

Abstract DH2023

Studying Parliamentary Economization using Topic Burstiness and Entropy

Introduction

This paper uses topic model burstiness and entropy to study the process of “economization” in Dutch postwar Lower House debates. According to studies into the history of economic thought, the postwar decades saw the gradual advance of “the economy” as the primary object of politics, and the intrusion of other political domains by economic values and frameworks. However, the historiography of economization almost exclusively focuses on intellectuals, scientific disciplines and economic institutions (Çalışkan & Callon, 2009; Mitchell, 1998; Schabas, 2009). The interaction between “economization” and democratic decision-making remains understudied.

This paper aims to contribute to this lacuna by studying the process of economization in proceedings of the Dutch Lower House in the period between 1945 and 1963. By studying the expected advance of economic language in parliament in this period, the impact of economization on democratic decision-making can be better understood.

Data

The paper uses proceedings from the Dutch Lower House. It focuses on the period between 1945 and 1963. The proceedings, consisting of curated reports of the speeches made in parliament, are tokenized and lemmatized. For training the topic model, speeches with fewer than ten words were pruned.

Methods

The paper uses Top2Vec, a relatively new method for topic modelling, to identify latent topics in parliamentary debates and generate document-topic distributions that can be used for subsequent analysis (See for recent examples of topic modelling in the Humanities: Viola & Verheul, 2020; Marjanen et al., 2020; Van Galen & Nicholson, 2018). Top2Vec generates word- and document embeddings and uses UMAP and HDBSCAN clustering to identify clusters of documents. After calculating cluster centroids, the most relevant terms for each topic are approximated using the cosine similarities between individual terms and topic/cluster centroids (Angelov, 2020).

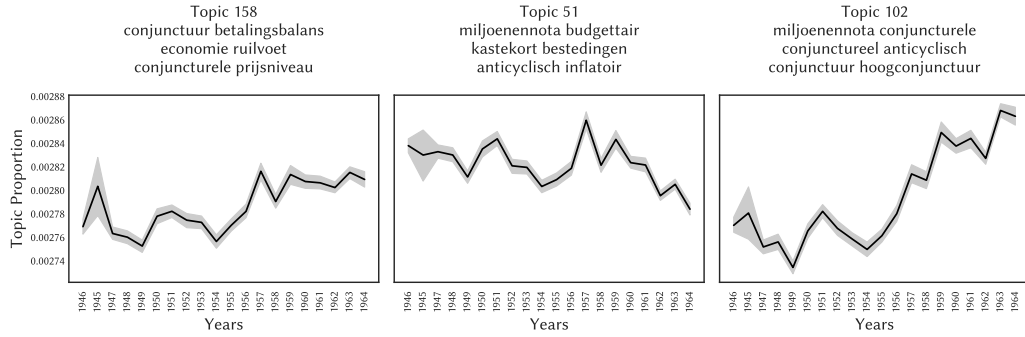
The paper focuses on three topics that deal explicitly with economic issues. They were manually selected based on the domain expertise of the author. Table 1 shows the topics and the corresponding terms.

Topic Index	Topic Terms (Translated)
158	business cycle, balance of payment, economy, trading, cyclical, price level, export position
51	budget, budget, cash float, investment, anti-cyclical, inflationary, cyclical
102	budget, cyclical, cyclical, anti-cyclical, business cycle, spending policy

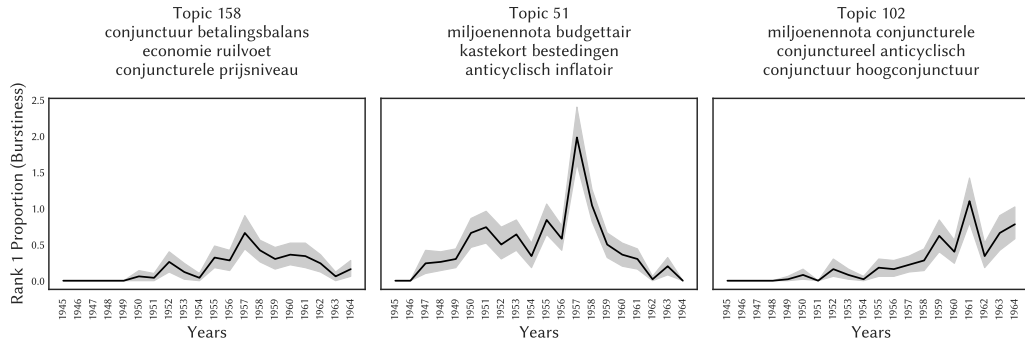
Table 1: Key terms for three manually selected economic topics.

