chintakayala. Public acceptability of personal carbon trading and carbon tax. *Ecological Economics*, 2010. Link. 5, 21

B. Bureau. Distributional effects of a carbon tax on car fuels in France. *Energy Economics*, 2011. Link. 16, 22

C. Camerer & R. Hogarth. The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework. *Journal of Risk and Uncertainty*, 1999. 77

T. A. Cameron. Individual option prices for climate change mitigation. *Journal of Public Economics*, 2005. Link. 81

S. Carattini, A. Baranzini, P. Thalmann, F. Varone, & F. Vöhringer. Green Taxes in a Post-Paris World: Are Millions of Nays Inevitable? *Environmental and Resource Economics*, 2017. Link. 5

S. Carattini, M. Carvalho, & S. Fankhauser. Overcoming public resistance to carbon taxes. *Wiley Interdisciplinary Reviews: Climate Change*, 2018. Link. 5, 21, 22

CGDD. Chiffres clés du climat France, Europe et Monde. Technical report, 2019. 21

R. Challier. Rencontres aux ronds-points. *La Vie des idées*, 2019. Link. 78

M. Clerc & V. Marcus. Elasticité-prix des consommations énergétiques des ménages. Technical report, Insee, 2009. 16

G. Cruces, R. Perez-Truglia, & M. Tetaz. Biased perceptions of income distribution and preferences for redistribution: Evidence from a survey experiment. *Journal of Public Economics*, 2013. Link. 6

T. Douenne. The vertical and horizontal distributive effects of energy taxes: A case study of a french policy. *The Energy Journal*, 2020. 15, 16, 22, 48

T. Douenne & A. Fabre. French attitudes on climate change, carbon taxation and other climate policies. *Ecological Economics*, 2020. 8, 15, 39, 71

S. Dresner, L. Dunne, P. Clinch, & C. Beuermann. Social and political responses to ecological tax reform in Europe: an introduction to the special issue. *Energy Policy*, 2006. Link. 20, 26

S. Dresner, T. Jackson, & N. Gilbert. History and social responses to environmental tax reform in the United Kingdom. *Energy Policy*, 2006. Link. 5

J. N. Druckman & M. C. McGrath. The evidence for motivated reasoning in climate change preference formation. *Nature Climate Change*, 2019. Link. 27, 77

Z. E. Gevrek & A. Uyduranoglu. Public preferences for carbon tax attributes. *Ecological Economics*, 2015. Link. 21

R. Golman, G. Loewenstein, K. O. Moene, & L. Zarri. The Preference for Belief Consonance. *Journal of Economic Perspectives*, 2016. Link. 78

W. M. Hanemann. Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses. *American Journal of Agricultural Economics*, 1984. Link. 80

C. I. Hovland, I. L. Janis, & H. H. Kelley. *Communication and persuasion; psychological studies of opinion change*. Yale University Press, New Haven, CT, US, 1953. 29

S.-L. Hsu, J. Walters, & A. Purgas. Pollution tax heuristics: An empirical study of willingness to pay higher gasoline taxes. *Energy Policy*, 2008. Link. 26

J. D. Jenkins. Political economy constraints on carbon pricing policies: What are the implications for economic efficiency, environmental efficacy, and climate policy design? *Energy Policy*, 2014. Link. 81

D. M. Kahan. Ideology, motivated reasoning, and cognitive reflection. *Judgment and Decision Making*, 2013. 27, 77

S. Kallbekken & H. Sælen. Public acceptance for environmental taxes: Self-interest, environmental and distributional concerns. *Energy Policy*, 2011. Link. 5, 6, 21

S. Kallbekken, S. Kroll, & T. L. Cherry. Do you not like Pigou, or do you not understand him? Tax aversion and revenue recycling in the lab. *Journal of Environmental Economics and Management*, 2011. Link. 20

D. Klenert, L. Mattauch, E. Combet, O. Edenhofer, C. Hepburn, R. Rafaty, & N. Stern. Making carbon pricing work for citizens. *Nature Climate Change*, 2018. Link. 5

M. J. Kotchen, K. J. Boyle, & A. A. Leiserowitz. Willingness-to-pay and policy-instrument choice for climate-change policy in the United States. *Energy Policy*, 2013. Link. 81

P. W. Kraft, M. Lodge, & C. S. Taber. Why People "Don't Trust the Evidence": Motivated Reasoning and Scientific Beliefs. *The ANNALS of the American Academy of Political and Social Science*, 2015. Link. 77

Z. Kunda. The Case for Motivated Reasoning. *Psychological Bulletin*, 1990. 4, 6, 77

I. Kuziemko, M. I. Norton, E. Saez, & S. Stantcheva. How elastic are preferences for redistribution? evidence from randomized survey experiments. *American Economic Review*, 2015. 6

A. T. Little. The Distortion of Related Beliefs. *American Journal of Political Science*, 2019. Link. 27, 77

A. Longo, A. Markandya, & M. Petrucci. The internalization of externalities in the production of electricity: Willingness to pay for the attributes of a policy for renewable energy. *Ecological Economics*, 2008. Link. 81

A. Millner & H. Ollivier. Beliefs, Politics, and Environmental Policy. *Review of Environmental Economics and Policy*, 2016. Link. 5, 28

J. J. Murphy, P. G. Allen, T. H. Stevens, & D. Weatherhead. A Meta-analysis of Hypothetical Bias in Stated Preference Valuation. *Environmental and Resource Economics*, 2005. Link. 81

J. L. M. Olea & C. Pflueger. A robust test for weak instruments. *Journal of Business & Economic Statistics*, 2013. 32

R. Rafaty. Perceptions of Corruption, Political Distrust, and the Weakening of Climate Policy. *Global Environmental Politics*, 2018. Link. 39

D. P. Redlawsk. Hot Cognition or Cool Consideration? Testing the Effects of Motivated Reasoning on Political Decision Making. *The Journal of Politics*, 2002. Link. 6, 77

P. Sapienza & L. Zingales. Economic experts versus average americans. *American Economic Review*, 2013. 11, 26

P. C. Stern, T. Dietz, & L. Kalof. Value Orientations, Gender, and Environmental Concern. *Environment and Behavior*, 1993. Link. 5

J. E. Stiglitz. Addressing Climate Change through Price and Non-Price Interventions. *European Economic Review*, 2019. Link. 27

B. Stokes, R. Wike, & J. Carle. Global Concern about Climate Change, Broad Support for Limiting Emissions. *Pew Research Center*, 2015. 8

D. Streimikiene, T. Balezentis, I. Alisauskaite-Seskiene, G. Stankuniene, & Z. Simanaviciene. A Review of Willingness to Pay Studies for Climate Change Mitigation in the Energy Sector. *Energies*, 2019. Link. 81

M. Thaler. The "Fake News" Effect: An Experiment on Motivated Reasoning and Trust in News. 2019. 6

P. Thalmann. The Public Acceptance of Green Taxes: 2 Million Voters Express Their Opinion. *Public Choice*, 2004. Link. 5, 6

UNEP. Emissions Gap Report 2018. Technical report, 2018. Link. 21

S. West & R. Williams. Estimates from a consumer demand system: implications for the incidence of environmental taxes. *Journal of Environmental Economics and Management*, 2004. 22

R. C. Williams, H. Gordon, D. Burtraw, J. C. Carbone, & R. D. Morgenstern. The Initial Incidence of a Carbon Tax Across Income Groups. *National Tax Journal*, 2015. Link. 22

J. Zhou. Boomerangs versus Javelins: How Polarization Constrains Communication on Climate Change. *Environmental Politics*, 2016. Link. 29

Appendices

A Raw data

Table A.1: Sample characteristics: quotas.

	Population	Sample
Sex		
woman	*0.52*	0.53
man	*0.48*	0.47
Age		
18-24	*0.12*	0.11
25-34	*0.15*	0.11
35-49	*0.24*	0.24
50-64	*0.24*	0.26
>65	*0.25*	0.27
Profession		
farmer	*0.01*	0.01
independent	*0.03*	0.04
executive	*0.09*	0.09
intermediate	*0.14*	0.14
employee	*0.15*	0.16
worker	*0.12*	0.13
retired	*0.33*	0.33
inactive	*0.12*	0.11
Education		
No diploma or *Brevet*	*0.30*	0.24
CAP or *BEP*	*0.25*	0.26
Bac	*0.17*	0.18
Higher	*0.29*	0.31
Size of town		
rural	*0.22*	0.24
<20k	*0.17*	0.18
20-99k	*0.14*	0.13
>100k	*0.31*	0.29
Paris area	*0.16*	0.15
Region		
IDF	*0.19*	0.17
Nord	*0.09*	0.10
Est	*0.13*	0.12
SO	*0.09*	0.09
Centre	*0.10*	0.12
Ouest	*0.10*	0.10
Occ	*0.09*	0.08
ARA	*0.12*	0.13
PACA	*0.09*	0.08

Table A.2: Household characteristics.

