Erasmus University Rotterdam

Erasmus School of Economics

Master Thesis Proposal
Econometrics and Management Science

Convergence of COVID policy: Testing Persuasion Bias

Author:
Greta Wackerbauer

Student number:
495066

Supervisor:
Robin Lumsdaine

Second Assessor:
Dennis Fok

Date: July 10, 2023

Abstract

This study examines convergence in policymakers' and advisors' opinions on 16 aspects of COVID-19 policy in the United Kingdom (UK) and the Netherlands. By applying the opinion formation framework by DeMarzo et al. (2003) empirical evidence is found on their predictions regarding bias and dimensionality of opinions. Textual data is used from COVID advisory body meeting notes and press conference transcripts between March 2020 and March 2021. The research takes a novel approach by assessing aspect-based opinions through sentiment and tone simultaneously. Additionally, it applies state-of-the-art topic modelling, sentiment analysis and emotion recognition techniques. As such, it contributes to the literature on textual analysis of policy communication. The main findings are that there is more convergence in opinions between policymakers and their respective advisors than convergence to an average opinion. Opinions remain at least two-dimensional throughout the observation period and obstacles to convergence are likely due to uncertainty and economic crises.

Keywords: BERTopic · Opinion mining · Persuasion Bias · Convergence

Contents

1 Introduction

During the COVID-19 pandemic, policymakers were confronted with a high level of uncertainty and risk surrounding their decisions regarding, among others, economic, healthcare, and social policies. In light of this, it is socially relevant to study the decision-making processes of different policymakers and how their opinions have evolved.

One potential framework to explain the opinion formation process in societies was proposed by DeMarzo et al. (2003). In their model agents are influenced by a cognitive bias called persuasion bias, wherein agents fail to adjust for repetition in the information they receive. The theoretical framework outlined by DeMarzo et al. (2003) suggests that, as a consequence of persuasion bias, opinions in a society will converge to a weighted average of all agents' opinions. These weights reflect the individual's influence and depend on their connections to others within the society or social network. This phenomenon is referred to as social influence. Moreover, the framework suggests that over time, differences in opinion across topics will become unidimensional. Thus, viewing opinions on any given topic as a scale, an agent's views may be to the left or right of the average opinion at the start, but over time their opinions will converge to be consistently to the left or right of the average on all topics. This scale could for instance be from conservative to liberal or from rigid to flexible.

This thesis aims to test the theories on opinion convergence proposed by DeMarzo et al. (2003), specifically their hypotheses regarding convergence of opinions and unidimensionality, in the case of COVID-19 policymakers. My main research question is thus

> *Is there evidence of convergence in policymaker's opinions regarding COVID-19 policies?*

with the following three sub-questions.

1. To what extent do opinions in transcripts from advisory body meetings and press conferences converge over time?
2. To what extent does advice from experts and political action converge over time, based on meeting notes, press conference transcripts, and enacted policies?
3. How does the dimensionality of each country's policy advice across 16 indicators of COVID policymaking evolve over time?

I will answer these questions by using aspect-based opinion mining (ABOM) on meeting notes from advisory bodies and transcripts from press conferences related to COVID-19 policy in the Netherlands and the United Kingdom (UK) between March 2020 and March 2021. ABOM is a natural language processing (NLP) method that extracts emotional sentiment toward different topics. This is scientifically relevant as it contributes to the literature on empirical evidence for the effects of persuasion bias and the influence of advisory bodies on political action. Additionally, it adds to the growing field of studies regarding policy-making during the COVID-19 pandemic.

The main findings can be summarized in three key results. First there is no robust evidence of convergence in the opinons of aspects in policymakers' or advisors' opinions towards an approximation of the weighted average opinion. Nevertheless, there is some indication in convergence of opinion rate of change regarding policy on vaccination and protection of the elderly. Second, convergence is more evident in opinions between policymakers and their respective advisors, especially regarding policy on facemasks. This policy aspect is found to be particularly divisive across countries but convergent groups are found that combine the policymakers' with their respective advisors. Finally, differences in opinions are found to be three to four dimensional. Uncertainty surrounding COVID and the number of press conferences and meetings likely explains this result. Changes in the dimensionality coincide with the announcements of strengthened or relaxed policies.

This research contributes to the literature on policy evaluation through the novel application of ABOM to policy communication to analyze sentiment and tone. The use of the framework derived by DeMarzo et al. (2003) also contributes to empirical evidence on persuasion bias. The findings suggest that coordination and communication between advisory bodies and political leaders has room for improvement. The investigation of the dimensionality of opinions over time also sheds light on possible obstacles to unidimensionality.

In the following, I will outline the relevant literature in Section 2. Next, I will discuss the considered data sets in Section 3. In Section 4 the methodological approach is described and Section 5 presents the results. Finally, Section 6 concludes.

2 Literature Review

In the following section, I will begin by reviewing the framework described in DeMarzo et al. (2003) and empirical evidence related to persuasion bias. Then, I will review previous applications of NLP and ABOM to policy communications. Finally, I will discuss the policy and development of the pandemic in the two countries investigated, the Netherlands and the UK.

2.1 Persuasion Bias and Unidimensionality

The two predictions of the model proposed by DeMarzo et al. (2003) are that (i) opinions converge to a weighted average of all opinions in a society and that (ii) differences in opinions will consistently fall on one side of the average opinion. The first is referred to as the persuasion bias hypothesis and the second is referred to as the unidimensionality hypothesis. For example, opinions on vaccines could be viewed on a spectrum from skeptical (left) to confident (right). Individuals will differ in their degree of hesitancy towards different vaccines at first, but as they begin to discuss their opinions in their social group their opinions will begin to converge towards each other. If someone who is extremely skeptical about a vaccine engages in a discussion with individuals that are very confident in a vaccine's efficacy, they may be swayed toward their opinion. As the discussion continues they may update their stance on other vaccines as well, characterizing themselves as hesitant towards all vaccines. That is, an individual will begin to modify their stance about different vaccines to be consistently to the left or right, that is more or less skeptical of the vaccine, than the average opinion in their social group.

The persuasion bias hypothesis is confirmed empirically if, within a group of individuals, opinions converge towards some weighted average of everyone's opinions. This convergence is expected to occur over time when individuals discuss opinions with each other and update their beliefs accordingly. This exchange of opinions can occur directly or indirectly, as not every individual listens to each other in every discussion. Indeed, as individuals update their opinions using this subset of people that they listened to, their opinion becomes biased as they fail to correct for repeated information in the discussion. For example, when a rumour is shared within an office, its credibility can become inflated when it repeated by different people. Let us assume person A shares a rumor with person B and with person C, person C will strengthen their belief in that rumour when person A shares this rumour with them in a later conversation. This is because they fail to adjust for the fact that the rumour originated from the same person. As a result, too much credibility is placed into this rumour and the office opinion becomes biased.

