
This is the **accepted version** of the article:

Savin, Ivan; Drews, Stefan; Maestre-Andrés, Sara; [et al.]. «Public views on carbon taxation and its fairness : a computational-linguistics analysis». *Climatic Change*, Vol. 162, issue 4 (Oct. 2020), p. 2107-2138 (2020). DOI 10.1007/s10584-020-02842-y

This version is available at <https://ddd.uab.cat/record/235173>

under the terms of the  **CC BY** COPYRIGHT license

Public views on carbon taxation and its fairness: A computational linguistics analysis

Ivan Savin^{a,b,c}, Stefan Drews^c, Sara Maestre-Andrés^c, and Jeroen van den Bergh^{c,d,e}

^a Corresponding author: ICTA, Edifici Z, UAB Campus, 08193 Bellaterra, Spain, ivan.savin@uab.cat.

^b Graduate School of Economics and Management, Ural Federal University, Yekaterinburg, Russian Federation

^c Institute of Environmental Science and Technology, Universitat Autònoma de Barcelona, Spain.

^d ICREA, Barcelona, Spain.

^e School of Business and Economics & Institute for Environmental Studies, VU University Amsterdam, the Netherlands.

Funding: This work was funded by a Recercaixa 2016 project titled ‘Understanding Societal Views on Carbon Pricing and an ERC Advanced Grant from the European Research Council (ERC) under the European Union’s Horizon 2020 Research and Innovation Programme [grant agreement n° 741087]. I.S. acknowledges financial support from the Russian Science Foundation [RSF grant number 19-18-00262].

Abstract

Carbon taxes evoke a variety of public responses, often with negative implications for policy support, implementation and stringency. Here we use topic modelling to analyze associations of Spanish citizens with a policy proposal to introduce a carbon tax. This involves asking two key questions, to elicit (1) citizens' associations with a carbon tax and (2) their judgment of the fairness of such a policy for distinct uses of tax revenues. We identify 11 topics for the first question and 18 topics for the second. We perform regression analysis to assess how respondents' associations relate to their carbon-tax acceptability, knowledge and socio-demographic characteristics. The results show that, compared to people accepting the carbon tax, those rejecting it show less trust in politicians, think that the rich should pay more than the poor, consider the tax to be less fair, and stress more a lack of renewable energy or low-carbon transport. Respondents accepting a carbon tax emphasize more the need to solve environmental problems and care about a just society. These insights can help policy makers to improve the design and communication of climate policy with the aim to increase its public acceptability.

Keywords: carbon pricing, fairness perception, policy acceptability, public opinion, structural topic modelling.

1. Introduction

There is considerable public resistance against carbon taxation, which creates a barrier to policy implementation, stringency and stability. Research has identified various public concerns about carbon taxes, including high perceived personal costs, regressive cost distribution, damages to the wider economy, skepticism about the policy inducing low-carbon behavior change, and intuitions that the ‘real’ objective of such policy is raising tax revenues (Carattini et al., 2018; Klenert et al., 2018). Most of this research analyzes public responses to closed-ended survey questions. An alternative approach is to let survey respondents freely formulate their views on specific issues in open-ended questions. This approach elicits attitudes and beliefs in a less restricted way than closed-ended questions with pre-formulated answers. It allows exploring more profoundly the cognitive content of individuals’ attitudes and beliefs. Given that people can access and express their thoughts, open responses likely represent what is personally important to them.

We use this approach to examine open-ended survey responses by the general public concerning a policy proposal to introduce a carbon tax in Spain. We intend to answer the following questions: (1) What are citizens’ associations with a carbon tax? (2) What do they consider fair or unfair regarding such a policy, considering distinct spending scenarios for the tax revenues? (3) How do these associations relate to respondents’ acceptability of, knowledge about, and perceived fairness of the tax, as well as socio-demographic characteristics?

Freely formulated associations, such as to terms like ‘climate change’, have typically been analyzed through manual coding of survey responses into thematic categories (Leiserowitz, 2006; Lorenzoni et al., 2006; Moloney et al., 2014). For our purposes, we depart from this technique and instead apply a more recent method known as topic modelling. This denotes a computer-based algorithmic approach to classify public responses to open-ended survey questions into distinct topics. In particular, we use structural topic modelling (STM, Roberts et al., 2014) which was developed specifically for analyzing short open-ended questions from surveys.

Compared to alternative topic modelling techniques, STM has been found generate better topics for short responses as are typical of open-ended survey questions (Roberts et al., 2014, pp.1071-73). STM has also been compared with human coding of survey questions and was found to provide similar results (Roberts et al., 2014, pp.1075). The advantage of STM is that it avoids potential inconsistency arising from subjective and idiosyncratic assessments by distinct human coders. In addition, it requires considerably less processing time and assures better reproducibility. STM has already seen application to political and sociological themes, including analyses of climate policy documents (Lesnikowski et al., 2019). Some studies have used STM to examine public perceptions of climate change in Norway and China asking questions such as “What comes to your mind when you hear the word ‘climate change?’” (Tvinnereim & Fløttum, 2015; Tvinnereim et al., 2017a). Other studies have asked people “What should be done about climate change?” (Tvinnereim et al., 2017b) or what people think about other individuals’ beliefs about climate change (Mildenberger and Tingley, 2017).

To our knowledge, we present the first study that applies topic modelling to textual responses regarding climate policy measures in general and a carbon tax in particular. We analyze associations with a carbon tax and its acceptability as well as with fairness of the policy. The latter aims to achieve a better understanding of the dimensions of fairness perceptions regarding carbon taxation (Maestre-Andrés et al., 2019a). This study also adds to the literature by focusing on public

views from a sample of Spanish citizens, which complements prior studies that have mostly been undertaken in other European countries or the US (Bristow et al., 2010; Carattini et al., 2017; Kotchen et al., 2017; Jagers et al., 2018; Beiser-McGrath and Bernauer, 2019; Douenne and Fabre, 2020; see Carattini et al. 2018 and Maestre-Andrés et al, 2019 for overviews). Spain currently does not have an explicit carbon tax, nor is there any serious public or political debate about introducing it (see Appendix A for further details on Spanish context). This makes it a somewhat distinct case compared to countries that are frequently studied.

By means of just two open questions, we collect a diversity of associations with regard to a carbon tax and its fairness perception, which we then classify into meaningful topics and measure popularity of these topics among people with different socio-demographic characteristics. Our analysis finds that those who criticize the carbon-tax policy tend to worry about a high tax burden, a lack of low-carbon goods and services, tax exemptions for companies, and policy effectiveness if implementation is not international. These concerns are more often expressed by older people and those who frequently use a car. These findings help to identify social groups to target by particular information communication so as to increase support for a carbon tax.

2. Data and methods

In this study we draw on two open-ended questions from an online survey conducted in August 2019 in Spain. The sample includes 2004 participants and is representative of the general population in terms of age, gender, geographical distribution and political orientation, while it has, as is not uncommon for online surveys (Craig et al., 2013), somewhat higher levels of education and income (Table C1 in Appendix C compares the survey sample with the general population of Spain). Using an online panel, the survey had a response rate of 58.7%. Respondents took on average (median) about 15 (13) minutes to finish, with a minimum of five and a maximum of 70 minutes.

The survey was conducted in Spanish language. To avoid possible inconsistencies due to translation, we conduct our topic modelling analysis on the original texts, while below we provide our results with translations into English. The two open-ended questions aim to capture thoughts about the general idea of a carbon tax. This avoids that respondents are steered or limited in their answers by specific information details in questions, instead letting them freely associate with the notions raised in the question. Indeed, media coverage about climate policy, in particular carbon taxation, often refrains from offering extensive details about the policy. Most citizens are merely exposed to the general idea of a carbon tax, as reflected by our question. This is consistent with prior research which shows that initial, spontaneous thoughts about an idea or object shape subsequent thinking about it (Lodge and Taber, 2013). This approach is also in line with recent literature that uses the STM method in the context of climate change, asking questions of a similar general nature (Tvinnereim and Fløttum, 2015; Tvinnereim et al., 2017a). Knowledge obtained in this way can help to better frame information campaigns about the policy.

The survey started by presenting to respondents the following contextual statement: “Under the Paris Agreement from 2015, each country, including Spain, must implement policies to reduce their CO₂ emissions, which contribute to climate change. One major proposal to achieve emissions reduction is by implementing a carbon tax on fossil fuels whose combustion is the main cause of CO₂ emissions.” After this, the first open-ended question was formulated as follows:

“What do you think of a policy proposal to introduce a carbon tax? We would like to ask you to take your time and write a few sentences. All types of responses are welcome.”

This question was placed right at the beginning to avoid that information provided in the course of the survey would influence responses to it. We formulated the question as general as possible to avoid limiting the respondents in their answers through any specific information, thus allowing them to freely associate with the notion of a carbon tax. This approach is in line with recent literature that uses STM in the context of climate change (Tvinnereim and Fløttum, 2015; Tvinnereim et al., 2017a). As we show in Section 3, posing the question in this way allowed us to elicit many different concerns of the respondents.

The average length of responses to this question is 23 words, the median 17, and the minimum one. The left plot in Figure C1 in Appendix C demonstrates the complete distribution of length of answers skewed strongly to the right indicating that shorter answers tend to have a much higher frequency than longer ones. Short responses are typical for open-ended questions and should not be confused with non-responses. In fact, they are meaningful, as can be seen by examples of three-word responses such as “I think it is correct” (“Me parece correcto”), “I think it is fine” (“Me parece bien”); of two-word responses: “ok” (“de acuerdo”), “it is correct” (“es correcto”), “very bad” (“muy mal”), “very good” (“muy bien”), and one-word responses (which are rare): “fine” (“bien”), “valid” (“valido”), “favorable” (“favorable”), “populism” (“populismo”). When we later measure prevalence of topics across responses (Section 3), the STM algorithm adds up the shares of each topic in each response individually. Hence, all responses, regardless of length, have an equal weight in the quantitative analysis, reflecting the simple principle that each respondent is equally important to the survey.¹

Subsequently, we asked survey respondents a series of closed questions on attitudes to, and perceptions of, effectiveness and fairness regarding the carbon tax, for six distinct scenarios of using the tax revenues:

1. Unspecified revenue uses
2. Return all the revenues to compensate low-income households.
3. Support the development of climate projects (e.g. investing in public transport, planting trees, subsidies for renewable energy).
4. Use half of the revenues to support the development of climate projects and the other half to compensate low-income households.
5. Return the revenues in equal amount to all households as compensation.
6. Use half of the revenues to support development of climate projects and the other half to compensate all households in equal amount.

In addition to these questions on specific revenues, right before the second open question, respondents were asked what plays a more important role in their judgments of the policy, “fairness” or “effectiveness”. Hence, respondents had at this time of the survey considered a range

¹ The average length of our responses is comparable to that in other studies using STM (Tvinnereim and Fløttum, 2015; Tvinnereim et al., 2017b).

of possible dimensions of fairness.² To understand what underlies fairness perceptions, we asked a second open-ended question to participants:

“What does fairness mean for you in the previous questions? We would like to ask you to take your time and write a few sentences about it. All types of responses are welcome.”

The average length of responses is 15 words, the median 9 and the minimum one. The right plot in Figure C1 shows the complete distribution, which is also skewed strongly to the right.

Before proceeding with the analysis, the textual responses were cleaned for capital letters (setting everything to lower case), accents, stop words (such as “we”, “and”, “like”) and words with a length of less than 3 letters. Afterwards, responses were automatically stemmed, i.e. inflected words were reduced to their word stems, base or root form (e.g., the words ‘polluting’ and ‘pollute’ are reduced to the term ‘pollut’). Words appearing three times or less in all responses were deleted as rare words are hard to classify in any topic. This resulted in answers to the first open-ended question containing 870 (unique) words and 17604 tokens (i.e. word occurrences with repetition). Answers to the second open-ended question contain 613 words and 10487 tokens.

To identify dominant topics in the responses we rely on computational linguistics analysis. In particular, we use the structural topic modelling (STM) method, which generates topics using a clustering algorithm identifying the co-occurrence of words across survey responses. Formally speaking, a topic model operates as Bayesian inference of words related to a given topic and the topics being discussed in given response, based on analyzing responses already observed. In other words, a topic model tries to backtrack from the responses a set of topics that are likely to have generated them.

The advantage of topic modelling (TM) over simple keyword search is that it considers words not in isolation but together with other words that they appear with. This way words can have different meanings depending on their contexts. In particular, it assumes that each word in the responses is generated through a two-step process: first, that each response has its own topic distribution, and a topic is randomly drawn from it; second, that each topic has its own word distribution, and a word is randomly drawn from this distribution for the topic selected in the first step. Each response is essentially a result of multiple repetitions of the statistical model describing these two steps. This reflects that survey responses exhibit multiple topics, where each response exhibits the topics in different proportion. TM discovers the topic distribution for each response and the word distribution of each topic iteratively, by fitting this two-step procedure to the observed responses until it finds the best model that describes the underlying data. Similar to cluster analysis, TM reduces the dimensionality of linguistic data from words to topics based on the co-occurrence of words in a collection of responses to infer the underlying topics in those texts and the weight of each topic in each individual response. For example, if we observe the word “always” in the topic that we label as “We need affordable low-carbon transport”, it implies that this word appeared more frequently and exclusively in combination with other words in this topic, and that people expressing to be in favor of low-carbon transport used it more often.

² In fact, providing respondents with multiple questions about use of carbon tax revenues can help to elicit a variety of associations related to fairness, such as fairness related to a particular revenue use, what would be a fair policy in general, or what the notion of fairness means to them exactly.

An important advantage of STM over classical topic modelling (e.g. Latent Dirichlet Allocation) is that it incorporates information about the survey responses, such as the respondents' knowledge about the matter, individual acceptability of the carbon tax and socio-demographic characteristics. Hence, instead of assuming that topical prevalence (i.e. the degree to which a single response belongs to a given topic) and topical content are constant across respondents, we use the additional information about the authors as updated Bayesian priors in which we expect to see variance between responses. This has proven to result in higher quality topic identification and to provide more opportunities to include closed-ended responses in the analysis (Roberts et al., 2014).