	Population	Sample
Household composition (mean)		
Household size	*2.36*	2.38
Number of adults	*2.03*	1.93
Number of c.u. (consumption units)	*1.60*	1.61
Energy source (share)		
Gas	*0.42*	0.36
Heating oil	*0.12*	0.09
Size of accommodation (m^2)		
mean	*97*	96
p25	*69*	66
p50	*90*	90
p75	*120*	115
Distance travelled by car (km/year)		
mean	*13,735*	15,328
p25	*4,000*	4,000
p50	*10,899*	10,000
p75	*20,000*	20,000
Fuel economy (L/100 km)		
mean	*6.39*	7.18
p25	*6*	5
p50	*6.5*	6
p75	*7.5*	7

SOURCES: Matched BdF, except for number of adults (ERFS) and heating oil (CEREN).

NOTE: After controlling the false discovery rate at 5%, *t*-tests reject that the sample mean is equal to the population mean for 12 of our 42 variables in Tables A.1 and A.2. Refer to Section 2.2.2 on page 9.

B Notations

To improve the understanding of our specifications in the regression tables, we adopt consistent notations throughout the paper. For questions where possible answers are "Yes"/"No"/"PNR", we define two kinds of dummy variables: the default ones correspond to *not "No"* answers, and we place a dot on dummy variables for *"Yes"*. For example, acceptance is denoted by A while approval is denoted by $\dot{A}$. Furthermore, for questions that are asked several times, namely, acceptance and win/lose category, an exponent is added to specify the step at which the question is asked. Table B.1 describes these exponents and the notations corresponding to the different notions of gain that we use. We use uppercase for binary and lowercase for continuous variables, and Greek letters denote objective notions, with a hat for our estimation of gains and without for the true (unknown) ones. To provide another example, the broad notion of self-interest at the initial step, i.e., the belief that one does not lose, is denoted by G^0, and the strict belief that one wins with the tax & targeted dividend is denoted by $\dot{G}^T$.

Table B.1: Notations for the different reforms and for gain notions.

Step:	Initial	after information: 1		with Targeting
Variants:	–	Progressivity	Feedback	–
Exponent	0	P	F	T

Gain	Subjective	True	Estimated
Numeric	g	γ	$\widehat{\gamma}$
Binary	$\dot{G}\,(g>0)$, $G\,(g\geq 0)$	Γ	$\widehat{\Gamma}$

NOTE: Refer to Section 2.3 on page 15.

C The use of official household survey data

The paper employs official survey data for two purposes: (*i*) computing the distribution of increases in fossil fuel expenditures and (*ii*) predicting the expected net gain of each respon-

dent based on their energy characteristics. Section C.1 presents the three official surveys from Insee (the French national statistics bureau) that we use. Section C.2 details the formulas needed to compute the value of the dividend and households' expected net gains from their expenditures. Section C.3 explains how by using two distinct surveys, we can obtain a simple formula to predict respondents' net gain simply based on their energy characteristics and then test the likelihood of making a correct prediction out-of-sample. Finally, Section C.4 displays the objective net gain of the policy by income decile to show that it is progressive.

C.1 Official household surveys from Insee

Consumer survey "Budget de Famille" The consumer survey (BdF 2011) is a household survey providing information on all households' revenues and expenditures, together with many sociodemographic characteristics. It was conducted in several waves from October 2010 to September 2011 on a representative sample of 10,342 French households. The main advantage of the BdF when studying the incidence of carbon taxation is that it reports expenditures on both housing and transportation energy. Housing energy expenses are determined from households' bills, and for most other goods, respondents report their expenditures over the past week. However, as explained in Douenne (2020), this data collection is problematic when examining the incidence of a tax on transportation energy since short-run fluctuations in consumption lead to overestimation of the heterogeneity in expenditures.

Transport survey "Enquête Nationale Transports et Déplacements" To overcome this limitation, the BdF is matched with the transport survey (ENTD 2008). The ENTD was conducted in several waves from April 2007 to April 2008 on a representative sample of 20,178 French households. It provides information on household characteristics, vehicle fleets and use over the past week, but most important, it provides information on the annual distances traveled with these vehicles. This last information enables us to recover the distribution of transport fuel expenditures without overestimating its spread. Such matching is not necessary for housing energy since it already represents consumption over long periods in the

BdF.

Housing survey "Enquête Logement" The housing survey (EL 2013) was conducted from June 2013 to June 2014 on a sample of 27,137 households in metropolitan France. It includes considerable information on households' characteristics, as well as their housing energy bills. The distribution of energy expenditures is very close to that of the BdF.

C.2 Formulas to compute monetary effects of carbon tax policy

To compute the monetary impact of a carbon tax increase on household h, we decompose current energy expenditures $E_h(\tau)$ as a product of current price $P(\tau)$ and current quantities consumed $Q_h(\tau)$, each being a function of the excise tax τ of which the carbon tax is a part:[29]

$$E_h(\tau) = P(\tau)\, Q_h(\tau)$$

Using a first-order approximation, we can then express small variations in expenditures as follows:

$$\frac{dE}{E}(\tau) = \frac{dP}{P}(\tau) + \frac{dQ}{Q}(\tau) \tag{10}$$

We can rewrite the variation in quantities as a function of the price variation:

$$\frac{dQ}{Q}(\tau) = e\frac{dP}{P}(\tau)$$

where $e = \frac{dQ_h}{dP} \cdot \frac{P}{Q_h}$ is the price elasticity of the energy good considered, which is assumed to be constant and identical across households. For all energy types, we can decompose the final price itself as follows:

$$P(\tau) = (p + i\tau)(1 + t)$$

where t is the value added tax (VAT) rate (assumed constant) that applies after excise taxes, i is the incidence of excise taxes on consumers (assumed constant), and $p + (i-1)\tau$ is the

[29] The French carbon tax "Contribution Climat Energie" is a component of existing taxes on energy products: TICPE for transport and heating oils, and TICGN for natural gas.

producer price as a function of τ.[30] When the carbon price changes so that the excise taxes vary from τ to some level τ', we therefore have the following:

$$\frac{\Delta P(\tau)}{P} = \frac{P(\tau') - P(\tau)}{P(\tau)} = \frac{(p + i\tau')(1+t) - (p+i\tau)(1+t)}{(p+i\tau)(1+t)} = \frac{i(\tau' - \tau)}{p+i\tau}$$

Thus, by carrying on the first-order approximation, we can express an increase in expenditures associated with a carbon price increase as follows:

$$\Delta E_h(\tau) = E_h(\tau)(1+e)\frac{\Delta P}{P} = E_h(\tau)(1+e)\frac{i(\tau' - \tau)}{p+i\tau} \tag{11}$$

We can replicate similar calculations to obtain the expected variations in the tax paid on energy by household h, ΔT_h. Starting from the expression for T_h—which is the sum of excise taxes and the VAT on the energy good—we have the following:

$$T_h(\tau) = Q_h(\tau)\left((1+t)\tau + t\big(p + (i-1)\tau\big)\right)$$

From this, we obtain the following:

$$\Delta T_h(\tau) = Q_h(\tau)\left(1 + e\frac{i(\tau' - \tau)}{p+i\tau}\right)\left(t\big(p + (i-1)\tau'\big) + (1+t)\tau'\right) - Q(\tau)\left(t\big(p + (i-1)\tau\big) + (1+t)\tau\right) \tag{12}$$

Finally, the net gain of a household h from a tax & dividend is written as follows:

$$\gamma_h(\tau) = N_h^a \cdot \frac{\sum_h \Delta T_h(\tau)}{N^a} - \Delta E_h^{transport}(\tau) - \Delta E_h^{housing}(\tau) \tag{13}$$

where γ_h denotes its net gain from the policy, N_h^a is the number of adults receiving the dividend in this household, N^a is the total number of adults receiving it, and $\Delta E_h^{transport}$ ($\Delta E_h^{housing}$) is the increase in their expenditures on transport (housing) energy. From households' energy expenditures and making assumptions for the elasticities and tax incidence, equations (11) to (13) enable us to obtain the value of dividends and the impact of the policy

[30] Hence, p is the producer price when $\tau = 0$.

on households' purchasing power. We use equation (13) to estimate the biases and objective distribution of net gains in Section 3 and the customized feedback in Section 4.

When asked to estimate the impact of the policy on their own purchasing power, respondents simply had to make an estimation over:

$$\Delta E_h(\tau) = E_h(\tau)(1+e)\frac{\Delta P}{P}$$

where, for simplicity, ΔP was given for transport fuels, and $\frac{\Delta P}{P}$ was given for housing energy. Thus, they were not required to make any specific assumption about existing taxes or tax incidence but simply to estimate their consumption and price elasticity.