Unless the group structure cancels out the biased weights for each piece of information, the opinion to which the group converges will differ from the truth. The time it takes for opinions to converge increases quadratically in the number of individuals involved. More specifically, DeMarzo et al. (2003) predict that when agents are fully rational and a network has strong connections, then opinions within this network will converge within N^2 rounds of exchange, where N is the number of individuals and a round of exchange is a discussion between at least two members of the group. While long-run differences in opinions can exist prior to this convergence, these differences become unidimensional, meaning that they fall consistently to the left or right of the average opinion on some arbitrary scale (i.e vaccine skepticism).

In practice, it is difficult to assess the weights assigned to each opinion and to identify what the underlying truth is, thus this model has mainly been studied in a lab setting. For example, in a lab experiment involving four agents, Brandts et al. (2015) find that opinions converge in a manner consistent with the framework of DeMarzo et al. (2003). In another lab experiment with the same number of agents, Corazzini et al. (2012) find evidence in line with a generalization of this model. Namely, where the weights used in averaging opinions depend not only on who is listening to an agent but also on whom they listen to. These experiments provide support for the hypotheses made by DeMarzo et al. (2003) in small groups. Additionally, the findings of Corazzini et al. (2012) suggest an individual's influence on the average opinion depends on the attention they receive and give to others.

While lab experiments allow for a more controlled test of the model proposed by DeMarzo et al. (2003), empirical studies have also sought to apply this framework. For instance, the unidimensionality hypothesis was tested empirically using survey responses on the 2007 election in France. In contrast to the aforementioned lab experiments, Page (2020) finds that opinions on the French political parties have not converged, despite a long period of political stability. Using a measure of reliability, dissimilarity, and PCA the dimensionality of differences in opinions is assessed. Page (2020) finds that 3-6 factors are sufficient to explain the variation across survey questions and the first factor captures a large portion of this variation. Thus, while one dimension is dominating the differences in opinions, no evidence of unidimensionality is found. In the context of COVID-19 policy, studies in the US have found evidence of polarisation in opinions on containment policies and in the perception of pandemic risk (Allcott et al., 2020; Bursztyn et al., 2020).

This lack of support for the unidimensionality hypothesis could be attributed to a number of factors that prevent or slow convergence. First, the individuals involved in discussions are not constant. For example, Page (2020) suggests that the voting population in France may change faster over time than political opinions can converge. Second, some opinions may be unchanging. Groups or organisations with strong agendas could influence others' opinions without changing their own stance. If these groups have a large influence it results in noisy signals that can prevent convergence (Golub & Jackson, 2010). For instance, Bursztyn et al. (2020) find a correlation between an individual's COVID-19 risk perception and behaviour and their preferred media outlet. Third, the importance and divisiveness of topics may change over time. If the key topics that are focused on in a society shift, or if new divisive topics emerge, opinions may remain multidimensional even if there is a tendency to converge (Page, 2020). Furthermore, on a global scale, communication involves multiple languages. Foerster (2018) proposes that when agents communicate in such a network, beliefs will fluctuate in the neighborhood of these languages.

Moreover, in a study of the consolidation of narratives in Swedish newspapers, Bertsch et al. (2021) find a strong positive correlation between consolidation and GDP growth. They find evidence of an overall convergence in opinions during periods of economic expansion and divergence during recessions. Important reference events are shown to improve consolidation. This suggests that convergence may be slower during economic recessions or times of uncertainty, while influential, global events could increase the rate of convergence.

These obstacles to opinion convergence and unidimensionality reveal complications in the opinion formation framework that lab conditions are shielded from. Under perfect conditions, policymakers' opinions should converge in the long run if they listen to one another. However, in a real world setting a lack of convergence does not falsify the persuasion bias or unidimensionality hypotheses. Instead, changes in opinion can be used to make inferences about who policymakers listen to and how they update their beliefs. In the short run, the dimensionality of differences in opinion on fixed topics is expected to decrease. Various global events and changes in policy could affect this dimensionality. Key events, like announcements or changes in policy, can reduce the dimensionality of differences in opinions, while economic hardship or other obstacles to convergence may increase it.

2.2 NLP in Policy Communication

In order to analyse these changes in opinions, textual analysis can be used to extract aspect-based opinions from textual data. First aspects in the text are identified and then opinions towards each aspect are determined. In this section, I provide an overview of previous work on extracting opinions from policy communications or other related forms of text.

Aspects are identified through topic modeling. A common approach in social science research is the structural topic model, which is an unsupervised machine learning method that can incorporate additional covariates about the text (Goyal & Howlett, 2021; Madrigal, 2023; Zhou et al., 2023). Similarly neural topic models, like Scholar, can also incorporate metadata included in documents (Dreier et al., 2022). Alternatively, a hybrid approach can be used that combines the unsupervised topic model with manual encoding to obtain more interpretable and relevant topics for different research objectives (Yarchi et al., 2021). Topics may also be identified based on keywords, document features, or word embeddings derived from transformers (S. Müller, 2020; Viehmann et al., 2022). When comparing several of these techniques, namely feature-based approached, transformer-based approaches, and neural networks, for stance detection on Tweets about German Covid-19 policy measures, Viehmann et al. (2022) finds that the transformer-based model, also known as Bidirectional Encoder Representations from Transformers (BERT)-based model, has the highest accuracy. In fact, BERT is the base model chosen by M. Müller et al. (2023) in their adapted COVID-Twitter-BERT. Thus, the BERT-based model is a state-of-the-art topic modeling technique that can be applied to COVID-related textual data.

There are many ways to represent opinions, including sentiment (Meyer-Gutbrod & Woolley, 2021), tone (Motta & Stecula, 2023), or stance (Viehmann et al., 2022). In most cases these values are derived using a dictionary containing words and their associated scores. One option is to use an existing dictionary, such as the NRC Emotional Lexicon (Meyer-Gutbrod & Woolley, 2021), the Lexicoder dictionary (S. Müller, 2020), the Linguistic Inquiry and Word Count dictionary (Motta & Stecula, 2023), or Valence Aware Dictionary and sEntiment Reasoner (VADER) (Cochrane et al., 2022). Alternatively, a domain-specific dictionary can be constructed (Madrigal, 2023; Zhou et al., 2023). Indeed, Cochrane et al. (2022) highlight the superior performance of dictionaries generated using word embeddings over other supervised machine learning techniques and dictionaries. They also highlight the performance of VADER and transformer-based models. These techniques can be applied to study opinions in various forms across many applications. For example, to study sentiment expressed in press briefings (Meyer-Gutbrod & Woolley, 2021), parliamentary speeches (Cochrane et al., 2022), or online discourse (Zhou et al., 2023).

