Exploring Topic-Metadata Relationships with the STM: A Bayesian Approach

Anonymous ACL submission

Abstract

The initial purpose of topic models was to identify latent topical clusters within unstructured text. Meanwhile, the focus of advanced studies has changed primarily to estimating the relationship between the discovered topical structure and theoretically relevant metadata. Methods used to estimate such relationships must take into account that the topical structure is not directly observed, but instead being estimated itself in an unsupervised fashion. In the Structural Topic Model (STM; Roberts et al., 2016), for instance, multiple repeated linear regressions of sampled topic proportions on metadata covariates are performed. This is done by using a Monte Carlo sampling technique known as the *method of composition*. In this paper, we propose two modifications of this approach: First, we implement a substantial correction to the model by replacing linear regression with the more appropriate Beta regression. Second, we provide a fundamental enhancement of the entire estimation framework by substituting the current blending of frequentist and Bayesian methods with a fully Bayesian approach instead. This allows for a more appropriate quantification of uncertainty. We illustrate our improved methodology by investigating relationships between Twitter posts by German parliamentarians and different metadata covariates related to their electoral districts.

1 Introduction

The rise in popularity of social media has led to an unprecedented increase in the supply of publicly available unstructured text data. Researchers often wish to examine relationships between observable metadata (e.g., characteristics of a document's author) and in-text patterns (Farrell, 2016; Kim, 2017). Probabilistic topic models identify such in-text patterns by producing a posterior distribution over different topics. Estimating relationships with observed metadata, however, is not trivial as the target variable is latent and itself being estimated from the text data itself.

Due to its popularity in the social sciences, in this work we focus on exploring and estimating topic-metadata relationships with the Structural Topic Model (STM; Roberts et al., 2016). The estimation of topic-metadata relationships in the `stm` package (Roberts et al., 2019), which implements the STM in R, combines Monte Carlo sampling with a frequentist linear regression. Even though this estimation technique is prone to producing predictions incompatible with standard definitions of probability, it is frequently applied in the literature (cf. Appendix A). This leads to implausibilities of two different forms: On the one hand, authors sometimes report negative expected topic proportions (e.g. Farrell, 2016; Moschella and Pinto, 2019), on the other hand, there are multiple cases where "only" the confidence bands are partly in negative ranges (e.g. Cho et al., 2017; Chandelier et al., 2018; Bohr and Dunlap, 2018; Heberling et al., 2019). In both cases, it is ignored that sampled topic proportions are confined to $(0, 1)$ by definition, which severely harms the interpretability of the model's results.

In this paper, we suggest two key modifications to the `stm` implementation (Roberts et al., 2019): First, our Beta regression approach is a natural correction of the linear regression approach, accounting for topic proportions being restricted to the interval $(0, 1)$. Second, we propose the use of a *Bayesian* estimation design within the method of composition to allow for a more coherent estimation and interpretation of topic-metadata relationships; in particular, we obtain a posterior predictive distribution of topic proportions at different values of metadata covariates.

We demonstrate the added value of our model corrections by analyzing Twitter posts of German politicians, gathered from September 2017 through April 2020. Politics has been particularly impacted by the rise of social media as evidenced by the Brexit vote and US presidential elections, with Twitter being extensively used for direct communication by politicians. We investigate relationships between latent topics in the tweets of German members of parliament (MPs) and corresponding metadata, such as tweet date or unemployment rate in the respective MP's electoral district. In doing so, we attempt to link the topics discussed to specific events as well as to socioeconomic characteristics of the MP's electoral districts.

2 Background

Topic models seek to discover latent thematic clusters, called topics, within a collection of discrete data, usually text. Besides identifying such clusters, topic models estimate the proportions of the discovered topics within each document. Many topic models build upon the well-known Latent Dirichlet Allocation (LDA; Blei et al., 2003), which is a generative probabilistic three-level hierarchical Bayesian mixture model that assumes a Dirichlet distribution for topic proportions. The Correlated Topic Model (CTM; Blei et al., 2007), for instance, builds on the LDA, but replaces the Dirichlet distribution with a logistic normal distribution to capture inter-topic correlations. The STM adopts this approach, but additionally incorporates document-level metadata into the estimation of topics:[1]

- For document $d \in \{1, \dots, D\}$ and topic $k \in \{1, \dots, K\}$, a topic proportion $\theta_{d,k}$ is drawn from a logistic normal distribution.[2]
- The parameters of the logistic normal distribution depend on document-level metadata covariates $\mathbf{x}_d$.

