Can We Reconcile French People with the Carbon Tax? Disentangling Beliefs from Preferences*

Thomas Douenne and Adrien Fabre†

May 2019

Abstract

Using a new survey and National households' survey data, we investigate French perception over carbon taxation. We find that French people largely reject a tax and dividend policy where revenues of the tax would be redistributed uniformly. However, their perception about the properties of the tax are biased: people overestimate the negative impact on their purchasing power, wrongly think the scheme is regressive, and do not perceive it as environmentally effective. Our econometric analysis shows that correcting these three bias would suffice to generate majority acceptance. Yet, we find that people's beliefs are persistent and their revisions biased towards pessimism, so that only a few can be convinced.

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

Keywords: Climate Policy; Carbon tax; Bias; Beliefs; Preferences; France; Perceptions

*Acknowledgments: We are grateful to Mouez Fodha, Fanny Henriet and Katheline Schubert for their comments and their help to get funding. We also thank seminar participants at the Paris School of Economics and an anonymous referee of the FAERE Working Paper series. We would also like to show our gratitude to Stefanie Stantcheva, Michael I. Norton and the Harvard Business School for allowing us to use their Qualtrics account. 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 had initially committed to an ambitious trajectory for the price of carbon. 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 hit 86.2€/tCO_2 by 2022. Yet, 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. This dramatic shift in French climate policy is the direct consequence of the popular protest of the "Yellow Vests", which started against the carbon tax. Several factors contributed to increase public's discontent against the tax. The increasing revenues from the carbon tax were mostly used to fund the budget rather than redistributed to households; the policy was concomitant with other unpopular tax reforms, and with high barrel prices. The challenge for economists and policy makers is now to disentangle the determinants of the disapproval, and understand to what extent the rejection reflects intrinsic preferences or biased perceptions over the properties of carbon pricing.

In order to explain French attitudes towards carbon taxation, we conducted a survey on a representative sample of 3,002 French households. We focus on a "Tax & Dividend" carbon tax increase with uniform lump sum compensation,[1] which allows to specify clearly the distributive effects of the policy, at variance with the policy abandoned by the government. The reform is approved by only 10% of respondents and disapproved by 70% (the rest do not know or do not want to answer). We analyze the perceptions of three well-known determinants of acceptance of the carbon tax: the impact on one's purchasing power, the progressivity of the scheme, and its environmental effectiveness. We compare subjective beliefs elicited from our new survey to objective impacts on respondents' purchasing power computed using National households' survey data. This comparison shows that people largely overestimate the negative impact of carbon taxation on their purchasing power. 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 policy would reduce pollution and fight climate change. These results suggest that rejection of carbon taxation does not commonly result from clashing preferences, such as a disinterest for the ecology or a dislike for price instruments, but rather from biased beliefs about the properties of the reform. To show this point, we investigate the effects on acceptance of "rectifying" the beliefs over each of these properties. We identify the effect of correcting one's belief by instrumenting each belief

[1] Such policy has been notably advocated in The Wall Street Journal by 3354 U.S. economists (2019). For a more detailed analysis of this policy, see Metcalf & Weisbach (2009).

with a random treatment: a piece of information randomly displayed or not, or a compensation targeted randomly to create exogenous variations in eligibility. A first result is that displaying information convince only a few people. For example, among those advantaged by the reform who wrongly believe they would lose, only 12% are convinced that they would gain when we disclose them our estimation. Worse, respondents revise their beliefs in a biased way, giving more weight to new information when it shows they would lose from the reform, i.e. when it provides them arguments against the tax. That being said, beliefs in self-interest, environmental effectiveness and progressivity all have a large effect on acceptance: about 40 percentage points (p.p.) for the first two and 27 p.p. for the latter, other things equal. Besides, we find that those who believe the policy is effective have a median willingness to pay in the typical range of the literature, at 100€/year per household, suggesting that French preferences for climate are not specific. Finally, in a hypothetical scenario where all biased beliefs could be corrected, we estimate an approval rate of 90%, confirming that the rejection is driven by erroneous perceptions.

A growing literature studies attitudes towards climate policies, as this issue is becoming critical in the public debate. For a thorough review of this literature, we refer to Carattini et al. (2018), and also signal the less recent Dresner et al. (2006a) and the more synthetic Klenert et al. (2018). Connected literature includes behavioral and political obstacles to environmental policies (Millner & Ollivier, 2016) and willingness to pay for climate mitigation (Jenkins, 2014). Closely related to our paper, Stern et al. (1993) is an early work proposing and testing a model of preferences for environmental quality aimed at disentangling egoistic from altruistic motives on the one hand, and beliefs from values on the other hand. Among all possible attitudes, they show that belief about consequences on self-interest is the only predictor of the willingness to pay Pigouvian taxes. Using a post-electoral 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 and Swedish respondents respectively, Bristow et al. (2010) and Brannlund & Persson (2012) document a higher approval when the reform addresses distributional issues. Baranzini & Carattini (2017) report that a majority of the people they interviewed in Geneva do not believe 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) shows that self-interest matters for acceptance, but less than concerns for environmental effectiveness or distributional effects.

In the present paper, we also study how these three motives affect acceptance. We contribute to the literature by providing evidence for causal effects where past studies essentially show correla-

tions, often relying on proxies such as fuel consumption to proxy self-interest (e.g. Thalmann, 2004; Kallbekken & Sælen, 2011). At variance with such shortcuts, we do not assume that people are fully rational nor have perfect information. Thus, our methodology allows disentangling erroneous beliefs from *pure* effects of preferences. The paper also quantifies biases regarding the costs of the carbon tax. To our knowledge, this study is the first to compare subjective beliefs about the private cost of the policy and objective data using micro-simulation, showing that 89% of people overestimate their private costs. We show evidence that these biases do not only explain the policy rejection, but also follow from it through motivated reasoning (Kunda, 1990). Another contribution of this paper is to show the persistence of people's beliefs over carbon tax policies, and the asymmetry and further pessimistic bias in their revision. We provide evidence on how these biases translate in attitudes towards policies, which echoes recent advances in the literature on beliefs' formation where beliefs enter directly utility (e.g. Bénabou & Tirole, 2002), leading to a selective processing of new information (Eil & Rao, 2011; Sharot et al., 2011).

The rest 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 measure the bias regarding the impacts of carbon taxation. In section 4, we highlight new biases in the way individuals revise these beliefs. In section 5, we estimate the effects on acceptance of rectifying biased beliefs. Section 6 concludes. Further results and methodological complements are reported in Appendix.

2 Survey and data

2.1 Our survey, "Beliefs & climate policies"

2.1.1 Survey data collection

The 3002 responses were collected in February and March 2019 through the survey company Bilendi. This company maintains a panel of French respondents whom they can email with survey links. Respondents are paid 3€ if they fully complete the survey. The respondents who choose to respond are first channeled through some screening questions that ensure that the final sample is representative along six socio-demographic characteristics: gender, age (5 brackets), education (4), socio-professional category (8), size of town (5) and region (9). The quotas are relaxed by 5% to 10% relative to actual proportions. Table A.1 in Appendix A shows that our sample is still extremely representative. Nonetheless, observations are weighted to correct small differences between sample

and population proportions. The median time for completion of the survey was 19 minutes. We made sure that all questions requiring some concentration were in the first half of the survey.

2.1.2 The survey

The full survey in French can be seen on-line.[2] It contains several random branches and treatments that are independent from one another: Figure 2.1 presents in a diagram the sequence of treatments (represented by ellipses) and questions (boxes).

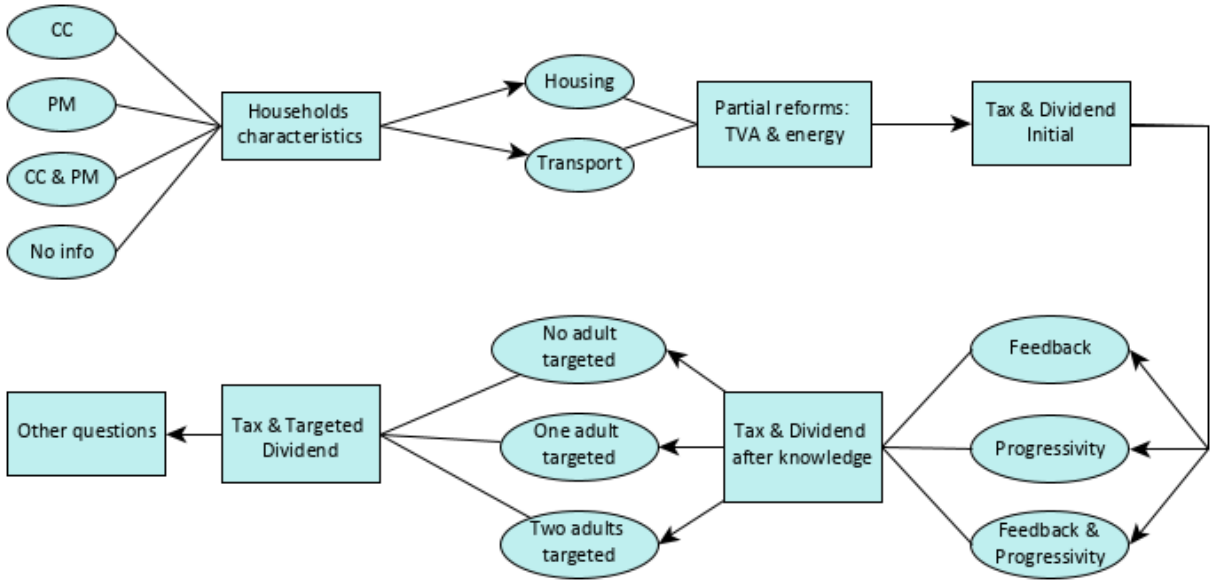


Figure 2.1: Diagram of the sequence of treatments (ellipses) and questions (boxes).