To implement the STM method, we used an STM package in R by Roberts et al. (2019). Since the method is aiming to optimize over a complex multimodal search space, its results are sensitive to initialization. In other words, it provides different answers for different starting values of the Bayesian priors. Therefore, the employed algorithm chooses starting values based on the robust estimator using the method of moments (Arora et al. 2013; Roberts et al. 2019).

For the first open-ended question we use the following additional information (covariates) about the responses to form better quality topics: age, gender, education, car use, self-perceived knowledge about a carbon tax, and acceptability of a carbon tax (Figure C2 in the Appendix).³ Subsequently, this is illustrated by econometrically estimating the effects of covariates on topic prevalence across survey responses (see Section 3). Prior literature on climate policy attitudes in general, and on carbon taxes specifically, suggests that these variables can explain variation in public responses (Drews and van den Bergh, 2016; Carattini et al., 2018). For the second open-ended question we use the same set of covariates except for the last one which is replaced by 'perceived fairness of a carbon tax' (the lower plot in Figure C2).⁴

The first step in the STM analysis is deciding upon the number of topics k . The method then assigns to each response a vector with k weights, expressing the degree to which the response belongs to a particular topic. These values sum up to one. In an extreme case, if all values of the vector but one are zeros, the response completely belongs to one topic. Typically, however, a response belongs to multiple topics. These weights are called "topic prevalences". We further estimate topic prevalence for the whole dataset to measure relative topic size (see topic proportions, e.g., Table 1).

Defining the number of topics k , we consider topic model performance on three criteria, namely 'heldout log-likelihood' of the models (accuracy of the model to predict words from a sample that has been excluded from the estimation step), exclusivity (conditional probability of seeing the topic given the words) and semantic coherence (the extent to which frequent words from the same topic tend to appear in the same responses; or how well words from the same topic co-occur within each response). Thus, we follow Roberts et al. (2014) who argues that semantically

³ Because of many missing values present in responses for income and political views (approx. 25% and 15% respectively), we do not use these as covariates.

⁴ As we demonstrate in Figure C3 in Appendix C, we do not have any strong correlation between any pair of covariates included in each of the two topic models. Since perceived fairness of carbon tax and its acceptability are strongly positively correlated, we decided to not include them in the same model. We further estimate the Variance Inflation Factor (VIF) which measures how much the variance of a regression coefficient is inflated due to multicollinearity in the model. In all estimations, we observe a VIF value between 1.04 and 1.14, which is far below the conservative benchmark of 5, suggesting absence of multicollinearity.

interpretable topics should consist of words that tend to co-occur within responses, and that their top keywords are unlikely to overlap with keywords from other topics.

The top panel in Figure 1 shows the performance of alternative model specifications (k ranges from 3 to 20⁵) for the first open-ended question according to these three criteria. Selecting the optimal number of topics is challenging, particularly since coherence tends to fall with the number of topics, while the prediction accuracy and exclusivity, in contrast, tends to rise.⁶ Next to prediction power, exclusivity and coherence, a fourth implicit consideration in selecting a topic number is the model’s complexity: the larger the number of topics the harder are they to interpret. For that reason, once the marginal gain of an additional topic declines, we stop (see e.g. Huang et al. 2017 using a similar rule). On this basis, we find the preferred number of topics for the first open question to be 11, as highlighted by red color in each of the three graphs. This number has the same exclusivity as models with 15 or more topics, while potential loss in ‘heldout log-likelihood’ is more than compensated by lower model complexity and maintained coherence. If we consider the same performance criteria for the second open-ended question (lower panel in Figure 1), we choose a higher number of topics, namely 18. This choice of k performs well on predictive power of the model (i.e. ‘heldout log-likelihood’) and achieves a higher exclusivity of topics while not sacrificing semantic coherence (it is as high as in the model with 11-12 topics).

⁵ We limit the maximum number of topics to 20 in view of the relatively small data sample we are working with. Note that Tvinnereim et al. (2017b), who used a larger survey sample, considered a range between 3 and 12 topics only.

⁶ This reflects the earlier finding by Chang et al. (2009) that topic models with higher held-out log-likelihood are sometimes harder to interpret.

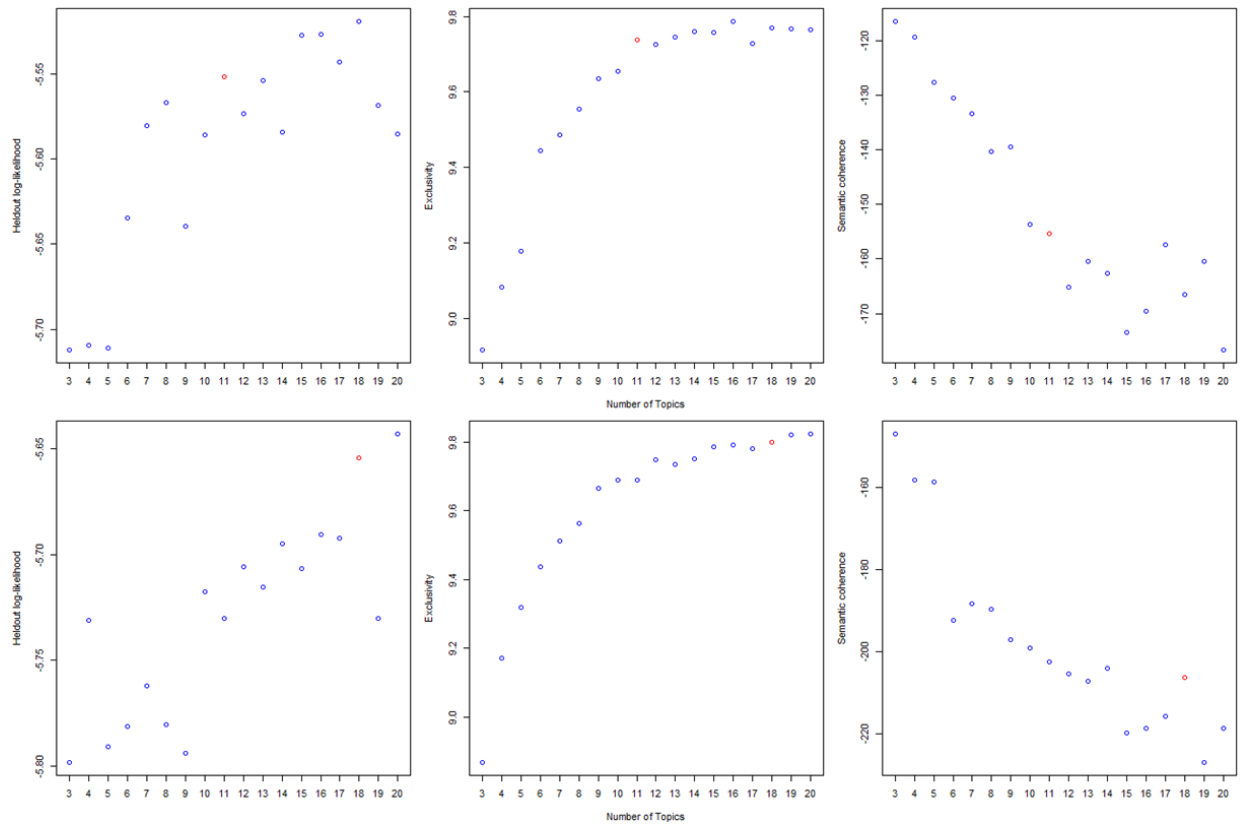


Figure 1. Model performance for distinct number of topics for the first (top panel) and the second (lower panel) open-ended questions

3. Results

3.1. Associations with the proposal to introduce a carbon tax

Table 1 offers a description of the 11 identified topics to the first open question about citizens' views on a carbon tax. It shows the most discriminating (frequent and exclusive) words by topics and the share of the text belonging to each of the topics (topic proportion), whereas Table 2 gives one illustrative statement per topic.⁷ After exploring the most frequent and exclusive words for each topic (see Table 1 and Figure 2) together with their illustrative responses, we come up with topic labels that reflect the main themes in a clear and concise way.

⁷ The statements come from the ten responses with the highest topic prevalence.

Table 1. Topics identified for responses to the first question

	Topic label	Most discriminating terms (in terms of frequency and exclusivity)	Topic proportion
1	We need measures to reduce emissions	measure, emission, reduce, big, take, level, possibl, contaminating, control, base <i>medida, emision, reducir, grand, tomar, nivel, posibl, contaminando, controlar, base</i>	8.8%
2	Stop climate change to save the planet	change, planet, climate, necessary, important, caring, unique, time, could, stop <i>cambio, planeta, climatico, necesario, important, cuidar, unica, tiempo, podria, frenar</i>	11.9%
3	There are already too many taxes	put, too much, upload, solution, leave, tax, bad, do, solution, breathe <i>poner, demasiado, subir, solucion, dejar, impuesto, mal, hacer, solucion, respirar</i>	11.5%
4	Develop renewable energy	energy, fuel, alternative, renew, fossil, clean, example, encourage, stimulate, solar <i>energia, combust, alternativa, renov, fosil, limpia, ejemplo, fomentar, incentivar, solar</i>	10.1%
5	This is a good policy	well, good, perfect, correct, pollut, appropriate, future, initiative, totally, idea <i>bien, buen, perfecto, correcta, contamin, correcto, futuro, iniciativa, totalment, idea</i>	9.9%
6	We need affordable low-carbon transport	car, vehicle, less, citizen, electric, politics, government, always, politician, help <i>coch, vehiculo, meno, ciudadano, electrico, politica, gobierno, siempr, politico, ayuda</i>	10.1%
7	A way to take more money from people	more, believe, money, should, think, pollute, equal, implement, new, class <i>mas, creo, dinero, deberia, pienso, contaminan, igual, implantar, nuevo, clase</i>	8.7%
8	Will encourage public transport and reduce pollution	environment, transport, better, public, improve, help, sufficient, collect, city, awareness <i>medio ambiente, transport, mejor, publico, mejorar, ayudar, suficient, recaudatorio, ciudad, conciencia</i>	6.7%
9	Good policy as long as applied to companies	company, particular, avoid, serve, insufficient, believe, apply, payment, seems, autonomous <i>empresa, particular, evitar, sirva, insuficient, opinar, aplicar, pago, parec, autónomo</i>	8.0%
10	It will only work if all countries participate	only, can, country, person, collect, great, create, missing, side, fine <i>solo, pue, pais, persona, recaudar, gran, crear, falta, lado, multa</i>	7.0%
11	It will just hurt final consumers	final, time, common, hold, resource, case, new, while, consumer, price <i>final, vez, pie, sosten, recurso, caso, nueva, mientras, consumidor, precio</i>	7.3%

Note: The terms shown are those that are the most frequent as well as exclusive to each topic. The original text was in Spanish. Here we report the Spanish terms and their English translation. Labels for each topic reflect the content of the terms and associated survey responses. Words are stemmed; e.g., the term ‘parec’ comprises ‘parecer’ and ‘parece’.

Table 2. Illustrative responses for topics constructed on the first open-ended question

<p>Topic 1. We need measures to reduce emissions All measures taken to reduce emissions seem good to me. <i>Todas las medidas que se tomen para reducir emisiones me parecen bien.</i></p>
<p>Topic 2. Stop climate change to save the planet All measures to eliminate polluting emissions are welcome. We must reverse climate change and achieve a clean and healthy planet. <i>Todas las medidas para eliminar las emisiones contaminantes son bienvenidas. Hay que revertir el cambio climático y conseguir un planeta limpio y saludable.</i></p>
<p>Topic 3. There are already too many taxes Very bad because we already have too many taxes. <i>Muy mal porque ya tenemos demasiados impuestos.</i></p>
<p>Topic 4. Develop renewable energy I don't believe in taxes, what they should do is facilitating the use of non-polluting alternative energies, hydrogen batteries, solar or wind energy. <i>No creo en los impuestos, lo que tendrían es que facilitar el uso de energías alternativas no contaminantes, pilas de hidrógeno, energía solar o eólica.</i></p>
<p>Topic 5. This is a good policy I think it is a good initiative. <i>Me parece una buena iniciativa.</i></p>
<p>Topic 6. We need affordable low-carbon transport It does not seem bad to me, but I would prefer help to buy hybrid, electric cars that consume less. In addition, the government should encourage the development of more efficient electric vehicles. <i>No me parece mal, pero preferiría ayudas para comprar coches híbridos, eléctricos y que consuman menos. Además el gobierno debería incentivar el desarrollo de vehículos eléctricos más eficientes.</i></p>
<p>Topic 7. A way to take more money from people I think it is an excuse to get more money from the taxpayer. <i>Creo que es una excusa para sacar más dinero al contribuyente.</i></p>
<p>Topic 8. Will encourage public transport and reduce pollution Correct policy to protect the environment. It promotes the use of public transport and / or less polluting ones. <i>Política correcta para proteger el medio ambiente. Fomenta el uso de transporte público y/ o menos contaminante.</i></p>
<p>Topic 9. Good policy as long as applied to companies I agree, as long as the pressure is more on enterprises and industry than on people. <i>Me parece bien, siempre y cuando sean las empresas y la industria a las que más se les exija y no tanto a las personas.</i></p>
<p>Topic 10. It will only work if all countries participate It would be fine, as long as all the countries signed it, since it is not worth doing it and neighboring countries taking advantage of that. <i>Pues estaría bien, siempre y cuando todos los países lo firmaran, ya que de nada sirve hacerlo tú y que el país de al lado se aproveche de eso.</i></p>
<p>Topic 11. It will just hurt final consumers I do not find it correct, since it is eventually paid by the final consumer and not by large corporations, which will increase the final price and its benefits with the excuse of the tax. <i>No lo encuentro correcto, ya que finalmente lo acaba pagando el consumidor final y no las grandes corporaciones, que con la excusa del impuesto aumentan el precio final y sus beneficios.</i></p>

While topics 1, 2, 5, 8 and 9 all seem to express positive associations with the idea to introduce a carbon tax, the remaining topics (3, 4, 6, 7, 10, 11) express negative or ambivalent views. To further validate our interpretation of the topics, we analyze the co-occurrence of the

topics in our responses, as shown in Figure D1 in Appendix D. In particular, we measure the correlation of topic prevalences in our responses. As one can see, topics 1, 2, 5 and 9 indeed tend to show up in the same responses exhibiting positive associations with the idea to introduce a carbon tax in Spain. Similarly, topics 3, 4, 6, 7 and 11 tend to co-occur in the responses with negative associations. Topic 10 on international coordination seems to also fall in the latter cluster of topics, although here the correlations do not follow the dichotomous structure as clearly as for the topics above. Lastly, topic 8 does not demonstrate any systematic positive or negative correlation with any of the two clusters of topics.