The concept of economization implies a gradual process that unfolds in a covert way and structurally transforms political discourse. Parts of this process are likely to be captured by diachronic topic proportions since economization is likely to reflect in an overall increase in economic language. However, economic issues can be increasingly discussed in parliament without the broader political context being economized". Therefore, economization pertains more to the intrusion of other policy domains by economic language. Instead of topic proportions, frequencies or sizes, their rankings thus appear as more relevant units of analysis. "Aneconomized" speech would be about any issue but the economy, but simultaneously maintains highly ranked economic topics. This, in turn, can be considered as a decline in "burstiness", since the process of economization would translate into economic topics becoming less explicitly debated on the one hand, but increasingly present as high ranking topics on the other. Burstiness can be defined as the degree to which a topic appears in "bursts". Existing work leverages this measure to detect bursty topics that indicate events, or to improve topic modelling algorithms (Doyle & Elkan, 2009; Qi et

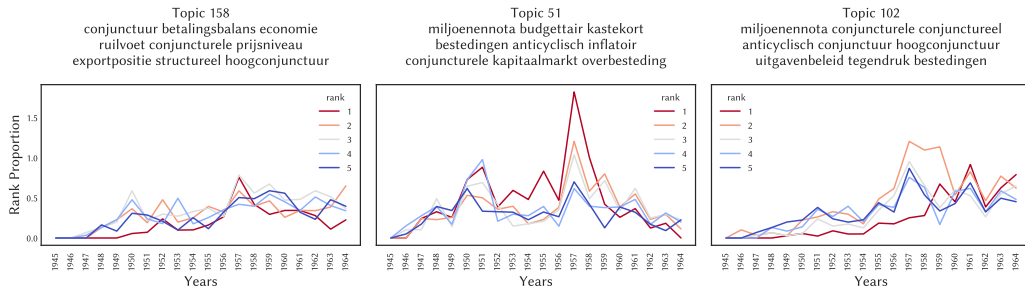
al., 2014). For this study, burstiness appears as a suitable target statistic. It can also be used for subsetting speeches with high-ranking economic topics. Economization would also imply a decreasing diversity in these speeches. The more political debate becomes economized, the similar speeches appear. Additional to burstiness, the paper therefore also considers relative entropy in subsetting speeches as a way of understanding economization.



(a) Diachronic proportions of economic topics aggregated on the year-level. Propotions are derived from normalized document-topic distributions.



(b) Bootstrapped diachronic burstiness as measured in rank 1 proportion. Burstiness is the ratio of speeches in which a topic is ranked first.



(c) Diachronic rank 1-5 proportions. The proportion of rank N is the ratio of speeches in which a topic is ranked N th.

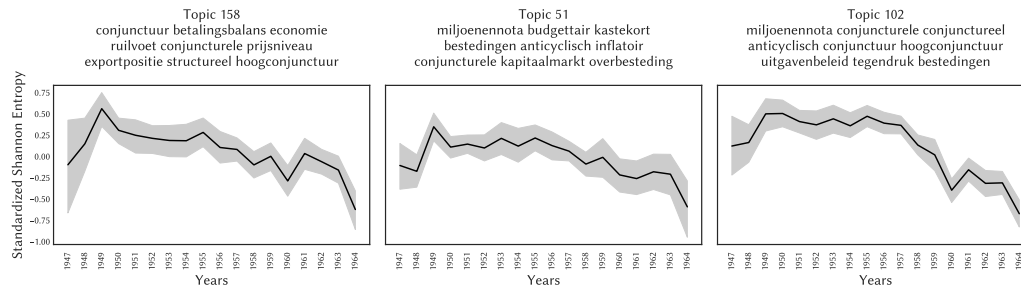
Figure 1: Topic Proportions, Rankings and Burstiness