For parameter estimation, the STM employs a variational EM algorithm, where in the E-step the variational posteriors are updated using a Laplace approximation (Wang and Blei, 2013; Roberts et al., 2016). In the M-step, the approximated Kullback-Leibler divergence is minimized with respect to the model parameters.

[1]Within the STM, document-level covariates can also be used to fine-tune topic-word distributions (Roberts et al., 2016), but we do not further discuss this here.

[2]The `stm` package provides several metrics to choose the hyperparameter *K*, as will be discussed in Section 5.2.

3 Estimating Topic-Metadata Relationships in the STM

The STM produces an approximate posterior distribution of topic proportions. A point estimate can be obtained for example as the mode of this distribution. Topic proportions are often used in subsequent analysis, for instance in order to estimate their relationship with metadata. We argue that the usual practice of simply regressing point estimates of topic proportions on document-level covariates is not adequate for estimating topic-metadata relationships. This approach ignores that topic proportions are themselves estimates, neglecting much of the information contained in their posterior distribution. In this section, we propose a method to adequately explore the relationship between topic proportions and metadata covariates.

One way to account for the uncertainty in topic proportions is the "method of composition" (Tanner, 2012, p. 52), which is a simple Monte Carlo sampling technique. Let y be a random variable with unknown distribution $p(y)$ from which we would like to sample and let z be another random variable with known distribution $p(z)$. If $p(y|z)$ is known, we can sample from

$$p(y) = \int p(y|z)p(z)dz,$$

using the following procedure:

1. Draw $z^* \sim p(z)$.
2. Draw $y^* \sim p(y|z^*)$.

Discarding z^*, the resulting y^* are samples from $p(y)$.[3]

In Roberts et al. (2016), the authors employ a variant of the method of composition established by Treier and Jackman (2008), which uses linear regression to obtain the conditional distribution $p(y|z)$. To demonstrate this variant, let $\boldsymbol{\theta}_{\bullet k} = (\theta_{1,k}, \dots, \theta_{D,k})^T \in (0,1)^D$ denote the proportions of topic k and let $\mathbf{X} := [\mathbf{x}_1 | \dots | \mathbf{x}_D]^T$ be the covariates for all D documents. Let further $q(\boldsymbol{\theta}_{\bullet k})$ be the approximate posterior distribution of topic proportions given observed documents and metadata, as produced by the STM. The idea now is to repeatedly draw samples $\boldsymbol{\theta}^*_{\bullet k}$ from $q(\boldsymbol{\theta}_{\bullet k})$ and subsequently perform a regression of each sample $\boldsymbol{\theta}^*_{\bullet k}$ on covariates $\mathbf{X}$ to obtain coefficient estimates $\hat{\boldsymbol{\xi}}$. Treier and Jackman (2008) view the asymptotic distribution of $\hat{\boldsymbol{\xi}}$ as posterior density

[3]Note that this method is an exact sampling method.

Algorithm 1: Method of composition with frequentist regression

repeat procedure m **times**:
  Draw $\boldsymbol{\theta}^*_{\boldsymbol{\cdot}k} \sim q(\boldsymbol{\theta}_{\boldsymbol{\cdot}k})$, where q is the approximate posterior of $\boldsymbol{\theta}_{\boldsymbol{\cdot}k}$.
  Regress $\boldsymbol{\theta}^*_{\boldsymbol{\cdot}k}$ on $\mathbf{X}$; store estimated regression coefficients $\hat{\boldsymbol{\xi}}$ and corresponding covariance matrix.
  Draw $\boldsymbol{\xi}^*$ from the (asymptotic) distribution of $\hat{\boldsymbol{\xi}}$.
  Predict topic proportions $\theta^*_{pred,k} = g(\mathbf{x}^T_{pred}\boldsymbol{\xi}^*)$ at new covariate values $\mathbf{x}_{pred}$.
end procedure

for $\boldsymbol{\xi}$, i.e., as $p(\boldsymbol{\xi}|\boldsymbol{\theta}^*_{\boldsymbol{\cdot}k}, \mathbf{X})$. Using samples $\boldsymbol{\xi}^*$ from this distribution, we can "predict" topic proportions $\theta^*_{pred,k} = g(\mathbf{x}^T_{pred}\boldsymbol{\xi}^*)$ at new covariate values $\mathbf{x}_{pred}$. (g is the regression response function: Identity function for linear regression; Logistic function for Beta regression.) Algorithm 1 summarizes the method. Note that sampling from the posterior of topic proportions in the first step of Algorithm 1 accounts for the uncertainty in $\boldsymbol{\theta}_{\boldsymbol{\cdot}k}$, while the uncertainty of the regression estimation itself is addressed by sampling from the (asymptotic) distribution of the regression coefficient estimator.