Priming on environmental issues Two blocks of information are randomly displayed or not: one on climate change and the other on particulate matter. This priming divides the sample in four groups, who receive either one block of information, the other, none or both of them. Climate change information includes temperature trends for long run future, worrying facts on current and expected impacts, and a claim that keeping global warming below 2 °C is technically feasible. Particulates information consists in the estimated impact on French mortality (48,000 deaths per year), life expectancy (9 months less), and the assertion that reducing fuel consumption would improve health. The time spent on each block is saved, and links to scientific references are provided to support the information.

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

Household characteristics In addition to the quotas strata, socio-demographic characteristics include zip code, household structure, income of the respondent and of their household. A block on energy usage contain questions that allow us to estimate the impact of a carbon tax increase on housing expenditures: surface of accommodation, heating type (collective or individual) and energy source; as well as on transport expenditures: number of vehicles, type(s) of fuel, distance traveled last year and average fuel economy. The distributions of answers are much in-line with official statistics, as shown in Table A.2 in Appendix A.

Partial tax reform One partial tax reform is randomly allocated to the respondent: it consists in an increase in the carbon tax by 50€/tCO_2 specific either to heating fuel and gas, or to gasoline and diesel. Partial reforms on housing and transport feature the same string of questions. We first ask whether the respondent thinks their household would lose more or less than average in case of an unspecified increase in indirect taxation, both for the VAT tax and for the energy considered (housing or transport). Then, we specify in a new block that the revenues of the tax would be distributed equally to each adult, entailing a yearly transfer of 50€ (resp. 60€) for the partial tax on housing (resp. transport).[3] We also provide the price increases implied by the tax: +13% (resp. +15%) for gas (resp. domestic fuel) on the one hand, and +0.11€ (resp. +0.13€) for a liter of gasoline (resp. diesel) on the other. We make clear at the beginning that this survey is conducted by two researchers in social sciences, but we present the policy starting with "The government studies..." to capture the effect of distrust in 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. Depending on their answer, we further ask them to estimate their expected gain (or loss) among 5 (or 6) intervals. The interval thresholds are tailored to each respondent, as they are computed in proportion of the number of consumption units of their household (as defined by Eurostat). Finally, respondents are asked to estimate their own elasticity as well as that of French people. To this end, we borrow the phrasing of Baranzini & Carattini (2017), and ask the expected decrease in consumption that would follow a 30% increase in the price of heating (or equivalently, an increase of 0.50€/L in fuel prices), among 5 brackets.

[3]We chose to redistribute per adult instead of consumption unit to make the scheme more understandable. We limited the number of beneficiaries to two per household to come closer to current welfare benefits that depend on the number of consumption units.

Tax & Dividend

Initial perceptions Our main reform of interest is an increase by 50€/tCO_2 of the French carbon tax, that concerns both housing and transport.[4] The revenues generated are redistributed equally, so that each adult receives a yearly lump sum compensation of 110€ (this figure was obtained using typical elasticities, see 2.2.1). After describing the reform, a first block of questions elicits the respondent's perceptions. Their subjective gain is asked in the same manner as for the partial tax. The priming that "scientists agree that a carbon tax would be effective in reducing pollution" is randomly displayed or not before asking whether the reform would be effective in reducing pollution and fighting climate change. Then, respondents are asked to pick the categories of losers and winners with the reform. Finally, we ask: "Would you approve this reform?" and let the respondent choose between "Yes", "No" and "PNR (I don't know, I don't want to answer)". In the following, we say that a respondent *approves* a reform if they respond "Yes", and that they *accept* the reform if they do *not* respond "No". Given the low rates of approval, we study primarily the acceptance to get tighter confidence intervals. We also provide results for approval, and obtain similar findings.

Opinion *after knowledge* To test the persistence of beliefs and measure the importance of self-interest and fairness motives in the acceptance of the reform, we provide some information on the effect of the reform. To a random half of the sample, 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 two-thirds of the respondents (the remaining half plus one third of the respondent with that priming on *progressivity*), we tell that: "In five cases over 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 traveled with an average consumption of [fuel economy] L for 100 km.)". Indeed, section 2.2.2 shows that we estimate correctly if a household wins or lose in 83% of cases.

As most respondents do not believe the claim that the reform was progressive, we ask to the second half of the sample whether they think it is, to analyze the effects of priming on compliant respondents. We also ask again the winning category (i.e. if the respondent's household would win, lose or be unaffected by the reform) and the approval to the reform. Finally, we let the respondent pick the reasons why this reform seems beneficial, and undesirable. We report these results in a

[4]Electricity is exempted from the French carbon tax "Contribution Climat Energie" (CCE) as it is already taxed on the EU-ETS, and has a low carbon content due to the large share of nuclear power.

companion paper, Douenne & Fabre (2019), where we also study the socio-demographic determinants for approval.

Tax & targeted dividend In this block, we ask for the winning category and for the approval of four alternative reforms. Each respondent deals with just one of them, which differs from the main reform only in the way revenues are recycled. Here, the payments, still equal among recipients, are targeted to adults whose income is below some threshold. The four thresholds correspond to the bottom 20%, 30%, 40% and 50% of the income distribution. They are computed using inflated deciles of individual income from the *Enquête sur les Revenus Socio-Fiscaux* (ERFS 2014) produced by Insee (the French national statistics bureau).[5] Respondents whose income lies between two thresholds are allocated randomly to a reform featuring one of them. 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 above 2220€/month (which is the 70th. percentile), the reform they face is determined by the income of the households' second adult. Finally, when both (or the only one) adults in the household earn more than 2220€/month, their reform is allocated randomly between the four variants. Table 2.1 describes the targeted reforms and the proportion of respondents allocated to each of them, along with the proportion one would expect from the *ERFS*. The two sets of figures match almost perfectly, indicating that our sample is representative along the income dimension.

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

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*

Other questions We do not detail the other questions of the survey, because we analyze them in a companion paper, Douenne & Fabre (2019). We scrutinize opinions on environmental policies, including other ways to recycle the revenues of a carbon tax. We measure the knowledge and perceptions of climate change; ask some specific questions over shale gas, over the influence of climate change on the choice to give birth and on willingness to change one's lifestyle. We study the

[5] Incomes entitled to the household rather to its members, such as certain welfare benefits, are divided equally among the two oldest adults of the household.

use, availability and satisfaction with public transportation and active mobility. We ask political preferences, including the positioning in relation to the Yellow Vests. Finally, we let the respondent express any comment in a text box.

2.1.3 Ensuring data quality

We took several steps to ensure the best possible data quality. We excluded the 4% of respondents who spend less than 7 minutes on the full survey. We confirm that response time is not significantly correlated with our variables of interest (such as approval or subjective gain). In order to screen out inattentive respondents, a test of quality of the responses was inserted, which asked to select "A little" on a Likert scale. The 9% of respondents who failed the test were also excluded from our final sample of 3002 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. An examination of these answers shows no significant correlation with our variables of interest, and suggests that these respondents have simply mistaken 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 incomes. We have divided these figures by 12.

2.1.4 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, while we put a dot on dummy variables for *"Yes"*. For example, acceptance is denoted A while approval is denoted $\dot{A}$. Furthermore, for questions that are asked several times, namely acceptance and winning category, an exponent is added to specify the step at which the question is asked. Table 2.2 describes these exponents as well as the notations corresponding to the different notions of gain that we use. Uppercase is used for binary and lowercase for continuous variables, Greek letters denote objective notions, with a hat for our estimation of gains and without for the true (unknown) ones. To give another example, the broad notion of self-interest at the initial step, i.e. the belief that one does not

lose, is denoted G^I, and the strict belief that one wins at Tax & Targeted dividend is denoted $\dot{G}^T$.

Table 2.2: Notations for the different reforms and for gain notions.

Step	Initial	After knowledge		with Targeting
Variants	–	Progressivity	Feedback	–
Exponent	I	P	F	T

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

2.2 Official households surveys

2.2.1 Eliciting objective aggregates and distributions

One of the goals of the paper is to compare respondents' perceptions with actual impact of a carbon tax on households' purchasing power. For this purpose, we use the database constructed by Douenne (2018) whose objective was to estimate the distributive effects of a carbon tax for French households. This database matches two households surveys produced by Insee: the consumer survey *Budget de Famille* (BdF 2011) and the transport survey *Enquête Nationale Transports et Déplacements* (ENTD 2008).

Consumer survey "Budget de Famille" The consumer survey *Budget de Famille* (BdF 2011) is a household survey providing information over all households' revenues and expenditures, together with many socio-demographic characteristics. It was conducted in several waves from October 2010 to September 2011, over a representative sample of 10,342 French households. The main advantage of BdF when studying the incidence of carbon taxation is that expenditures in both housing and transportation energies are reported. Consumption of housing energies is taken from households' bills, and for most other goods respondents report their expenditures over the past week. As explained in Douenne (2018), this data collection is problematic when looking at the incidence of a tax on transportation energies, as short-run fluctuations in consumption lead to overestimate the variation in expenditures.

Transport survey "Enquête Nationale Transports et Déplacements" To overcome this limitation, BdF is matched with the transport survey *Enquête Nationale Transports et Déplacements* (ENTD 2008). ENTD was conducted in several waves from April 2007 to April 2008, over a representative sample of 20,178 French households. It provides information on households characteristics, their vehicle fleet and use over the past week, but most importantly it gives information on annual distances travelled with these vehicles. This last information enables to recover the distribution of transport fuel expenditures without over-estimating its spread. Such a matching is not necessary for housing energies as these already represent consumption over long periods in BdF. Finally, data from National Accounts are used to make the data representative of the year 2017.

