

Comments to the Author

This manuscript explores the roles that beliefs, attitudes, and biases play in determining support for carbon taxes in France. To explore these effects, it collects original survey data from a representative sample of French residents. The survey randomizes a variety of “treatments”, including information about the incidence of various hypothetical tax policies and environmental damages from fuel consumption. The survey results lead to a variety of findings: most individuals underestimate their net gains (i.e. overestimate their net losses) from a carbon tax; additional information has little impact on beliefs for those who think they will be net losers; many individuals incorrectly believe the tax will be regressive (and may simply be unable to understand progressivity in this context); and a belief that one will be a net winner from the tax policy strongly relates to supporting the policy. While many of the results are descriptive (which is not to say that they are uninteresting), the last result is informed by survey-generated quasi-experimental variation in the impacts of hypothetical policies.

The paper tackles an important political economy/environment question using a very carefully designed and executed survey. Several of the results are of policy interest, and the exposition is generally clear. My comments are ordered more or less in the order of the paper, but I believe the important issues to address are 4, 6, and 7.

1. In Section 3.1, the paper analyzes what factor correlate with bias in subjective gains (i.e. the absolute difference between what people think they will gain/lose from the policy, and what they will actually gain/lose from the policy). Since we do not have strong priors about which factors should matter — economic theory is of limited use here since people do not appear to have rational expectations — this exercise amounts to a bit of a fishing expedition. Indeed, most of the factors seem weakly related at best to the dependent variable. **In this context it could be helpful to adjust for multiple testing, for example by controlling the family wise error rate (FWER) or false discovery rate (FDR).** Also, in Table 3.1, what is “ecologist”? I assume that term is roughly equivalent to “environmentalist” in the US (i.e. someone with strong preferences for environmental protection), but the word ecologist refers to an actual scientist.

2. I’m not certain what to conclude from Section 3.2 (perceived environmental effectiveness of the policy). The authors write, “A more plausible explanation for perceived ineffectiveness is that people do not believe that the policy would be sufficient to substantially affect pollution and climate change. Taking respondents’ average anticipated elasticities for transport and housing energies (that are fairly accurate), the tax should reduce French greenhouse gas [by]...0.8% of French annual emissions, 0.01% of global ones, and does not suffice to reach the official objective of carbon neutrality in 2050.” If the respondents anticipated elasticities are correct, as the paper says, then it sounds like they are arguably correct in perceiving that the policy will be environmentally ineffective. But the paper refers repeatedly to “perceived ineffectiveness”, a phrase which strongly suggests that the policy is not actually ineffective, but merely perceived as ineffective. It’s not clear that the perceptions are actually wrong here.

3. Equation (1), in Section 4.1, specifies “correct updating” (U) as a function of initial beliefs and other factors. U ranges from 1, if the respondent correctly updates from an invalid belief



to a valid one, to -1 , if the respondent incorrectly updates from a valid belief to an invalid one. With no update, $U = 0$. However, since the regression sample is conditioned on the initial perceived gain being of the opposite sign from the (estimated) true gain, then respondents in the sample must have initially been incorrect. Thus a regression sample respondent can either not update at all, in which case $U = 0$, or can update from incorrect to correct, in which case $U = 1$. Thus it appears that for the actual estimation sample, the dependent variable collapses down to binary. Is that correct? If so, it could be helpful to state it. Also, an easier way to write “ $\text{sgn}(g) = \text{sgn}(\gamma)$ ” would simply be “ $g * \gamma < 0$ ”; this expression does not require defining a function.

4. The quasi-experimental portion of the paper implements a regression discontinuity design (RDD) that uses income as the forcing variable. In this context income is a relevant forcing variable because the hypothetical policies redistribute tax revenues only to people below the 20th, 30th, 40th, or 50th income percentiles (different respondents get assigned to different income cutoffs). This is a clever estimation strategy, but I have several concerns about the specification and execution.

First, this is a RDD that theoretically should not work, because the treatment does not change sharply at the threshold given that incomes are stochastic over time. Simply put, as the RDD bandwidth around the income threshold approaches zero, the gains to being below the threshold for a household that is within that tiny bandwidth also go to zero. The counterargument is that, as long as there is a first stage result (i.e. households think they gain from being right below the income threshold), then it doesn't matter whether the design is theoretically invalid — maybe it works because people don't have any foresight. Still, it would be more comforting if the design worked in both theory and practice.

Second, because there are multiple income thresholds (from the 20th percentile to the 50th percentile), I don't think you want to restrict the coefficients on income to be constant. The CEF could change at different thresholds, so you should probably interact the terms in the summation in equation (2) with income threshold indicators as well. Additionally, it would be standard practice in a RDD to at least drop the lowest and highest income households from the sample (e.g. everyone below the 10th percentile and above of the 60th percentile).

Finally, the “mixing” of a RDD and IV appears as if it is just a standard “fuzzy RDD”. Is there something else here that I'm missing?

5. The manuscript seemingly ignores another source of random variation in net gains, which is the assignment of income thresholds to individuals. My understanding from Section 2.1.2 is that an individual falling between two thresholds gets randomly assigned to one of them; e.g. an individual at the 23rd percentile gets randomly assigned a threshold of the 20th or 30th percentile. This procedure implies that, conditional on income bin, whether you have a net gain or loss is randomly assigned. On the positive side, this represents true random variation, not just quasi-random variation. On the negative side, the exclusion restriction may not hold if people have strong preference on redistribution (and thus intrinsically prefer policies that are more or less redistributive, regardless of whether they personally benefit). Still, it seems like this variation could be worth exploiting, even if it comes with a caveat.

6. At the end of Section 5.1, the manuscript compares its results to those from other papers. In short, it finds that self interest (i.e. changing support in response to becoming a winner from the policy) is more important than has been suggested in previous papers. But I'm not sure that the results are directly comparable to those in other papers, because the "treatment," i.e. the change in net benefits from becoming a winner, is enormous in this case. The paper codes winning as binary, but presumably virtually everyone who qualifies at a given threshold is a winner, especially when the threshold is at 20% or 30% (if you tax 100% of the population and recycle the revenue to 20% of the population, it's very difficult to lose if you're in that 20%!). Furthermore, the winners are generally "winning" by several hundred percent of their baseline carbon tax burden. As such I'm not sure the results are comparable to other studies, in which the net gains of winners and losers do not differ nearly as strongly.

7. In Section 5.2 (environmental effectiveness), the relationship between the priming instruments and beliefs about the policy's environmental effectiveness is not obvious. The priming information is about the risks of pollution or climate change, and it's not clear why that should affect beliefs about whether the carbon tax would be effective at reducing pollution or slowing climate change. It is thus not too surprising that the instruments are weak. In fact, the first stage is even weaker than Appendix Table F.2 suggests because the authors also did pretesting in choosing to exclude Z_PM from the first-stage equation (based on what they saw in the data). One robustness check would be to use LIML, which is generally less sensitive to weak instruments.

8. Section 5.3 uses observational data to estimate the relationship between policy support and believed progressivity of the policy, absent a compelling instrument for progressivity. While believed progressivity is strongly predictive of policy support, this relationship could represent reverse causality. In fact, I believe that is basically what was argued in Section 4 — people who are motivated to oppose (support) a policy will look for reasons to affirm their opposition (support), even if these reasons have little basis in fact.

9. Some typos I noticed:

On p. 16, "...shown to be as least as accurate..."

On p. 30, "random discontinuity design" should be "regression discontinuity design"

On p. 31, "We precise that eligibility..."