To visualize topic-metadata relationships, Roberts et al. (2016) generate multiple "predictions" $\theta^*_{pred,k}$ and calculate empirical quantities such as the mean and quantiles. Calculating mean and credible intervals in such a Bayesian fashion implicitly assumes a (posterior predictive) distribution for $\theta^*_{pred,k}$. This distribution, however, directly depends on the regression - which is frequentist as implemented in the `stm` package. We address this point in detail in Section 4.2.

4 Methodological Improvements

While we agree with performing Monte Carlo sampling of topic proportions in order to integrate over latent variables, we aim to address two flaws:

- **Inadequate modeling of proportions**: The method of composition is implemented in the R package `stm` via the `estimateEffect` function, which employs a linear regression in the second step of Algorithm 1 (implying $g = id$ in the last step). This implementation ignores that topic proportions are naturally restricted to the interval $(0, 1)$. As a consequence, when using the `estimateEffect` function, we frequently observed predicted topic proportions outside of $(0, 1)$, as is exemplarily shown for one specific topic-covariate combination in Figure 1.
- **Mixing Bayesian and frequentist methods**: The method of composition used by Treier and Jackman (2008) and Roberts et al. (2016) mixes Bayesian and frequentist methods. As described in Section 3, a frequentist regression is used inside the method of composition, yet estimates are obtained in a Bayesian manner via calculation of empirical mean and quantiles. Recall that according to Treier and Jackman (2008), $\boldsymbol{\xi}^*$ can be considered a sample from the posterior of regression coefficients. However, the coefficients resulting from a frequentist regression do not have any distribution because the frequentist framework assumes them to be fixed parameters. As a consequence, one cannot sample from the distribution of regression coefficients, which is why Treier and Jackman (2008) sample $\boldsymbol{\xi}^*$ from the distribution of coefficient *estimators*. This distribution, however, only exists by making frequentist assumptions.

In Sections 4.1 and 4.2 below we further discuss these problems and present corrections and alternatives, all of which are implemented in the R package `stmprevalence`. [4]

4.1 Frequentist Beta Regression

As noted above, the linear regression approach is often used carelessly in the literature, neglecting that topic proportions are non-negative by definition. Farrell (2016) and Moschella and Pinto (2019), for instance, produce figures containing negative expected topic proportions, while Cho et al. (2017), Chandelier et al. (2018), Bohr and Dunlap (2018), and Heberling et al. (2019) display confidence bands partly covering negative values.

We correct the approach employed within the `stm` package by replacing the linear regression with a regression model that assumes a dependent

[4] Source code in supplementary material; will be made available on GitHub upon publication.

Algorithm 2: Method of composition with Bayesian Beta regression

repeat procedure m **times**:
  Draw $\boldsymbol{\theta}^*_{\boldsymbol{\cdot}k} \sim q(\boldsymbol{\theta}_{\boldsymbol{\cdot}k})$, where q is the approximate posterior of $\boldsymbol{\theta}_{\boldsymbol{\cdot}k}$.
  Perform a Bayesian Beta regression of $\boldsymbol{\theta}^*_{\boldsymbol{\cdot}k}$ on $\mathbf{X}$ using normal priors centered around zero.
  Draw $\theta^*_{pred,k} \sim p(\boldsymbol{\theta}_{pred,k}|\boldsymbol{\theta}^*_{\boldsymbol{\cdot}k}, \mathbf{X}, \mathbf{x}_{pred})$, i.e., conditional on sample $\boldsymbol{\theta}^*_{\boldsymbol{\cdot}k}$.
end procedure

