Topic correlation networks inferred from open-ended survey responses reveal signatures of ideology behind carbon tax opinion

Maximilian Puelma Touzel*
Mila Quebec Artificial Intelligence Institute,
Department of Computer Science and Operations Research,
Université de Montréal
puelmatm@mila.quebec

Erick Lachapelle
Department of Political Science,
Université de Montréal
erick.lachapelle@umontreal.com

Abstract

Ideology can often render policy design ineffective by overriding what, at face value, are rational incentives. A timely example is carbon pricing, whose public support is strongly influenced by ideology. As a system of ideas, ideology expresses itself in the way people explain themselves and the world. As an object of study, ideology is then amenable to a generative modelling approach within the text-as-data paradigm. Here, we analyze the structure of ideology underlying carbon tax opinion using topic models. An idea, termed a topic, is operationalized as the fixed set of proportions with which words are used when talking about it. We characterize ideology through the relational structure between topics. To access this latent structure, we use the highly expressive Structural Topic Model to infer topics and the weights with which individual opinions mix topics. We fit the model to a large dataset of open-ended survey responses of Canadians elaborating on their support of or opposition to the tax. We propose and evaluate statistical measures of ideology in our data, such as dimensionality and heterogeneity. Finally, we discuss the implications of the results for transition policy in particular, and of our approach to analyzing ideology for computational social science in general.

1 Introduction

Mitigating the worst effects of climate change requires that our society shift away from burning fossil fuels as its primary energy source. Additional, parallel transitions in other areas of society are likely needed to achieve this [1]. One such area is economics, where one transition feature is internalizing the cost of pollution. A price on carbon is an instrument to do this that economists agree is simple, flexible, and can be easily set on schedule to rise as needed. Several countries have instituted such pricing systems, but a major obstacle in wider adoption appears to be a lack of public support.

Public support for carbon pricing typically falls along ideological lines, with conservatives on the right of the political spectrum opposing the policy This is in spite of rebate programs that funnel collected taxes back to most citizens (8/10 households in Canada are estimated to currently receive

*mptouzel.github.io

Tackling Climate Change with Machine Learning: workshop at NeurIPS 2021.

more than they pay as a result of the policy). Recent work has studied the effects of rebates on people's support for the policy [2]. In particular, survey participants tend to underestimate the size of the rebate they receive.

What is the effect of communicating the factual benefits of the tax? When participants were then shown how much they received, their support of the policy remained unchanged. Such behaviour is not necessarily irrational: strong priors can make rational decision-making insensitive to new data [3]. After being shown the evidence, the participants' belief that they were net losers (paid out more than they received) did change, however: surprisingly, it strengthened! These results suggest that there is a broader value system being recruited here that, if better understood, could inform the design of more effective communication of the policy's benefits.

We hypothesized that oppose responses arise from a well-worn 'tax is bad' ideology, involving a handful of correlated ideas ('distrust of government', 'unfairness', etc.) that mutually enforce each other. Here, we used Canadian survey data [2] and a generative *bag-of-words* model of word responses to infer the topic structure underlying the three types of responses (*oppose*, *support*, and *not sure*). We find moderate, significantly significant evidence for this hypothesis across a set of mutually independent metrics, agnostic to semantics.

2 Related Work

Many measures of corpus analysis study frequency of word usage[2], which does not expose how the same words can be used when talking about different things. Other broadly used approaches such as sentiment analysis classify responses into only a few affective classes ('like'/'dislike'). By formulating a rich latent topic structure, topic models address both these limitations. Topc models are now an established approach to understanding human-generated responses in the social sciences [4]. The Structural Topic Model in particular has been applied to understand open-ended responses on a carbon tax in Spain [5], Norway [6], and the US [7]. Here, we make a similar application to data obtained in Canada. Unlike these previous works we focus on topic-topic correlations as a means to interrogate ideology.