Word clouds indicating dominant terms for the resulting topics are presented in Figure 2. One of the advantage of word clouds over indicating most discriminating words as in Table 1 is that one can distinguish between frequent – with large font size – and exclusive – with dark color – words. For example, the words “help” and “improve” are not so frequent but very exclusive words for topic 8, while words “agreement” and “proposal” are frequent but not so exclusive for topic 2. These nuances are not visible from an aggregated measure. The word clouds also show that while some topics are largely dominated by few if not single words (like “good” and “always” in topic 5 or “more” and “money” in topic 7), others constitute combinations of words with more equal distribution of weights among them (such as topics 2, 4 and 6). These combinations of words provide insights that go beyond what one can learn from the topic label. For example, in topic 2 the words “important” and “necessary” come out strongly, which suggests the urgency expressed by respondents when they talk about the need to stop climate change.



Figure 2. Word clouds of eleven topics generated from the first open-ended question

Note: The font size corresponds to the probability (weight) of the respective word given the topic, while the color of the word corresponds to its exclusivity (the darker the color, the more exclusive are the words).

Now we proceed with analyzing factors explaining the variation of topic prevalences among the respondents. To this end, we provide the results of regression analysis of topic proportions over the six explanatory variables we used in building our topic model. In particular, a linear regression model was specified for each of the 11 topics (indexed by k) as follows:

$$\begin{aligned} \text{Topic Prevalence}_k \sim & \text{Constant}_k + \text{Age} + \text{Gender} + \text{Education} + \text{Car Use} \\ & + \text{Perceived knowledge of carbon tax} + \text{Acceptability of carbon tax} + \text{Residual}_k \end{aligned}$$

These results deliver several insights that can be understood as follows. Whenever we observe a prevalence of topic k to substantially differ across certain covariate, it implies that respondents characterized with different values of that covariate significantly differ in expressing that topic. Figure 3 displays the resulting regression lines with the most steep and significant coefficients (see Table D1 in Appendix D for more details). Incidentally, since the covariates are used by STM to form a topic model, this can explain the strong association between prevalence of some of the resulting topics and the observed characteristics of the respondents,

First of all, older people in Spain compared to younger ones tend to more strongly believe that there are already too many taxes (T3) and that carbon tax will work only if applied in many countries simultaneously (T10).⁸ Younger people, in turn, are more likely to find a carbon tax to be a good initiative as long as companies will carry the tax burden (T9). Concerning gender, we find least variation between the topics. The only exception is T9, showing that women compared to men are more likely to find a carbon tax a better idea which is consistent with results obtained from the closed-ended questions (Maestre-Andrés et al., 2019b).

More educated people, compared to less educated ones, are more likely to judge the carbon tax critically as they believe it will mainly affect final consumers (T11) and that Spain lacks sufficient renewable energy sources to allow for a switch to low-carbon consumption (T4). Less educated people in contrast are more likely to consider the carbon tax to be a good policy (T5) but also stress the existence of already many other taxes (T3).

Regular car users and rural inhabitants that are highly dependent on transport⁹ are more likely to stress the need to develop affordable low carbon transport alternatives to gasoline cars (T6) and consider the carbon tax as a way to retrieve more money from people (T7). Furthermore, they are much less likely to talk about the need to stop climate change (T2) and the necessity for countries to act together (T10) than those who never use a car or use it very infrequently.

People who evaluate their own knowledge about carbon tax as high (which is a small fraction – see Figure C2) tend to stress more the need to develop low carbon transport and renewable energy (T4 and T6), while those who think they do not understand the policy talk a lot more about having too many taxes (T3) and the need to stop climate change (T2).

Finally, people who completely accept the policy are more likely to mention the need to stop climate change in general (T2), as well as to judge the carbon tax as a good policy (T5). Those who completely oppose the policy express a general tax aversion (T3) and the need to develop low-carbon transport (T6).

⁸ Henceforth, we refer to topic # as T#.

⁹ The correlation between car use and number of inhabitants is -0.21 indicating that frequent car users often come from rural areas.

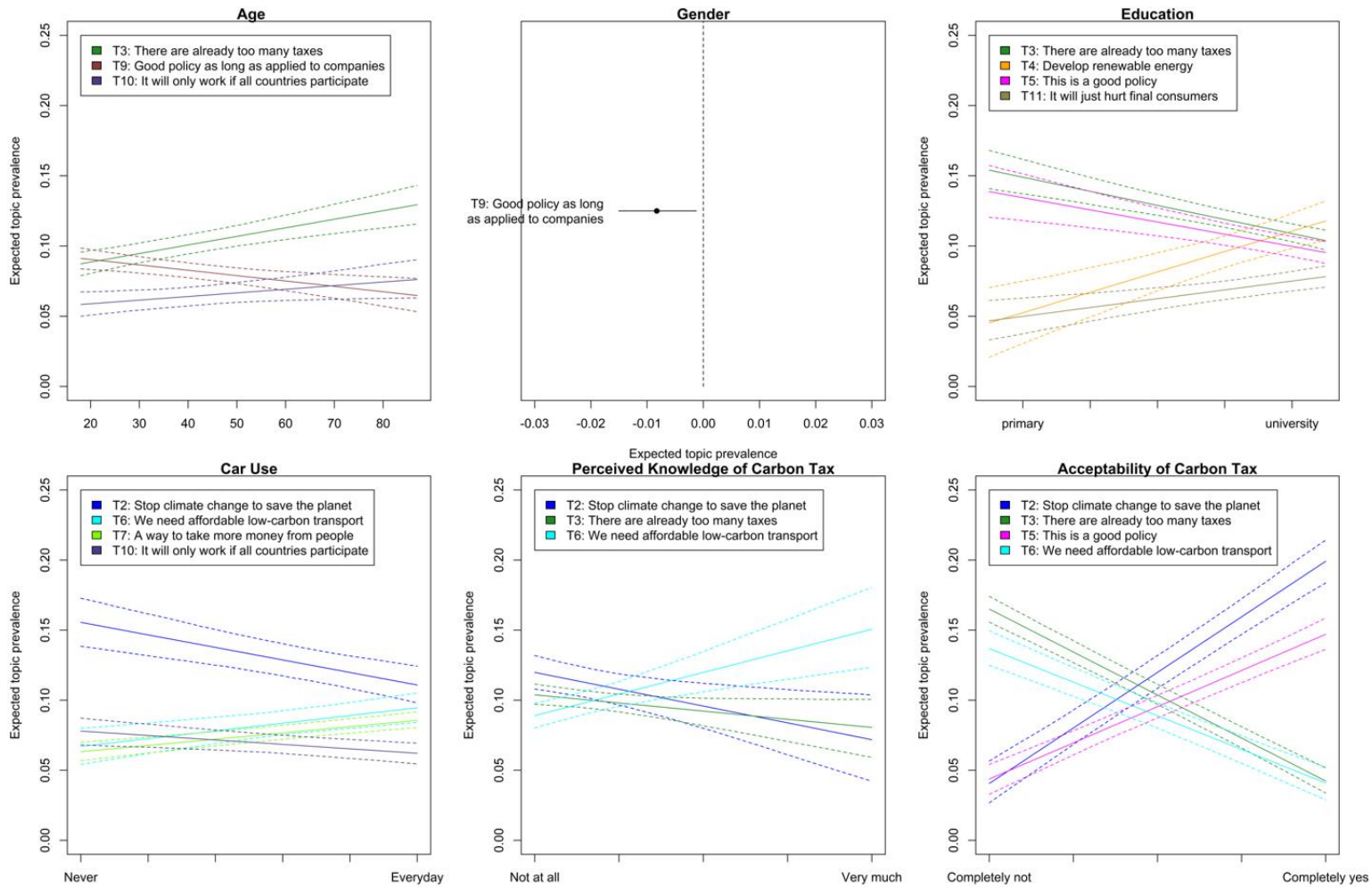


Figure 3. Effect of covariates on topic prevalence for associations with carbon tax implementation

Note: Values generated by a regression where the outcome variable is the proportion of each public response dedicated to each topic, given the selected STM model. The panel shows point estimates and confidence intervals of the effects of selected covariates on topic prevalence, holding all other covariates constant. The plot for gender shows mean difference in topic proportions between male and female (a positive value on the X-axis indicates a larger prevalence for men). Only a subset of topics demonstrating most variability across the respective covariate is displayed. Results for all topics are available in Appendix C. Confidence intervals plotted as dashed lines indicate the 95% uncertainty range and include both regression and measurement uncertainties associated with the STM model.

We tested our results for robustness by replicating the procedure on the first open question after excluding people who completed the survey very quickly (“speeders”), as they may have not thought as well about their responses. In particular, we discarded 218 (approx. 11%) respondents who used less than eight minutes on the whole survey (i.e. less than half of the time that our respondents on average required). The results are reported in Appendix D demonstrating that all our conclusions hold.

3.2. Associations with the fairness of a carbon tax

We now conduct the same analysis for the responses of the survey participants to the question of how they understand “fairness” in relation to carbon taxes. Table 3 and 4 show the most discriminating words and illustrative statements, respectively.

Table 3. Topics identified for responses to the second question

	Topic label	Most discriminating terms (frequency and exclusivity)	Topic proportion
1	The rich should pay more than the poor	always, poor, same, rich, pay, pollute, disadvantaged, return, payment, company <i>siempr, pobr, mismo, rico, pagamo, contaminan, desfavorecido, devuelvan, paga, empresa</i>	4.8%
2	Spend money on low-income households and renewable energy	energy, renew, receive, less, car, home, user, use, think, change <i>energia, renov, reciban, meno, coch, hogar, usuario, utilicen, creo, cambiar</i>	7.2%
3	Give everyone what they deserve	every, give, correspond, deserv, moral, respecting, distribution, say, day, give <i>cada, dar, correspond, merec, moral, respetando, reparto, decis, día, dando</i>	6.7%
4	Do not compensate the families who pollute	family, return, reason, bad, should, future, responsib, vehicle, improvement, sense <i>familia, devolv, razon, mal, deberia, futuro, respons, vehiculo, mejora, sentido</i>	4.2%
5	Justice does not exist in Spanish politics	exist, justice, any, depend, country, injustice, politician, concept, good, should <i>exist, justicia, ningun, depend, pai, injusticia, politico, concepto, buen, debemo</i>	6.1%
6	Use revenues to reduce emissions and save the planet	planet, emission, health, most, equitable, relationship, earth, save, avoid, issue <i>planeta, emision, salud, mayoria, equitativo, relacion, tierra, salvar, evitar, tema</i>	4.5%
7	Unfair tax	right, town, project, part, today, some, good, rest, cause, alternative <i>correcto, pueblo, proyecto, part, hoy, alguna, buen, resto, causa, alternativa</i>	5.1%
8	Do not benefit just one social class	class, low, income, high, average, benefit, people, economy, burden, fine <i>clase, baja, renta, alta, media, beneficiar, pue, economia, carga, multa</i>	4.3%

9	Be fair	fair, be, return, mean, benefit, impartial, technology, now, time, invest <i>justo, ser, devuelva, significa, benefica, ecuanim, tecnologia, ahora, tiempo, invertir</i>	6.5%
10	To not harm the poor, compensate them	function, equitable, economic, treat, redistribution, independently, purchasing, access, earn, people <i>funcion, equitativa, economica, tratar, redistribucion, independientemente, adquisitivo, acced, cobra, gen</i>	5.2%
11	Equal rights and compensation	law, person, distribute, equitably, comply, proportion, receive, right, obligation, same <i>ley, persona, repartir, equitativamente, cumplir, proporcion, recibir, derecho, obligacion, misma</i>	6.6%
12	Equality of basic living conditions and opportunities	equality, equal, condition, common, one, world, all, opportunity, good, equally. <i>igualdad, igual, condicion, comun, una, mundo, toda, oportunidad, bien, igualment</i>	7.3%
13	Use revenues to reduce pollution and make life better	do, equity, pollution, better, do, consumer, can, in addition, type, collected <i>hacer, equidad, contaminacion, mejor, haga, consumidor, pueda, adema, tipo, recaudado</i>	6.1%
14	Contribute to a just society	social, society, nobody, action, just, inequality, thing, correct, difference, harm <i>social, sociedad, nadi, accion, justa, desigualdad, cosa, correcta, diferencia, perjudiqu</i>	5.1%
15	Who pollutes more should pay more	possib, hurt, pay, more, see, result, pocket, pollut, need, beneficiary <i>posibl, perjudicado, pagar, mas, vean, salga, bolsillo, contamin, necesidad, beneficiado</i>	7.3%
16	Compensate those most in need	money, help, needy, benefit, carbon, support, compensate, solve, need, weak <i>dinero, ayudar, necesitado, benefici, carbono, apoyar, compensar, solucionar, necesita, debil</i>	5.9%
17	Use revenues for environment and climate	environment, change, climate, favor, said, harm, fight, care, humanity, context <i>medio ambiente, cambio, climatico, favorec, dicho, daño, combatir, cuidar, humanidad, contexto</i>	4.9%
18	A good level of economic wellbeing	level, must, have, go, individual, economic, solution, group, utopia, live <i>nivel, debe, tener, van, individuo, economico, solucion, grupo, utopia, vivir</i>	2.4%

Note: The terms shown are those that are the most frequent as well as exclusive to each topic. The original text was in Spanish. Here we report the Spanish terms and their English translation. Labels for each topic reflect the content of the terms and associated survey responses. Words are stemmed; e.g., the term ‘deserv’ comprises ‘deserved’ and ‘deserve’.