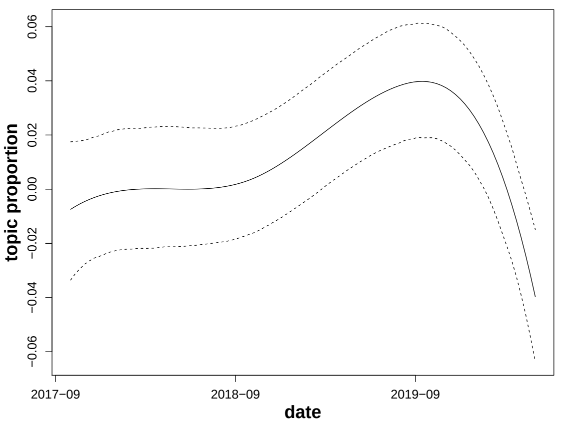

Figure 1: Mean prediction and 95% confidence intervals of topic "Climate Protection" over time, generated using `estimateEffect` from the R package `stm`.

variable in the interval $(0, 1)$. As shown by Atchison and Shen (1980), the Dirichlet distribution is well suited to approximate a logistic normal distribution, though inducing less interdependence among the different topics. When employing a Dirichlet distribution, the univariate marginal distributions are Beta distributions. We thus perform a separate Beta regression for each topic proportion on $\mathbf{X}$, using a logit-link.[5] This approach now again corresponds to Algorithm 1, but with g being the logistic sigmoid function in this case.[6]

4.2 Bayesian Beta Regression

Treier and Jackman (2008) and the authors of the STM consider $\boldsymbol{\xi}^*$ to be samples from the posterior of regression coefficients. While it is possible to view frequentist regression from a Bayesian perspective, it implies assuming a uniform prior distribution for regression coefficients $\boldsymbol{\xi}$ - which is rather implausible. More generally, the mixing of Bayesian and frequentist frameworks within the method of composition lacks theoretical foundation, especially when employing an *asymptotic* distribution of regression coefficient estimators. This applies to the model of Treier and Jackman (2008) as well as to the Beta regression presented in Section 4.1. Furthermore, note that when using a frequentist regression, the estimated uncertainty is with respect to the prediction of the mean of topic proportions. However, when exploring topic-metadata relationships it might be preferable to examine the variation of individual topic proportions among documents at different values of metadata covariates.

Therefore, we propose to replace the frequentist regression in Algorithm 1 by a Bayesian Beta regression with normal priors centered around zero. This enables modeling topic-metadata relationships in a fully Bayesian manner while preserving the methodological improvements from Section 4.1. Algorithm 2 summarizes this approach. By drawing $\theta^*_{pred,k}$ at covariate values $\mathbf{x}_{pred}$, we obtain samples from the posterior predictive distribution

$$p(\boldsymbol{\theta}_{pred,k}|\boldsymbol{\theta}^*_{\boldsymbol{\cdot}k}, \mathbf{X}, \mathbf{x}_{pred}) = \int p(\boldsymbol{\theta}_{pred,k}|\mathbf{x}_{pred}, \boldsymbol{\xi})p(\boldsymbol{\xi}|\boldsymbol{\theta}^*_{\boldsymbol{\cdot}k}, \mathbf{X})d\boldsymbol{\xi},$$

where $p(\boldsymbol{\xi}|\boldsymbol{\theta}^*_{\boldsymbol{\cdot}k}, \mathbf{X})$ denotes the posterior distribution of regression coefficients. This allows displaying the (predicted) variation of topic proportions at different covariate levels. As before, quantities of interest, such as the mean and quantiles, are obtained by averaging across samples; now, however, these samples are generated within a fully Bayesian framework.

5 Application [7]

In this section, we first apply the STM to German parliamentarians' Twitter data and subsequently demonstrate both the built-in (`stm`) and

[5] Note that the distribution of regression coefficient estimators is asymptotically normal for Beta regression (Ferrari and Cribari-Neto, 2004, p. 17).

[6] While runtime for estimating Beta regressions is considerably longer in relative terms, it is still short in absolute terms, which is why runtime concerns can be disregarded for the practical use of our approach.

[7] Source code in supplementary material; Will be made available on GitHub upon publication.