3 Method

Topic models are generative probabilistic models that generate words in a response from an underlying set of topics, each given as the set of usage frequencies of words in a given vocabulary. Topic models are typically *bag of word* models, which eschew grammar and syntax to focus only on the content and prevalence of words. We exploit the availability of rich metadata by picking a topic model with rich latent structure of word usage statistics: the *Structural Topic Model* (STM) [8]. Like the correlated topic model [9], it uses a logistic normal distribution to define the topic weights on a document and can thereby exhibit arbitrary topic-topic covariance via the covariance matrix parameters of the logistic normal distribution. Unlike the CTM, it also allows for meta-data to skew the word and topic prevalence[3].

3.1 Data Processing, Inference, & Analysis

We analyzed a dataset of responses of 3,313 survey participants from across Canada in 2019 on the question of support for the recently implemented carbon tax [2]. French responses were first translated into English using the Google Translate API. The response corpus was pre-processed through spell-checking, removal of stop words, and reduction to word stems. We built a Python interface to a well-established and computationally efficient STM inference suite written in the R programming language [10].

The conjugacy of Dirichlet prior used in LDA allows for the efficient variational Bayes algorithm. While more expressive, the logistic normal distribution used in STM is not conjugate to the multinomial distribution, making efficient inference less straightforward. The STM package nevertheless has a highly optimized parameter learning approach. It uses variational expectation maximization,

[2]This includes more refined frequencies such as TF-IDF

[3]Metadata skews topic prevalence via the mean of the normal distribution in the logistic normal

with the expectation made tractable through a Laplace approximation. Accuracy and performance is improved by integrating out the word-level topic and through a spectral initialization procedure. For details see [8].

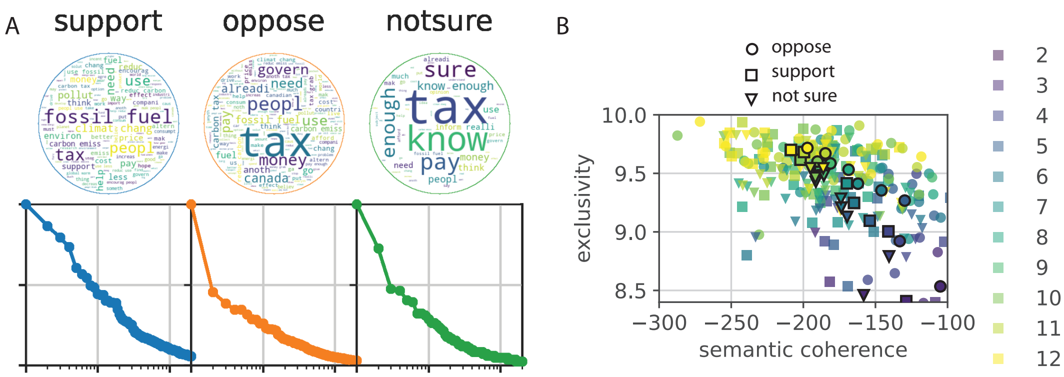

Figure 1: (a) Word Cloud (top) and normalized word frequency rank histogram (bottom) for the three responses. (b) Topic quality. Exclusivity plotted against semantic coherence for topic number $K = 2, \ldots, 12$ for the 3 response types.

As with other topic models, an STM takes the number of topics, K, as fixed parameter. We analyzed results across a range of the topic number to ensure validity of our conclusions. We assessed topic quality across different topic numbers by plotting *exclusivity* (high when a topic's frequent words are exclusive to that topic) versus *semantic coherence* (high when a topic's frequent words co-occur often).