While Cochrane et al. (2022) find that emotional arousal is captured less accurately then sentiment in speech transcripts, this research takes a novel approach at modeling emotion. It further builds on the work by Zhou et al. (2023) that compares the degree of politicization in online COVID-19 communication stemming from experts, politicians and governmental agencies.

2.3 COVID policy

Globally, COVID-19 has caused large-scale disruptions and uncertainties in economic development. Policies include lockdowns, travel bans, and containment regulations. In addition, fiscal policies are focused on mitigating the socio-economic impact of the pandemic. However, there is heterogeneity across countries in policy effectiveness and individuals' perceptions of policies as well as considerable variation of perception for groups and individuals over time (Georgarakos & Kenny, 2022). Han et al. (2022) find that containment policies are the most effective in reducing virus transmission.

The two countries considered here make an interesting case study as the UK and the Netherlands are similar in terms of historical, social, and political factors (Todd et al., 2022).

In a study of the effectiveness of COVID-19 policies across 68 countries, Han et al. (2022) find that the UK is the country with the largest indirect effect on other countries. This makes it a particularly relevant country to consider in terms of policy opinions and opinion formation. The measures announced by the two countries in response to the COVID-19 pandemic follow each other closely and are shown along with key actions of the World Health Organisation (WHO) in Figure 2.1.

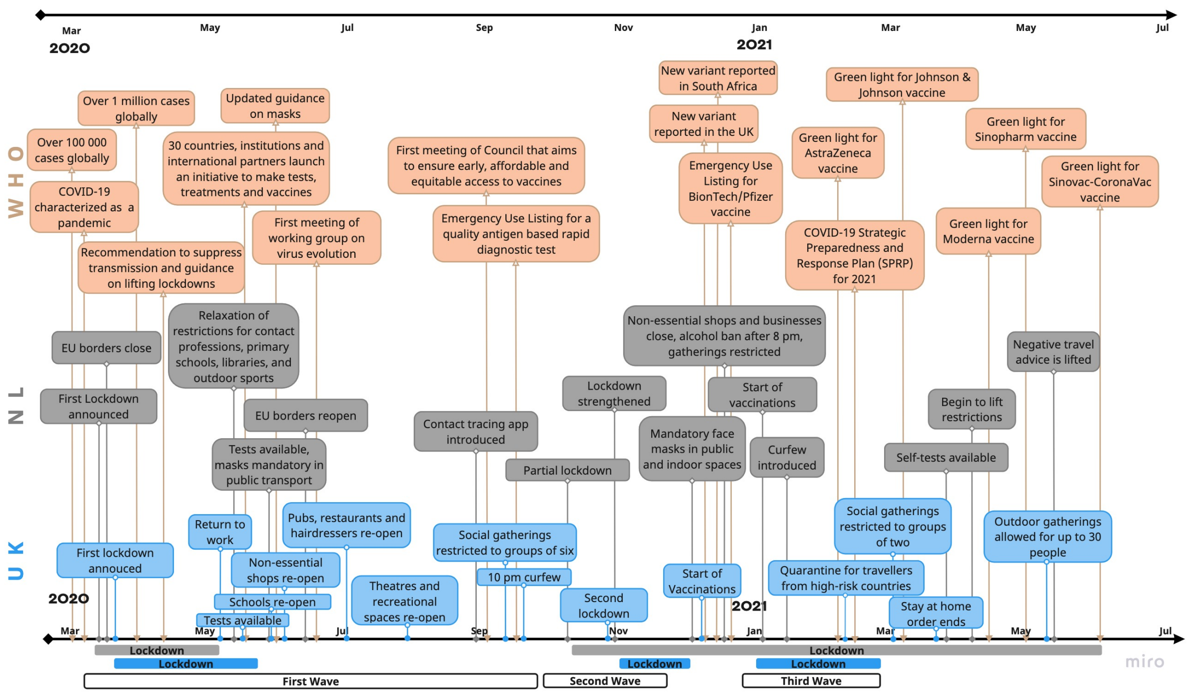

Figure 2.1. Timeline of events in the UK and NL and key actions of the WHO

More specifically, both countries start to lockdown prior to the WHO guidance on suppressing transmission of the virus. In the Netherlands, this first lockdown lasts from March 15th until May 11th, 2020. The UK enforces their lockdown from the 23rd of March until June 1st. In both countries, containment policies stay in place even after the lockdown ends and are relaxed step by step (Ramaekers et al., 2023; "Timeline of UK government coronavirus lockdowns and restrictions", 2022). Tests also become available in early June 2020 and mid-May for the Netherlands and the UK respectively. Throughout the pandemic, both countries make use of stay-at-home orders, workplace closures, and school closures. Additionally, both countries make distinctions in policy for essential shops or businesses and for contact professions.

During the second and third wave of COVID-19 cases, the strategies start to differ. On the one hand, the Netherlands enforced a partial lockdown in October, that was gradually strengthened through mandatory face masks, restrictions on gatherings, alcohol bans, shops closing and a curfew ("Coronavirus tijdlijn", n.d.). On the other hand, the UK restricted social gatherings and introduced a curfew as early as September but didn't announce the second lockdown until November 2020. This lockdown was lifted after 4 weeks. Shortly after the WHO received reports of new variants being discovered in the UK and South Africa. A third lockdown was entered in response to the third wave of COVID-19 cases between the 6th of January and the 1st of March 2021, but a stay at home order and restrictions on social gatherings remained in force until the 29th of March and the 17th of May, respectively. Meanwhile, the Netherlands remained in a lockdown but also began to lift restriction in April 2021.

Key actions of the WHO include launching initiatives and forming councils that focus on international cooperation and coordination in developing treatments, tests and vaccines for COVID-19 ("Timeline: WHO's COVID-19 response", n.d.). They also issued Emergency Use Listings, which assess the risk of unlicensed medical products and treatments with the aim of increasing their availability. The first emergency use listing for a vaccine was granted for the BionTech (also known as Pfizer) vaccine on the 31st of December 2020, which the UK medicine regulators were the first to approve already on the 8th of December (Hale et al., 2021). Vaccinations thus started in early December in the UK while they began a month later in January of 2021 in the Netherlands. In line with their commitment to ensure early, affordable and equitable access to vaccines, the WHO issued emergency use listings for 5 other vaccines between December 2020 and June 2021.