Table 4. Illustrative responses for topics constructed on the second open-ended question

<p>Topic 1. The rich should pay more than the poor May the poor not always pay the mistakes of the rich. <i>Que no paguen siempre los pobres los errores de los ricos.</i></p>
<p>Topic 2. Spend revenues on low-income households and renewable energy Climate change is a serious problem, a tax is fair if its revenues are used to create new renewable energy sources; but it would also be fair to use a small part of them to help low-income households afford the renewal of energy sources. <i>El cambio climático es un problema grave, si el impuesto se destina a crear nuevas fuentes de energía renovables es justo, pero también sería justo que una pequeña parte de este impuesto se destine a los hogares de bajos ingresos para que puedan afrontar la renovación de fuentes de energía.</i></p>
<p>Topic 3. Give everyone what they deserve Moral principle that inclines to act and judge respecting the truth and giving each one what corresponds. <i>Principio moral que inclina a obrar y juzgar respetando la verdad y dando a cada uno lo que le corresponde.</i></p>
<p>Topic 4. Do not compensate the families who pollute Even if it is reasonable that people with a low income use polluting vehicles, it is not reasonable to compensate them. They should, for example, use public transport or bicycles <i>Que sea razonable, si gente con bajos ingresos utiliza vehículos contaminantes no es razonable compensarlos. Deberían, por ejemplo, utilizar transporte público o bicicletas.</i></p>
<p>Topic 5. Justice does not exist in Spanish politics Justice is a dictionary word but there is no justice in this country with the politicians we have. <i>Justicia es palabra de diccionario pero no existe justicia en este país con los políticos que tenemos.</i></p>
<p>Topic 6. Use revenues to reduce emissions and save the planet The income obtained from a tax to reduce emissions should be used for the same purpose, to reduce emissions, improve air quality and ensure the health of the planet. <i>Los ingresos obtenidos por un impuesto para reducir emisiones deben destinarse al mismo fin, reducir emisiones, mejorar la calidad del aire y garantizar la salud del planeta.</i></p>
<p>Topic 7. Unfair tax I think it is an unfair tax. Simply. <i>Lo veo un impuesto injusto. Simplemente.</i></p>
<p>Topic 8. Do not benefit just one social class Justice would be the equitable distribution of the tax burden between the most disadvantaged social strata and the middle and upper classes. <i>Justicia sería el reparto equitativo de las cargas impositivas entre las capas sociales más desfavorecidas y clases medias y altas.</i></p>
<p>Topic 9. Be fair Be just. <i>Ser justo.</i></p>
<p>Topic 10. To not harm the poor, compensate them It does not exacerbate the situation of the people currently in the worse situation. <i>Que no suponga empeorar la situación de las personas actualmente en peor situación.</i></p>
<p>Topic 11. Equal rights and compensation Redistribute equitably and have the same rights. <i>Repartir equitativamente y tener los mismos derechos.</i></p>
<p>Topic 12. Equality of basic living conditions and opportunities It means fair conditions for everyone regarding equal opportunities. <i>Significa una equidad de condiciones para todas las personas en lo que se refiere a igualdad de oportunidades.</i></p>

<p>Topic 13. Use revenues to reduce pollution and make life better I would not find 'fair' that money generated by a 'pollution' tax (in my view, it would be a punishment for polluting), would be used in other things that were not aiming to reduce that pollution and seek alternatives as soon as possible to avoid precisely that tax. <i>No encontraría 'justo' que un dinero generado por un impuesto por 'contaminación' (a mi modo de ver, sería como un castigo por contaminar) se utilizara en otras cosas que no fueran el revertir esa contaminación y buscar cuanto antes alternativas para evitar precisamente ese impuesto.</i></p>
<p>Topic 14. Contribute to a just society Let us all participate in accordance with our income in a more just and equal society with good relations among all humans <i>Que todos participemos en funcion de nuestra renta en una sociedad mas justa e igualitaria y de calidad entre todos los humanos</i></p>
<p>Topic 15. Who pollutes more should pay more Equity exists in terms that the one who pollutes the most, pays the most. <i>Que exista equidad en cuanto a que el que más contamine pague más.</i></p>
<p>Topic 16. Compensate those most in need Help the most affected groups and help them to adapt to the new measures. <i>Ayudar a los colectivos más afectados y ayudarles a adaptarse a las nuevas medidas.</i></p>
<p>Topic 17. Use revenues for environment and climate If it is a tax to fight climate change, it would be fair that the revenues are used, effectively, to generate programs that help fighting climate change. <i>Si se trata de un impuesto para combatir el cambio climático, lo justo es que lo recaudado se emplee, de forma efectiva, en generar programas que ayuden a combatir el cambio climático.</i></p>
<p>Topic 18. A good level of economic wellbeing To maintain a minimum level of equality. If some people have a lot in excess and others lack a lot, bring these limits closer, for example. Squeeze more the one who has more and even pressing the one who has less, who has to contribute as well, help him (her) to live in an honest and dignified way. <i>Mantener un nivel mínimo de igualdad. Si a gente le sobra mucho y otra gente le falta mucho. Acercar estos límites, por ejemplo. Apretar más al que más tiene y aun apretando al que menos tiene, que también ha de aportar, que también se le ayude a él a vivir de una forma honrada y digna.</i></p>

Overall, 18 topics were generated, which gives us an even more diverse picture than for the analysis of the first open-ended question. Here we again examined the illustrative responses and discriminating words to derive appropriate topic labels. These indicate that while certain topics stress equity, equality and fairness more generally (topics 3, 8, 9, 11, 12, 14), others highlight the different tax revenue uses (topics 2, 4, 10, 16), the need to protect the environment (topics 6, 13, 17) and the issue of distribution of the tax burden in the economy (topics 1, 8, 15). Less frequently were stressed the problem of unfairness in Spain (T5) and unfairness of the carbon tax in particular (T7).

To reduce complexity and to understand how these topics relate to one another, we again look on the topic co-occurrence in our responses by measuring the correlation of their prevalences. The result is shown in Figure D2 in Appendix D. While the picture is more complex compared with just two clusters of topics observed in the first open-ended question, here we also see several clusters of topics appearing together. Topics 2, 4, 16 and 17 form one such cluster. All these topics discuss different possibilities to redistribute revenues, mostly in favor of poor people. Another cluster is formed by topics 6, 9 and 13. Two of them emphasize the need to use carbon tax revenues to limit pollution, while one is about fairness in general. In addition, topics 1 and 15 form a strong pair (i.e. correlation close to 0.5): people stressing that the rich should pay more than the poor are likely to complement this argument by suggesting that richer people and firms polluting more

should also pay more. A similar strong pair is formed by topics 11 and 12 which both address equality: equal rights and equitable compensation next to equal basic living conditions and opportunities. This pair is in addition frequently accompanied by topic 14 on a just society (correlation with those two topics in the order of 0.3). A less strong but consistent pair of topics is 5 and 7: people who express the idea that justice does not exist in Spanish politics tend to stress that carbon tax is unfair.

Figures D1 and D2 show one way how STM adds value compared to classical topic modelling, namely because it accounts for correlation between topics. This is in line with the idea that people tend to connect certain topics (Blei and Lafferty, 2007). For example, people evaluating the possibility to introduce a carbon tax as a good instrument can condition this on the tax being applied to companies. In this way, STM yields more accurate topic classification. An additional step we make in Figures D1 and D2 is to demonstrate related correlations between the topics rather than looking on the topics in isolation. This helps to better understand the topics and arrive at an overall picture of public perception.

Word clouds for the resulting topics are presented in Figure 4. Again, while some topics are largely dominated by few words (like “justice” in topic 5 or “more” in topic 15), in other topics frequencies of occurrence for constituent words are more equally spread.



Figure 4. Word clouds of eighteen topics generated from the second open-ended question

Note: Font size corresponds to the probability (weight) of the respective word for each topic, while the color of the word corresponds to its exclusivity (the darker the color, the more exclusive are the words).

In Figure 5 we demonstrate variation of topic prevalence across several respondent characteristics. This is obtained by a regression analysis of topic proportions over the six explanatory variables (here instead of acceptability we use fairness perception) summarized in the following regression equation:

$$\begin{aligned} \text{Topic Prevalence}_k \sim & \text{Constant}_k + \text{Age} + \text{Gender} + \text{Education} + \text{Car Use} \\ & + \text{Perceived knowledge of carbon tax} + \text{Fairness perception of carbon tax} \\ & + \text{Residual}_k \end{aligned}$$

Again, whenever we observe a prevalence of topic k to substantially differ across certain covariates, it implies that respondents with different values of that covariate significantly differ in expressing that topic (see Table D2 in Appendix D for more results). Figure 5 demonstrates that older people tend to assess fairness in Spanish politics much more critically than younger ones (T5) and more strongly propose a redistribution of revenues to low-income households and renewable energy consumption (T2). In contrast, younger respondents are more likely to talk about fairness in general (T9) and equality among people (T12).

In terms of gender, similar to the first question, we see less heterogeneity among respondents. While women tend to more strongly stress fairness (T9), men worry more about a carbon tax potentially biasing benefits to a particular social class (T8) and exhibit less trust in politicians (T5).

More educated people tend to more strongly favour redistributing tax revenues to poor households (T10) and warn against redistributing revenues to people with high carbon footprint (T4). Less educated people – similar to younger respondents – stress more the issues of fairness and equality (T9 and T12) and – here rather in line with older respondents – critically consider the fairness in Spanish politics (T5). People who use the car much more frequently tend to more strongly stress the issue of fairness (T9).

People with high self-assessed knowledge about carbon tax fear more strongly the risk of disproportionately high benefits for a particular social class (T8) and – perhaps in line with their pro-environmental behavior – stress more the need to care for the environment (T17), while those respondents who think they know little about a carbon tax talk about fairness in general (T9) and warn against redistributing money to families with high carbon footprint (T4).

Finally, people who judge the carbon tax to be a fair policy stress more the need to take care for the environment (T6 and T17), while those who find it most unfair most logically express so (T7), fear disproportional benefits for some social class, particularly the rich people (T8 and T1) and have little trust in the Spanish government (T5).

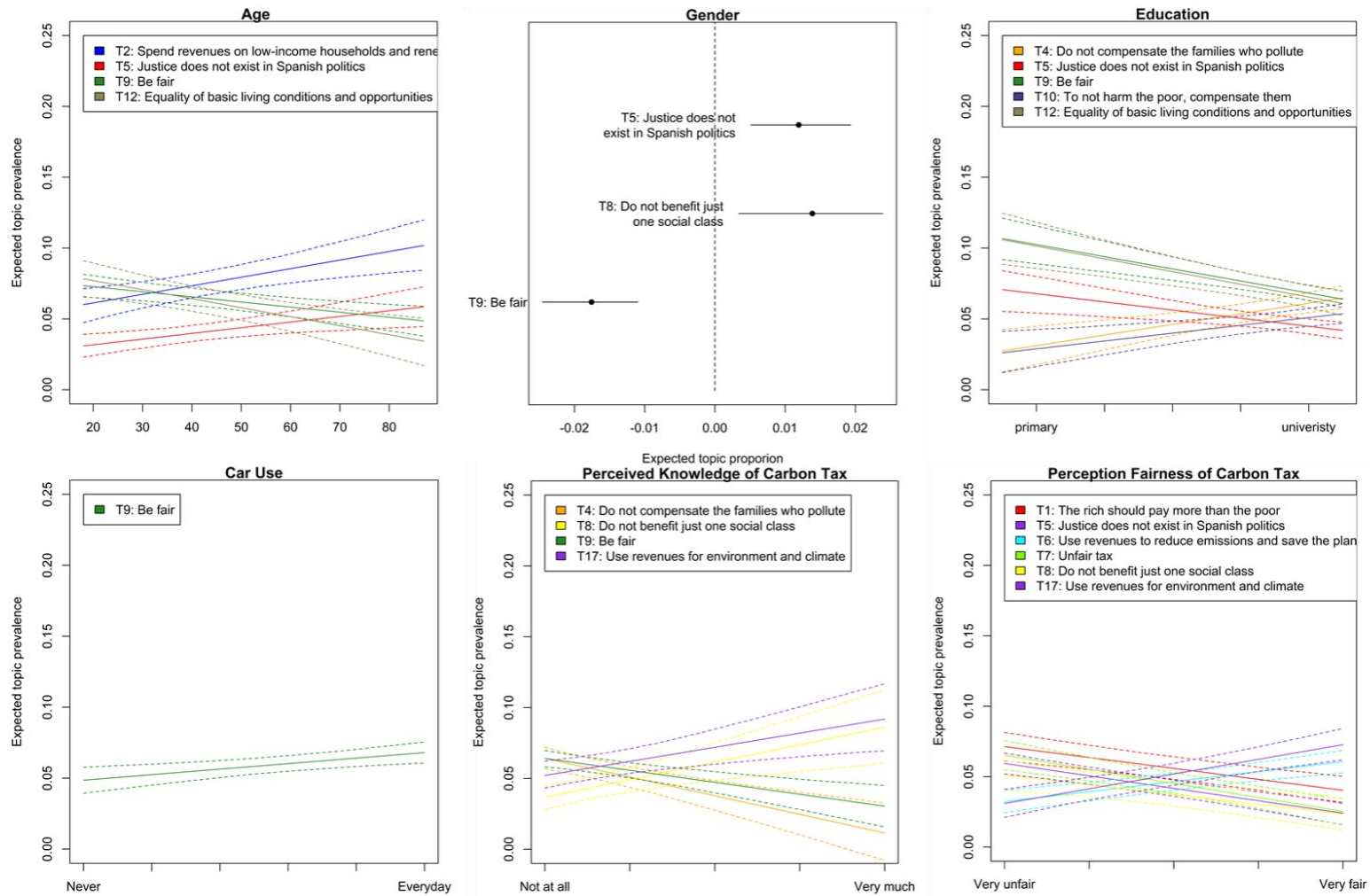


Figure 5. Effect of covariates on topic prevalence for associations with fairness with regard to carbon tax revenue use in Spain

Note: Values generated by a regression where the outcome variable is the proportion of each public response dedicated to each topic, given the selected STM model. The panel shows point estimates and confidence intervals of the effects of selected covariates on topic prevalence, holding all other covariates constant. The plot for gender shows mean difference in topic proportions between male and female (a positive value on the X-axis indicates a larger prevalence for men). Only a subset of topics demonstrating most variability across the respective covariate is displayed. Results for all topics are available in Appendix C. Confidence intervals plotted as dashed lines indicate the 95% uncertainty range and include both regression and measurement uncertainties associated with the STM model.