Computing tax incidence and revenues From this dataset, we are able to compute the distribution of households' *objective* net gains in purchasing power after the policies proposed. We also use it to determine the total tax revenue to be redistributed lump sum. The formulas used for these computations are given in Appendix B.1. Our computations use typical elasticities found in the literature on French households: -0.4 for transport and -0.2 for housing, as well as an incidence borne at 80% by consumers.[6]

2.2.2 Computing households expected net gains

Simulating expected gains for the feedback Households are asked about yearly distance traveled and average fuel consumption of their private vehicles. From this information, it is then possible to compute the expected cost of a carbon tax on transport fuels. For housing energies, the collected information does not enable such simple calculations. Instead, we use the housing survey *Enquête Logement* (EL 2013) that provides information on households expenditures in housing energies as well as many demographic and energetic characteristics. The survey was conducted between June 2013 and June 2014 over a sample of 27,137 households in metropolitan France. This enables to compute the expected cost from the carbon tax on housing energies, and regress it on households' characteristics. The results are provided in Appendix B.2, where they are shown to be as least as accurate to the ones obtained from alternative prediction methods and specifications. The distribution of housing energy expenditures is very comparable to the one of BdF. Adding this estimated cost to the one simulated for transportation energies, subtracting the amount received in

[6] These values correspond to the short run uncompensated price elasticities estimated by Douenne (2018), and are in line with previous findings on French households (e.g. Clerc & Marcus, 2009; Bureau, 2011).

lump sum transfers, and dividing by the number of consumption units, we estimate the impact of the policy on the purchasing power of each household.

Assessing feedback's accuracy The previous estimation could have also been conducted with BdF data. Still, running this estimation on the housing survey enables to test the accuracy of our prediction out-of-sample: for each household in BdF, we compute the impact of the policy on their purchasing power and compare it to the prediction from our simulations. Figure C.1 in Appendix C.2 shows how the probability that our prediction is correct depends on objective gains. For five households out of six, we correctly predict whether their purchasing power would increase or decrease through the policy. This ratio is made symmetrical: among households in BdF predicted to win, 83.4% were actual winners, while among those predicted to lose, 83.4% were actual losers. Since the simulation and its test are done on different households, there is no concern for over-fitting: the probability that our feedback is correct is not overestimated. 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.

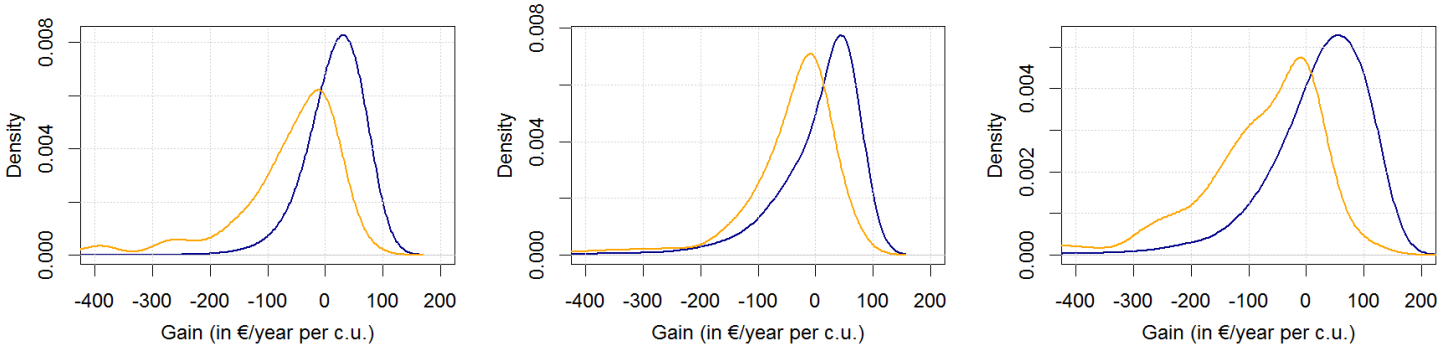
3 Perceptions

3.1 Self-interest

Over-estimation of policy costs Figure 3.1 plots the kernel density of expected net gains for objective data from BdF, and subjective beliefs from our survey. Figure 3.2 compares the CDF of objective vs. subjective net gains.[7] It clearly appears from these figures that on average, respondents overestimate the cost of the policy, even in the extreme case of perfectly inelatic expenditures. This result holds both for the carbon tax and for partial carbon taxes on transports and housing energies. The average net gains from the carbon tax on transport, housing and both, are respectively 18€ per consumption unit (c.u.), 6€ per c.u. and 24€ per c.u. from BdF data. Extrapolating from our survey, we instead find average subjective net gains of respectively −61€, −43€ and −89€. The median gap of 116€ between objective and subjective gains indicates a substantial bias towards loss

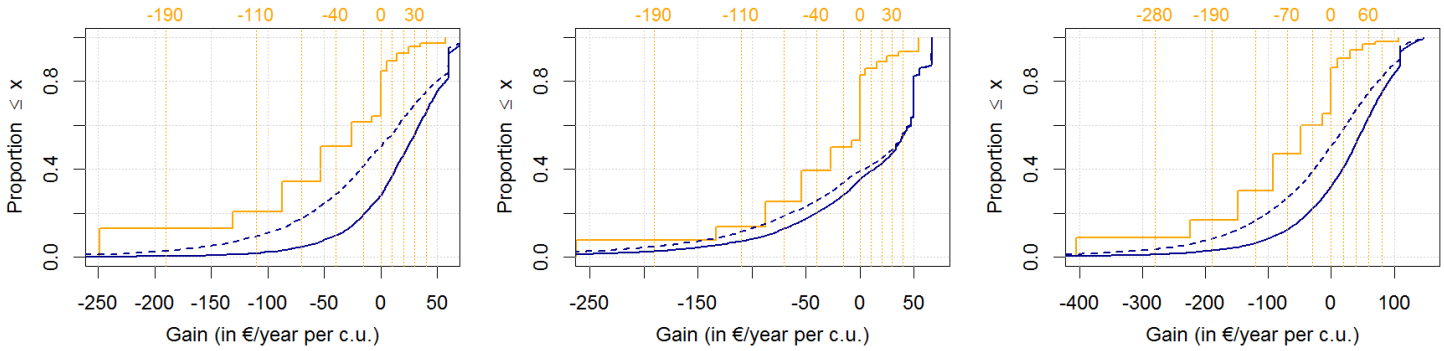
[7]The subjective intervals are translated into numerical values, assuming that the distribution within each interval are the same than for BdF data. Within each bin, we draw values that match the actual distribution for the PDF, while we simply take the actual average for the CDF. Among the several methods that we tried to assign numerical values, all realistic ones yield identical results, and we find an overestimation of policy costs even in the most conservative one (taking the maximal bounds of intervals).

from typical respondents. This bias is common, as we find that 89% of respondents underestimate their gain of purchasing power using our households' specific estimation. This proportion remains as high as 77 % when assuming inelastic expenditures, which provides a lower bound on the share who underestimate their net gain in utility. In other words, while 70% of households should benefit (in monetary terms) from the compensated carbon tax, only 14% think they would (and 22% see themselves unaffected).[8]



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

Figure 3.1: Distribution of objective (dark blue) vs. subjective (orange) net gains from our Tax & Dividend.



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

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

Figure 3.2: CDF of objective (dark blue) vs. subjective (orange) net gains from our Tax & Dividend.

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

Heterogeneity in bias In order to characterize profiles of individuals more likely to mis-perceive their gains, we regress mis-perception over many respondents' characteristics. Mis-perception is defined as a gap between objectively estimated and subjective net gains beyond 110€ per c.u., because our estimation differs from true objective gain by more than 110€ in only 5% of cases. This definition insures that the 55% of respondents with a mis-perception have in fact a large bias. Besides, other definitions for the bias yield very similar results. The results given in Table 3.1 show that mis-perception is largely idiosyncratic: controlling for a large set of variables[9] (column 1), the R^2 remains small (0.06). Still, we identify several variables having a significant effect on mis-perception. In particular, females appear to mis-perceive their gain 4 percentage points (p.p.) more on average, other things equal. Ecologists on the other hand are about 6 p.p. less likely to display a large bias. Interestingly, while the standard left/right political leaning has no significant effect, the position towards the Yellow Vests appears as the most critical determinant in mis-perception. Relative to respondents who declared to be opposed to the movement, those who declared to "understand", "support" or "be part" of it are more likely to mis-perceive their gains. This effect is increasing with the degree of adhesion, up to 20 p.p. for individuals who declared to be part of the movement. These last two results (difference with ecologists and supporters of Yellow Vests) suggest that mis-perceptions could be explained by "motivated reasoning" (Kunda, 1990), a tendency to form beliefs in accordance with preferences. This hypothesis is supported by the results from column (3), where position towards the policy are included: people who approve the policy are 28 p.p. less likely to mis-perceive their gains relative to those who do not accept it, and 10 p.p. less relative to those who do not know. We can think that the degree of support of the policy is what determines most of the bias (explaining e.g. why *ecologist* loses its explanatory power when we control for the support), and that the Yellow Vests variables remain significant only because they capture different *degrees* of rejection of the tax (which our Yes/No question cannot do).

Overall, typical biases are large and closely related to one's convictions. However, causality between beliefs and rejection can go in both directions. Section 4 provides evidence that some people think they lose because they oppose the tax, while section 5 shows that perceived outcomes strongly influences approval.

[9]The control variables used throughout the paper are described in Appendix E.

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)
Sex: Female	0.036^{*}	0.030	0.042^{**}
	(0.020)	(0.020)	(0.019)
Ecologist	-0.064^{**}	-0.061^{**}	-0.025
	(0.026)	(0.026)	(0.026)
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)
Controls: Socio-demo, political leaning	✓	✓	✓
Observations	3,002	3,002	3,002
R^2	0.061		0.098

*p<0.1; **p<0.05; ***p<0.01

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 to fight climate change. Our findings confirm this result: among our survey respondents, only 17% answered "Yes" when asked whether our Tax & Dividend would be effective in reducing pollution and fighting climate change, 66% answered "No", 18% that they did not know.