C.3 Predicting gains and losses

As explained in Section 2.3, to estimate respondents' bias and provide customized feedback on their win/lose category, we need to estimate the increase in their housing energy expenditures, $\Delta E_h^{housing}$, based on their energy characteristics. To do so, we use the housing survey *Enquête Logement* (EL 2013), which provides information on household expenditures on housing energy and many demographic and energy characteristics. It enables us to compute $\Delta E^{housing}$ and regress it on household characteristics. We can then use the coefficients obtained to compute $\widehat{\Delta E}^{housing}$ (and thus obtain $\widehat{\gamma}$) for any household. The specification we chose is as follows:

$$\Delta E_h^{housing} = \beta_0 + \beta_1 \chi_h^G + \beta_2 \chi_h^F + \beta_3 \sigma_h + \varepsilon_h \tag{14}$$

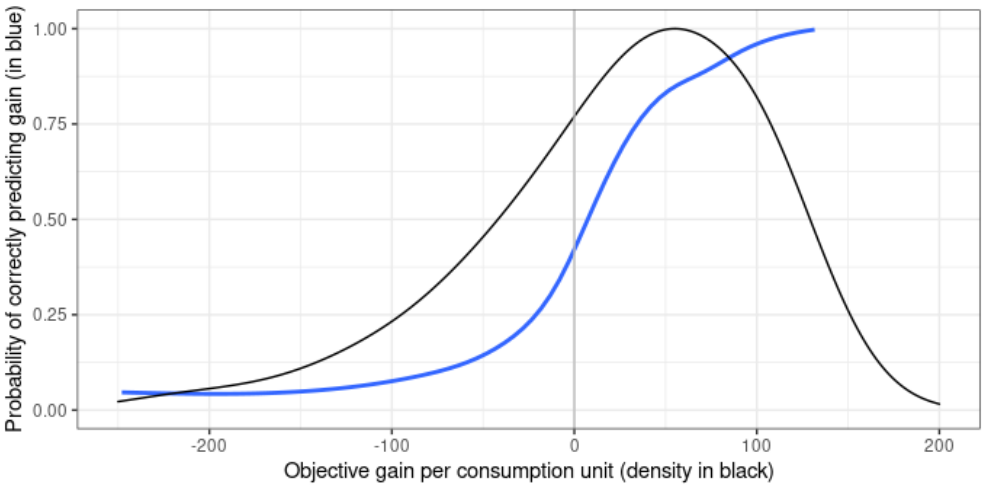
where χ_h^G (resp. χ_h^F) is a dummy variable equal to 1 if the household uses gas (res. heating oil) for heating and σ is the size of the household's accommodation in square meters. The results are provided in Table C.1 together with those of other specifications. The last row of the Table C.1 shows the out-of-sample error rate, computed with the consumer survey, as explained in Section 2.3. All specifications yield a similar error rate of 15-17%. Given the concern that respondents could make mistakes when reporting their accommodation size in

the entry field, we used the first specification in our survey since it does not rely as heavily as the others on the accommodation size. To balance the error rates for losing households that are mistakenly estimated to be winners and for winners who are mistakenly estimated to be losers, we add a constant of 16.1 in our estimation for the yearly net gain, which is thus the sum of 16.1 plus 110 times one or two (depending on the number of adults) minus increases in transport and housing energy expenditures. We selected the OLS as our prediction method for the estimation of the net gain because it compares well to alternative methods. We also classified winners and losers using a decision tree and obtained a very close error rate: 17.4%. Finally, statistical matching provided an error rate of 17.7%.

Table C.1: Determinants of housing energy expenditures.

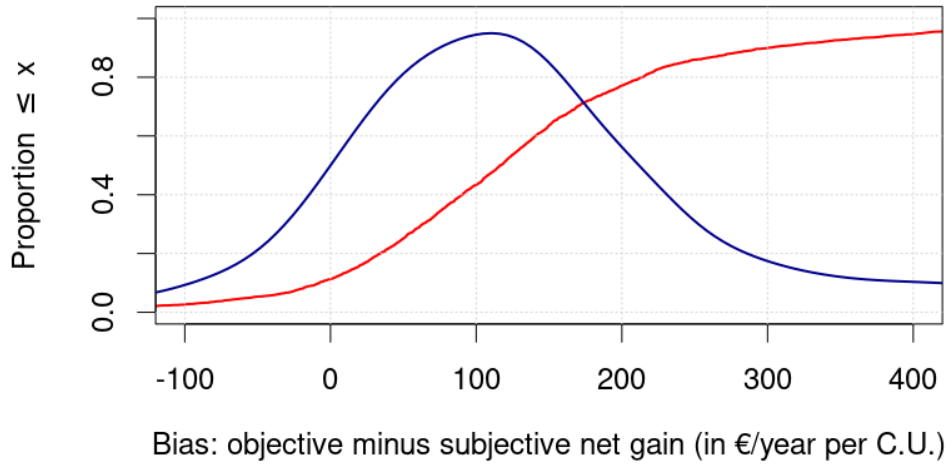
	Increase in housing energy expenditures (€/year)		
	(1)	(2)	(3)
Constant	−55.51		−0.634
	(1.237)		(1.489)
Housing energy: Gas	124.6		1.173
	(1.037)		(2.323)
Housing energy: Heating oil	221.1	129.8	130.4
	(1.719)	(3.752)	(4.002)
Accommodation size (m^2)	0.652		0.024
	(0.012)		(0.015)
Accommodation size × Gas		1.425	1.397
		(0.007)	(0.024)
Accommodation size × Heating oil		0.945	0.922
		(0.029)	(0.032)
Observations	26,729	26,729	26,729
R^2	0.545	0.716	0.599
Error rate	0.166	0.155	0.155

NOTE: The increase in energy expenditures is directly computed from households' energy bills in the housing survey, based on equation (11) in the Appendix C.2. See the discussion in the main text, Section 2.3.2 on page 16.



NOTE: The black curve corresponds to the density of households' objective net gains in the consumer survey, and the blue curve corresponds to the probability that our net gain estimation correctly predicts the win/lose category. As shown by the blue curve, households in the consumer survey who would gain 100€ per c.u.—as directly computed from their energy bills—were predicted to be winners based on their energy characteristics in 96% of cases. See the discussion in the main text, Section 4.1.1 on page 22.

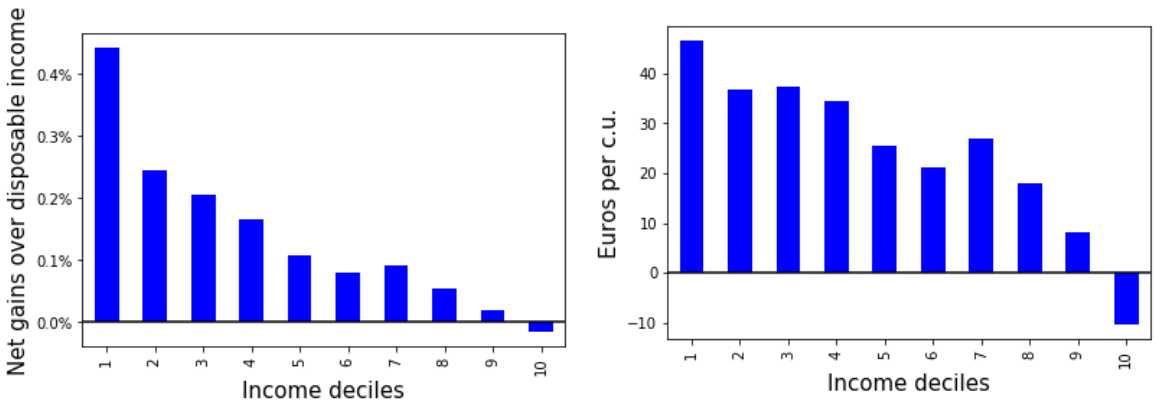
Figure C.1: Probability that our net gains estimation correctly predicts the win/lose category.



NOTE: The red curve indicates that for 11% of the respondents, their objective gains are lower than their subjective ones; meanwhile, for 23% of them, they are higher by at least 200€. The blue curve indicates that the most common bias is an underestimation of gains by approximately 100€. See the discussion in the main text, Section 3.1 on page 17.

Figure C.2: CDF (in red) and PDF (in blue) of the bias.

C.4 Distributional effects



NOTE: Net gains are defined in equation (13). They correspond to the dividend minus the increase in expenditures (ΔE), not in taxes (ΔT). Although the latter would sum to zero when considering the population in aggregate because the reform is budget neutral, the former does not because fossil fuel expenditures adjust downwards following the increase in the carbon tax. See the discussion in the main text, Section 3.3 on page 21.

Figure C.3: Average net gain of the carbon tax and dividend policy by income decile (computed using Insee data).

D Beliefs and persistence

D.1 Elasticities

To see how perceived elasticities relate to the perception of effectiveness, we run two separate regressions where the dependent variable E is equal to 0 if the respondent does not perceive the policy to be environmentally effective and 1 otherwise. In the first regression, we regress the perceived effectiveness on the perceived elasticity for housing; and in the second, we regress it on the perceived elasticity for transport energy. Table D.1 below reports the results with and without control variables. They all consistently indicate that perceived elasticities are correlated with beliefs about the policy's effectiveness since for both sectors a respondent anticipating an elasticity of -1 is (on average) 6 p.p. more likely to perceive the tax & dividend policy to be effective than one anticipating no elasticity. Although significant, the magnitude of the effect is modest, showing that the perceived ineffectiveness of tax instruments should not be attributed to small subjective elasticities.