3.3. Associations between topics from the two questions

Finally, we examine topic co-occurrence between the two open-ended questions. In particular, we measure covariance between topic prevalences for the two questions. This will help us to establish a link between associations of people concerning carbon tax in general (Q1) and the fairness perception (Q2). The result of this exercise is presented in Figure D3 in Appendix D.

Several observations are in order. First, those who think that the carbon tax is just another tax to raise money (T7 for Q1) or to increase tax burden (T3 for Q1) have little trust in (Spanish) politicians, think that the rich should pay more than the poor, consider it generally unfair, or express little concern for climate change. This observation is in line with the literature on carbon pricing demonstrating that countries with greater public distrust of politicians tend to have higher greenhouse gas emissions, weaker climate policies, and lower public support (Klenert et al., 2018; Maestre-Andrés et al. 2019a). It is also consistent with public trust in political institutions having been documented to be rather low in Spain (Torcal, 2014). Second, similar concerns are expressed by participants who stress the lack of low-carbon transport (T6 in Q1) and the risks for final consumers to carry the whole tax burden (T11 in Q1). Third, people accepting the carbon tax emphasize the need to combat environmental problems in their responses to Q1, tend to talk about fairness and a just society in response to Q2.

4. Conclusions

This article has examined public perceptions about a carbon tax in Spain by analyzing open-ended survey responses on carbon pricing and its fairness. Using structural topic modelling, a considerable variety of topics was identified for both questions. Most notably, citizens who oppose carbon taxes often argue that too many taxes already exist, that a carbon tax is just a new way of taking money out of their pockets, and that instead renewable energy and low-carbon transport should be promoted. Furthermore, some question the effectiveness of such an initiative if applied nationally only. In contrast, those who express positive views do so by stressing their general concern for climate change and the need to take action, their considerations that the policy is ‘good’ or reasonably good depending on whether it is applied to companies and not directly to consumers. A few respondents also believe that a carbon tax will encourage public transport and reduce pollution. These findings largely resonate with those of the literature which is mostly based on the analysis of closed-ended questions (e.g. Carattini et al. 2018). An advantage of our approach is that we obtain a rich palette of associations with a carbon tax, without any restriction on the type of response.

Although the idea of a ‘fair’ carbon tax has received considerable attention, fairness remains an ambiguous qualification (Maestre-Andrés et al., 2019a). Here we identified a large number of topics clarifying how people understand ‘fairness’ in relation to carbon taxation and how its revenues are spent. Some topics stress avoidance of regressive policy effects and protecting sensitive groups such as low-income households through financial transfers and assuring that those who pollute more should pay more. In contrast, other topics reflect a concern that the use of tax revenues should not benefit just one social class but should be distributed equitably. Finally, some consider the policy to be fair if it contributes directly, or indirectly via spending of revenues, to reducing CO₂ emissions and slowing down climate change. In particular this last dimension of fairness has not been sufficiently recognized in prior studies using closed-ended survey questions.

While it is known that perceived environmental effectiveness increases policy acceptability directly (Maestre-Andrés et al., 2019a), it may thus also do so indirectly by increasing fairness perceptions.

A main finding of calculating correlations between the responses to the two questions is that people who reject a carbon tax, in comparison with those who accept it, tend to show less trust in politicians, emphasize more that the rich should pay more than the poor, think more that there is a lack of renewable energy and low-carbon transport, and express less concern for climate change. On the other hand, people accepting a carbon tax more strongly stress the need to solve environmental problems and show more care about a just society. Overall, these results help to clarify what people have in mind when they respond to closed-ended, single-item survey questions capturing perceived fairness of the policy, demonstrating an additional, complementary value of the method. Future research using closed-ended questions may draw on our derived topics to test the effects of different fairness principles on policy acceptance.

Our study provides further insights into how topic prevalence varies across individual characteristics. For example, older, less educated and car-dependent respondents more frequently voiced the concern that they pay too many taxes already. This might be related to some people's worry about a tax that is supposed to trigger a behavioral response when they cannot see a way to reduce their car dependency or change to a more expensive low-carbon vehicle. Those who oppose carbon taxes most often point to the need to develop alternative low-carbon transport, which suggests that mobility is key to understanding public resistance. With regard to fairness, returning at least part of the revenues to citizens appears as a strategy that can enhance public acceptability, as indicated by the three topics relating to revenue use (2, 10 and 16) comprising almost 20 percent of responses related to this issue. Moreover, it seems that exempting industries or companies from a carbon tax – a common practice in many carbon price schemes – may be felt by many citizens as unfair, which has recently also been found in research on related energy policy (Andor et al., 2019). Taken together, the findings can be useful for policy communicators, for example, in framing the content of campaigns and targeting specific groups (Lachapelle, 2017). This, however, will depend on different national contexts, which is why it would be fruitful to extend research on carbon taxation with topic modelling to other countries.

References

- Andor, M.A., Frondel, M., Sommer, S., 2018. Equity and the willingness to pay for green electricity in Germany. *Nat Energy* 3, 876–881. <https://doi.org/10.1038/s41560-018-0233-x>
- Arora S., Ge R., Halpern Y., Mimno D., Moitra A., Sontag D., Wu Y., Zhu M. (2013) A practical algorithm for topic modeling with provable guarantees, Dasgupta S., McAllester D. (Eds.), *Proceedings of the 30th International Conference on Machine Learning*, Volume 28 of *Proceedings of Machine Learning Research*, PMLR, Atlanta, Georgia, USA, 280-288
- Beiser-McGrath, L.F., Bernauer, T., 2019. Could revenue recycling make effective carbon taxation politically feasible? *Sci. Adv.* 5, eaax3323. <https://doi.org/10.1126/sciadv.aax3323>
- Blei D.M., Lafferty J.D. (2007) A correlated topic model of science, *The Annals of Applied Statistics* 1(1)m 17-35
- Bristow, A.L., Wardman, M., Zanni, A.M., Chintakayala, P.K., 2010. Public acceptability of personal carbon trading and carbon tax. *Ecological Economics* 69, 1824–1837. <https://doi.org/10.1016/j.ecolecon.2010.04.021>
- Carattini, S., Baranzini, A., Thalmann, P., Varone, F., Vöhringer, F., 2017. Green Taxes in a Post-Paris World: Are Millions of Nays Inevitable? *Environmental and Resource Economics*. <https://doi.org/10.1007/s10640-017-0133-8>
- Carattini, S., Carvalho, M., Fankhauser, S., 2018. Overcoming public resistance to carbon taxes. *Wiley Interdisciplinary Reviews: Climate Change* e531. <https://doi.org/10.1002/wcc.531>
- Chang, J., Boyd-Graber, J., Gerrish, S., Wang, C., Blei, D. M. (2009). Reading tea leaves: How humans interpret topic models. In *Advances in neural information processing systems NIPS*, 288–296
- Craig B.M., Hays R.D., Pickard A.S., Cella D., Revicki D.A., Reeve B.B., 2013. Comparison of US panel vendors for online surveys, *Journal of Medical Internet Research* 15(11): e260
- Douenne, T., Fabre, A. 2020. French attitudes on climate change, carbon taxation and other climate policies. *Ecological Economics* 169, 106496.
- Drews, S., van den Bergh, J.C.J.M., 2016. What explains public support for climate policies? A review of empirical and experimental studies. *Climate Policy* 16, 855–876. <https://doi.org/10.1080/14693062.2015.1058240>
- Huang, A., Lehavy, R., Zang, A. Y., and Zheng, R. (2017). Analyst information discovery and interpretation roles: A topic modeling approach. *Management Science*, 64(6):2833-2855
- Jagers, S.C., Martinsson, J., Matti, S., 2018. The impact of compensatory measures on public support for carbon taxation: an experimental study in Sweden. *Climate Policy* 1–14. <https://doi.org/10.1080/14693062.2018.1470963>
- Klenert, D., Mattauch, L., Combet, E., Edenhofer, O., Hepburn, C., Rafaty, R., Stern, N., 2018. Making carbon pricing work for citizens. *Nature Climate Change* 8, 669–677. <https://doi.org/10.1038/s41558-018-0201-2>
- Kotchen, M.J., Turk, Z.M., Leiserowitz, A.A., 2017. Public willingness to pay for a US carbon tax and preferences for spending the revenue. *Environmental Research Letters* 12, 094012. <https://doi.org/10.1088/1748-9326/aa822a>
- Lachapelle, E., 2017. Communicating about Carbon Taxes and Emissions Trading Programs. *Oxford Research Encyclopedia of Climate Science*. <https://doi.org/10.1093/acrefore/9780190228620.013.431>

- Leiserowitz, A. (2006). Climate Change Risk Perception and Policy Preferences: The Role of Affect, Imagery, and Values. *Climatic Change*, 77(1–2), 45–72. doi:10.1007/s10584-006-9059-9
- Lesnikowski, A., Belfer, E., Rodman, E., Smith, J., Biesbroek, R., Wilkerson, J.D., Ford, J.D., Berrang-Ford, L., 2019. Frontiers in data analytics for adaptation research: Topic modeling. *Wiley Interdisciplinary Reviews: Climate Change* 10, e576. <https://doi.org/10.1002/wcc.576>
- Lodge, M., Taber, C.S., 2013. *The Rationalizing Voter*. Cambridge University Press.
- Lorenzoni, I., Leiserowitz, A., De Franca Doria, M., Poortinga, W., & Pidgeon, N. F. (2006). Cross-National Comparisons of Image Associations with “Global Warming” and “Climate Change” Among Laypeople in the United States of America and Great Britain. *Journal of Risk Research*, 9(3), 265–281. doi:10.1080/13669870600613658
- Maestre-Andrés, S., Drews, S., van den Bergh, J., 2019a. Perceived fairness and public acceptability of carbon pricing: a review of the literature. *Climate Policy* 19, 1186–1204. <https://doi.org/10.1080/14693062.2019.1639490>
- Maestre-Andrés, S., Drews, S., Savin, I., van den Bergh, J., 2019b. Carbon tax acceptability with information provision and mixed revenue uses. Working paper, ICTA-UAB.
- Mildenberger, M., Tingley, D., 2017. Beliefs about Climate Beliefs: The Importance of Second-Order Opinions for Climate Politics. *British Journal of Political Science* 1–29. <https://doi.org/10.1017/S0007123417000321>
- Moloney, G., Leviston, Z., Lynam, T., Price, J., Stone-Jovicich, S., & Blair, D. (2014). Using social representations theory to make sense of climate change: What scientists and nonscientists in Australia think. *Ecology and Society*, 19(3). doi:10.5751/ES-06592-190319
- Roberts, M.E., Stewart, B.M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S.K., Albertson, B., Rand, D.G., 2014. Structural Topic Models for Open-Ended Survey Responses. *American Journal of Political Science* 58, 1064–1082. <https://doi.org/10.1111/ajps.12103>
- Roberts, M.E., Stewart, B.M., Tingley, D. 2019. stm: An R Package for Structural Topic Models. *Journal of Statistical Software* 91 (2), doi: 10.18637/jss.v091.i02.
- Torcal, M., 2014. The Decline of Political Trust in Spain and Portugal: Economic Performance or Political Responsiveness? *American Behavioral Scientist* 58, 1542–1567. <https://doi.org/10.1177/0002764214534662>
- Tvinnereim, E., Fløttum, K., 2015. Explaining topic prevalence in answers to open-ended survey questions about climate change. *Nature Climate Change* 5, 744–747. <https://doi.org/10.1038/nclimate2663>
- Tvinnereim, E., Liu, X., Jamelske, E.M., 2017a. Public perceptions of air pollution and climate change: different manifestations, similar causes, and concerns. *Climatic Change* 140, 399–412. <https://doi.org/10.1007/s10584-016-1871-2>
- Tvinnereim, E., Fløttum, K., Gjerstad, Ø., Johannesson, M.P., Nordø, Å.D., 2017b. Citizens’ preferences for tackling climate change. Quantitative and qualitative analyses of their freely formulated solutions. *Global Environmental Change* 46, 34–41. <https://doi.org/10.1016/j.gloenvcha.2017.06.005>

Appendix to “Public views on carbon taxation and its fairness: A computational linguistics analysis”.

Appendix A. Context of carbon taxes in Spain

Spain has historically not had an explicit carbon tax on all products but has high levels of taxes on energy use. In 2014, Spain implicitly priced over 30% of the country’s total carbon emissions above €30/ CO₂-ton, with the highest taxes generally being placed on the transportation sector (OECD, it does not specify the year). In 2019, Spain again priced a similar percentage of around 30% of emissions at an average of €15/CO₂-ton, covering a mere 3% of its total emissions output (Pettigrew, 2020). Moreover, as a member of the European Union, Spain is covered by the EU ETS. According to the World Bank (2019), Spain has a carbon tax since 2014 (\$17 per ton of CO₂), however this is a tax aimed at reducing fluorinated greenhouse gases (F-gases) (see the entry for Spain in the Carbon Pricing Dashboard of the World Bank: https://carbonpricingdashboard.worldbank.org/map_data). This can be interpreted as implicit carbon price, but to call them an explicit carbon tax would be inaccurate. Moreover, there is hardly any discussion about carbon pricing in Spanish media or society. For example, an analysis of coverage of climate change mitigation in Spanish newspapers shows that although “economic incentives” receive some attention when discussing solutions, carbon taxation overall does not stand out as an important issue (Fernández-Reyes and Iménez Gómez, 2019). Finally, note that literature reviews of carbon taxation tend not to mention Spain (e.g., Haites, 2018).