An explanation sometimes encountered to explain perceptions of ineffectiveness is that most people believe that energy consumption is little elastic (Kallbekken & Sælen, 2011; Carattini et al., 2018). To test this hypothesis, we regress a binary variable E equal to 0 if the respondent does not

perceive the policy as environmentally effective and 1 otherwise, on their subjective price elasticity for French people. As respondents were randomly assigned to transport or housing, we run a separate regression for both types of energies. Table C.1 in Appendix C.3 reports results with and without control variables. They all consistently indicate that a respondent anticipating an elasticity of -1 is (on average) 6 p.p. more likely to perceive the policy as 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 reduced to small subjective elasticities. Indeed, among respondents who perceive the policy as environmentally ineffective, almost half anticipate responses to price signals larger than the literature.[10]

A more plausible explanation for perceived ineffectiveness is that people disbelieve that the policy would be sufficient to affect *substantially* pollution and climate change. Taking respondents' average anticipated elasticities for transport and housing energies, the tax should reduce French greenhouse gas (GhG) emissions by 5.7 Mt of CO_2 equivalent (CO_2e) each year, according to simulation from BdF data. This reduction corresponds to 0.8% of French annual emissions, 0.01% of global ones, and does not suffice to reach the official objective of a decrease in emissions by 1.8% *each year* to reach carbon neutrality in 2050.[11] Thus, although respondents do anticipate responses to price incentives, an interpretation of our results is that they do not perceive a 50€/tCO_2 national carbon tax as a proportionate reaction to climate change.

3.3 Progressivity

It is often argued that a critical barrier to 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). A large literature has shown that carbon taxation alone is regressive (Poterba, 1991; Metcalf, 1999; Grainger & Kolstad, 2010), meaning that it is more costly for poorer households as a share of their resources. However, it has also been shown that redistributing its revenue through uniform lump sum transfers — a mechanism known as flat-

[10] Overall, average subjective elasticities are close to these estimates for transports (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 around -0.2 for housing and -0.4 for transport.

[11] The computations ignore the carbon tax paid by firms and its effects. In 2014, French GhG consumption based emissions were equal to 712 MtCO_2e (CGDD, 2019). 2017 global emissions were 53.5 GtCO_2e (UNEP, 2018). Alterea evaluates at 1.8% each year the reduction required to meet France National Low-Carbon Strategy (CGDD, 2015).

recycling — can make the policy progressive (West & Williams, 2004; Bento et al., 2009; Williams et al., 2015), including for France (Bureau, 2011; Douenne, 2018). Figure B.2 in Appendix B.3 displays the average net gain by income decile for our Tax & Dividend. It clearly appears from this figure that lower income households would gain more than richer households, both in relative and in absolute terms. Yet, only 19% of respondents think the policy would benefit poorest households, against 60% who declare it would not, and 21% who do not know. These results show that beliefs about the distributive effects of carbon taxation are biased, which was not shown in previous studies.

4 Are beliefs persistent?

The previous section has shown that people's low acceptance of our Tax & Dividend coincided with biased beliefs about the properties of the scheme. As knowledge about these properties has been shown decisive for acceptance (Carattini et al., 2018), it is important to assess to what extent beliefs are persistent. In the following, we test respondents' reaction to information about their gains, environmental effectiveness and progressivity. If their biased views simply reflect a lack of knowledge, we should expect them to revise their beliefs after new information is provided. We show that on average people update less than they should, indicating persistent biased beliefs. Moreover, we show that revisions are asymmetric and strongly biased towards pessimism.

4.1 Self-interest

Conservatism and pessimism After telling respondents that given their characteristics, they have 5 chances over 6 to "win" or "lose" from the policy, we observe that only 39% of them agree with our feedback ($G^F = \widehat{\Gamma}$), far from the 83% expected for a correct update. This result is consistent with findings of the literature that people tend to update like "conservative" Bayesians (Edwards, 1968). The full transition matrices of people's beliefs are given in Tables C.2 and C.3 in Appendix C.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 that, for the 24% of individuals who receive a "lose" feedback ($\widehat{\Gamma} = 0$), the *ex post* belief is on average consistent with a perfect updating. If anything, these people would rather tend to agree *too much* with our noisy signal, especially when excluding people who initially consider themselves as

unaffected (i.e. focusing on $g^I \neq 0$), for whom 83% does not lie in the 95% confidence interval.[12] The observed conservatism on the whole sample therefore comes from respondents who receive a "win" feedback ($\widehat{\Gamma} = 1$). Among the 60% of respondents who initially thought they would lose in this group, only 12% endorse the "win" feedback. This is in sharp contrast with the respondents who initially thought they would win and receive a "lose" feedback, since 82% of them endorse our prediction. Thus, pessimistic beliefs are persistent, but optimistic ones are not.

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

	Aligned with feedback: $G^F = \widehat{\Gamma}$	
	$\widehat{\Gamma} = 1$ (75.8%)	$\widehat{\Gamma} = 0$ (24.2%)
Initial belief winner ($g^I > 0$) (14.0%)	78.8% [73.2%; 83.4%]	81.5% [65.0%; 91.3%]
Initial belief unaffected ($g^I = 0$) (21.7%)	21.6% [17.6%; 26.2%]	44.9% [33.5%; 56.8%]
Initial belief loser ($g^I < 0$) (64.3%)	12.2% [10.3%; 14.5%]	93.9% [90.9%; 96.0%]
Initial belief affected ($g^I \neq 0$) (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 binomial probabilities are given in brackets.

Table C.4 in Appendix C.2 conducts the same analysis for the 28% of respondents whose gain is largely positive or largely negative, i.e. above 110€ per c.u. in absolute terms. For such respondents, our out-of-sample prediction of winning category is correct in 99% of cases, as can be seen on Figure C.1 in Appendix C.2. The alignments with our feedback are similar for the whole sample and for these respondents for whom we are sure to make a correct prediction. The similarity of alignments for different prediction accuracy rules out the possibility that some respondents do not update because

[12] We observe that respondents revise less their belief when they feel "unaffected", indicating that some of them may have misunderstood this category for "I don't know".

their private information would be *truly* more accurate than our prediction.

There are several ways to rationalize the conservatism and pessimism we observe. The simplest is that respondents perceive our information as biased. They may think we wrongly estimate their likelihood to win, and that we are too optimistic. A second interpretation is that they give too much value to their private information relative to the base rate one, a phenomenon known as base rate neglect (Grether, 1980). That is to say, pessimistic winners might be over-confident in seeing themselves as specific so that they partly discard the new information, e.g. by thinking they are part of the one sixth for whom our prediction is erroneous. Another explanation is that uncertainty makes people see their possible gains as a distribution (see Stiglitz, 2019). Then, instead of reporting the average of this distribution, people subject to risk-aversion would reason with conservative estimates for their gains, even more so if they believe that this survey can influence policy makers. In the end, these interpretations are of course not mutually exclusive, and empirically not distinguishable as they all translate in too much pessimism in the revision of their beliefs.

Determinants of correct updating To handle the notion of *correct updating*, we define a variable U which equals $+1$ if the respondent adopts a feedback that invalidates their initial belief, -1 if they update against the feedback that confirms it, or 0 if they do not update. Over the sub-sample of respondents who should have updated because their initial winning category is not aligned with our feedback, we regress the *correct updating* U over the initial belief not to lose G^I and a vector of characteristics $\mathbf{C}$:

$$U_i = \delta_0 + \beta_U G_i^I + \beta_{\mathbf{C}}\mathbf{C} + \epsilon_i \,|\, \dot{G_i^I} = -\widehat{\Gamma}_i, \tag{1}$$

The high value for β_U reported in column (1) of Table 4.2 again proves that, among those who should have updated, those who initially think they win (the optimistic losers) update significantly more correctly than those who do not (the pessimistic winners). Many other specifications have been tested, depending on how unaffected are treated and whether controls are included or not, or with the winning category after the feedback as covariate, and they all confirm a win/lose asymmetry in updating.

Beyond this asymmetry, column (2) shows that some respondents' characteristics are correlated with correct updating. Relative to unemployed and inactive people, retired, active and students update more correctly, the latter being 25 p.p. more likely to correctly revise their beliefs when invalidated than unemployed and inactive. Similarly to perceptions of net gains, position towards

Table 4.2: Asymmetric updating of winning category

	Correct updating (U)		
	(1)	(2)	(3)
Constant	0.120***	−0.041	−0.150
	(0.012)	(0.190)	(0.189)
Winner, before feedback ($\dot{G}$)	0.695***	0.685***	0.646***
	(0.078)	(0.080)	(0.080)
Initial tax: PNR (I don't know)			0.163***
			(0.031)
Initial tax: Approves			0.158***
			(0.046)
Retired		0.143*	0.146*
		(0.080)	(0.079)
Active		0.165***	0.175***
		(0.055)	(0.054)
Student		0.249***	0.234***
		(0.076)	(0.075)
Yellow Vests: PNR		−0.048	−0.043
		(0.047)	(0.047)
Yellow Vests: understands		−0.090***	−0.063*
		(0.034)	(0.034)
Yellow Vests: supports		−0.101***	−0.059*
		(0.035)	(0.036)
Yellow Vests: is part		−0.172***	−0.137**
		(0.062)	(0.062)
Among invalidated	✓	✓	✓
Controls: Socio-demo, politics, estimated gains		✓	✓
Observations	1,365	1,365	1,365
R^2	0.055	0.111	0.133

*p<0.1; **p<0.05; ***p<0.01

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

the Yellow Vests is significantly correlated with correct updating. The magnitude of this effect is large, as people who are part of the movement are 17 p.p. less likely to correctly update than people who oppose it. This can be linked to the higher distrust of the Yellow Vests towards the government, documented in Algan et al. (2019), that could apply to information provided by researchers regarding policies. Column (3) includes disapproval as a covariate: it indicates that, as for the initial bias in subjective gains, disapproving the reform is associated with a less correct update by 16 p.p. and partly explains the effect of the Yellow Vests.