A brief overview of the development of the pandemic in the UK and the Netherlands (NL) is given below. Figures 2.2 shows the cumulative number of confirmed cases as a percentage of the population (Hale et al., 2021).

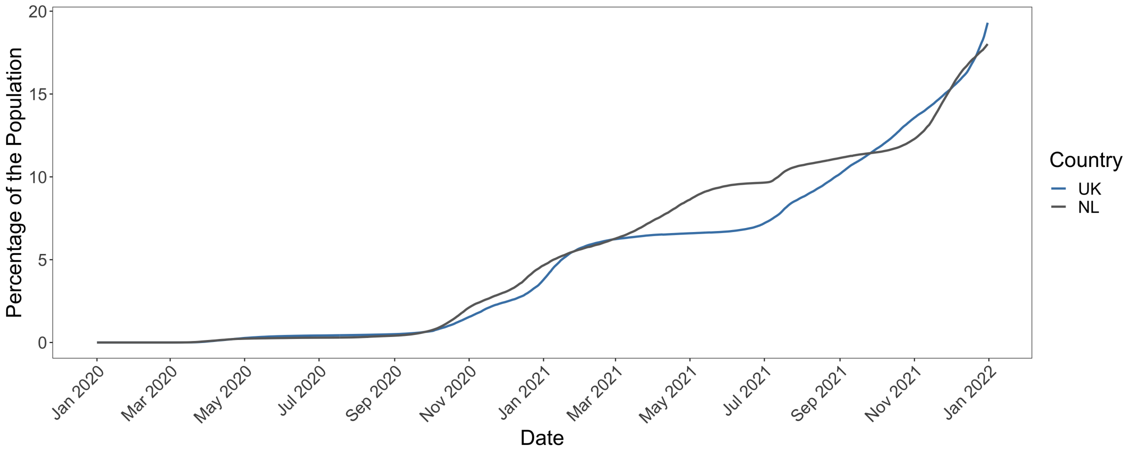

Figure 2.2. Confirmed cases as a percentage of the population in the UK and NL

In the UK there is a stagnation in the number of confirmed cases between January and July 2021 while the growth resembles exponential growth elsewhere. In contrast, the confirmed cases in the Netherlands do not show a stagnation but the growth rate is otherwise similar. However, it is important to note that the number of confirmed cases depends heavily on the number of tests performed and thus only provides an indication of the state of the pandemic. The percentage of the population that has received vaccinations in the Netherlands and the UK over the course of 2020 and 2021 is shown in Figue 2.3 (Hale et al., 2021).

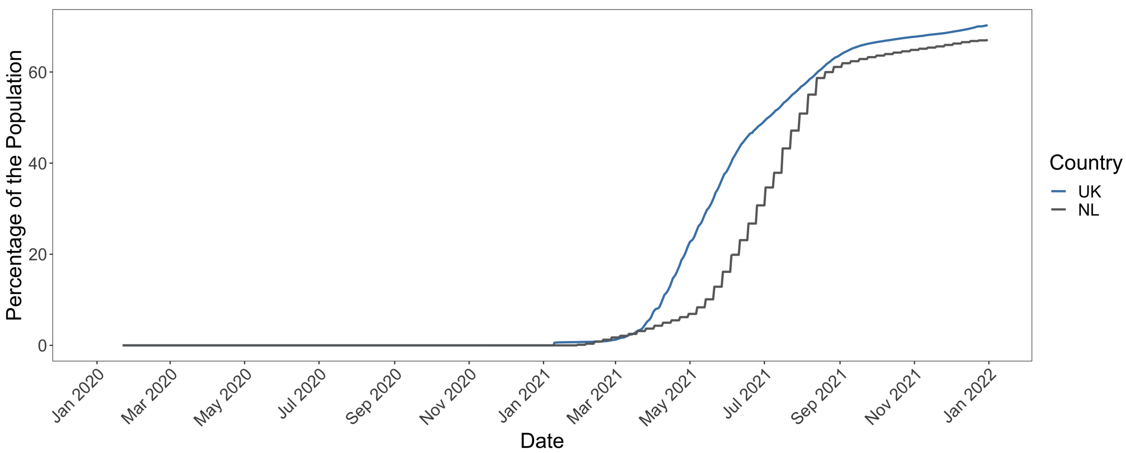

Figure 2.3. Vaccinated percentage of the population in the UK and NL

The percentage of the population that is fully vaccinated for both countries increases from March 2021 onwards and levels off at around 70%. The vaccinations show more rapid growth in the UK while the Netherlands has slightly slower, step-like growth. This is likely due to the fact that vaccinations became available earlier in the UK than in the Netherlands. Individuals above the age of 80 were eligible to get the vaccine from the 8th of December 2020 in the UK while in the Netherlands it was made available from the 25th of January 2021 (Hale et al., 2021). Individuals who were considered at risk and above the age of 16 were eligible to get the vaccine after the 16th of February in the UK while in the Netherlands at-risk individuals above the age of 20 were eligible after the 20th of February. The age floor for vaccine eligibility continued to be lower in the UK than in the Netherlands for the general population until the 12th of June 2021 (Hale et al., 2021).

3 Data

The data used in this research consists of textual data from meeting notes and press conferences between March 2020 and March 2021 ($t = 1, ..., 242$), as well as ordinal scores for 16 policy indicators and 4 indices on the governmental response constructed by Hale et al. (2021). This observation period is limited by the dates of the coronavirus press conferences held in the UK, the first of which was held on the 3rd of March 2020 and the last of which was held on the 23rd of June 2021. I restrict the window further to exclude observation after the 31st of March 2021 due to the reduced frequency of press conferences and meetingn notes from April 2021 onwards. In order to maintain consistency, only these press conferences marked as being focused on COVID-19 are used. I focus on two countries that are members of the Organization for Economic Cooperation and Development (OECD), namely the Netherlands and the UK. These countries have a large range of publicly available information regarding their policies and their decision-making processes.

First, the UK has made transcripts from 133 press conferences publicly available between March 2020 and March 2021 (GOV.UK, 2022) Accompanying these transcripts meeting minutes are made available from the Scientific Advisory Group for Emergencies (SAGE) which contain situation updates and policy advice from a group of scientific experts. During this observation period, this group met 74 times.