References

- Fernández-Reyes, R., and Iménez Gómez, I. 2019. Media coverage of climate change mitigation in the Spanish press. In: *La comunicación de la mitigación ante la emergencia climática*. ISBN 9788418167034, p. 207-246.
- Haites, E. (2018). Carbon taxes and greenhouse gas emissions trading systems: what have we learned? *Climate Policy* 18(8): 955-966.
- Pettigrew S. 2020. Spanish carbon pricing the lowest in the EU. <https://www.climatecorecard.org/2020/03/spanish-carbon-pricing-the-lowest-in-the-eu/>
- OECD, it does not specify the year. Effective carbon rates. Pricing CO₂ through taxes and emission trading systems. Available at <https://www.oecd.org/tax/tax-policy/environmental-tax-profile-spain.pdf>
- World Bank. 2019. “State and Trends of Carbon Pricing – 2019.” Washington, DC.

Appendix B. Survey questions used in this study

Under the Paris Agreement from 2015, each country, including Spain, must implement policies to reduce their CO₂ emissions, which contribute to climate change. One major proposal to achieve emissions reduction is by implementing a carbon tax on fossil fuels whose combustion is the main cause of CO₂ emissions.

1. Do you know how a carbon tax works? (Response options: not at all, a little, somewhat, a lot, very much)

2. Here are several sentences about carbon taxation. Can you tell us whether you think they are true or false? (Response options: “true”, “false”, “do not know”)

- A carbon tax is levied on the carbon content of fossil fuels, such as coal and oil.
 - A carbon tax mandates all producers and consumers which low-carbon technology they should adopt.
 - A carbon tax makes renewable energy sources, such as solar electricity, more expensive than fossil fuels.
 - A carbon tax imposes a legally binding limit on the amount of CO₂ emissions that firms and consumers are allowed to emit.
- A carbon tax allows reducing other, existing taxes such as VAT or labour taxes.
- A carbon tax will raise the price of coal and reduce the price of gasoline.

The following information is provided to half of the sample (group 1) about how a carbon tax works:

“A carbon tax is a charge on fossil fuels in proportion to the amount of carbon they contain as this determines how many CO₂ emissions result from their combustion. This will, for instance, raise the price of coal more than that of gasoline and the latter more than that of natural gas. Producers and consumers are then stimulated to switch to renewable energy, save energy on heating, alter fuel-based transport, etc. Because fuel prices alter, the prices of other products and services throughout the economy will change as well: the ones that generate considerable CO₂ in production will become more expensive, while prices are likely to alter little or remain the same for products and services that cause little or no CO₂ during production. A significant carbon tax thus encourages all firms and household to shift to goods and services that use fewer high-carbon energy sources during their production.”

3. How effective do you think a carbon tax is for reducing CO₂ emissions? (Response options: Very ineffective, ineffective, neither ineffective nor effective, effective, very effective).

4. How fair or unfair do you consider a carbon tax? (Response options: very unfair, somewhat unfair, neither unfair nor fair, somewhat fair, very fair).

5. Who do you think should carry most of the burden of the carbon tax? (Response options: businesses, consumers, both, none).
6. How do you think a carbon tax will affect you personally? (Response options: I would be much worse off, I would be somewhat worse off, I would be neither worse off nor better off , I would be somewhat better off , I would be much better off).
7. How do you think a carbon tax will affect low-income households? (Response options: they would be much worse off, they would be somewhat worse off, they would be neither worse off nor better off, they would be somewhat better off , they would be much better off).
8. How trustworthy do you think politicians are in implementing the carbon tax properly? (Response options: very untrustworthy, untrustworthy, neither untrustworthy not trustworthy, trustworthy, very trustworthy).
9. How acceptable do you find a carbon tax? (Response options: completely unacceptable, somewhat unacceptable, neither unacceptable nor acceptable, somewhat acceptable, completely acceptable).
10. Which of these two objectives do you think is the main purpose of a carbon tax? (Response options: to generate revenues, to change behaviour of producers and consumers, do not know).

Carbon taxes generate revenues which can be used for different purposes. As an illustrative example, for instance, according to one estimate, a (low) carbon tax of €5/ton CO₂ would already generate approximately €1.3 billion of additional government revenues. To put this number in perspective: Spain's expenditures for education were €2.6 billion in 2018.

Here we present five options to use the revenues:

- Return all the revenues to compensate low-income households.
- Support the development of climate projects (e.g. investing in public transport, planting trees, subsidies for renewable energy).
- Use half of the revenues to support the development of climate projects and the other half to compensate low-income households.
- Return the revenues in equal amount to all households as compensation.
- Use half of the revenues to support development of climate projects and the other half to compensate all households in equal amount.

11. How effective do you think the carbon tax is for reducing CO₂ emissions if its revenues are used to [the question is repeated for each of the five revenue uses mentioned above]? (Response options: Very ineffective, ineffective, neither ineffective nor effective, effective, very effective).
12. How fair or unfair do you consider a carbon tax if its revenues are used to [the question is repeated for each of the five revenue uses mentioned above]? (Response options: very unfair, somewhat unfair, neither unfair nor fair, somewhat fair, very fair).
13. How do you think a carbon tax affects you personally if its revenues are used to [the question is repeated for each of the five revenue uses mentioned above]? (Response options: I would be much worse off, I would be somewhat worse off, I would be neither worse off nor better off, I would be somewhat better off, I would be much better off).

14. How you think a carbon tax affects low-income households if its revenues are used to [the question is repeated for each of the five revenue uses mentioned above]? (Response options: they would be much worse off, they would be somewhat worse off, they would be neither worse off nor better off, they would be somewhat better off, they would be much better off).

15. How acceptable do you find the carbon tax if its revenues are used to [the question is repeated for each of the five revenue uses mentioned above]? (Response options: completely unacceptable, somewhat unacceptable, neither unacceptable nor acceptable, somewhat acceptable, completely acceptable).

16. What percentage of the total carbon tax revenues (100%) would you prefer to allocate for each of the 3 proposed options? Please make sure that the total amount is equal to 100%.	% of revenue allocated
Support the development of climate projects	
Return the revenues to compensate low-income households	
Return the revenues in equal amount to all households as compensation	

17. Which factor– effectiveness or fairness – played a stronger role in your decision on how to allocate the revenue generated by the carbon tax? (Response options: only effectiveness, more effectiveness than fairness, equally effectiveness and fairness, more fairness than effectiveness, only fairness, neither effectiveness nor fairness).

18. How many people are in your household?

19. Could you tell us in which of the following ranges your net monthly income of your household falls? (Responses options: No income, €1000 or less, between €1001-€2000, between €2001-€3000, between €3001-€4000, more than €4001, I prefer not to answer).

20. What is the highest level of studies you have completed? (Response options: less than 5 years of school, primary, secondary, bachelor/medium professional formation, superior professional formation, university).

21. How concerned are you about climate change? (Response options: not at all, a little, somewhat, much, very much).

22. Where would you situate yourself ideologically? Use a scale ranging from 1 to 10, where 1 is ‘left-wing’ and 10 is ‘right-wing’?

23. Which political party did you vote in the last general elections of the 28th April 2019?

24. How often do you use a car? (Response options: never, less than once a month, few times a month, once a week, few times a week, everyday).

25. How many minutes do you travel by car on an average working day? (Response options: none, less than 30 minutes, between 30 and 60 minutes, between 61 (1 hour) and 90 minutes (1:30 hours), Between 91 minutes (1:30hours) and 120 minutes (2 hours), more than 120 minutes (2hours or more).

Appendix C. Descriptive statistics

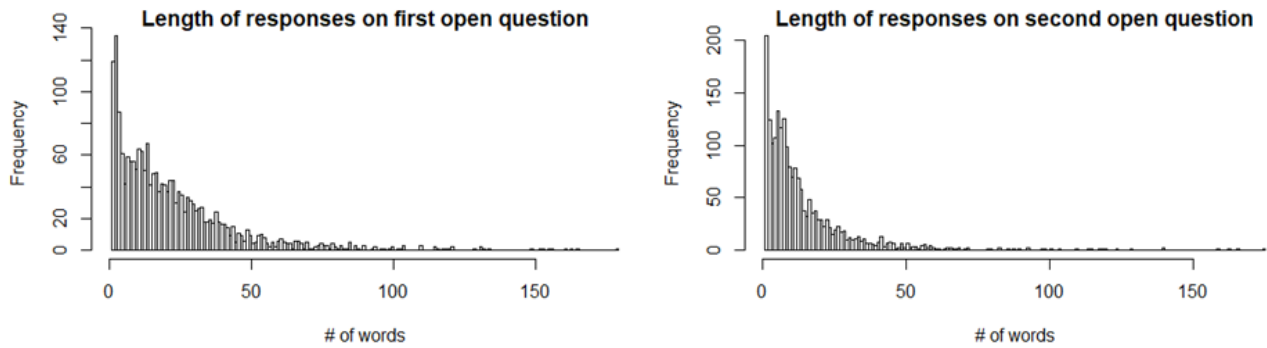


Figure C1. Length of responses to the two open-ended questions
 Note: on the X-axis the shortest length of response is 1.

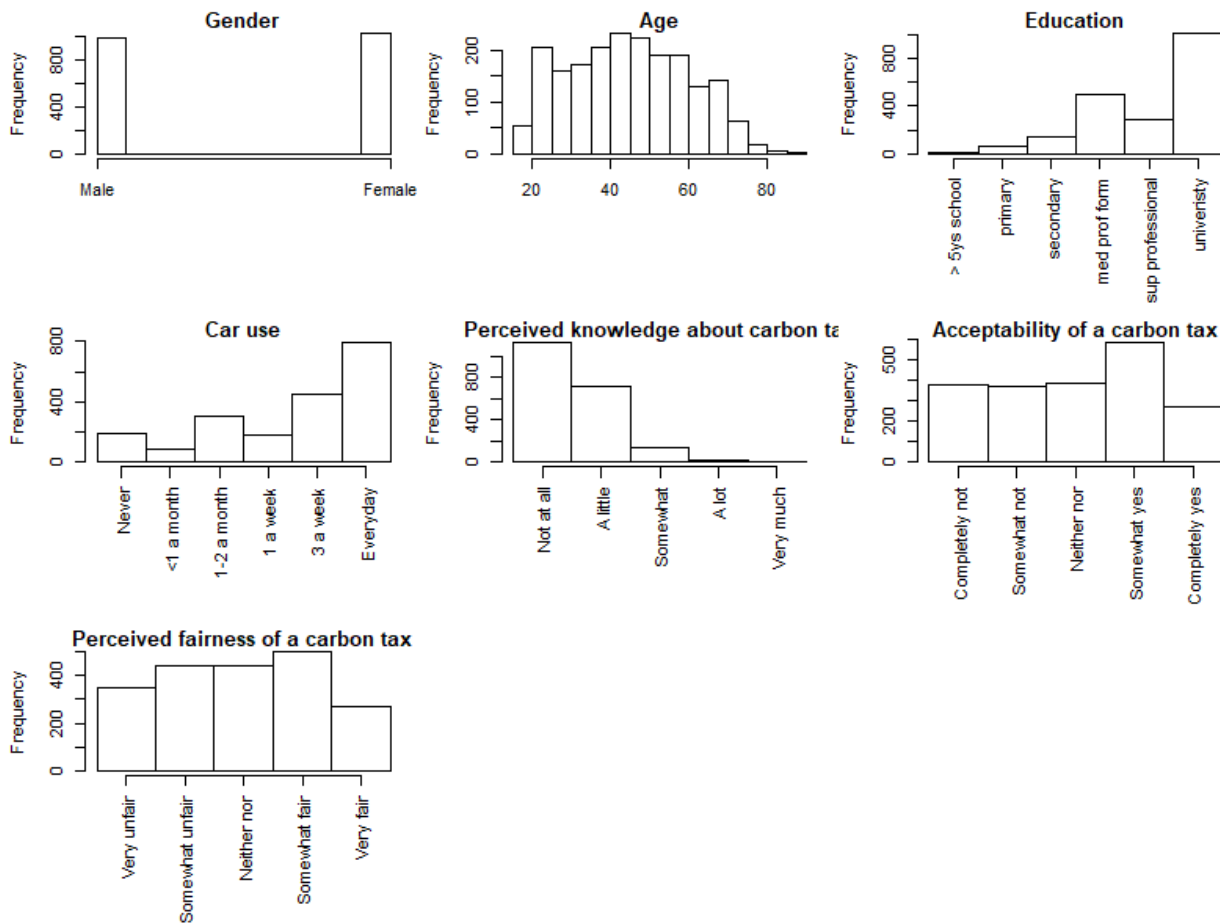


Figure C2. Distribution of values of covariates for the open-ended questions
 Note: for education "med prof form" means university entrance level or medium professional formation and "sup professional" means superior professional formation.

Table C1. Descriptive statistics on sociodemographic variables.

Variables	Description	Mean (SD) or %	Spanish population
Gender	Dummy: female	51.1%	51.1%
Age	18 to 87 years old	45.15 (15.18)	44.2
Household size	Number of people living in respondent's house	2.96 (1.21)	2.5
Monthly household income	1 (No income) to 6 (More than 4000€)	3.73 (1.12) translating into a mean between €2200 and 2700	2295 euros per month
Education	1 (Less than 5 years of school) to 6 (University)	5.02 (1.16); 89.92% of the sample have medium professional or higher studies	60.2% have a medium professional or higher studies (CIS, 2019)*
Political orientation	1 (left-wing) to 10 (right-wing)	4.47 (2.45)	4.5 (2.1) (CIS, 2019)*

Notes: Sampling was done by using quotas on age, gender and geographical distribution, making the survey sample representative of the general population on these characteristics. The rest of variables are compared with census data from the Spanish National Institute of Statistics (www.ine.es) unless other source is indicated. The results show that the sample is also representative for the other covariates, except for education and income.