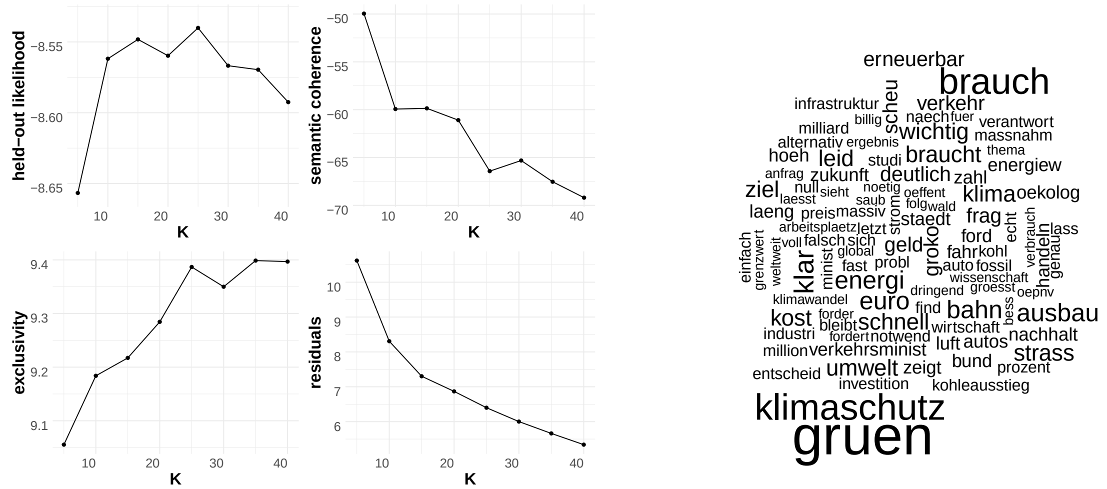

Figure 2: Left: Model evaluation metrics for hyperparameter K (number of topics). Right: Word cloud for the topic labeled as "Climate Protection".

our (`stmprevalence`) methods to explore topic-metadata relationships.

5.1 Data[8]

For all German MPs during the 19th election period (starting on September 24, 2017), we gathered personal information such as name, party affiliation, and electoral district from the official parliament website as well as Twitter profiles from the official party websites, using *BeautifulSoup* (Richardson, 2007). Next, after excluding MPs without a public Twitter profile, we used *tweepy* (Roesslein, 2020) to scrape all tweets by German MPs from September 24, 2017 through April 24, 2020. We also gathered socioeconomic data, such as GDP per capita and unemployment rate, as well as 2017 election results on an electoral-district level. Text preprocessing, such as transcription of German umlauts, removal of stopwords, and word-stemming, was performed with *quanteda* (Benoit et al., 2018).[9]

We define a document d as the concatenation of an individual MP's tweets during a single calendar month to achieve a sufficient document length. Our final data set includes 10,998 monthly MP-level documents, each one associated with 90 covariates.

5.2 Model Fitting and Global-level Analysis

Before fitting the STM, we need to decide on the number of topics, K. To do so, we use the following four model evaluation metrics: *held-out likelihood*, *semantic coherence*, *exclusivity*, and *residuals*. The held-out likelihood approach is based on document completion. The higher the held-out likelihood, the more predictive power the model has on average (Wallach et al., 2009). Semantic coherence means that words characterizing a specific topic also appear together in the same documents (Mimno et al., 2011). Exclusivity, on the other hand, indicates to which degree words characterizing a given topic *only* occur in that topic. Finally, the residuals metric, which is based on residual dispersion, indicates a (potentially) insufficiently small value of K whenever the residual dispersion is larger than one (Taddy, 2012).

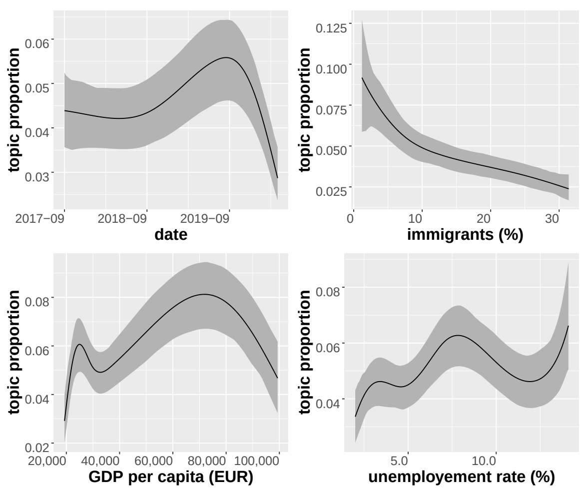

Figure 3: Mean prediction and 95% confidence intervals of topic "Climate Protection" for different document-level covariates, obtained using a frequentist Beta regression from the R package `stmprevalence`.

[8] Raw data: https://figshare.com/s/7a728fcb6d67a67fc3d6.

[9] An in-depth discussion of topic model preprocessing and its application to Twitter data can be found in Lucas et al. (2015).

