

Referee report for Douenne and Fabre

1. Summary

Using a survey in the aftermath of the Yellow Vests movement in France, the authors investigate the determinants of the support for a carbon tax and dividend. The majority of people (70%) oppose a carbon tax and dividend, and the opposition is driven by three types of beliefs: a belief that the policy would impose financial losses on individuals (self-interest), a belief that the policy is not effective for the environment, and a belief that the policy is regressive. Providing people information in these three domains has small effects on their beliefs. Furthermore, people's belief updating is biased: when people who think they would lose financially from a carbon tax and dividend are told that they would benefit, almost nobody believes it. People who support the Yellow Vests are less likely to update their beliefs, which is evidence for motivated reasoning. Using a survey experiment and an IV strategy, authors are able to show that beliefs about financial impacts on self and about environmental effectiveness do have large causal effects on the support for the carbon tax and dividend: moving from not thinking one benefits to thinking one does benefit increases support by about 50 percentage points, and the effect of moving from thinking the policy is ineffective to thinking that the policy is effective is of a similar size. Overall, the authors conclude that, if people's beliefs could be moved, a carbon tax and dividend would be supported by the overwhelming majority of French people. However, due to motivated reasoning among other factors, people's beliefs are very hard to change.

2. Main comments

This is a great study that I enjoyed reading. The carbon tax has overwhelming support among economists. Yet, the public is generally far from being as supportive. Understanding the reasons for this gap is important to design effective policy that also has a chance of being politically acceptable. This study's key strength is to experimentally manipulate three different types of beliefs and assess the impact of such information on the support for a carbon tax in France, a country that recently saw a mass social movement against the carbon tax. The exercise allows authors to inform the policy debate about the carbon tax. Their findings suggest that the carbon tax is unlikely to gain broad support in France in the short run, because people have three strong reasons to oppose it (self-interest, inefficiency, regressivity), and people's beliefs are relatively inflexible due to motivated reasoning among other factors.

While overall the analysis is sound, the explanations are sometimes confusing and insufficient to get a clear idea of what the authors actually did.

2.1 Provide more context about French politics and the public's beliefs

Given that this journal is read by an international audience, it is necessary to give more context. Are the French opposed to the carbon tax because they don't think the environmental problem is important? Are they opposed to any environmental policy or to carbon taxes only? How important are political parties in French environmental policy in general and carbon taxes in particular?

Say a bit more about the Yellow Vests movement, as well as the extent to which each of the French parties supports a carbon tax and why parties support or oppose it. This is important because in other countries like the US, opposition to a carbon tax is largely explained by political motivations.

Similarly, while I understand that the authors expand more on this topic in Douenne and Fabre (2020), it would be useful to say more about French beliefs about climate change and the public's environmental policy preferences and how these beliefs and preferences compare to other developed countries.

Overall, this background information is important to be able to draw sharper policy conclusions and should occupy at least two or three paragraphs.

2.2 Clarify the exercise leading to the financial win / loss predictions for each household

The exercise predicting financial gains and losses for each household was very confusing.

First, it is not clear why several different surveys need to be used, including surveys that seem to contain the same information (for example Budget de Famille and Enquete Logement). Data sources also come from different years, including 2008, 2011 and 2017: what did the authors do to make these different survey years reasonably comparable?

Second, the authors need to explain in more detail (step by step) their prediction of gain/ loss and more carefully justify why their prediction is believable, as well as describe the limits of the prediction. This is important because their analysis of belief updating is predicated on these predictions being reasonably accurate ("bias" is defined relative to these predictions). The appendix C.1 and C.2 is not very reader friendly at this point because it is all in the weeds and lacks a roadmap: first say what the goal is, and justify each methodological step with respect to that goal.

Because the explanation is not very clear, this statement on p. 16 that "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" is not clear at all. What are the two sides of this comparison and how are they computed?

Then following on p. 16 the authors write "For five households out of six, we correctly predict whether their purchasing power would increase or decrease through the policy. We make this ratio symmetrical": what does it mean to "make" the ratio symmetrical and why is it necessary to make it symmetrical?

Appendices C1 and C2 should be re-written more clearly. In the main text, there should be a couple of formulas for the prediction that are key and serve to explain how the prediction works, then the data can be discussed a bit more more succinctly in the main text and refer to the appendix for details.

2.3 Bayesian updating and the relevance of motivated reasoning

The authors do not devote enough space in the main text discussing whether people's belief updating could be rational under some assumptions (and which specific assumptions). There is some discussion of this in the appendix E but it needs to be brought back in the main text, and also discussed together with potential limitations of the author's predictions regarding losses and gains. In the appendix E itself, the writing can be confusing: "Column (2) in Table E.1 somehow validates this hypothesis, as even controlling for the bias, a higher subjective gain is significantly associated with a higher propensity to

update correctly (while those who feel unaffected update less correctly all else being equal).” The link between this sentence and the hypothesis discussed in the paragraph right before it is unclear.

One potential explanation for why motivated reasoning seems so pervasive here is that the financial stakes are fairly low, about a couple hundred euros per year (Figure 3.1). Unfortunately, the relevance of stakes probably cannot be tested directly in the data since even people who stand to lose a lot only lose about 400 euros, so there may not be enough power to distinguish between low and high stake choices. So the authors should discuss this and cite relevant literature.

2.4 Clarify the RDD / IV strategy

Explain the rationale behind choosing different thresholds for the tax & targeted dividend on p. 11. My understanding is that the overall idea is to randomly choose, for each person, whether they get the dividend or not; this is implemented by allocating the higher or lower threshold within that person’s income decile. Strangely enough, this is not the case for the very low or very high income people: why?

Also, while the methodology is reminiscent of an RDD, it seems more similar to a stratified randomized controlled trial where people are randomized into treatment depending on their income decile. So I would rather call the authors’ exercise a stratified survey experiment rather than RDD. Could the authors further discuss this?

Finally, are there any differences in the LATE between lower and higher income people? Maybe there is not enough power, but it would be interesting to know.

3. Additional comments

p. 4 “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, all else equal.” In the context of the introduction, it is not clear what beliefs mean. One guesses beliefs are probably dichotomous variables, but it would be worth clarifying because the formulation makes it sound as if they are continuous (i.e. how effective the policy is, how progressive, etc.).

p. 11 “Tax & Targeted Dividend In this block, we ask for the winning category and for the approval of four alternative reforms.” What is the “winning category” here?

p. 18: Figure 3.1 is missing the caption for the second sub-figure, which is probably (b) Housing energies.

p. 22 “These results show that beliefs about the distributive effects of carbon taxation are biased, which has not been shown in previous studies.” Actually, Carattini et al. 2017 show using a survey in Switzerland that providing evidence on the progressive effects of the carbon tax and dividend considerably increases support for the policy. It would be worth discussing differences with that study, since here the information about progressivity did not change people’s level of support.

p. 24 Table 4.1 needs a note that fully explains all notation. The same applies for all tables that use notation such as G^f .

p. 30: typo “random discontinuity design”

p. 36 “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 one's 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." Obviously, no, you cannot be confident that the effect can be properly isolated. Be more cautious and say that these controls hopefully account for most confounds, or something along these lines.

p. 48 Table A1 and A2, add a t-test for the differences and briefly comment in the text about the largest differences (e.g. there seem to be some differences by education).

p. 62 Figure D1 needs a note with more explanation of what's going on in it. Same applies for many Tables and Figures in the paper. The tables and figures should be roughly understandable without having to look back at the text.