* Source: Centro de Investigaciones Sociológicas (CIS), Barómetro de Septiembre 2019.

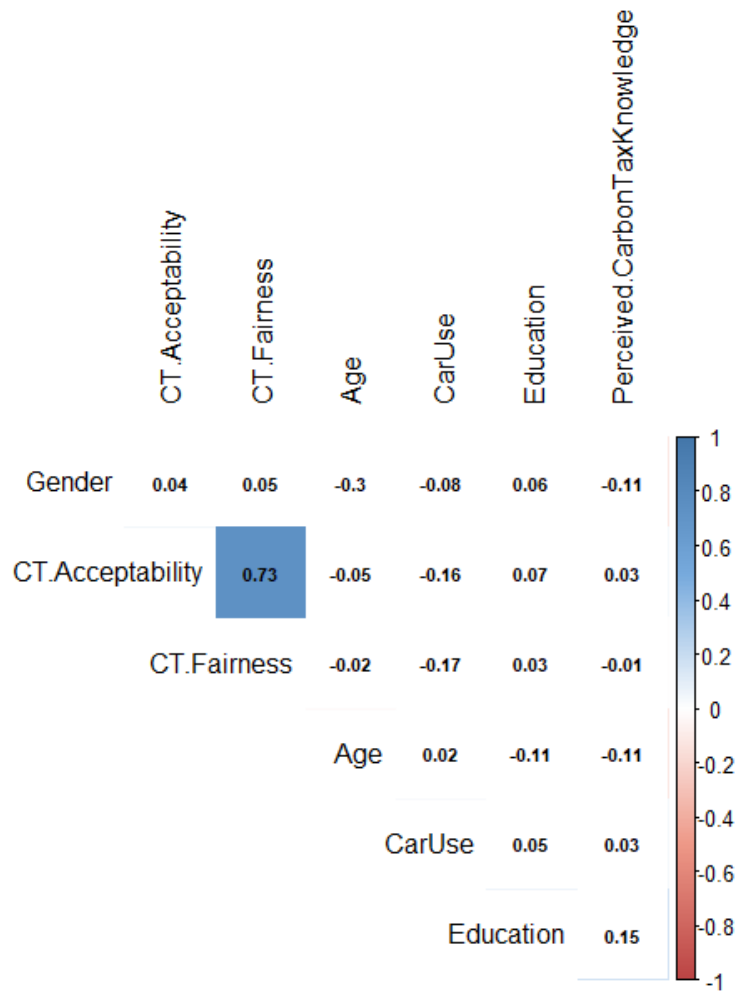


Figure C3. Testing pairwise correlations among covariates in our STM models.

Note: Colors indicate correlations significant at 5% or higher.

Appendix D. Further results

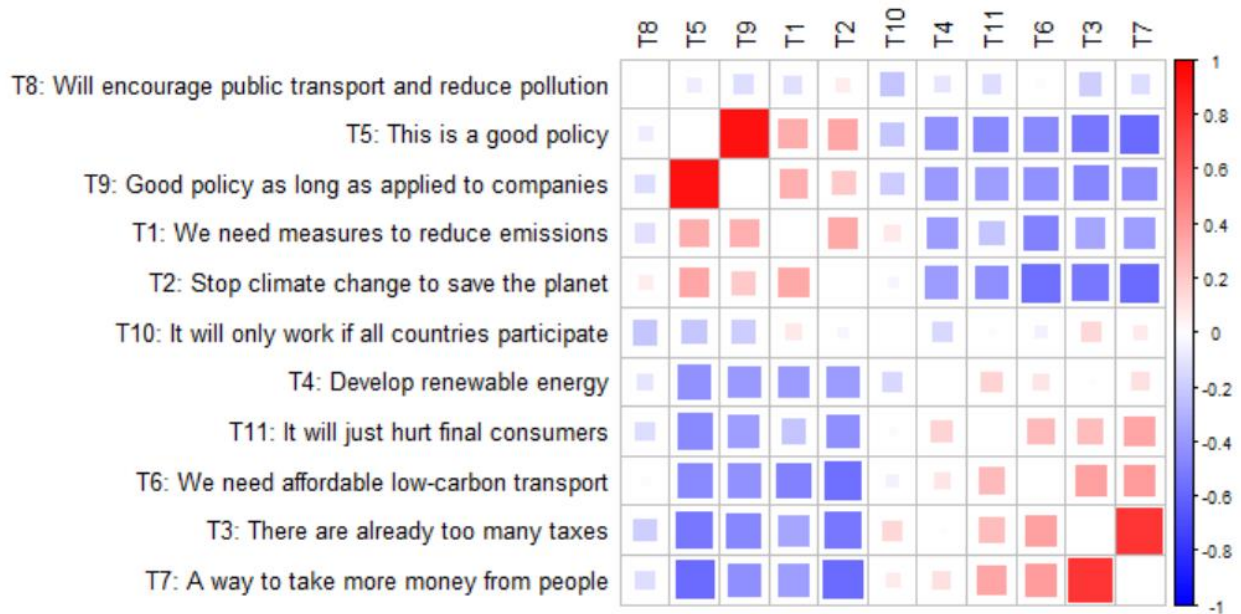


Figure D1. Topic co-occurrence for the first open-ended question

Note: The order of topics results from hierarchical clustering which positions more correlated topics closer together.

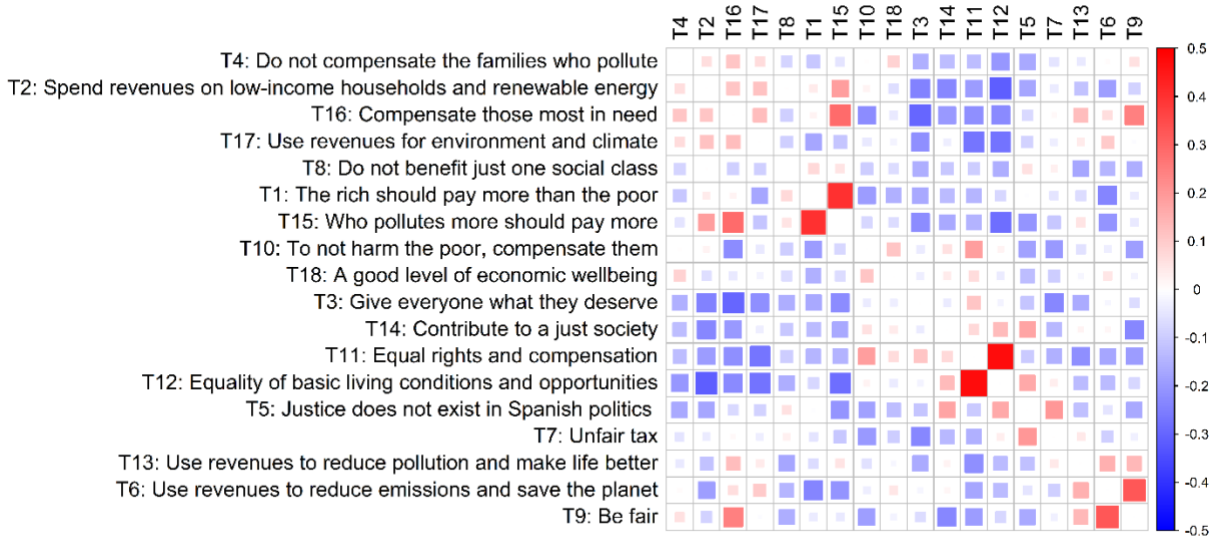


Figure D2. Topic co-occurrence for the second open-ended question

Note: The order of topics results from hierarchical clustering that positions more correlated topics closer together.

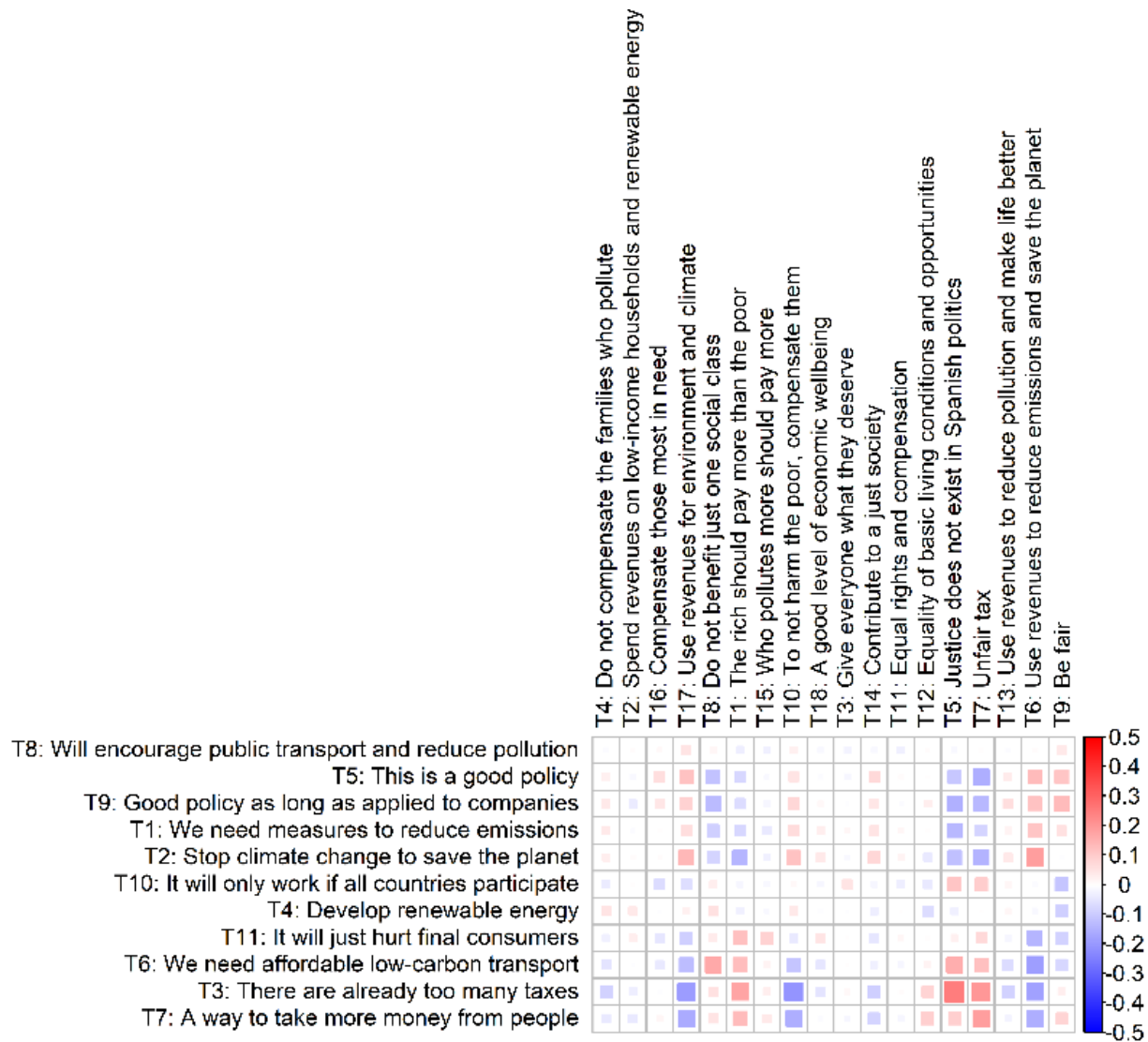


Figure D3. Topic co-occurrence between the first and the second open-ended questions
Note: the order of topics is adopted from Figures D1 and D2 so that topics co-occurring more often for each question appear next to each other on the respective axes.

Table D1. Results of the regression analysis for the STM model based on the first open question

	Intercept	Age	Gender	Education	Car use	Perceived carbon tax knowledge	Carbon tax acceptability
Topic 1	0.0597***	-0.0002	0.0088*	0.0011	-0.0009	-0.0073**	0.0110***
Topic 2	0.0556**	0.0001	-0.0125*	0.0040	-0.0095***	-0.0121***	0.0396***
Topic 3	0.2108***	0.0006***	0.0058*	-0.0101***	0.0024**	-0.0054***	-0.0306***
Topic4	0.0048*	0.0001	-0.0042	0.0142***	0.0010	0.0070	-0.0078***
Topic 5	0.0597***	0.0003	0.0071	-0.0086***	-0.0016	-0.0067**	0.0255***
Topic 6	0.1491***	0.0001	-0.0093	-0.0027	0.0060***	0.0151***	-0.0239***
Topic 7	0.1123***	-0.0004**	0.0027	-0.0018	0.0045***	0.0018	-0.0114***
Topic 8	0.0930***	-0.0002	-0.0003	-0.0041**	0.0008	-0.0016	0.0031*
Topic 9	0.0638***	-0.0003***	0.0086***	-0.0007	-0.0016*	-0.0043*	0.0109***
Topic 10	0.0755***	0.0002***	-0.0067*	0.0025	-0.0032***	0.0062**	-0.0043***
Topic 11	0.0711***	-0.0003*	0.0001	0.0061***	0.0022*	0.0074**	-0.0122***

Note: Asterisks ***, **, and * denote 1%, 5%, and 10% significance, respectively. Coefficients indicate whether prevalence of respective topics changes with the value of the covariates.