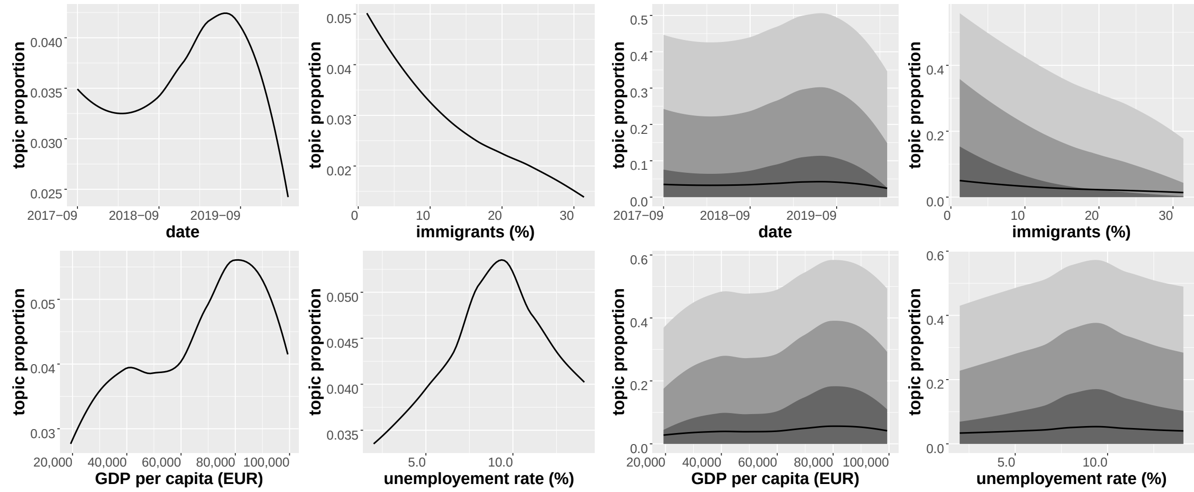

Figure 4: Left: Mean prediction of topic "Climate Protection" for different document-level covariates, obtained using a Bayesian Beta regression from the R package `stmprevalence`. Right: 95% (light grey), 90% (grey), and 85% (dark grey) quantiles of the posterior predictive distribution of topic "Climate Protection".

The left part of Figure 2 shows these four metrics for a grid of K between five and 40 with step size five. Both $K = 15$ and $K = 20$ seem to be good choices. Given the better interpretability for models with fewer topics, we choose $K = 15$.

After fitting the model we label all topics manually with human interpretable labels, using, i.a., a word cloud as displayed in the right part of Figure 2. To obtain an overview of the model output, we can conduct different global-level analyses, such as inspecting global topic proportions $\bar{\theta}_k = \frac{1}{D}\sum_{d=1}^{D}\theta_{d,k}$ or creating a network graph.

5.3 Topic-Metadata Relationships

Moving from global- to document-level, we now visualize relationships between document-level topic proportions $\theta_{d,k}$ and covariates $\mathbf{x}_d$. Specifically, we examine the extent to which German MPs discussed the topic "Climate Protection" over time and in relation to several socioeconomic variables regarding their respective electoral districts.

To demonstrate the shortcomings of the approach implemented in the `stm` package, we first apply the `estimateEffect` function to produce "naïve" estimates for the relationship between estimated topic proportions and document-level covariates. Figure 1 shows the estimated proportion of climate protection over time, peaking during the UN Climate Action Summit 2019 held in September 2019. As can be observed, `estimateEffect` produces predicted topic proportions outside of $(0, 1)$. This is due to using a linear regression, which places no restrictions on the range of the dependent variable.

Next, we evaluate the results when replacing the linear regression by a Beta regression, which restricts the dependent variable to the $(0, 1)$-interval. The top left plot of Figure 3 shows that the overall trend over time is similar to the one in Figure 1, yet the range is shifted and no negative values are observed. In addition, Figure 3 depicts the relationship of the climate protection topic with three socioeconomic covariates, for all of which we only obtain non-negative values. On average, the higher the share of immigrants in an electoral district, the less frequently MPs associated with this district tend to discuss climate-related subjects. For GDP per capita, we notice an increase until around EUR 70k, but for very high incomes this trend is reversed. The unemployment rate shows an ambiguous relationship, with rather large fluctuations.