Explanatory mechanisms The previous results suggest that conservatism in beliefs revision does not simply follows from people's cognitive difficulties when dealing with Bayes' rule. The determinants identified ("win" vs. "lose" feedback, attitude towards the policy and the Yellow Vests) rather indicate that beliefs formation and revision are intrinsically linked to preferences, a mechanism known as "motivated reasoning" in the literature (Kunda, 1990). Closely related is the well documented "confirmation bias", the tendency to weigh information differently depending on whether it confirms or not a prior belief. More recently, studies have also shown evidence of a "good news – bad news effect" (Eil & Rao, 2011; Sharot et al., 2011), meaning that people tend to give more weight to information that satisfies their ego or self-interest. Our results go *a priori* in the opposite direction since people discard what we could consider "positive" news ($\widehat{\Gamma} = 1$) but not negative ones ($\widehat{\Gamma} = 0$). Still, this effect is consistent with motivated reasoning as the lack of update has been shown much stronger for people who disapprove the policy and for supporters of the Yellow Vests. Thus, people's positioning does not only bias their initial perception about the policy, but also their processing of new information about it.

4.2 Environmental effectiveness

Table C.5 in Appendix C.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 — around 5 p.p. — seems modest given that the question immediately follows the priming. The effects of information on climate change (CC) or particulates (PM) are smaller, and only CC is significant, which is understandable as they were displayed at the very beginning of the survey and do not mention any environmental policy. As suggested by Millner & Ollivier (2016), given the complexity of the mechanisms at play, drawing a causal link between causes and consequences of environmental

problems requires considerable cognitive efforts, making it difficult to convince about effectiveness of policies that decentralize efforts to tackle pollution. Finally, we observe that our primings have no significant effect on beliefs over causes and consequences of climate change. Overall, these primings appear insufficient to change most people's mind about climate change and carbon tax effectiveness.

4.3 Progressivity

Table C.6 in Appendix C.4 shows the absence of 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 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: 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 the information is associated with a lower perceived regressivity by 5 p.p. A possible explanation for the strong belief in regressivity is that people view the tax as regressive (relative to income) and the transfer as neutral (in absolute values), and mistakenly conclude that their combination is regressive. Anyhow, without a deep explanation of the underlying mechanisms, the progressivity of the policy remains unintuitive to most people, and we cannot convince them easily.

5 Motives for acceptance

Our results clearly indicate that, as of today, a carbon tax is unlikely to be accepted in France. However, we have also shown that people largely display wrong 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 true preferences or wrong perceptions. The question we address is whether correcting biased beliefs would be sufficient for a carbon tax to be accepted.

5.1 Self-interest

Identification challenge While three quarters of the respondents are expected to win from our Tax & Dividend, 62% of these winners consider that they would not win and disapprove the policy. We want to estimate to what extent knowing they would win would lead them to approve the reform. Because respondents thinking they would win might differ in many respects from those thinking they would not, we cannot simply regress approval on perception of winning.

Main identification strategy In order to identify the effect *ceteris paribus* of self-interest on acceptance, we exploit exogenous variations in gains and losses. To do so, we consider Tax & targeted dividend, where respondents are randomly affected to a compensation scheme to which they are eligible or not (see section 2.1.2). Our methodology mixes a random discontinuity design (RDD) and an instrumental variable (IV) strategy. Let us denote by $I_{1,i}$ respondent's i income, and by $I_{2,i}$ the income of the second adult of their household if their is one, which we control for with a dummy variable S_i equal to 1 when there is a single adult. Let $T_{1,i}$ and $T_{2,i}$ be two binary variables indicating whether these individuals would be eligible to the transfer.[13] We also denote by G_i^T a dummy variable equal to 0 if respondent i thinks they would lose from Tax & targeted dividend, and 1 otherwise. Similarly, A_i^T is a dummy variable equal to 0 if respondent i disapproves this policy and 1 otherwise. We can then write a two stage least-square model, with the following first stage equation:

$$G_i^T = \alpha_0 + \alpha_1 T_{1,i} + \alpha_2 T_{2,i} + \alpha_c c_i + \alpha_S S_i + \sum_{j=1}^{2} \left(\alpha_{1,j} I_{1,i}^j + \alpha_{2,j} I_{2,i}^j \right) + \eta_i \tag{2}$$

where second order terms for income bring more flexibility. We precise that eligibility is defined using income thresholds c_i that are randomly allocated to households (see section 2.1.2) and introduced as fixed effects to control for threshold specific preferences. Formally, we define eligibility of adult $k \in \{1; 2\}$ as:

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

[13] As explained in section 2.1.2, we explicitly limit the number of beneficiaries to two by household.

Finally, the second stage writes:

$$A_i^T = \beta_0 + \beta_1 \widehat{G_i^T} + \beta_c c_i + \beta_S S_i + \sum_{j=1}^{2} \left(\beta_{1,j} I_{1,i}^j + \beta_{2,j} I_{2,i}^j \right) + \epsilon_i \tag{4}$$

where $\widehat{G_i^T}$ denotes the fitted value of G_i^T from the first stage regression. As can be seen from first stage results in Appendix D.1, eligibility of both respondents and households' second adults are positively correlated with beliefs of winning, so both instruments are relevant. The exclusion restriction states that conditional on income, being eligible affects approval solely through beliefs of winning. The RDD procedure employed in the first stage ensures that this is the case: conditional on income, eligibility is random, and it should affect acceptance only through self-interest.

Alternative specifications for robustness To further ensure the robustness of the results, we estimate several alternative specifications. First, we run the same RDD + IV design adding control variables (specification 2). In particular, we control for initial acceptance of our Tax & Dividend as this should explain most of the variation in the dependent variable. Second, we compare our results with a simple OLS where we control for relevant variables (3). Third, we use a logit model (4) to ensure that imposing linearity does not bias the results. Finally, we exploit a methodology similar to the main specification but applied to the feedback, without (5) and with (6) control variables. Indeed, as our estimation of gains for this feedback is a continuous variable $\widehat{\gamma}$, but the feedback itself is a binary variable $\widehat{\Gamma}$, we run an RDD to predict the belief of winning after feedback G^F. We then exploit variations around the threshold of zero net gain (which are random conditional on net gain) to explain acceptance A^F. This alternative two stage least square writes:

$$G_i^F = \alpha_0 + \alpha_1 \widehat{\Gamma}_i + \sum_{j=1}^{2} \alpha_{1,j} (\widehat{\gamma}_i)^j + \eta_i \tag{5}$$

$$A_i^F = \beta_0 + \beta_1 \widehat{G_i^F} + \sum_{j=1}^{2} \beta_{1,j} (\widehat{\gamma}_i)^j + \epsilon_i \tag{6}$$

where $\widehat{G_i^F}$ denotes the fitted value of G_i^F from the first stage regression. The identification assumption of this second IV states that conditional on estimated net gains ($\widehat{\gamma}$), receiving a win feedback ($\widehat{\Gamma} = 1$) affects approval solely through self-interest. Finally, we also investigate alternative versions of the previous models where we estimate the effect to "win" instead of "not to lose", and on "approval" instead of "acceptance".

Table 5.1: Effect of self-interest on acceptance.

	Targeted Acceptance (A^T)				Feedback Acceptance (A^F)	
	IV		*OLS*	*logit*	*IV*	
	(1)	(2)	(3)	(4)	(5)	(6)
Believes does not lose	0.571***	0.567***	0.443***	0.431***	0.517***	0.434***
	(0.092)	(0.092)	(0.014)	(0.018)	(0.170)	(0.135)
Initial tax Acceptance (A^I)		0.339***	0.360***	0.342***		0.428***
		(0.033)	(0.026)	(0.034)		(0.055)
Controls: Incomes	✓	✓	✓	✓		✓
Controls: Estimated gain		✓	✓	✓	✓	✓
Controls: Target of the tax	✓	✓	✓	✓		
Controls: Socio-demo, other motives		✓	✓	✓		✓
Observations	3,002	3,002	3,002	3,002	1,968	1,968
R^2	0.033	0.302	0.470		0.044	0.526

*p<0.1; **p<0.05; ***p<0.01

NOTE: Standard errors are reported in parentheses. For logit, average marginal effects are reported and not coefficients. The list of controls can be found in Appendix E.

Results First stage regression results are given in Appendix D.1. The effective F-Statistics (Olea & Pflueger, 2013) range from 37 to 57, indicating that both targeted transfers and feedback are strong instruments. Table 5.1 provides the second stage results for the six main specifications, and additional specifications can be found in Appendix D.2. Overall, the estimated effects of self-interest indicate that believing not to lose increases acceptance by about 50 p.p. The results obtained for the local average treatment effect (LATE) on compliers in IV regressions — i.e. on people who recognize they would not lose because of the treatment — are significantly higher than for the average treatment effect (ATE) estimated with OLS (57 vs. 44 p.p. for Tax & targeted dividend). Given the wide set of control variables used, and in particular our powerful control A^I, one can be rather confident that the OLS estimate is unbiased, and the difference due to the specificity of compliers in the LATE. This explanation holds for both the targeted scheme and acceptance after feedback, and is supported by the findings of section 4.1 for the latter. Indeed, as respondents most likely to revise their beliefs after a "win" feedback are also more in favor of the tax, the IV coefficient for compliers is logically higher than the point estimate found with OLS. The comparisons

of columns 1 with 2, and 5 with 6 show that adding control variables does not significantly affect the coefficient for Tax & targeted dividend, but does for acceptance after feedback, where the controls seem to capture part of the specificity of compliers. Overall, both methods yield similar effects: 57 p.p. with targeted transfers and 52 p.p. with the feedback. Finally, results from the logit regression (4) confirm that the assumptions of the linear probability model (LPM) do not bias the results.[14]

By isolating *beliefs* about self-interest *ceteris paribus*, we showed that this motive has a large effect on carbon tax acceptance. This result confirms previous findings of the literature (Stern et al., 1993; Thalmann, 2004; Baranzini & Carattini, 2017), but contrasts with the results of Kallbekken & Sælen (2011) who found that self-interest plays a limited role in Norway. Their different result could come from their methodological approach that relies on proxies to capture self-interest in regressions, or on differences between preferences of French and Norwegian people.