Second, in the Netherlands, a similar team of scientific experts was formed called the outbreak management team (OMT)(RIVM, n.d.). This group met 48 times in the considered period and meeting notes were published by the National Institute for Public Health and the Environment (Rijksinstituut voor Volksgezondheid en Milieu, RIVM). Additionally, 86 press conferences related to COVID were held during this period by Prime Minister Rutte and Deputy Prime Minister Kaag (Ministerie van Algemene Zaken, 2022).

The frequency of publication of these four textual data sources is shown in figure 3.1.

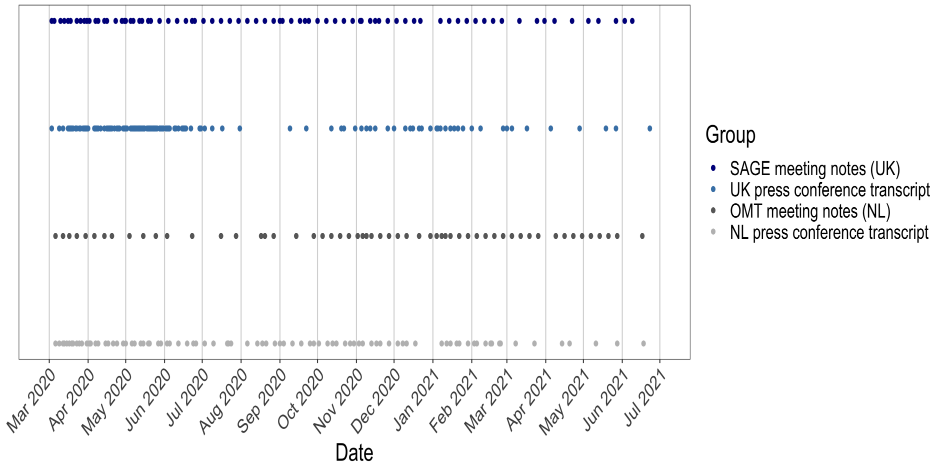

Figure 3.1. Frequency of publication meeting notes and transcripts

Moreover, international organizations such as the WHO and the OECD also released advice regarding pandemic policies. The World Health Organisation held 126 press conferences between March 2020 and March 2021 while the Secretary General of the OECD released 197 policy responses and statements (OECD, 2021; Organization, n.d.). The WHO, on the one hand, plays an important role in preventing and advising on public health emergencies and thus potentially impacts the convergence of policymakers' opinions. The OECD, on the other hand, can serve as a proxy for the average weighted opinion in OECD countries in the absence of textual data from all member countries. Each published transcript, piece of policy advice or released meeting note is hereafter referred to as a document.

Summary statistics for the number of documents per month and the number of sentences per document for each source are shown in Table 3.1. These six sources will hereafter be referred to groups.

Table 3.1. *Summary statistics*

	WHO	OECD	SAGE	OMT	UK	NL
Total numb. of Documents	126	197	74	48	133	86
Avg. numb. of documents per month	9.692	15.154	5.6923	3.6923	11.083	6.6154
	(2.462)	(10.93)	(2.250)	(1.377)	(10.48)	(2.599)
Avg. numb. of sentences per document	269.627	139.53	72.919	115.625	57.985	323.256
	(35.815)	(56.890)	(29.888)	(49.462)	(23.960)	(60.349)

Note. Avg. stands for average, numb. stands for number.
Note. Standard deviations are shown in round brackets

Finally, the ordinal scores are taken from the Oxford COVID-19 Government Response Tracker created and maintained by the Blavatnik School of Government, a department of the

University of Oxford (Hale et al., 2021). The ordinal scores represent the degree to which a policy targets a given indicator, with 0 meaning there is no policy related to this indicator and incremental increases in stringency as the score increases. A list of the 16 indicators and their description is shown in Table A.1 in the Appendix. They can be grouped into eight indicators on containment policies, 2 indicators on economic policies, and 6 indicators on health policies. Hale et al. (2021) average these indicators to construct three indices that take values between 0 and 100, the government response index (using all 16 indicators), the Stringency index (using the containment indicators, and an indicator on public health campaigns), and the Economic support index (using the economic indicators) shown in Figures 3.2-3.4. I supplement these indices with a fourth index focusing on health policies using the same aggregation method as outlined by Hale et al. (2021) (using the health indicators). After January 2021, some containment indicators are differentiated (see A.1), meaning that different policies applies to vaccinated and non-vaccinated people. In these cases, the dotted line represents the index based on policies for non-vaccinated individuals, while the solid line continues to show the index for policies pertaining to vaccinated individuals.