Table D.1: Effect of subjective elasticities on perceived environmental effectiveness.

	Environmental effectiveness: not 'No'			
	(1)	(2)	(3)	(4)
Price elasticity: Housing	−0.062		−0.055	
	(0.032)		(0.032)	
Price elasticity: Transport		−0.056		−0.060
		(0.030)		(0.030)
Controls: Sociodemo, energy, incomes, estimated gains			✓	✓
Observations	1,501	1,501	1,501	1,501
R^2	0.003	0.002	0.089	0.090

NOTE: Environmental effectiveness refers to the belief that the policy would be effective at reducing pollution and fighting climate change. Price elasticities for housing and transport are elicited from respondents' expected reduction in energy consumption of French people following a given increase in energy prices. For more details, see the discussion in the main text, Section 3.2 on page 20. The list of controls can be found in the Appendix F.

D.2 Self-interest

Table D.2: Transition matrix after telling respondents they are expected to *win* (75.8%).

Before \ *After*	**Winner** (25%)	**Unaffected** (28%)	**Loser** (47%)
Winner (16%)	79%	13%	8%
Unaffected (24%)	22%	63%	15%
Loser (60%)	12%	18%	70%

NOTE: See the discussion in the main text, Section 4.1.1 on page 22.

Table D.3: Transition matrix after telling respondents they are expected to *lose* (24.2%).

Before \ *After*	**Winner** (3%)	**Unaffected** (12%)	**Loser** (86%)
Winner (7%)	16%	3%	81%
Unaffected (15%)	5%	50%	46%
Loser (78%)	1%	5%	94%

NOTE: See the discussion in the main text, Section 4.1.1 on page 22.

Table D.4: Share with new beliefs aligned with feedback, among those with a large gain or loss ($|\widehat{\gamma}| > 110$).

	Aligned with feedback: $G^F = \widehat{\Gamma}$	
	win ($\widehat{\Gamma} = 1$) (81.6%)	lose ($\widehat{\Gamma} = 0$) (18.4%)
Initial belief winner ($g > 0$) (19.4%)	77.6% [68.5%;84.7%]	78.4% [43.2%;94.5%]
Initial belief unaffected ($g = 0$) (28.2%)	20.7% [14.8%;28.1%]	32.7% [14.7%;57.7%]
Initial belief loser ($g < 0$) (52.3%)	10.8% [7.3%;15.8%]	92.2% [84.5%;96.3%]
Initial belief affected ($g \neq 0$) (70.8%)	32.7% [27.7%;38.1%]	91.1% [83.5%;95.4%]
All (100%)	28.9% [24.8%;33.3%]	83.0% [74.8%;88.9%]

NOTE: The 95% confidence intervals for binomial probabilities are given in brackets. The table reads as follows: among those who initially think that they would win ($g^0 > 0$) but are told they are expected to lose ($\widehat{\Gamma} = 0$), 78.4% agree that they would lose ($G^F = 0$). Compared to Table 4.1, this table focuses on the subsample of 546 respondents with a large gain or loss ($|\widehat{\gamma}| > 110$)) who received feedback. See the discussion in the main text, Section 4.1.1 on page 22.

D.3 Environmental effectiveness

Table D.5: Effect of information interventions on beliefs about environmental effectiveness

	Environmental effectiveness			
	not "No"			"Yes"
	OLS		*logit*	*OLS*
	(1)	(2)	(3)	(4)
Info on Environmental Effectiveness (Z_E)	0.043	0.063	0.052	0.059
	(0.017)	(0.018)	(0.018)	(0.014)
Info on Climate Change (Z_{CC})	0.044	0.041	0.043	0.029
	(0.024)	(0.024)	(0.024)	(0.018)
Info on Particulate Matter (Z_{PM})	0.039	0.029	0.037	0.017
	(0.024)	(0.024)	(0.024)	(0.019)
$Z_{CC} \times Z_{PM}$	−0.040	−0.033	−0.042	−0.005
	(0.035)	(0.034)	(0.033)	(0.027)
Controls: Sociodemo		✓	✓	✓
Observations	3,002	3,002	3,002	3,002
R^2	0.003	0.047		0.075

NOTE: See the discussion in the main text, Section 4.2 on page 28.

D.4 Progressivity

Table D.6: Effect of information on perceived progressivity

	Progressivity: not "No" (P)		
	(1)	(2)	(3)
Constant	0.419	0.435	0.052
	(0.022)	(0.033)	(0.319)
Information on progressivity (Z_P)	−0.021	0.050	0.051
	(0.027)	(0.040)	(0.041)
Large bias ($\lvert\widehat{\gamma}-g\rvert > 110$)		−0.028	−0.040
		(0.045)	(0.045)
Interaction $Z_P \times (\lvert\widehat{\gamma}-g\rvert > 110)$		−0.130	−0.117
		(0.055)	(0.055)
Controls: Sociodemo, politics			✓
Observations	1,444	1,444	1,444
R^2	0.0004	0.018	0.094

NOTE: A large bias is defined as a difference between the subjective (g) and objectively estimated ($\widehat{\gamma}$) net gain larger than 110€/year per c.u. See the discussion in the main text, Section 4.3 on page 28.

E Additional specifications for the estimation of acceptation motives

Table E.1: Effect of self-interest on acceptance: second stages of alternative specifications

	Targeted Dividend (A^T)			After Feedback (A^F)		
	Acceptance	Approval		Acceptance	Approval	
	(1)	(2)	(3)	(4)	(5)	(6)
Believes wins	0.574	0.357		1.131	0.609	
	(0.136)	(0.117)		(0.298)	(0.233)	
Believes does not lose			0.343			0.347
			(0.113)			(0.133)
Controls: Incomes (piecewise continuous) estimated gains, sociodemo, other motives	✓	✓	✓	✓	✓	✓
Controls: Policy assigned	✓	✓	✓			
Subsample: [p10; p60] (A^T) or $\lvert\widehat{\gamma}\rvert < 50$ (A^F)	✓	✓	✓	✓	✓	✓
Effective F-Statistic	21.3	21.3	15.6	11.4	11.4	21.3
Observations	1,969	1,969	1,969	757	757	757
R^2	0.321	0.217	0.217	0.541	0.518	0.518

NOTE: See the results of the main specifications, Table 5.2 on page 34. As in the latter table, the source of exogenous variation in the belief used in first stages for the targeted dividend is the random assignment of the income threshold, which determines eligibility for the dividend. The first stage for the non-targeted dividend instead exploits the discontinuity in the win/lose feedback when the net gain switches from negative to positive.

Table E.2: Effect of self-interest on acceptance: the role of income

	Acceptance of Tax & Targeted Dividend (A^T)				
	(1)	(2)	(3)	(4)	(5)
Believes does not lose (G^T)	0.773	0.556	0.549	0.535	0.502
	(0.222)	(0.133)	(0.133)	(0.133)	(0.130)
Income above 35th percentile ($\mathbb{1}_{I>p35}$)	0.343				
	(0.508)				
$G^T \times \mathbb{1}_{I>p35}$	−0.392				
	(0.311)				
Initial policy Acceptance (A^0)	0.387	0.353	0.354	0.356	0.359
	(0.058)	(0.041)	(0.041)	(0.041)	(0.040)
Percentile with additional income slope change		30	40	50	60
Controls: Incomes (piecewise continuous)	✓	✓	✓	✓	✓
estimated gains, sociodemo, other motives					
Subsample: [p10; p60]; Controls: Policy assigned	✓	✓	✓	✓	✓
Effective F-statistic	5.5	15.3	15.2	15.2	16.1
Observations	1,969	1,969	1,969	1,969	1,969
R^2	0.571	0.321	0.321	0.321	0.321

NOTE: See the results of the main specifications, Table 5.2 on page 34. The source of the exogenous variation in the belief used in the first stage is the random assignment of the income threshold, which determines eligibility for the dividend.

Table E.3: Effect of believing in environmental effectiveness on support: second stages of alternative specifications

	Initial Tax & Dividend			
	Acceptance (A^0)			
	LIML	*OLS*	*IV*	*OLS*
	(A1)	(A2)	(A3)	(A4)
Environmental effectiveness: "Yes"	0.643	0.367		
	(0.320)	(0.020)		
Environmental effectiveness: not "No"			0.479	0.413
			(0.230)	(0.015)
Instruments: info E.E. & C.C.	✓		✓	
Controls: Socio-demo, other motives	✓	✓	✓	✓
Effective F-Statistic			6.0	
Observations	3,002	3,002	3,002	3,002
R^2	0.295	0.295	0.218	0.379

NOTE: Standard errors are reported in parentheses. The list of controls can be found in Appendix F, and the main results in Table 5.4 on page 37. As in the latter Table, the dependent variable corresponds to either initial approval (answer "Yes" to support of the policy) or acceptance (answer not "No"). The first stage exploits the information randomly displayed about climate change (C.C.) and the effectiveness of carbon taxation (E.E.) as exogenous instruments.