Finally, we display the results from the fully Bayesian approach discussed in Section 4.2. As can be seen in the left plot of Figure 4, the predicted progressions of mean topic proportions at different covariate values are mostly similar to those obtained with the frequentist Beta regression, yet the range is compressed and shifted downwards. In addition to the empirical mean, the right plot of Figure 4 depicts different empirical quantiles of the posterior predictive distribution of topic proportions. Here we can see that topic proportions at different covariate values vary starkly for different MPs. In general, we find that a fully Bayesian approach enables a much more comprehensive analysis of topic-metadata relationships because it allows for displaying the variation of individual topic proportions observed in the data.

6 Conclusion and Outlook

To explore topic-metadata relationships while accounting for the probabilistic nature of topic proportions, the R package `stm` implements repeated linear regressions of sampled topic proportions on metadata covariates by using the method of composition. In this paper, we identified shortcomings of and proposed improvements upon this original implementation, applying latter ones to a dataset containing Twitter posts by German MPs. Our methods are equally applicable to other topic models and beyond.

Several possibilities exist to build upon our explorative methods. For instance, to make inference in a Bayesian setting, our approach could be used in combination with MCMC-based methods. If the goal is to make causal inference beyond explorative purposes, one must take into account that the estimation of topic proportions induces additional dependence across documents. Developing methods to identify underlying causal mechanisms is the subject of current research (see e.g. Egami et al., 2018).

References

J Atchison and Sheng M Shen. 1980. Logistic-normal distributions: Some properties and uses. *Biometrika*, 67(2):261–272.

Kenneth Benoit, Kohei Watanabe, Haiyan Wang, Paul Nulty, Adam Obeng, Stefan Müller, and Akitaka Matsuo. 2018. quanteda: An r package for the quantitative analysis of textual data. *Journal of Open Source Software*, 3(30):774.

David M Blei, John D Lafferty, et al. 2007. A correlated topic model of science. *The Annals of Applied Statistics*, 1(1):17–35.

David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022.

Jeremiah Bohr and Riley E Dunlap. 2018. Key topics in environmental sociology, 1990–2014: results from a computational text analysis. *Environmental Sociology*, 4(2):181–195.

Marie Chandelier, Agnes Steuckardt, Raphael Mathevet, Sascha Diwersy, and Olivier Gimenez. 2018. Content analysis of newspaper coverage of wolf recolonization in france using structural topic modeling. *Biological conservation*, 220:254–261.

Isaac Cho, Ryan Wesslen, Alireza Karduni, Sashank Santhanam, Samira Shaikh, and Wenwen Dou. 2017. The anchoring effect in decision-making with visual analytics. In *2017 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pages 116–126. IEEE.

Naoki Egami, Christian J Fong, Justin Grimmer, Margaret E Roberts, and Brandon M Stewart. 2018. How to make causal inferences using texts. *arXiv preprint arXiv:1802.02163*.

Justin Farrell. 2016. Corporate funding and ideological polarization about climate change. *Proceedings of the National Academy of Sciences*, 113(1):92–97.

Silvia Ferrari and Francisco Cribari-Neto. 2004. Beta regression for modelling rates and proportions. *Journal of applied statistics*, 31(7):799–815.

J Mason Heberling, L Alan Prather, and Stephen J Tonsor. 2019. The changing uses of herbarium data in an era of global change: An overview using automated content analysis. *BioScience*, 69(10):812–822.

In Song Kim. 2017. Political cleavages within industry: Firm-level lobbying for trade liberalization. *The American Political Science Review*, 111(1):1.

Christopher Lucas, Richard A Nielsen, Margaret E Roberts, Brandon M Stewart, Alex Storer, and Dustin Tingley. 2015. Computer-assisted text analysis for comparative politics. *Political Analysis*, 23(2):254–277.

David Mimno, Hanna M Wallach, Edmund Talley, Miriam Leenders, and Andrew McCallum. 2011. Optimizing semantic coherence in topic models. In *Proceedings of the conference on empirical methods in natural language processing*, pages 262–272. Association for Computational Linguistics.

Manuela Moschella and Luca Pinto. 2019. Central banks' communication as reputation management: How the fed talks under uncertainty. *Public Administration*, 97(3):513–529.

Leonard Richardson. 2007. Beautiful soup documentation. *April*.

Margaret E. Roberts, Brandon M. Stewart, and Edoardo M. Airoldi. 2016. A model of text for experimentation in the social sciences. *Journal of the American Statistical Association*, 111(515):988–1003.

Margaret E. Roberts, Brandon M. Stewart, and Dustin Tingley. 2019. stm: An R package for structural topic models. *Journal of Statistical Software*, 91(2):1–40.