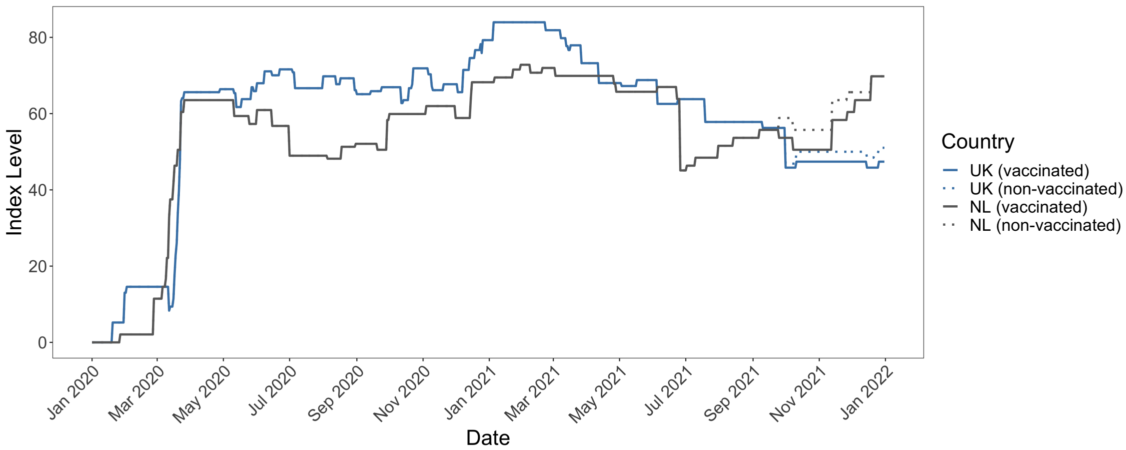

Figure 3.2. Government Response Index

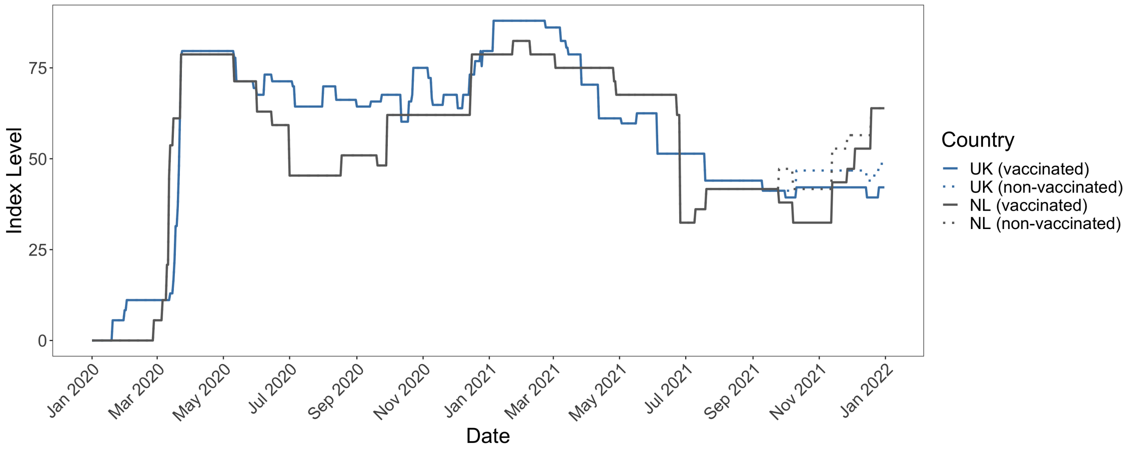

Figure 3.3. Stringency Index

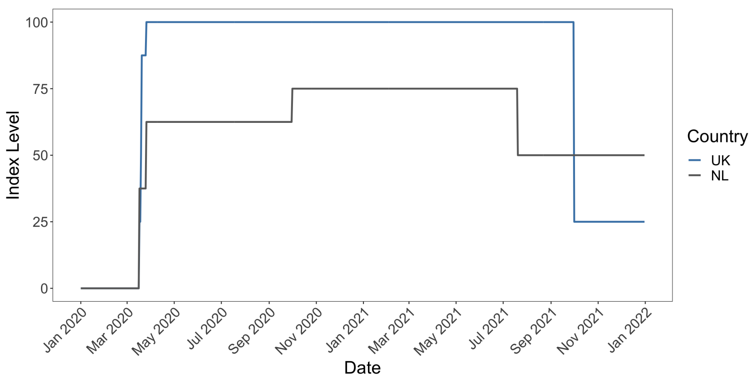

Figure 3.4. Economic Support Index

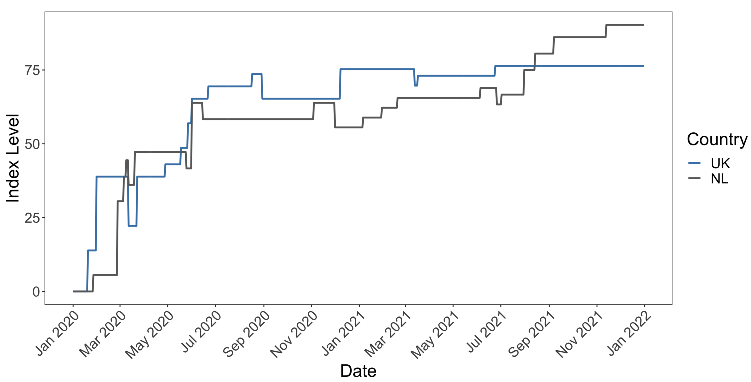

Figure 3.5. Health Policy Index

4 Methodology

Using ABOM, policymakers' opinions towards different topics can be extracted from transcripts and meeting notes. This procedure can be split into three tasks. First, the text is classified into the desired topics. Then the next two tasks take place at a document level, aspect by aspect. The second task is the construction of sentiment scores and the third is emotion recognition. An outline of the required steps is shown in Section 4.1 and each step is elaborated on in the following Subsections.

4.1 Pipeline

An illustration of the NLP methodology is shown in Figure 4.1.

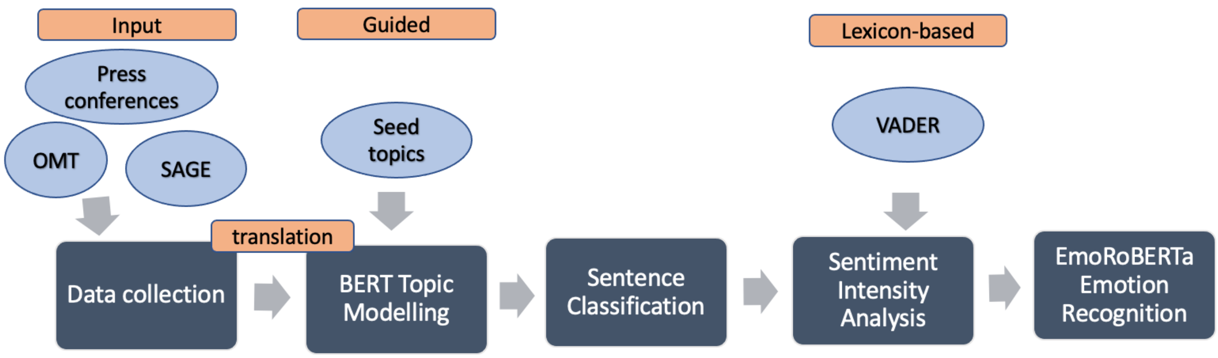

Figure 4.1. Machine Learning Pipeline

First, the text data is collected and time-stamped. The transcripts and meeting notes from the Netherlands are translated from Dutch to English using DeepL in order to use the same lexicon for sentiment analysis. While multi-lingual NLP is possible, a lexicon for policy-related sentiment is not available in Dutch thus the analysis is conducted in English. Next, the text is split into sentences. Using the set of sentences in the press conferences and meeting notes in the UK and NL the BERTopic model extracts topics using guided Machine Learning (see Section 4.3). After the topics have been extracted and the sentences have been classified, unrelated sentences are removed. Then, for each document, the sentences are grouped by topic. Finally, the sentiment analysis model uses a lexicon to calculate polarity and tone is extracted using a transformer based sentence classification model fine-tuned for emotion recognition. This pipeline results in a polarity score and five emotion scores for each topic at the document level. These six scores are hereafter referred to as measures of opinion.

4.2 Text Processing

As mentioned above, the original text is bilingual and translated using DeepL, which is a free neural machine translation (NMT) system. It has been shown to slightly outperform other NMTs when translating idioms from Spanish to English and when translating a script from French to English (Hidalgo-Ternero, 2020; Yulianto & Supriatnaningsih, 2021).