We focussed on the statistical properties of the set of posterior mean estimates (obtained by averaging posterior samples), $\hat{\boldsymbol{\theta}}_d = (\hat{\theta}_{d,1}, \ldots, \hat{\theta}_{d,K})$, of the topic mixture weight vector, one for each response, indexed by d. We used four intuitive and independent characteristics of a data cloud (see fig:fourmeasures:

1. *Size*: normalized generalized covariance, $\prod_{k=1}^{K-1} \lambda_k^{\frac{1}{K-1}}$, using the eigenvalues λ_k of the covariance matrix (excluding the 0 eigenvalue associated with the mode orthogonal to the simplex). Assuming unimodality, smaller volumes imply more compact and therefore more similar mixtures. This measures the heterogeneity over the participant group.
2. *Location*: average distance from center, $(H_{\max} - \bar{H})/H_{\max}$, where i.e. document-averaged entropy $\bar{H} := \frac{1}{D}\sum_{d=1}^{D} H(\hat{\boldsymbol{\theta}}_d)$, and normalized by the maximum $H_{\max} = \log K$. Smaller values imply stronger average preferences for some topics. This measures the degree to which the participant group raises specific topics.
3. *Eccentricity*: effective dimension $(\sum_k \lambda_k)^2/(\sum_k \lambda_k^2)$, where λ_k is the kth eigenvalue of the correlation matrix $\Sigma_{\hat{\theta}}$, normalized by K. This is measures how many degrees of freedom the heterogeneity has and thus how expressive is the group's topic mixing. Lower dimensionality implies more constrained heterogeneity.
4. *Direction*: the number of positive correlations, above some small threshold, $c_{\text{thresh}} = 0.01$, normalized by the maximum, $K(K-1)$. This measures how much synergy exists among the recruited topics.

3.2 Results

We find that indeed oppose responses appear more focussed on the word 'tax' than the support responses Figure 1(a). The results of topic quality (Figure 1(b)) show that topic-averaged values give a linear trade-off between the two, with topic number setting where along the trade-off the model resides. Quality is highest in *oppose* responses, followed by *support* and *not sure*. We attribute this to the singular focus that *oppose* responses have on the word *tax*, as compared to the much more diffuse responses (in word use) of *support* responses. Note, however, that the topic variance in each dimension is at least as large as the difference in topic means at fixed topic number so the aforementioned ranking is only clear after averaging over topics. Semantic coherence bottomed out at around 10 topics, but there wasn't strong evidence for a single number of topics. We continued analyzing models with 2 to 12 topics.

Next, we looked at topic-topic correlations. We compared results for extreme choices of prior (parametrized by $\sigma \in [0, 1]$) on the Logistic Normal covariance: a uniform prior ($\sigma = 0$) and an independence prior ($\sigma = 1$). We found that results for $\sigma = 0$ were more noisy the support responses much more variable. With finite σ, support responses were pushed to low correlation, while the correlations for oppose responses persisted, suggesting they are truly in the latent structure of the data.

The results for the four geometric properties for these two cases is shown in 2. The two most salient differences are that oppose responses are less expressive in how they mix topics (see (c)) and they have more synergy (positive correlation) (d). This is consistent with our hypothesis of a rigid ideology underlying carbon tax opposition. This was true across values of σ (not shown).