F Control variables

Sociodemo: *respondent's income, household's income, sex, age* (5 categories)*, employment status* (9 categories)*, socioprofessional category* (8 categories)*, region of France* (10 categories)*, size of town* (5 categories)*, diploma* 4 categories*, household size, number of people above 14, number of adults, number of c.u., income per c.u., smokes, favored media for news* (5 categories)*.*

Politics: *extreme left, left, center, right, extreme right, interest in politics* (3 categories)*, conservative, liberal, humanist, patriot, environmentalist, apolitical.*

Political leaning: *extreme left, left, center, right, extreme right, indeterminate.*

Energy: *heating mode* (collective vs. individual)*, heating energy* (7 categories)*, annual distance travelled, fuel economy, diesel* (binary)*, gasoline* (binary)*, number of vehicles.*

Incomes: *income of respondent, income of the second adult, income of respondent squared, income of the second adult squared, dummy for absence of second adult.*

Incomes (piecewise continuous): *income percentile of respondent* (I_1)*, income percentile of the second adult* (I_2)*, dummy for absence of second adult,* $\min(I_1 - 20, 0)$, $\min(I_1 - 70, 0)$, $\min(I_2 - 20, 0)$, $\min(I_2 - 70, 0)$.

Estimated gains: *simulated net gain, squared simulated gain.*

Online Appendix

G Questionnaire (*For online publication*)

Information intervention

1. [No information] Welcome to this survey.
   It was conceived by two social science researchers. It lasts about 15-20 minutes.

2. [Info PM] Welcome to this survey.
   It was conceived by two social science researchers. It lasts about 15-20 minutes.

   Before starting, please read carefully the information below on particulate matter pollution:

   - particulate matter is responsible for 48,000 deaths in France each year;
   - particulate matter reduces the life expectancy of French people by 9 months;
   - reducing fuel consumption would reduce the health problems associated with particulate matter.

   Source: France Public Health Report (2016)

3. [Info CC] Welcome to this survey.
   It was conceived by two social science researchers. It lasts about 15-20 minutes.

   Please read carefully the information below on climate change.

   - Climate change is already responsible for 150,000 deaths annually.
   - If greenhouse gas emissions continue on their current trend, the average global temperature increase will be +5°C in 2100 and +8°C in 2250.

- A rapid transition to renewable energies is technically possible and would contain global warming at +2°C.

According to scientists, in the absence of ambitious measures:

- a large proportion of species face an increased risk of extinction,
- natural disasters will intensify (hurricanes, heat waves, droughts, floods, forest fires, etc.);
- by 2100, 270 million more people would be flooded each year due to sea-level rise;
- violent conflicts and migration flows can be expected to increase.

Sources: Burke et al (2009), Hinkel et al (2014), IPCC Report (2014), Meinshausen et al (2011), Patz et al (2005)

Sociodemographics

4. What is your postal code?

5. What is your gender (in the sense of civil status)?
   Female; Male

6. What is your age group?
   18 to 24 years old; 25 to 34 years old; 35 to 49 years old; 50 to 64 years old; 65 years old or more

7. What is your employment status?
   Permanent; Temporary contract; Unemployed; Student; Retired; Other active; Inactive

8. What is your socioprofessional category? (Remember that the unemployed are active workers).
   Farmer; Craftsperson, merchant; Independent; Executive; Intermediate occupation; Employee; Worker; Retired; Other Inactive

9. What is your highest degree?
   No diploma; Brevet des collèges; CAP or BEP [secondary]; Baccalaureate; Bac +2 (BTS, DUT, DEUG, schools of health and social training, etc.); Bac +3 (licence...) [bachelor's]; Bac +5 or more (master's, engineering or business school, doctorate, medicine, master, DEA, DESS, etc.)

10. How many people live in your household? Your household includes: you, your family members who live with you, and your dependents.

11. What is your net **monthly** income (in euros)? **All income** (before withholding tax) is included here: salaries, pensions, allowances, APL [housing allowance], land income, etc.

12. What is the net **monthly** income (in euros) **of your household**? **All income** (before withheld taxes) is included here: salaries, pensions, allowances, APL [housing allowance], land income, etc.

13. In your household, how many people are 14 years old or older (**including yourself**)?

14. In your household, how many people are over the age of majority (**including yourself**)?

Energy characteristics

15. What is the surface area of your home? (in m²)

16. What is the heating system in your home?
    Individual heating; Collective heating; PNR (Don't know, don't say)

17. What is the main heating energy source in your home?
    Electricity Town gas; Butane, propane, tank gas; Heating oil; Wood, solar, geothermal, aerothermal (heat pump); Other; PNR (Don't know, don't say)

18. How many motor vehicles does your household have?
    None; One; Two or more

19. [Without a vehicle] How many kilometers have you driven in the last 12 months?

20. [One vehicle] What type of fuel do you use for this vehicle?
    Electric or hybrid; Diesel; Gasoline; Other

21. [One vehicle] What is the average fuel economy of your vehicle? (in Liters per 100 km)

22. [One vehicle] How many kilometers have you driven your vehicle in the last 12 months?

23. [At least two vehicles] What type of fuel do you use for your main vehicle?
    Electric or hybrid; Diesel; Gasoline; Other

24. [At least two vehicles] What type of fuel do you use for your second vehicle?
    Electric or hybrid; Diesel; Gasoline; Other

25. [At least two vehicles] What is the average fuel economy of all your vehicles? (in Liters per 100 km)

26. [At least two vehicles] How many kilometers have you driven in all your vehicles in the last 12 months?

Partial reforms [transport / housing]

27. Do you think that an increase in the VAT would result in a loss of more purchasing power for your household than for the average French household?
    Yes, much more; Yes, a little more; As much as the average; No, a little less; No, a lot less; PNR (Don't know, don't say)

28. Do you think that an increase in [fuel taxes / taxes on gas and heating oil] would cause your household to lose more purchasing power than an average French household?
    Yes, much more; Yes, a little more; As much as the average; No, a little less; No, a lot less; PNR (Don't know, don't say)

29. The government is studying a fuel tax increase whose revenues would be redistributed to all households, regardless of their income. This would imply:

- [an increase in the price of gasoline by 11 cents per liter and diesel by 13 cents per liter / a 13% increase in the price of gas, and a 15% increase in the price of heating oil];
- an annual payment of [60 / 50]€ to each adult, or [120 / 100]€ per year for a couple.

In terms of purchasing power, would your household be a winner or a loser with such a measure?
Winner; Unaffected; Loser

30. [*Winner* selected] **According to you, your household's purchasing power would increase:**
From 0 to [10·c.u.] € per year; From [10·c.u.] to [20·c.u.] € per year; From [20·c.u.] to [30·c.u.] € per year; From [30·c.u.] to [40·c.u.] € per year; More than [40·c.u.] € per year

31. [*Loser* selected] **According to you, the purchasing power of your household would decrease by:**
From 0 to [15·c.u.] € per year; From [15·c.u.] to [40·c.u.] € per year; From [40·c.u.] to [70·c.u.] € per year; From [70·c.u.] to [110·c.u.] € per year; From [110·c.u.] to [160·c.u.] € per year; More than [160·c.u.] € per year

32. If fuel prices increased by 50 cents per liter, by how much would **your household** reduce its fuel consumption?
0% - [I already consume almost none / I am already not consuming]; 0% - [I am constrained on all my trips / I will not reduce it]; From 0% to 10%; From 10% to 20%; From 20% to 30%; More than 30% - [I would change my travel habits significantly / I would change my consumption significantly]

33. In your opinion, if [fuel prices increased by 50 cents per liter / gas and heating oil prices increased by 30%], by how much would **French people** reduce their consumption on average?

From 0% to 3%; From 3% to 10%; From 3% to 10%; From 10% to 20%; From 20% to 30%; More than 30%

34. Do you think that an increase in taxes on gas and heating oil would cause your household to lose more purchasing power than the average French household?
Yes, a lot more; Yes, a little more; As much as average; No, a little less; No, a lot less; PNR (Don't know, don't say)

Tax & dividend: initial

35. The government is studying an increase in the carbon tax whose revenues would be redistributed to all households, regardless of their income. This would imply:

- an increase in the price of gasoline by 11 cents per liter and diesel by 13 cents per liter;
- an increase of 13% in the price of gas, and 15% in the price of heating oil;
- an annual payment of 110€ to each adult, or 220€ per year for a couple.