Moreover, the three NLP tasks outlined above make use of the context around the words in order to derive insights. The chosen transformer-based topic and emotion models draw insight from the context a word is used in and thus filtering out stopwords would remove information the model could otherwise take advantage of. The chosen sentiment analysis model is a lexicon-based approach that incorporates semantic and grammatical rules. While it is common to remove punctuation and overly common words (stopwords) in lexicon-based sentiment analysis techniques, this is not appropriate here as punctuation and words like "really" or "but" influence the detected sentiment and its intensity. Thus no stopwords need to be removed in this case. Since the analysis is partially performed on a sentence level, it is important, however, to ensure that sentences can be identified. Punctuation marks are unreliable for this purpose when a text contains references to figures, statistics, ellipses, citations or links. Thus all documents are preprocessed to remove punctuation used in any of the five mentioned patterns. Stopwords are only removed when constructing the topic labels as they would not serve as informative descriptors of topic content (see Section 4.3).

4.3 Content classification

To distinguish between different aspects of policy discussed in the transcripts or meeting notes a topic modeling approach is used. The model used here is based on a Bidirectional Encoder Representation from Transformers (BERT) model. BERT is a pre-trained model that can be fine-tuned for domain or task-specific applications. For example, M. Müller et al. (2023) use COVID-19-related tweets to fine-tune BERT-LARGE and find that their COVID-Twitter-BERT (CT-BERT) outperforms BERT-LARGE in the COVID-19 domain but also in other commonly used classification datasets. BERTopic employs Sentence-BERT (SBERT) which is fine-tuned for detecting similarities across sentences (M. Grootendorst, 2022). SBERT produces numerical representations of sentences called embeddings. These embeddings can then be used in conjunction with dimensionality reduction and clustering techniques to construct topic clusters. The dimensionality reduction algorithm used here is Uniform Manifold Approximation and Projection (UMAP) and it reduces the SBERT embeddings to 5-dimensional vectors (McInnes et al., 2018). Next, clusters are formed using Hierarchical Density-based Spatial Clustering of Applications with Noise (HDBSCAN) with a minimum cluster size of 100 to ensure that a sufficient number of documents contain sentences classified in this topic (McInnes et al., 2017). To construct labels for these clusters a bag-of-words approach is used. Namely, all sentences within a cluster are combined into a single document to construct cluster-based Term Frequency-Inverse

Document Frequency scores (c TF-IDF) which represent word importance for each cluster by taking into account how frequently a given word occurs within a cluster and how infrequently it occurs outside of it (M. Grootendorst, 2022). Stopwords are not removed when the embeddings are produced such that the context of each word is not manipulated, but they can be removed when constructing the c TF-IDF scores to ensure that topics are meaningful and interpretable. Topics are constructed on a sentence level as BERT performs better on short-form text input and requires a large number of observations for effective dimensionality reduction (M. Grootendorst, 2022).

Furthermore, to ensure only policy-relevant topics are derived a hybrid guided approach is used. Guided Topic Modelling allows the researcher to specify a set of topics which the model will attempt to find. If an insufficient number of documents is found for any of these suggested topics no such topic will be identified. Additionally, the number of topics found in the final model is not restricted to the number or suggest topics. Instead the Guided Topic Model will produce document clusters with topic representations that were converged to the suggest topics but no topics were imposed or restricted (M. P. Grootendorst, n.d.). Using the 16 policy indicators as a guide, a list of words representing the types of policies captured by each indicator are used to form these guide topics. For instance, the topic indicator C7 regarding restriction of internal movement is encoded to take the value 1 when travel between regions or cities is advised against and 2 when national movement restrictions are in place. Correspondingly, the words trip, visit, movement and national were used. I choose to use four words for each indicator as too many words risk obscuring the topic and too few may result in imprecise topics. Additionally, after the initial topics are extracted, they are merged manually to obtain 16 topics and one outlier topic. This combination of computational and manual analysis is recommended for the classification of political attitudes by Yarchi et al. (2021). This approach allows the analysis of the aspects most relevant for the research in question, rather than the aspects most common in the documents. The suggested topics are shown in Table 4.1.

Table 4.1. *Guide Topics*

Indicator	Name	Suggestion
C1	School Closures	school, children, parents, education
C2	Workplace Closing	work, business, remote, close
C3	Cancel Public Events	event, cancel, venue, distance
C4	Restrictions on Gatherings	gathering, bubble, group, crowd
C5	Public Transportation	transport, public, bus, train
C6	Stay at Home Order	curfew, stay, home, lockdown
C7	Restrictions on Internal Movement	trip, visit, movement, national
C8	International Travel Controls	international, travel, passport, flight'
E1	Income Support	income, support, cash, money
E2	Debt/Contract Relief for Households	debt, relief, aid, package
H1	Public Information Campaigns	information, campaign, public, health
H2	Testing Policy	test, result, PCR, testing
H3	Contact tracing	contact, trace, source, app
H6	Facial coverings	mask, face, cover, mouth
H7	Vaccination Policy	vaccine, dose, vaccinate, pfizer
H8	Protection of Elderly People	elderly, protect, vulnerable, nursing

4.4 Sentiment Analysis

Sentiment analysis is performed using the Sentiment Intensity Analyzer from the Natural Language Toolkit (NLTK) library in Python. This sentiment analysis method is used with the VADER lexicon that combines valence word scores from a dictionary with heuristic rules (Hutto & Gilbert, 2014). Valence is the degree to which a word conveys positive or negative emotions. The heuristic rules adjust the intensity of a sentiment by taking punctuation, negation, intensification, and capitalization into account (Hutto & Gilbert, 2014). The VADER lexicon is most suited for the analysis of social media text but outperforms other lexicons in other domains as well. This approach is superior to other dictionary-based approaches as it combines word scores with a set of rules that captures information from the surrounding words and grammatical structures.

For each input, this sentiment analyzer returns a compound polarity score which is the sum of the word valence scores after adjusting for heuristic rules normalized to be between -1 and 1. Taking all sentences in a document classified as belonging to a given topic as the input, this compound score is the aspect-based polarity score.

4.5 Emotion Recognition

Next, emotion recognition is performed using EmoRoBERTa (Kamath et al., 2022). This is a BERT-based model fine-tuned for emotion classification using Reddit comments with 28 different labels. While this allows for a more accurate representation of human emotion, this level of granularity is not necessary for the purpose of representing opinions through tone. Thus, I aggregate the resulting probabilities into six groups in order to simplify the analysis. To aggregate these labels the probabilities for the grouped labels are summed. This approach makes the simplifying assumption that each emotion is independent. The six groups and their label are shown in Table 4.2. Note that since the probabilities sum to one across the 28 classes, I focus only on the first five emotions, Affection, Excitement, Uncertainty, Frustration and Sadness.