Table D2. Results of the regression analysis for the STM model based on the second open question

	Intercept	Age	Gender	Education	Car use	Perceived carbon tax knowledge	Perceived carbon tax fairness
Topic 1	0.0565***	0.0002	0.0109**	-0.0026**	0.0022*	-0.0031	-0.0077***
Topic 2	0.0017	0.0006***	0.0119**	0.0029	-0.0011	0.0050	0.0020
Topic 3	0.0742**	0.0001	-0.0025	0.0048	-0.0028	0.0001	-0.0028
Topic4	0.0018	0.0001	0.0060	0.0077***	0.0015	-0.0131***	0.0035**
Topic 5	0.1196***	0.0004***	-0.0120***	-0.0058***	-0.0025**	0.0021	-0.0089***
Topic 6	0.0588***	-0.0002*	0.0007	-0.0009	-0.0026**	-0.0048**	0.0071***
Topic 7	0.1084***	-0.0002	-0.0052	-0.0030	-0.0008	0.0071**	-0.0100***
Topic 8	0.0577***	0.0002	-0.0132***	0.0005	0.0006	0.0115***	-0.0071***
Topic 9	0.0809***	-0.0004***	0.0169***	-0.0087***	0.0037***	-0.0078***	0.0021*
Topic 10	0.0432***	-0.0003**	-0.0043	0.0053***	-0.0032**	0.0006	0.0048**
Topic 11	0.0477***	0.0001	-0.0058	0.0030*	0.0014	-0.0051*	0.0010
Topic 12	0.1600***	-0.0006***	-0.0041	-0.0084***	0.0001	-0.0024	-0.0034*
Topic 13	0.0400**	-0.0002	-0.0030	0.0037*	0.0011	-0.0007	0.0033**
Topic 14	0.0025	0.0002	-0.0042	0.0016	0.0001	0.0043	0.0036
Topic 15	0.0637***	0.0001	-0.0033	0.0008	0.0022**	-0.0023	-0.0011
Topic 16	0.0560***	-0.0001	0.0035	-0.0034**	0.0020**	-0.0018	0.0022**
Topic 17	0.0160	0.0003**	0.0123**	0.0001	-0.0028*	0.0106***	0.0105***
Topic 18	0.0208**	-0.0002***	0.0028	0.0024**	0.0007	-0.0013	0.0011

Note: Asterisks ***, **, and * denote 1%, 5%, and 10% significance, respectively. Coefficients indicate whether prevalence of respective topics changes with the value of the covariates.

To test our results for robustness, we replicated the procedure on the first open question after excluding people who completed the survey very quickly (“speeders”), as they may have not

thought as well about their responses. In particular, we discarded 218 (approx. 11%) respondents who required less than eight minutes on the whole survey. For the first open question this reduced the sample to 843 unique terms and 16707 tokens. Testing the resulting dataset on the optimal number of topics we find that 12 topics fits the data better than 11 as by adding just one more topic produces better results on all three selection criteria (Figure D4)

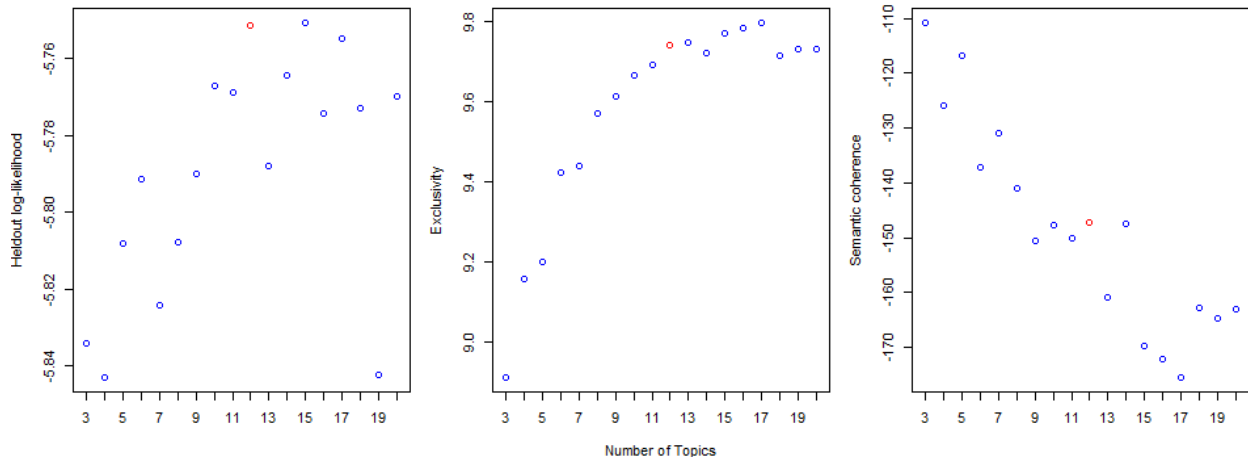


Figure D4. Model performance for distinct number of topics for the first (open-ended question after deleting the speeders).

After forming the model with 12 topics using the same six covariates we used before (age, gender, education, car use, perceived level of knowledge about carbon tax and acceptability of carbon tax), we report the most discriminating terms of the resulting topics and their proportions in Table D3. Recognizing that many of these topics look very similar to those reported in Table 1 based on the full data sample, we tried to match the closest equivalents by assigning to a topic X from Table 1 a topic X* in Table D3.¹⁰

For 7 out of 11 topics we could find a unique equivalent. One should stress at this point that an equivalent does not imply a perfect one-to-one relationship, but that a considerable fraction of discriminating terms between the topics is overlapping, while open textual responses with high topic prevalence convey the views consistent with the topic label. For topics 2 and 7 from Table 1 we could find not one but two equivalents in Table D3 (i.e. these have been seemingly split in two distinct clusters each), while two other topics, 9 and 11, have been merged into a single cluster. Thus, we clearly see how views expressed in topics from Table 1 can be observed also in Table D3, while also topic proportions of equivalent topics are similar.

¹⁰ Henceforth, we will use X* as a part of topic label to indicate the close relationship.

Table D3. Topics identified for responses to the first question after deleting “speeders”

	Topic label	Most discriminating terms (frequency and exclusivity)	Topic proportion
1*	We need measures to reduce emissions	measure, emission, reduce, implement, time, environmental, start, transition, effect, effective <i>medida, emission, reducir, implementar, tiempo, medioambiental, empezar, transicion, efecto, efectiva</i>	7.5%
2*	Stop climate change to save the planet	reduce, change, climatic, contribute, action, medium, help, positive, improvement, proposal <i>reducir, cambio, climatico, contribuir, accion, medio, ayud, positiva, mejora, propuesta</i>	10.3%
2**	Stop climate change to save the planet	planet, necessary, unique, duty, awareness, idea, way, shape, essential, serious <i>planeta, necesario, unica, debemo, conciencia, idea, manera, forma, imprescindible, serio</i>	7.2%
3*	There are already too many taxes	solution, tax, collection, paid, more, raise, put, excessive, excuse, exist, <i>solucion, impuesto, recaudatorio, pagamo, mas, subir, poner, excesivo, excusa, existen,</i>	11.7%
4*	Develop renewable energy	energy, fuel, alternative, renew, fossil, clean, promote, solar, source, sun, <i>energia, combust, alternativa, renov, fosil, limpia, fomentar, solar, fuent, sol</i>	9.2%
5*	This is a good policy	seem, well, even, correct, perfect, future, initiative, good, raised, correct, <i>parec, bien, aunqu, correcto, perfecto, futuro, iniciativa, buena, recaudado, correcta</i>	10.9%
6*	We need affordable low-carbon transport	class, rich, low, hybrid, buy, ecological, worker, method, automobile, pay <i>clase, rico, baja, hibrido, comprar, ecologico trabajadora, metodo, automovil, pagan</i>	7.8%
7*	A way to take more money from people	money, take out, want, fair, lower, base, like, fatal, contributor, collection <i>dinero, sacar quieren, justo, bajar, base, gusta, fatal, contribuyent, recaudatoria</i>	3.2%
7**	A way to take more money from people	pay, political, bad, government, people, salary, problem, do, worker, collect <i>pagar, politico, mal, gobierno, pueblo, sueldo, problema, hacen, trabajador, cobrar</i>	6.1%
8*	Will encourage public transport	car, vehicle, less, transport, help, public, gasoline, city, plane, information	7.8%

	and reduce pollution	<i>coch, vehiculo, meno, transport, ayuda, publico, gasolina, ciudad, avion, informacion</i>	
9* + 11*	Shall be applied to companies and not hurt final consumers"	always, company, grand, final, consumer, user, ordinary, pay, multinational, interest <i>siempr, empresa, grand, final, consumidor, usuario, pie, pagan, multinacional, interes,</i>	12.4%
10*	It will only work if all countries participate	each, level, country, sufficient, society, contamination, important, high, adequate, must <i>cada, nivel, pais, suficiente, sociedad, contaminacion, important, alto, adecuado, debe</i>	5.9%

Note: The terms shown are those that are the most frequent as well as exclusive to each topic. The original text was in Spanish. Here we report the Spanish terms and their English translation. Labels for each topic reflect the content of the terms and associated survey responses. Words are stemmed; e.g., the term ‘parec’ comprises ‘parecer’ and ‘parece’.

To further ensure that the new topics reflect similar findings concerning how expressed views differ among the population depending their age, gender and other covariates we used to form the topics, we produce Figure D5 reporting results of the regression analysis, where we establish associations between topic proportions and their covariates. As one can see, Figure D5 produces a result similar to the one in Figure 3. For example, the view that there are already too many taxes (Topics 3 and 3* in the older and newertopic models) are expressed comparatively more by older people, with less education, lower level of knowledge about the carbon tax and lower acceptability of htis policy initiative. Furthermore, examaning significance of other regression coefficients reported in Table D2, we find only car use for topic 10* („It will only work if all countries participate“) to be weakly significant (at 10% and not at 1% as it was for topic 10 in Figure 3).

Based on this evidence we conclude that our results reported in the main papers are robust to exclusion of respondents who completed the survey questionanire very quickly.

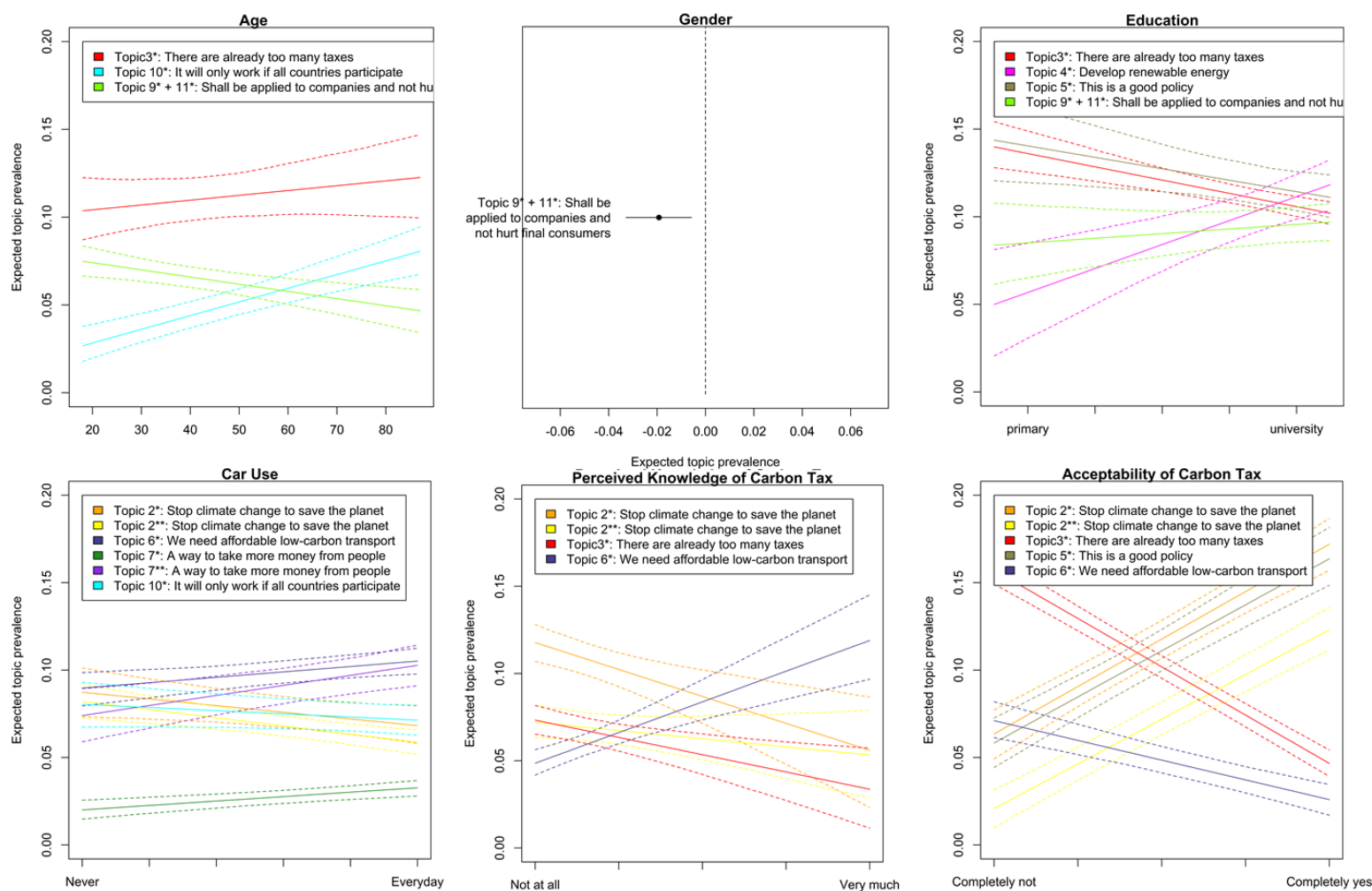


Figure D5. Effect of covariates on topic prevalence for associations with carbon tax implementation after deleting the speeders.

Note: Values generated by a regression where the outcome variable is the proportion of each public response dedicated to each topic, given the selected STM model. The panel shows point estimates and confidence intervals of the effects of selected covariates on topic prevalence, holding all other covariates constant. The plot for gender shows mean difference in topic proportions between male and female (a positive value on the X-axis indicates a larger prevalence for men). Only a subset of topics corresponding to those depicted in Figure 3 in the main text is displayed. Confidence intervals plotted as dashed lines indicate the 95% uncertainty range and include both regression and measurement uncertainties associated with the STM model.