In terms of purchasing power, would your household win or loser with such a measure?
Win; Be unaffected; Lose

36. [*Winner* selected] **According to you, your household's purchasing power would increase by:**
From 0 to [20·c.u.] € per year; From [20·c.u.] to [40·c.u.] € per year; From [40·c.u.] to [60·c.u.] € per year; From [60·c.u.] to [80·c.u.] € per year; More than [80·c.u.] € per year

37. [*Loser* selected] **According to you, the purchasing power of your household would decrease by:**
From 0 to [30·c.u.] € per year; From [30·c.u.] to [70·c.u.] € per year; From [70·c.u.] to [120·c.u.] € per year; From [120·c.u.] to [190·c.u.] € per year; From [190·c.u.] to [280·c.u.] € per year; More than [280·c.u.] € per year

38. [[empty] / Scientists agree that a carbon tax would be effective in reducing pollution.] Do you think that such a measure would reduce pollution and fight climate change?
Yes; No; PNR (Don't know, don't say)

39. In your opinion, which categories would lose [[blank] / purchasing power] with such a measure? (Several answers possible)
No one; The poorest; The middle classes; The richest; All French people; Rural or peri-urban people; Some French people, but not a particular income category; PNR (Don't know, don't say)

40. In your opinion, what categories would gain purchasing power with such a measure? (Several answers possible)
No one; The poorest; The middle classes; The richest; All French people; Urban dwellers; Some French people, but not a particular income category; PNR (Don't know, don't say)

41. Would you approve of such a measure?
Yes; No; PNR (Don't know, don't say)

Tax & dividend: after information

42. [Feedback] We always consider the same measure. As a reminder, it would imply:

- an increase in the price of petrol by 11 cents per liter and diesel by 13 cents per liter;
- an increase of 13% in the price of gas, and 15% in the price of heating oil;
- an annual payment of 110€ to each adult, or 220€ per year for a couple.

In five out of six cases, a household with the same characteristics as yours would **[win / lose]**. (The characteristics taken into account are: heating with [source] for a dwelling of [size] m²; [distance] km covered with an average consumption of [fuel economy] liters per 100 km).

Based on this estimate, do you now think that your household would be:
Winner; Unaffected; Loser

43. [Info on progressivity] On average, this measure would increase the purchasing power of the poorest households and decrease that of the richest who consume more energy.

    In view of this new information, do you think this measure would benefit the poorest?
    Yes; No; PNR (Don't know, don't say)

44. [No info on progressivity] Do you think this measure would benefit the poorest?
    Yes; No; PNR (Don't know, don't say)

45. In view of the above estimate, would you approve of such a measure?
    Yes; No; PNR (Don't know, don't say)

46. Why do you think this measure is beneficial? (Maximum three responses)
    Contributes to fight climate change; Reduces the harmful effects of pollution on health; Reduces traffic congestion; Increases my purchasing power; Increases the purchasing power of the poorest; Fosters France's independence from fossil energy imports; Prepares the economy for tomorrow's challenges; For none of these reasons; Other (specify):

47. Why do you think this measure is unwanted? (Maximum three answers)
    Is ineffective in reducing pollution; Alternatives are insufficient or too expensive; Penalizes rural areas; Decreases my purchasing power; Decreases the purchasing power of some modest households; Harms the economy and employment; Is a pretext for raising taxes; For none of these reasons; Other (specify):

Tax & targeted dividend

48. The government is studying an increase in the carbon tax whose revenues would be redistributed **to the [20 / 30 / 40 / 50]% of the poorest French people only**. This would imply:

- an increase in the price of gasoline by 11 cents per liter and diesel by 13 cents per liter;
- an increase of 13% in the price of gas, and 15% in the price of heating oil;
- an annual payment of [550 / 360 / 270 / 220]€ for each adult earning less than [780 / 1140 / 1430 / 1670]€ per month (welfare benefits included, before withholding tax);
- no compensation for the others.

We estimate that in your household, [number of recipients] persons would receive this payment.

In terms of purchasing power, would your household win or lose with such a measure?
Win; Be unaffected; Lose

49. Would you approve such a measure?
    Yes; No; PNR (Don't know, don't say)

Other questions The survey is completed by other attitudinal questions, treated in our companion paper, Douenne & Fabre (2020). Hereafter, we only describe the questions that are used in the present paper.

50. Please select "A little" (test to check that you are attentive).
    Not at all; A little; A lot; Completely; PNR (Don't know, don't say)

51. Do you smoke regularly? *Yes; No*

52. How much are you interested in politics?
    Almost not; A little; A lot

53. How would you define yourself? (Several answers possible)
    Extreme left; Left; Center; Right; Extreme right; Liberal; Conservative; Liberal; Humanist; Patriot; Apolitical; Environmentalist

54. How do you keep yourself informed of current events? Mainly through...
    Television; Press (written or online); Social networks; Radio; Other

55. What do you think of the Yellow Vests? (Several answers possible)
    I am part of them; I support them; I understand them; I oppose them; PNR (Don't know, don't say)

56. The survey is nearing completion. You can now enter any comments, comments or suggestions in the field below.

H Profile of the Yellow Vests (*For online publication*)

Table H.1: Positioning towards Yellow Vests, per category.

	Opposed	Understands	Supports	Is part	PNR
Extreme-left (2%)	6%	26%	51%	12%	5%
Left (20%)	17%	36%	36%	5%	7%
Center (13%)	49%	30%	15%	2%	6%
Right (16%)	40%	32%	20%	3%	6%
Extreme-right (9%)	11%	28%	47%	10%	5%
Indeterminate (40%)	19%	32%	30%	4%	13%
Liberal (5%)	48%	26%	18%	2%	6%
Conservative (2%)	22%	28%	30%	10%	11%
Humanist (11%)	21%	35%	29%	5%	10%
Patriot (8%)	21%	27%	39%	7%	6%
Apolitical (21%)	21%	31%	32%	4%	12%
Environmentalist (15%)	17%	39%	27%	5%	12%
Rural (21%)	20%	31%	34%	6%	9%
<20k (17%)	24%	28%	34%	6%	9%
20-100k (14%)	22%	33%	32%	4%	9%
>100k (31%)	29%	34%	26%	3%	8%
Paris (17%)	28%	33%	25%	4%	11%
No diploma or *Brevet* (30%)	21%	29%	34%	5%	10%
CAP or *BEP* (24%)	23%	28%	36%	6%	7%
Baccalauréat (17%)	22%	35%	29%	4%	11%
Higher (29%)	32%	21%	36%	3%	8%
Age: 18–24 (12%)	23%	34%	27%	4%	12%
Age: 25–34 (15%)	21%	33%	28%	7%	11%
Age: 35–49 (24%)	25%	32%	29%	5%	9%
Age: 50–64 (24%)	21%	32%	36%	4%	7%
Age: $\geq$ 65 (25%)	32%	30%	28%	3%	7%
Income decile: 1	25%	33%	26%	3%	14%
Income decile: 2	18%	31%	35%	5%	11%
Income decile: 3	17%	31%	32%	7%	12%
Income decile: 4	15%	33%	37%	6%	9%
Income decile: 5	21%	29%	36%	5%	8%
Income decile: 6	26%	33%	29%	6%	7%
Income decile: 7	25%	36%	28%	4%	7%
Income decile: 8	31%	31%	28%	3%	8%
Income decile: 9	39%	32%	20%	3%	6%
Income decile: 10	47%	29%	15%	3%	6%
Female (52%)	21%	34%	29%	5%	12%
Male (48%)	29%	30%	31%	5%	6%
Average	*25%*	*32%*	*30%*	*5%*	*9%*

NOTE: The percentages in parenthesis express the weighted share of each category from our sample. See discussion in the main text, Section 2.1 on page 7.

I Support rates for Tax & Dividend policies (*For online publication*)

Table I.1: Support for Tax & Dividend policies at different stages of the survey.

	"Would you approve of this reform?"		
	"Yes"	"No"	"PNR"
Initial stage (A^0)	10.4%	70.3%	19.3%
After feedback (A^F)	16.8%	63.0%	20.2%
Targeted dividend (A^T)			
bottom 20% (A^T)	19.1%	63.2%	17.7%
bottom 30%	15.0%	66.0%	19.0%
bottom 40%	17.3%	67.6%	15.1%
bottom 50%	12.8%	73.3%	13.9%
all	16.1%	67.6%	16.2%

NOTE: The table reads as follows: at the initial stage, 10.4% of the respondents approved of a Tax & Dividend. After receiving customized feedback (either win or lose), 16.8% of them approved it. When the dividend targets only people below the bottom 20% (to which the respondent or its spouse may be eligible or not), 19.1% of them approve it. Refers back to Paragraph 2.2.2.

J Heterogeneity in pessimism and motivated reasoning (*For online publication*)

J.1 Heterogeneity in pessimism

To better understand the determinants of the pessimistic updating of the win/lose category, we investigate the heterogeneity in updating. To measure *correct updating*, we define a variable U that equals $+1$ if the respondent adopts feedback that invalidates their initial belief, 0 if they do not update, or -1 if they initially felt *unaffected* but update as being against the feedback. Over the subsample of *invalidated* respondents who should have updated because their initial win/lose category is not aligned with our feedback ($g_i \cdot \widehat{\gamma}_i \leq 0$), we regress the

correct updating, U, on the initial belief not to lose, G^0, and a vector of characteristics, $\mathbf{C}$:

$$U_i = \delta_0 + \beta_U G_i^0 + \beta_{\mathbf{C}} \mathbf{C} + \varepsilon_i \quad \text{for } i : g_i \cdot \widehat{\gamma}_i \leq 0, \tag{15}$$

The high values for β_U reported in columns (1-3) of Table J.1 again prove that among those who should have updated, those who initially thought that they would win (the optimistic losers) update significantly more correctly than those who did not think so (the pessimistic winners). Beyond this asymmetry, columns (2-5) show that some respondent characteristics are correlated with correct updating. Relative to unemployed and inactive people, retired, active, and students are more likely to update correctly, the latter being 22 p.p. more likely to correctly revise their beliefs when they are invalidated than unemployed and inactive people (column 2). The categories of respondents who initially displayed the largest bias appear also to be less likely to update correctly. Indeed, people who are part of the Yellow Vests movement are 14 p.p. less likely to correctly update than people who oppose it, even when controlling for disapproval of the policy, which itself decreases the likelihood of correctly updating by 18 p.p. The reason why previous characteristics affect the likelihood of correctly updating remains unclear. It is possible that they are correlated with people's uncertainty about their net gains. Alternatively, the Yellow Vests' greater distrust of the government (documented in Algan et al., 2019) could also apply to information on policies provided by researchers. Finally, these results also indicate that motivated reasoning may be at play.