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 acceptance. To identify this effect, we estimate a two-stage least squares (2SLS) where the first stage uses random information to explain 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}) are poorly correlated with beliefs of effectiveness, we restrict the set of instruments to our primings on the scientific consensus (Z_E) and climate change (Z_{CC}). As discussed in section 4.2, these instruments are significantly related to our endogenous variable, yet potentially weak as our primings do not have a large effect on people's beliefs. If we denote E a dummy variable equal to 0 if the respondents thinks the policy is not environmentally effective and 1 otherwise, we can write a 2SLS model as follows:

$$E_i = \alpha_0 + \alpha_1 Z_{E,i} + \alpha_2 Z_{CC,i} + \alpha_{\mathbf{C}} \mathbf{C_i} + \eta_i \tag{7}$$

$$A_i^I = \beta_0 + \beta_1 \widehat{E}_i + \beta_{\mathbf{C}} \mathbf{C_i} + \epsilon_i \tag{8}$$

[14] Separation occurs in logistic regressions for self-interest and effectiveness. We made sure this does not lead to biased coefficients by conducting likelihood ratio tests and by running penalized regressions (Firth, 1993; Heinze & Ploner, 2003).

where $\widehat{E}_i$ denotes the fitted value of E_i from the first stage regression, and $\mathbf{C}$ a vector of characteristics.

Alternative specifications for robustness checks To ensure the robustness of our results, we estimate the previous 2SLS with control variables (1). Acknowledging that our primings could affect acceptation motives other than effectiveness alone, this potential bias should disappear when controlling for many variables including other motives. In addition, we estimate an OLS (2) model to compare the LATE of the first specifications with an ATE, and a logit model (3) to check the robustness of assumptions underlying the linear probability model. Finally, we also estimate two modified versions of our main specifications, where we switch from broad to strict definitions for environmental effectiveness (4) and for tax approval (5).

Table 5.2: Effect of believing in environmental effectiveness on acceptance

	Tax Acceptance (A^I)				Tax Approval ($\dot{A}^I$)
	IV	*OLS*	*logit*	*IV*	*IV*
	(1)	(2)	(3)	(4)	(5)
Environmental effectiveness: not "No"	0.479**	0.391***	0.370***		
	(0.230)	(0.015)	(0.018)		
Environmental effectiveness: "Yes"				0.505**	0.416**
				(0.242)	(0.168)
Instruments: info E.E., C.C. & P.M.	✓			✓	✓
Controls: Socio-demo, other motives incomes, estimated gains	✓	✓	✓	✓	✓
Observations	3,002	3,002	3,002	3,002	3,002
R^2	0.218	0.390		0.218	0.161

*p<0.1; **p<0.05; ***p<0.01

NOTE: Standard errors are reported in parentheses. For logit, average marginal effects are reported and not coefficients. The list of controls can be found in Appendix E.

Results The first stage regressions results can be found in Appendix D.1. Because of the relatively modest responses to our primings, the instruments are rather weak (effective F-statistic of 6 in column 1), a problem that is alleviated in the case of strict definitions (11 in columns 4 and 5). Given the exogeneity of our instruments, the only concern is a potential bias towards OLS, which

would entail too conservative estimates in our case. Table 5.2 reports the results of the second stages and alternative specifications. They all consistently indicate a strong positive effect of beliefs about environmental effectiveness on tax acceptance. Overall, the effects are statistically significant, and their size comparable to the one of self-interest, at 48 p.p. for our main IV regression (column 1). The LATE is again somewhat higher than the ATE estimated with OLS (2) — 48 vs. 39 p.p. — which is likely due to differences between compliers and other respondents: people who are most likely to change their mind following our information might also be more willing to accept the policy. Still, the coefficients obtained for the ATE remain large and indicate that beliefs over the policy's environmental effectiveness are critical for acceptance. Results of the logit model close to the ones of OLS confirm that the LPM does not induce a bias in the estimation (3). Lastly, the effect is relatively close with the strict definition of the dependent variable — 42 p.p. in (5), showing that the coefficient is not driven by a correlation between "PNR" responses to tax effectiveness and approval.

The critical role found for beliefs about environmental effectiveness is in line with findings of the literature (Sælen & Kallbekken, 2011; Kallbekken & Sælen, 2011; Baranzini & Carattini, 2017), although previous studies did not properly identify a causal effect. Our results thus confirm that convincing people about the environmental effectiveness of carbon tax would largely increase acceptance of this instrument.

5.3 Progressivity

Identification challenge and strategies As informing respondents does not convince them that our Tax & Dividend is progressive (see section 4.3), we cannot identify the causal effect of understanding the progressivity on acceptance using an IV estimation. Thus, we estimate how belief in progressivity correlates with acceptance using simple OLS and logit regressions. Controlling for many respondents' characteristics and other acceptance motives, one can be confident that the effect of progressivity is properly isolated. We focus on the acceptance question *after knowledge*, i.e. after asking whether the reform is progressive or not.[15] Table 5.3 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 winning category. On the contrary, columns (5)-(6) use

[15] As self-interest and effectiveness were made salient from the *initial* questions, treating progressivity *after knowledge* is the only way to make results comparable across motives.

strict definitions for both approval and the covariates, where only *"Yes"* (or *"win"*) answers activate the dummy variables.

Table 5.3: Effect of beliefs over progressivity on acceptance. Covariates refer either to broad (1-4) or strict (5-6) definitions of the beliefs, where strict dummies do not cover "PNR" or "Unaffected' answers.

	Acceptance (A^P) on *not "No"*				Approval ($\dot{A}^P$) on *"Yes"*	
	OLS		*logit*		*OLS*	
	(1)	(2)	(3)	(4)	(5)	(6)
Progressivity (P)	0.223***	0.237***	0.560***	0.544***	0.228***	0.482***
	(0.038)	(0.044)	(0.023)	(0.019)	(0.041)	(0.023)
Winner (G^P)	0.332***	0.332***			0.303***	
	(0.020)	(0.020)			(0.019)	
Effective (E)	0.258***	0.259***			0.244***	
	(0.023)	(0.023)			(0.020)	
($G^P \times E$)	0.127***	0.127***			0.126***	
	(0.034)	(0.034)			(0.037)	
Interaction: winner ($P \times G^P$)	0.183***	0.183***			0.098**	
	(0.050)	(0.050)			(0.048)	
Interaction: effective ($P \times E$)	0.172***	0.172***			0.281***	
	(0.057)	(0.057)			(0.059)	
Income (I, in k€/month)	0.017	0.018			0.037**	
	(0.022)	(0.022)			(0.018)	
Interaction: income ($P \times I$)		−0.008			−0.019	
		(0.013)			(0.014)	
$P \times G^P \times E$	−0.400***	−0.399***			−0.314***	
	(0.072)	(0.072)			(0.083)	
Controls: Socio-demo, incomes, estimated gains	✓	✓			✓	
Observations	3,002	3,002	3,002	3,002	3,002	3,002
R^2	0.460	0.460	0.162		0.391	0.130

*p<0.1; **p<0.05; ***p<0.01

NOTE: Standard errors are reported in parentheses. For logit, average marginal effects are reported and not coefficients. The list of controls can be found in Appendix E.

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 *ceteris paribus* is lower. The marginal effect of progressivity at the sample mean — i.e. accounting for the average marginal effect of interaction terms — is 27 p.p., slightly lower than the marginal effect obtained for self-interest (40 p.p.) and environmental effectiveness (31 p.p.). The effect obtained for the latter motive is lower than the OLS estimate found in section 5.2, because here acceptance is taken at a later step in the survey and not right after asking about environmental effectiveness, making it less salient. Besides, one might worry that perceived progressivity would be hard to disentangle from beliefs over net gains, as the latter is influenced by the former for a given income. To address this dependency, we include the interaction between progressivity and income as a covariate (2, 5). Although the coefficient is negative, in accordance with intuition, the effect is low and not significant. 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.

These results show that progressivity is a motive for acceptance almost as determinant as self-interest and effectiveness. This concern for distributional effects is consistent with the findings of Kallbekken & Sælen (2011), Brannlund & Persson (2012) and Gevrek & Uyduranoglu (2015), but contrasts with the results of Baranzini & Carattini (2017), who found that among Genevan, the concern exists but does not affect acceptance of environmental policies.

5.4 Complementarity between motives

5.4.1 Combined effects

Table 5.3 (column 1) shows that the combined effect of the three motives together — i.e. adding up their coefficients and the ones of their interactions — is as high as 90 p.p., confirming that these motives are the critical beliefs needed to accept the policy, and that they complement (rather than substitute) one another. Considering *approval* when respondents *do believe* that the policy is effective, progressive and satisfies their self-interest (5), the results are quite similar although it appears that the two altruistic motives are more complementary: their interaction term is high, and taken together they increase approval by 74 p.p. (marginal effect at the sample mean), against 64 p.p. for the combination of progressivity and self-interest, and 69 p.p. for self-interest and effectiveness.

The combined effect on approval of the three motives with their strict definition reaches 97 p.p. Thus, in a scenario where everyone would believe in the effectiveness and progressivity of the policy, and the 70% of households who win would know it, our estimation indicates that the approval rate would be as high as 90%. These results suggest that the carbon tax aversion that appears at first sight is exclusively linked to biases in beliefs, and that if theses biases were corrected for, carbon pricing could become close to consensual.