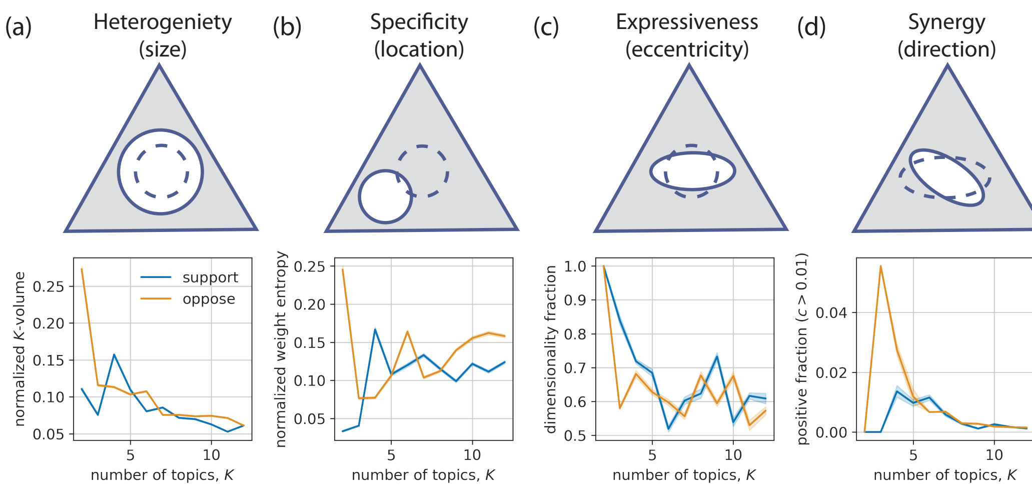

Figure 2: Geometric features of topic mixture weight data for independence prior $\sigma = 0.6$. Top row: Schematic of the variation in each feature in the simplex (triangle). Bottom row: The respective plot for the support (blue) and oppose (orange) responses. Lines are posterior means estimates and errorbars are standard deviation of 100 posterior samples.

3.3 Discussion

Here, we presented results using topic-topic correlation structure to infer properties of the semantic network of ideas used when justifying support of or opposition to the carbon tax. The results are suggestive rather than conclusive of our initial hypothesis that there exists a constrained, positively correlated set of topics underlying oppose responses. Further validation is need to bring these results into focus. We believe that running our analysis over demographics more likely to exhibit this ideology (conservative party voters, right-leaning individuals etc.) would strengthen the result. While we have chosen analysis that is independent of topic semantics, inspecting the topic meanings might shed further light on the nature of the ideology. Corroboration of the results shown here would motivate new longitudinal experiments to asses the elastic nature of belief change. Counter to the prevailing thinking in effective communications research, we conjecture that informational interventions targeting the neighborhood of topics recruited by oppose responses might be more effective than devoting all resources on a single issue.

Acknowledgments and Disclosure of Funding

MPT woudl like to acknowledge helpful discussions with Matto Mildenberger, Kathryn Harrison, and Dhanya Sridhar. MPT was funded under the CERC for Autonomous AI awarded to Irina Rish.

References

[1] J. G. Speth. The transition to a sustainable society. *Proceedings of the National Academy of Sciences of the United States of America*, 89(3):870–872, 1992.

[2] Matto Mildenberger, Erick Lachapelle, Kathryn Harrison, and Isabelle Stadelmann-Steffen. Limited impacts of carbon tax rebate programmes on public support for carbon pricing. *Nature Climate Change*, 12(2):141–147, 2022.

[3] Samuel J Gershman. How to never be wrong. *Psychonomic Bulletin and Review*, 26(1):13–28, 2019.

[4] J. Grimmer, M.E. Roberts, and B.M. Stewart. *Text as Data: A New Framework for Machine Learning and the Social Sciences*. Princeton University Press, 2022.

[5] Ivan Savin, Stefan Drews, Sara Maestre-Andrés, and Jeroen van den Bergh. Public views on carbon taxation and its fairness: a computational-linguistics analysis. *Climatic Change*, 162(4):2107–2138, 2020.

[6] Endre Tvinnereim, Kjersti Fløttum, Øyvind Gjerstad, Mikael Poul Johannesson, and Åsta Dyrnes Nordø. Citizens' preferences for tackling climate change. Quantitative and qualitative analyses of their freely formulated solutions. *Global Environmental Change*, 46(May):34–41, 2017.

[7] Marina Povitkina, Sverker Carlsson Jagers, Simon Matti, and Johan Martinsson. Why are carbon taxes unfair? Disentangling public perceptions of fairness. *Global Environmental Change*, 70(September), 2021.

[8] Margaret E Roberts, Brandon M Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson, and David G Rand. Structural topic models for open-ended survey responses. *American Journal of Political Science*, 58(4):1064–1082, 2014.

[9] David M. Blei and John D. Lafferty. Correlated topic models. In *Proceedings of the 18th International Conference on Neural Information Processing Systems*, NIPS'05, page 147–154, Cambridge, MA, USA, 2005. MIT Press.

[10] Margaret E Roberts, Brandon M Stewart, and Dustin Tingley. Stm: An R package for structural topic models. *Journal of Statistical Software*, 91:1–40, 2019.