J.2 Motivated reasoning

The previous results suggest that conservatism in belief revision does not simply follow people's cognitive difficulties when dealing with Bayes' rule. The greater likelihood of correct updating for those who support the reform indicates that political views and identity can shape belief formation. Indeed, the more people oppose the tax, the less likely they are to correctly update, as shown in columns (2-5) of Table J.1. From columns (4-5), we also see that this result is entirely driven by the "pessimistic winners": the updating of people who wrongly think that they will win does not depend on their approval, which is another indica-

Table J.1: Heterogeneity in updating.

	Correct updating (U)				
	(1)	(2)	(3)	(4)	(5)
Constant	0.120	−0.036	−0.011	−0.073	0.707
	(0.012)	(0.190)	(0.192)	(0.192)	(1.007)
Winner, before feedback ($\dot{G}$)	0.695	0.551	0.563		
	(0.078)	(0.083)	(0.083)		
Initial tax: PNR (I don't know)		0.179	0.186	0.199	0.113
		(0.032)	(0.067)	(0.033)	(0.155)
Initial tax: Approves		0.176	−0.031	0.216	−0.162
		(0.046)	(0.115)	(0.049)	(0.185)
Diploma × Initial tax: PNR			−0.003		
			(0.025)		
Diploma × Initial tax: Approves			0.072		
			(0.037)		
Subjective gain (g)		0.0004	0.0004	0.001	−0.001
		(0.0002)	(0.0002)	(0.0003)	(0.004)
Subjective gain: unaffected ($g = 0$)		−0.127	−0.126	−0.208	−0.331
		(0.033)	(0.033)	(0.033)	(0.219)
Bias about gain ($g - \widehat{\gamma}$)		−0.00005	−0.0001	−0.001	−0.0003
		(0.0001)	(0.0001)	(0.0003)	(0.0002)
Diploma (1 to 4)		0.014	0.009	−0.001	0.148
		(0.013)	(0.014)	(0.013)	(0.078)
Retired		0.130	0.127	0.108	0.124
		(0.079)	(0.079)	(0.080)	(0.435)
Active		0.166	0.165	0.160	0.113
		(0.054)	(0.054)	(0.054)	(0.365)
Student		0.224	0.229	0.183	0.402
		(0.075)	(0.075)	(0.074)	(0.526)
Yellow Vests: PNR		−0.045	−0.047	−0.031	0.013
		(0.047)	(0.047)	(0.048)	(0.246)
Yellow Vests: understands		−0.065	−0.066	−0.059	0.141
		(0.034)	(0.034)	(0.034)	(0.170)
Yellow Vests: supports		−0.063	−0.063	−0.050	−0.156
		(0.036)	(0.036)	(0.036)	(0.206)
Yellow Vests: is part		−0.141	−0.142	−0.106	−0.985
		(0.061)	(0.061)	(0.063)	(0.367)
Includes "pessimistic winners"	✓	✓	✓	✓	
Includes "optimistic losers"	✓	✓	✓		✓
Controls: sociodemo, politics, estimated gains		✓	✓	✓	✓
Observations	1,365	1,365	1,365	1,265	100
R^2	0.055	0.144	0.146	0.115	0.696

NOTE: Omitted variables are *Unemployed/Inactive* and *Yellow Vests: opposes*. The list of controls can be found in Appendix F.

tion that the revision in beliefs is driven by a rejection of the tax. This is not to say that few people seek to reach accurate beliefs. It still could be the case that informing any respondent that they would win makes them revise their subjective gain by, say, 100€ upwards, leading only those with small subjective losses to discover that they would win.[31] One can actually

[31]Those with small subjective gains who discover that they lose would similarly correctly update. Those with large subjective losses would not update while virtually no respondent has a large subjective gain. This would explain the asymmetry in the updates.

see from the positive and statistically significant effect of *subjective gain (g)* that such an accuracy motive is at play. However, this effect remains small relative to those indicative of policy support, indicating the presence of another mechanism, such as motivated reasoning. Column (3) further shows that the effect of approving the policy on correct updating is even stronger for more educated people—since the interaction term between approval and diploma is positive and significant—even capturing the entire effect of initial policy approval.

The previous findings are comparable to empirical evidence from Kahan (2013) that politically motivated reasoning about climate change is not a reasoning deficiency but rather a reasoning adaptation following the interest that individuals have in conveying "their membership in and loyalty to affinity groups central to their personal well-being". In our case, the position relative to the Yellow Vests proxies for the groups that respondents identify with, and the differentiated updating along this spectrum can be interpreted as motivated reasoning. In addition, the hypothesis that motivated reasoning follows from a rational adaptation purpose can explain our finding that better educated people are *more* prone to motivated reasoning since they are better able to formulate specious reasoning and reconcile antagonistic information and ideas. To the best of our knowledge, this result is the first case for rational motivated reasoning in the context of climate policies, complementing the findings of Druckman & McGrath (2019) that this mechanism can explain polarization around beliefs on climate change.[32]

Building upon the cognitive and social mechanisms described by Kraft et al. (2015) and documented by, e.g., Redlawsk (2002), we hypothesize the following narrative as one of the possible channels through which aversion to the carbon tax became entrenched. The Yellow Vests first gathered to defend their interest (above all, their purchasing power), and a side effect of the daily interactions on roundabouts was to bring material and emotional support to

[32]This evidence provides empirical support for various models of endogenous belief formation. For example, Little (2019) formalizes the idea that directional motives may override accuracy motives and that people update their auxiliary beliefs (in our case, the win/lose category) to preserve their consistency with their core beliefs (here, rejection of the tax). Admittedly, one might expect the importance of accuracy motives relative to directional motivated reasoning to increase in a higher stakes environment. However, this hypothesis cannot be tested in our setup, and previous literature does not provide conclusive evidence on the matter (Kunda, 1990; Camerer & Hogarth, 1999).

the protesters (Challier, 2019). A group identity soon developed, which crystallized shared beliefs and affects such as a rejection of carbon taxation. This group identity gained support from a large majority of the population, notably through social networks. Now, due to the loyalty to the group and affects that have entered their subconscious, Yellow Vests supporters instinctively oppose any carbon tax and are prone to find excuses to cope with contradictory messages, e.g., by denying the reliability of these messages (Golman et al., 2016). Admittedly, such a narrative falls short of explaining the majority rejection among those who oppose the Yellow Vests (which may originate from pessimistic perceptions more than tax aversion), but it illustrates how pessimistic beliefs can be so persistent among Yellow Vests supporters.

Overall, the persistent pessimism is consistent with people forming their beliefs in a motivated way. Nevertheless, other mechanisms—such as a distrust of the government—may play a key role. Further research with a different design is needed to determine the relative importance of these different mechanisms.

K Relation between support and belief in progressivity (*For online publication*)

Specifications used As noticed in Section 5.3, the ambiguous responses to our information intervention on progressivity do not allow us to perform an IV estimation to identify the causal effect of this motive. To explore how respondents' beliefs about progressivity relate to their support for the policy, we therefore estimate simple OLS and logit regressions. Even though we control for many variables, including beliefs over other motives of support, we may suspect that the coefficients obtained remain biased by omitted variables or reverse causality. They should therefore be taken as partial correlations and not causal estimates.

We focus on the acceptance question *after information*, i.e., after asking whether the reform is progressive or not. Table K.1 presents the results of different regressions, depending on the set of controls and on the choice of variables. Columns (1)-(4) report regressions of

acceptance on the broad definition of motives of acceptance: answers *not "No"* to progressivity, effectiveness and *not "lose"* to win/lose category. On the contrary, columns (5)-(6) use strict definitions for both approval and the covariates, where only *"Yes"* (or *"win"*) answers activate the dummy variables.

Table K.1: Support of the Tax & Dividend in function of beliefs in each motive.