5.4.2 Willingness to pay

For respondents who believe in effectiveness of our Tax & Dividend, we are able to infer their willingness to pay (WTP) for climate mitigation by studying the acceptance rate in 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 5.1 indicates that this WTP is about 60€/year per c.u., as this corresponds to the subjective loss below which a majority accepts the policy. This WTP is computed only among people who believe that the tax is not ineffective, as 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 5.1 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 not framed as a good to buy nor as a 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 ask directly respondents to select their preferred option for climate mitigation, be it in a in a contingent valuation method (Berrens et al., 2004; Cameron, 2005; Kotchen et al., 2013) or in 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 sub-sample, and subjective gains are endogenous with 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 by 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.

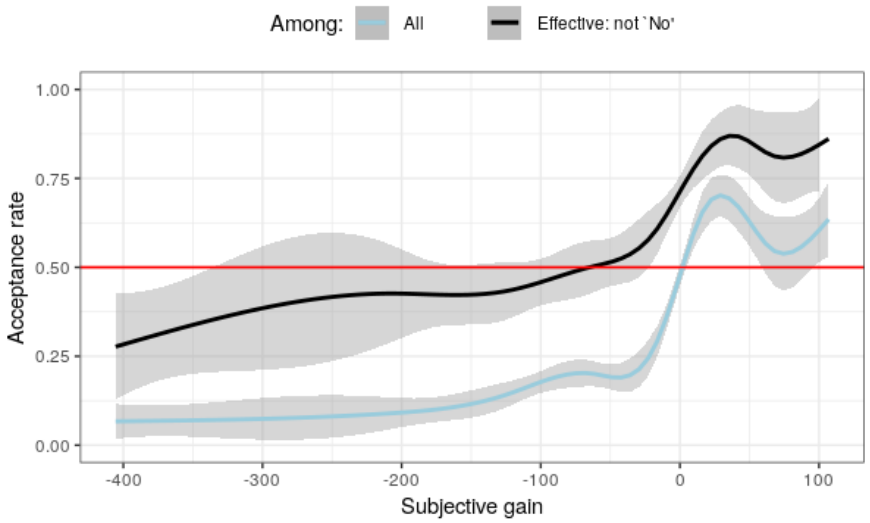


Figure 5.1: Acceptance rate by subjective gain, informing on the willingness to pay for climate mitigation.

6 Conclusion

In this paper, we investigate the relation between beliefs and acceptance of a carbon tax and dividend policy. Our analysis is based on a new survey and consumer survey data, enabling to compare subjective beliefs with objective impacts on French households. We find that 70% disapprove the proposed policy, which can be explained by biased beliefs about its properties. 89% of our survey respondents overestimate the negative impact on their purchasing power, and most of them do not perceive it as environmentally effective nor progressive. Both their beliefs and revisions after facing new information appear too pessimistic. Biases in revisions are stronger for people opposed to the tax, indicating that beliefs about tax impacts are partly shaped by motivated reasoning. At the same time, we find that acceptance is causally determined by beliefs. In particular, when believing they would not lose or that the policy is effective, people are about 40 p.p. more likely to accept it. The effect of believing in progressivity is also large, at 27 p.p. Given the complementarity between motives, our results suggest that a majority of people can accept our Tax & Dividend on purely altruistic ground, provided they believe in the fairness and efficacy of the reform. Adding to this the effect of self-interest for the 70% of households expected to win, we also find that if all biased beliefs could be corrected, the approval rate could reach 90%.

However, our treatment that provides authentic arguments in favor of the scheme never convinces

more than 12% of people. We can relate this conservatism to a strong distrust of the government, documented e.g. in Alesina et al. (2018). This result opens three main challenges. First, as it is unlikely that this trust issue 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, in which we assess both knowledge and beliefs about climate change, and preferred policies of French people. Second, as trust in government needs to be restored in the longer run, it is crucial to understand what causes the distrust and how it can be overcome. Third, it is important to assess to what extent our results are driven by the French context and influenced by the Yellow Vests movement. Although rejection of the tax may be lower in a different country, biases in perceptions are probably common everywhere. Thus, a lesson must be learned for policy design and implementation, to avoid a new carbon tax debacle *à la Française*.

References

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.

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

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

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

R. Bénabou & J. Tirole. Self-Confidence and Personal Motivation. *The Quarterly Journal of Economics*, 2002.

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.

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.

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

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

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

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

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

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

CGDD. France National Low-Carbon Strategy. Technical report, Ministry of Ecology, 2015.

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

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

T. Douenne. The vertical and horizontal distributive effects of energy taxes: A case study of a french policy. *FAERE Working Paper*, 2018.

T. Douenne & A. Fabre. French attitudes over climate change and climate policies. *mimeo*, 2019.

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.

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

. economists. Economists' Statement on Carbon Dividends. *The Wall Street Journal*, 2019.

W. Edwards. Conservatism in human information processing. In *Formal representations of human judgement*. Wiley, New York, b. kleinnmutz edition, 1968.

D. Eil & J. M. Rao. The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself. *American Economic Journal: Microeconomics*, 2011.

D. Firth. Bias reduction of maximum likelihood estimates. *Biometrika*, 1993.

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

C. A. Grainger & C. D. Kolstad. Who Pays a Price on Carbon? *Environmental and Resource Economics*, 2010.

D. M. Grether. Bayes Rule as a Descriptive Model: The Representativeness Heuristic. *The Quarterly Journal of Economics*, 1980.

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

G. Heinze & M. Ploner. Fixing the nonconvergence bug in logistic regression with SPLUS and SAS. *Computer Methods and Programs in Biomedicine*, 2003.

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.

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.

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

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.

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

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.

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

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.

G. E. Metcalf. A Distributional Analysis of Green Tax Reforms. *National Tax Journal*, 1999.

G. E. Metcalf & D. Weisbach. The Design of a Carbon Tax. *Harvard Environmental Law Review*, 2009.

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

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.

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

J. M. Poterba. Is the Gasoline Tax Regressive? *Tax Policy and the Economy*, 1991.

H. Sælen & S. Kallbekken. A choice experiment on fuel taxation and earmarking in Norway. *Ecological Economics*, 2011.

T. Sharot, C. Korn, & R. Dolan. How unrealistic optimism is maintained in the face of reality. *Nature Neuroscience*, 2011.

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

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

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.

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

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

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

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.

Appendices

A Raw data

Table A.1: Sample characteristics: quotas stratas.

	Population	Sample
gender		
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*	0.09
ARA	*0.12*	0.13
PACA	*0.09*	0.09

Table A.2: Households' characteristics.

	Population	Sample
Household composition (mean)		
Household size	*2.36*	2.38
Number of adults	*2.03*	1.93
c.u.	*1.60*	1.61
Energy source (share)		
Gas	*0.42*	0.36
Fuel	*0.12*	0.09
Accomodation surface (m^2)		
mean	*97*	96
p25	*69*	66
p50	*90*	90
p75	*120*	115
Distance traveled 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*	7.25
p25	*6*	5
p50	*6.5*	6
p75	*7.5*	7

Sources: Matched BdF; except for number of adults (ERFS) and domestic fuel (CEREN).

B Estimation for feedback

B.1 Formulas to compute monetary effects of carbon tax policy

In order to compute the monetary impact of a carbon tax increase, we decompose current energy expenditures $E(\tau)$ as a product of current price $P(\tau)$ and current quantities consumed $Q(\tau)$, each being a function of the excise tax τ within which the carbon tax is comprised:[16]

$$E(\tau) = P(\tau) Q(\tau)$$

All variations in expenditures can then be expressed as:

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

from which we can write the effect of a price change on quantities consumed:

$$Q(\tau') = Q(\tau)\left(1 + e\frac{dP}{P}(\tau)\right)$$

where $e = \frac{dQ}{dP} \cdot \frac{P}{Q}$ is the price elasticity of the energetic good considered, that is here assumed constant. For all energies, the final price can itself be decomposed as:

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

where t is the value added tax (VAT) rate that applies after excise taxes, i the incidence of excise taxes on consumers assumed constant, and $p + (i-1)\tau$ the producer price as a function of τ and for a given value of t.[17] When the carbon price changes so that the excise taxes varies from τ to some level τ', we therefore have:

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

Thus, following a carbon price increase, one can express the associated increase in expenditures for each energy as:

$$E(\tau') - E(\tau) = E(\tau)(1+e)\frac{dP}{P} = E(\tau)(1+e)\frac{i(\tau' - \tau)}{p + i\tau}$$

We can replicate similar calculations to obtain the expected variations in tax revenue T. Starting from its expression — which is the sum of excise taxes and the VAT over this tax — we have:

$$T(\tau) = Q(\tau)((1+t)\tau + t(p + (i-1)\tau))$$

[16] The French carbon tax "Contribution Climat Energie" is a component of existing taxes on energetic products: TICPE for transport and domestic fuels, TICGN for natural gas.

[17] Hence p is the producer price for a given value of t, when $\tau = 0$.

from which we obtain:

$$T\left(\tau'\right)-T\left(\tau\right)=Q\left(\tau\right)\left(1+e\frac{i\left(\tau'-\tau\right)}{p+i\tau}\right)\left[t\left(p+\left(i-1\right)\tau'\right)+\left(1+t\right)\tau'\right]-Q\left(\tau\right)\left[t\left(p+\left(i-1\right)\tau\right)+\left(1+t\right)\tau\right]$$

Following the literature, we assume price elasticities of -0.4 for transport fuels and -0.2 for housing energies. For the tax incidence on consumers, we assume a value of 0.8. These values were used to compute aggregate variations in tax revenue and determine the level of lump sum transfer per adult that a budget neutral policy would finance. When asked to estimate the impact of the policy on their own purchasing power, respondents simply had to make an estimation over:

$$E\left(\tau'\right)-E\left(\tau\right)=E\left(\tau\right)\left(1+e\right)\frac{dP}{P}$$

where for simplicity dP was given for transport fuels, and $\frac{dP}{P}$ for housing energies. 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.