Preliminary Results

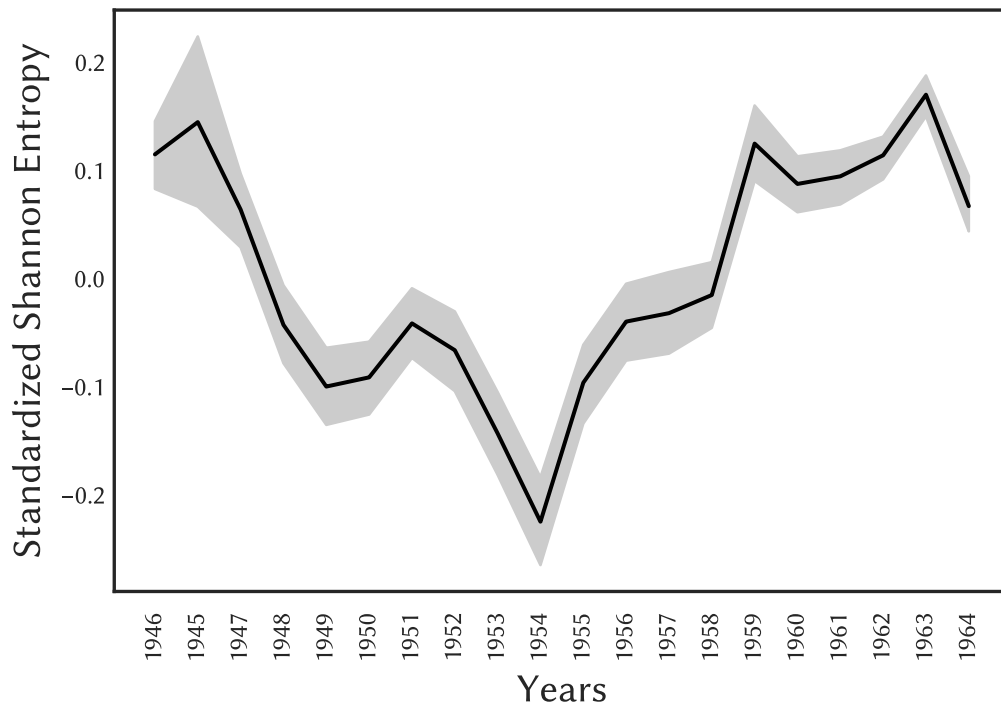
Preliminary results show complex dynamics at work in the evolution of economic topics. Figure 1a shows the overall increase in proportion for two economic topics, and the modest decline for another (topic 51) from 1957 onwards. As visible in Figure 1b, the burstiness, measured in the proportion of speeches in which an economic topic is ranked highest, does not fully conform to the diachronic topic proportion. Topic 51 for example, is extremely bursty in 1957, a peak that is less visible in its diachronic proportion. Similarly, topic 158 becomes bursty much later than proportional growth sets in. To understand the burstiness dynamics, Figure 1c shows the diachronic proportions of multiple ranks. In the case of topic 51, for example, approximately 1.8% of the speeches have this topic as their highest scoring one in the year 1957. This peak of rank 1 contrasts with the peak for rank 2 for topic 102 in the same year.

The different rankings can also be used to create speech subsets. In these subsets, diversity as measured by entropy is a good measure to extract signals of economization. Figure 2a shows the Shannon entropy over time for the three economic topics. Surprisingly, they show a comparable trend of modest decline. This differs from the overall topic entropy, visualized in Figure 2b. This might be a sign of economization, as speeches where the economy is highly ranked become less diverse and language becomes more the same.

All in all, these first investigations do not suggest a process of linear economization based on burstiness signals. However, entropy measures do show a declining diversity across economic topics. In the final paper, the results will be tested for robustness and example speeches will be used to find explanations for the patterns observed in these initial visualizations.



(a) Diachronic Shannon Entropy of economic topics.



(b) Diachronic mean Shannon Entropy for all topics.

Figure 2: Shannon Entropy

Literature

Angelov, D. (2020). Top2Vec: Distributed Representations of Topics (arXiv:2008.09470). arXiv.

- Çalışkan, K., & Callon, M. (2009). Economization, part 1: Shifting attention from the economy towards processes of economization. *Economy and Society*, 38(3), 369–398.
- Doyle, G., & Elkan, C. (2009). Accounting for burstiness in topic models. *Proceedings of the 26th Annual International Conference on Machine Learning*, 281–288.
- Marjanen, J., Zosa, E., Hengchen, S., Pivovarova, L., & Tolonen, M. (2020). Topic modelling discourse dynamics in historical newspapers (arXiv:2011.10428). arXiv.
- Mitchell, T. (1998). Fixing the Economy. *Cultural Studies*, 12(1), 82–101.
- Qi, X., Huang, Y., Chen, Z., Liu, X., Tian, J., Huang, T., & Wang, H. (2014). Burst-LDA: A new topic model for detecting bursty topics from stream text. *Journal of Electronics (China)*, 31(6), 565–575.
- Schabas, M. (2009). Constructing “The Economy”. *Philosophy of the Social Sciences*, 39(1), 3–19.
- Van Galen, Q., & Nicholson, B. (2018). In Search of America. *Digital Journalism*, 6(9), 1165–1185.
- Viola, L., & Verheul, J. (2020). Mining ethnicity: Discourse-driven topic modelling of immigrant discourses in the USA, 1898. *Digital Scholarship in the Humanities*, 35(4), 921–943.