	Support (after information)					
	Broad definition of variables (*not "No"*)				Strict definitions (*"Yes"*)	
	OLS			*logistic*	*OLS*	
	(1)	(2)	(3)	(4)	(5)	(6)
Progressivity (P)	0.223	0.214	0.560	0.544	0.228	0.482
	(0.038)	(0.039)	(0.023)	(0.019)	(0.041)	(0.023)
Winner (G^1)	0.332	0.264			0.303	
	(0.020)	(0.018)			(0.019)	
Effective (E)	0.258	0.112			0.244	
	(0.023)	(0.021)			(0.020)	
$(G^1 \times E)$	0.127	0.054			0.126	
	(0.034)	(0.030)			(0.037)	
Interaction: winner $(P \times G^1)$	0.183	0.144			0.098	
	(0.050)	(0.044)			(0.048)	
Interaction: effective $(P \times E)$	0.172	0.090			0.281	
	(0.057)	(0.050)			(0.059)	
Income (I, in k€/month)	0.017	0.025			0.037	
	(0.022)	(0.019)			(0.018)	
Interaction: income $(P \times I)$		−0.009			−0.019	
		(0.012)			(0.014)	
$P \times G^1 \times E$	−0.400	−0.320			−0.314	
	(0.072)	(0.063)			(0.083)	
Initial policy Acceptance (A^0)		0.467				
		(0.016)				
Controls: Sociodemo	✓	✓			✓	
Observations	3,002	3,002	3,002	3,002	3,002	3,002
R^2	0.460	0.586	0.162		0.391	0.130

NOTE: Standard errors are reported in parentheses. For the logit, the average marginal effects are reported and not the coefficients. The list of controls can be found in Appendix F. The covariates and dependent variables refer either to broad (1-4) or strict (5-6) definitions of the beliefs, where strict dummies do not cover "PNR" or "Unaffected" answers. See the discussion in the main text, Section 5.3 on page 37.

Results On average, believing that the reform is *not regressive* is associated with a higher *acceptance* rate by 56 p.p. (column 3) while believing it is *progressive* is associated with a higher *approval* rate by 48 p.p. (6). However, when one introduces other motives of acceptance and their interactions as covariates, with households characteristics as controls, one observes that the effect of progressivity is lower: its marginal effect at the sample mean - i.e., accounting for the average marginal effect of interaction terms - is 27 p.p.[33] To disentangle the link between beliefs over net gains and progressivity, we also include the interaction between progressivity and income as a covariate (2, 5). Although the coefficient is negative, in accordance with intuition, the effect is small and not significant. Adding the powerful control of initial policy acceptance in column (2) has a negligible influence on the effect of progressivity of 24 p.p. (instead of 27 p.p.), which validates our choice of the preferred specification (1). Despite the powerful control, column (2) is not our preferred specification because the effect of environmental effectiveness is mostly captured by the covariate "initial policy acceptance" since the information intervention on climate change predated the initial question on acceptance. Finally, using the strict definitions of beliefs and approval yields a smaller correlation (6) but similar results when accounting for relevant controls (5), showing that the effects are not driven by a correlation between "PNR" answers. Overall, although these results are not causal, they suggest that the belief that the tax is progressive is associated with a higher support, all else equal.

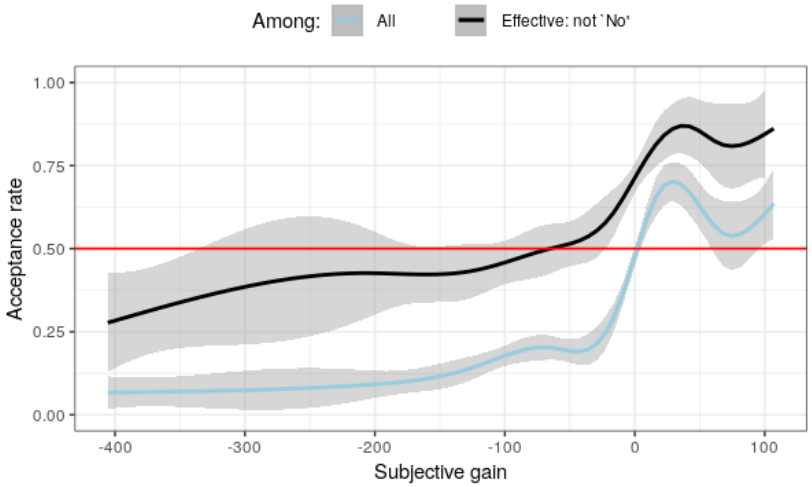
L Willingness to pay (*For online publication*)

For respondents who believe in the effectiveness of our Tax & Dividend, we are able to infer their willingness to pay (WTP) for climate mitigation by studying the acceptance rate as a function of subjective gain. We adopt a common practice in the literature and define the WTP as the monetary loss that the *median* agent is willing to incur (Hanemann, 1984). Figure L.1 indicates that this WTP is about 60€/year per c.u. since this corresponds to the

[33] Although these results are not causal, they show that 90% of those who believe in the three motives approve of the policy, along with 65-75% of those who believe in two of them.

subjective loss below when a majority accepts the policy. This WTP is computed only among people who believe that the tax is not ineffective since it would make little sense to assume that some people are willing to pay for an instrument that does not achieve its expected goal. Indeed, Figure L.1 shows that the "WTP" of the whole sample is zero, meaning that the median person accepts the policy only when they personally gain from it. Our method has several advantages. First, it can be interpreted as a willingness to accept as much as a willingness to pay because our instrument is neither framed as a good to buy nor as damage to be compensated for, and net gains do not distinguish cost increases from payments received. Second, our method is more akin to revealed preferences - and hence probably less biased (Murphy et al., 2005) - than previous ones because most studies directly ask respondents to select their preferred option for climate mitigation, be it using a contingent valuation method (Berrens et al., 2004; Cameron, 2005; Kotchen et al., 2013) or a discrete choice experiment (Longo et al., 2008; Alberini et al., 2018). Still, our estimation has two notable limitations relative to the literature: it relies on a non-representative subsample, and subjective gains are endogenous from acceptance.

To compare our estimation with those of the literature, expressed per household, we have to multiply our WTP by the average number of consumption units in households: 1.6. The WTP per household we get, 96€, lies in the typical range of the literature (Jenkins, 2014; Streimikiene et al., 2019), suggesting that the protests against carbon taxation encountered in France do not reflect specific preferences for environmental policies.



NOTE: The black curve indicates that a majority of those who did not answer "No" to the question on the effectiveness of the policy accepted the reform when their subjective gain was above -60€ per c.u. For the whole sample (blue curve), this majority acceptance is reached only when subjective gains are positive. This refers back to Section 2.2.

Figure L.1: Acceptance rate by subjective gain, informative of the willingness to pay for climate mitigation.

M Ensuring data quality (*For online publication*)

We took several steps to ensure the best possible data quality. We excluded the 4% of respondents who spent less than 7 minutes on the full survey. We confirm that our main results are robust to choosing another cutoff than 7 minutes (see Table M.1). In order to screen out inattentive respondents, a test for the quality of the responses was inserted, which asked respondents to select "A little" on a Likert scale. The 9% of respondents who failed the test were also excluded, which yields a final sample of 3,002 respondents. Also, when the questions about a reform were spread over different pages, we recalled the details of the reform on each new page. We checked for careless or strange answers on numerical questions, such as income or the size of the household. We flagged 10 respondents with aberrant answers to the size of the household (and capped it to 12) and up to 273 respondents with inconsistent answers, such as a household income smaller than individual income, or a fuel economy higher than 90 liters per 100 km. Being flagged and response time are not significantly

correlated with our variables of interest such as policy support or subjective gain (the correlation is always from -1% and 3%). An examination of the flagged answers suggests that these respondents simply misunderstood the question. Among these inconsistent answers, 58 respondents have answered more than 10,000€ as their monthly income (despite the word "monthly" being in bold and underlined), with answers in the typical range of French annual incomes. We have divided these figures by 12.

Table M.1: Robustness of main results to the exclusion of poor quality answers.

	Acceptance (A^T)			Correct updating (U)		
	all	> 11 min	not flagged	all	> 11 min	not flagged
Believes does not lose (.53)	0.526	0.547	0.558			
	(0.134)	(0.137)	(0.153)			
Winner, before feedback (.55)				0.542	0.532	0.553
				(0.083)	(0.085)	(0.091)
Initial tax: Approves (.18)				0.180	0.213	0.197
				(0.046)	(0.049)	(0.049)
Original regression: Table (column)	5.2 (1)	5.2 (1)	5.2 (1)	J.1 (2)	J.1 (2)	J.1 (2)
Effective F-statistic	15.2	14.5	11.8			
Whole sample size	2777	3165	2729	2777	3165	2729
Observations	1,978	1,825	1,826	1,370	1,261	1,242
R^2	0.320	0.318	0.326	0.142	0.150	0.155

NOTE: Two of our main results are checked on three alternative sampling restrictions: (1) inclusion of answers < 7 min, (2) exclusion of the 10% of answers < 11 min, and (3) exclusion of flagged (inconsistent) respondents. Weights have been recalculated for each sample. The estimates on the original sample are reported next to variable name. See the original Tables for more details. The correlation between our main variables of interest and response time or being flagged is always below 3%. Standard errors are reported in parentheses. This refers back to Section 2.2.2.