B.2 Predicting gains and losses

We regress the increase in housing energy expenditures on households' characteristic using EL 2013 survey. Table B.1 presents several specifications for such regression, and its last row shows the out-of-sample error rate, computed with BdF data. All specifications yield a similar error rate of 15-17%. Fearing that respondents could make mistakes when filling the accommodation size in the entry field, we used the first specification in our survey, as it does not rely as heavily as the others on the accommodation size. In order to balance the error rate for losing households that are mistakenly estimated winners, and winners who are mistakenly estimated losers, we add a constant of 16.1 in our estimation of 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 OLS as our prediction method for the estimation of net gain because it compared well with respect to alternative methods. We also classified winners and losers using a decision tree, and obtained an very close error rate: 17.4% (see Figure B.1). Finally, statistical matching provided an error rate of 17.7%.

Table B.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: Fuel 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 × Fuel 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

*p<0.1; **p<0.05; ***p<0.01

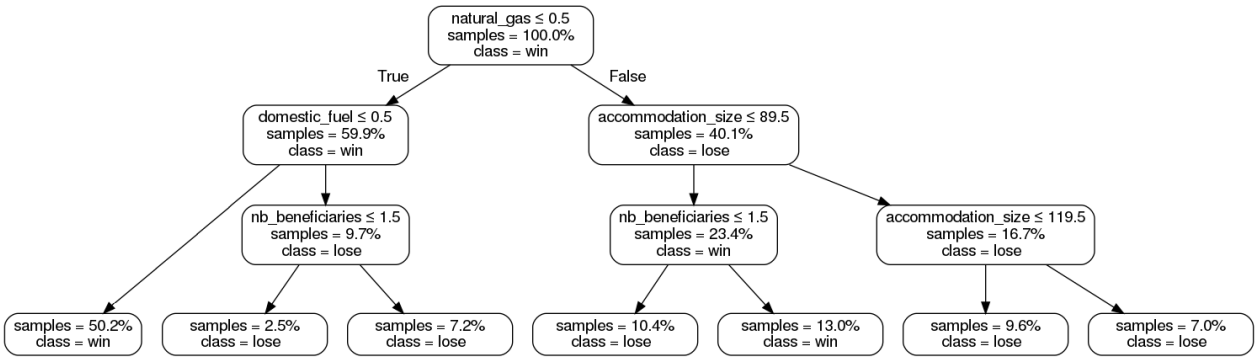


Figure B.1: Decision tree that classifies households between winners and losers.

B.3 Distributive effects

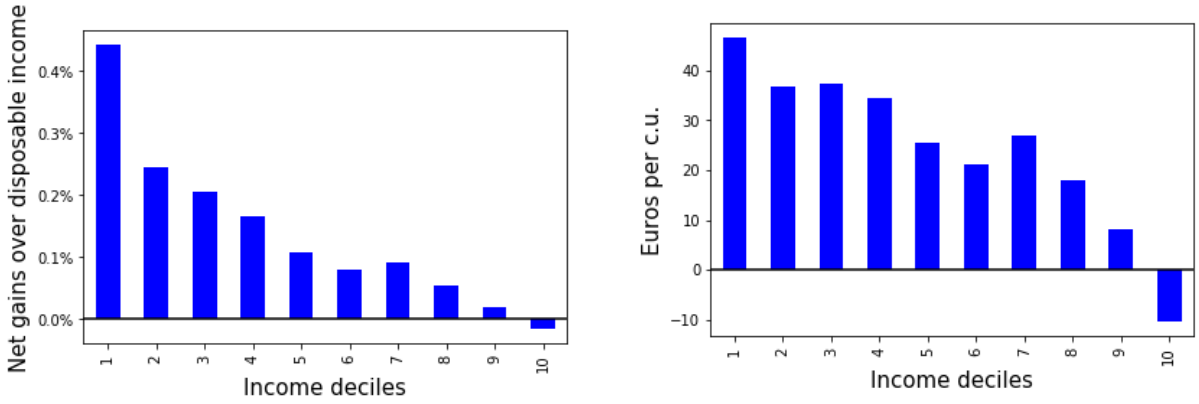


Figure B.2: Average cost of the carbon tax and dividend policy, by income decile.

C Beliefs and persistence

C.1 Elasticities

Table C.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: Transports		−0.056*		−0.060**
		(0.030)		(0.030)
Controls: Socio-demo, energy incomes, estimated gains			✓	✓
Observations	1,501	1,501	1,501	1,501
R^2	0.003	0.002	0.089	0.090

*p<0.1; **p<0.05; ***p<0.01

C.2 Self-interest

Table C.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%

Table C.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%

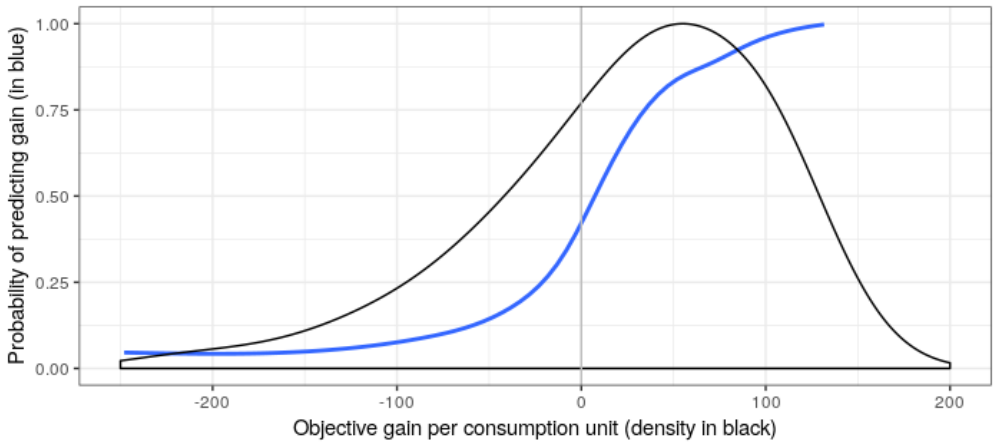


Figure C.1: Probability that our estimation of net gains correctly predicts the winning category.

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

	Aligned with feedback: $G^F = \widehat{\Gamma}$	
	$\widehat{\Gamma} = 1$ (81.6%)	$\widehat{\Gamma} = 0$ (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.

C.3 Environmental effectiveness

Table C.5: Effect of primings 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: Socio-demographics		✓	✓	✓
Observations	3,002	3,002	3,002	3,002
R^2	0.003	0.047		0.075

*p<0.1; **p<0.05; ***p<0.01

C.4 Progressivity

Table C.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: Socio-demo, politics			✓
Observations	1,444	1,444	1,444
R^2	0.0004	0.018	0.094

*p<0.1; **p<0.05; ***p<0.01

D Estimation of acceptation motives

D.1 Two stage least squares: first stage results

Table D.1: First stage regressions results for self-interest

	Believes does not lose			
	Targeted tax (G^T)		After feedback (G^F)	
	(1)	(2)	(5)	(6)
Transfer to respondent (T_1)	0.268***	0.227***		
	(0.028)	(0.027)		
Transfer to spouse (T_2)	0.180***	0.174***		
	(0.031)	(0.030)		
$T_1 \times T_2$	−0.190***	−0.161***		
	(0.038)	(0.037)		
Initial tax Acceptance (A^I)		0.163***		0.333***
		(0.033)		(0.038)
Simulated winner ($\widehat{\Gamma}$)			0.217***	0.210***
			(0.036)	(0.035)
Controls: Incomes	✓	✓		✓
Controls: Estimated gain		✓	✓	✓
Controls: Target of the tax	✓	✓		
Controls: Socio-demo, other motives		✓		✓
Effective F-Statistic	44.093	40.834	37.966	57.866
Observations	3,002	3,002	1,968	1,968
R^2	0.082	0.177	0.131	0.319

*p<0.1; **p<0.05; ***p<0.01

Table D.2: First stage regressions results for environmental effectiveness

	Environmental effectiveness	
	not "No"	"Yes"
	(1)	(4,5)
Info on Environmental Effectiveness (Z_E)	0.062***	0.059***
	(0.017)	(0.014)
Info on Climate Change (Z_{CC})	0.030*	0.028**
	(0.017)	(0.013)
Controls: Socio-demo, other motives, incomes, estimated gains	✓	✓
Effective F-Statistic	5.886	11.145
Observations	3,002	3,002
R^2	0.121	0.123

*p<0.1; **p<0.05; ***p<0.01

D.2 Additional specifications

Table D.3: Effect of self-interest on acceptance: second stages of alternative specifications.

	Targeted Tax			After Feedback		
	Acceptance	Approval		Acceptance	Approval	
	(1)	(2)	(3)	(4)	(5)	(6)
Believes wins	0.636***	0.382***		0.764***	0.612***	
	(0.100)	(0.080)		(0.251)	(0.197)	
Believes does not lose			0.351***			0.414***
			(0.073)			(0.133)
Controls: Incomes	✓	✓	✓	✓	✓	✓
Controls: Estimated gain				✓	✓	✓
Controls: Target of the tax	✓	✓	✓			
Observations	3,002	3,002	3,002	1,968	1,968	1,968
R^2	0.033	0.018	0.018	0.044	0.017	0.017

*p<0.1; **p<0.05; ***p<0.01

E Control variables

Socio-demographics: *respondent's income, household's income, sex, age* (5 categories), *employment status* (9 categories), *socio-professional 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, ecologist, apolitical.*

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

Energy: *heating mode* (collective vs. indivual), *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.*

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