Joshua Roesslein. 2020. Tweepy: Twitter for python! *URL: https://github.com/tweepy/tweepy*.

Matt Taddy. 2012. On estimation and selection for topic models. In *Artificial Intelligence and Statistics*, pages 1184–1193.

Martin A Tanner. 2012. *Tools for statistical inference*. Springer.

Shawn Treier and Simon Jackman. 2008. Democracy as a latent variable. *American Journal of Political Science*, 52(1):201–217.

Hanna M Wallach, Iain Murray, Ruslan Salakhutdinov, and David Mimno. 2009. Evaluation methods for topic models. In *Proceedings of the 26th annual international conference on machine learning*, pages 1105–1112.

Chong Wang and David M Blei. 2013. Variational inference in nonconjugate models. *Journal of Machine Learning Research*, 14(Apr):1005–1031.

Appendix

A Exemplary figures with implausible predictions

To demonstrate the importance of our proposed corrections of the STM, we collected figures from a selection of research papers where using the original implementation led to implausible estimates.

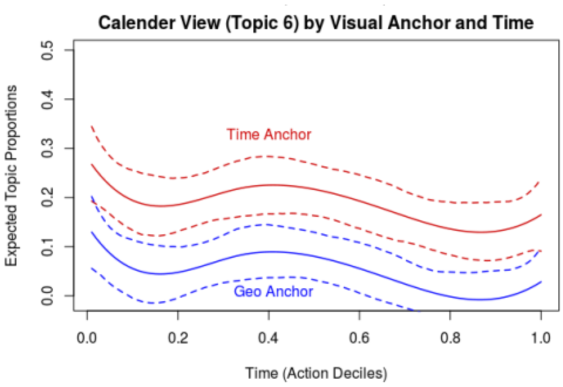

Figure 5: Example of negative confidence bands for covariate effects (Cho et al., 2017).

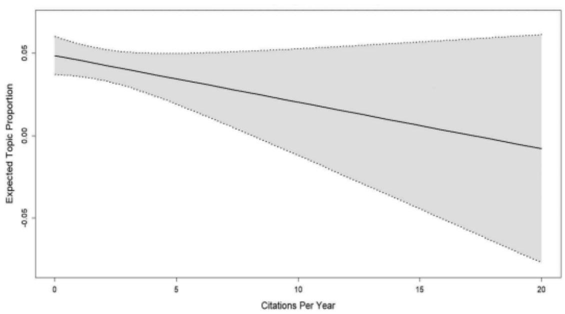

Figure 6: Example of negative confidence bands *and* negative covariate effects (Bohr and Dunlap, 2018).

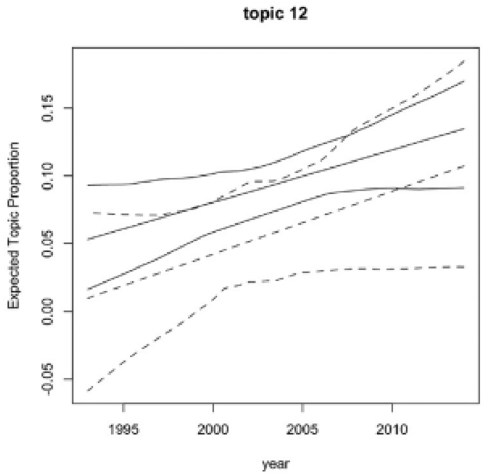

Figure 7: Example of negative confidence bands for covariate effects (Chandelier et al., 2018).

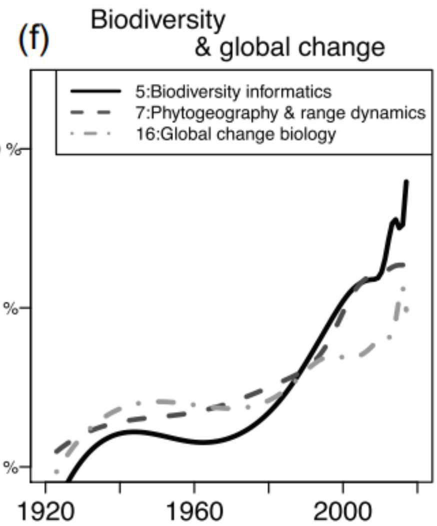

Figure 8: Example of negative covariate effects (Heberling et al., 2019).

Figure 9: Example of negative confidence bands *and* negative covariate effects (Moschella and Pinto, 2019).