Table 4.2. *Aggregated Emotion Labels*

Affection	Excitement	Uncertainty	Frustration	Sadness	Neutral
Admiration	Amusement	Realization	Annoyance	Sadness	Neutral
Caring	Curiosity	Surprise	Anger	Remorse	
Gratitude	Excitement	Nervousness	Disappointment	Grief	
Love	Optimism	Confusion	Disapproval	Embarrassment	
Approval		Fear	Disgust,		
Joy					
Pride					
Relief					
Desire					

4.6 Testing for Convergence

To determine to what extent the opinions of experts converge over time I apply these natural language processing (NLP) techniques described above to transcripts of press conferences in the UK and NL and the meeting notes of the OMT and SAGE. As these countries are engaged in an exchange with a much larger network of countries it is infeasible to derive the average weighted opinion to which the persuasion bias model predicts these opinions should converge. As a proxy, the opinions reflected in the documents published by the OECD are used. To test whether the opinions extracted from each group converge the log t test developed by Phillips and Sul (2007) is used on the extracted measures of opinion separately (sentiment and each of the 5 emotions). To correct for multiple testing I use the Bonferroni-Holm correction procedure (Abdi, 2010).

The test was developed for assessing convergence and transition paths of economies and has been applied to GDP in OECD countries, GDP in US states and cost of living in US cities (Phillips & Sul, 2007, 2009). Phillips and Sul (2007) use the ratio of two series X_{it} and X_{jt} to define convergence as

$$\lim_{k\to\infty} \frac{X_{it+k}}{X_{jt+k}} = 1. \tag{4.1}$$

This condition can also be expressed based on the factor loadings δ_{it} using a time-varying factor representation of X_{it} as $\delta_{it}\mu_t$, where μ_t is the common time-varying factor.

$$\lim_{k\to\infty} \delta_{it} = \delta.\forall i. \tag{4.2}$$

Under the condition that $\delta_{it} = \delta$ the cross-sectional variance ratio H_1/H_t while diverge to ∞,where H_t is a scalar defined as

$$H_t = N^{-1}\sum_{i=1}^{N}(h_{it}-1)^2,\ h_{it} = \frac{X_{it}}{N^{-1}\sum_{i=1}^{N} X_{it}}, \tag{4.3}$$

$i = 1, ..., N$ and $t = 1, ..., T$. Thus, to test whether the series converge the logarithmic regression shown in Equation (4.4) is estimated using ordinary least squares (OLS) estimation. In the logarithmic regression,

$$\log\frac{H_1}{H_t} - 2\log\log(t) = \alpha + \beta\log t + \epsilon_t,\ t = [rT], [rT]+1, ...T,\ r > 0, \tag{4.4}$$

the dependent variable is specified so that it will increase to ∞ under the null hypothesis of convergence and it will decrease to $-\infty$ under the alternative. The regression starts at $t = [rT]$ to focus on long run convergence. Monte Carlo simulations suggest that $r = 0.2$ is optimal considering the size and power of the test for moderately sized samples with $T > 100$(Du, 2017). The coefficient of interest, β, can be interpreted as the convergence rate. If it is larger than 2 it suggests the series converge in their level, and if it is positive but smaller than 2 it suggests that there is conditional convergence, meaning the series converge in their first differences. After constructing the T dimensional cross-sectional variances, the null hypothesis can be tested using a one-sided t-test of H_0': $\beta \geq 0$. This t-test is made robust by using a heteroskedasticity and autocorrelation consistent (HAC) estimation method for the standard errors, namely fixed quadratic spectral bandwith estimation (Sichera, Pizzuto, et al., 2019). Thus the complete null hypothesis H_0 is $\delta_{it} = \delta$ & $\beta \geq 0$. OLS estimate of β is consistent as $N \to \infty$ and $T \to \infty$ To perform this test, daily observations are imputed using a simple moving average, where missing values are calculated as the average of all observations within an optimal window (Moritz & Bartz-Beielstein, 2017). This avoids imposing a linear relationship and excessive sensitivity to outliers.

If the null hypothesis is rejected for the full sample of units N, a four-step recursive clustering algorithm developed by Phillips and Sul (2009) is used to find so-called 'convergence clubs', or sub-groups of converging units. In this algorithm, the units are first ordered according to their last observation. Second, the test for convergence is repeated on the subgroup of $2 < k < N$ units, iteratively dropping units until the test is not rejected to find a core group of converging units. Third, the test is repeated, adding each remaining unit to this core group one by one to detect any missing members of the convergence club. If the test statistic is larger than the critical value corresponding to a 5% significance level the unit is added to the club. Finally, If any units remain that were not added, the procedure is repeated to find a second group of converging units. Note, no correction for multiple testing is performed for these recursive tests. If no further group of converging units is found the algorithm is stopped and the remaining units are classified as diverging.

Next, to determine to what extent the opinions of experts converge with political action, I compare the aspect-based sentiment and tone scores from the meeting notes to those of the press conference transcripts for each country. Using the same test as outlined above, I compare the opinion in press conferences and meeting notes for the UK and NL.

To gain insight into the convergence of opinion to political action and its influence on opinion the opinion measures are compared to the policy indicators. To do so the ordinal scores are transformed to the polarity scale (-1-1) or the probability scale (0-1). In the former case the scores are inverted such that an increase in policy stringency is reflected as a decrease in polarity.

4.7 Factor Analysis

The third research question, which focuses on the dimensionality of each group's policy advice, can be answered using factor analysis. Persuasion bias suggests that long-term differences in opinions will converge to being either to the 'right' or 'left' of the average. Using the OECD sentiment as a measure of the average opinion in the network of developed nations, a relative difference of opinion can be constructed for each group's aspect-based opinion measures. This difference is defined as

$$d^t_{ikl} = \frac{x^t_{ikl} - x^t_{OECD,kl}}{N^{-1}\sum_j ||\mathbf{x}^t_j - \mathbf{x}^t_{OECD,l}||}, \tag{4.5}$$

where the subscript i refers to the group ($i = NL, UK, SAGE, WHO$), k refers to the relevant indicator ($k = 1, ..., 16$), l refers to the opinion measure ($l = 1, ..., 6$),and t refers to the period of exchange ($t = 1, ..., T$) (DeMarzo et al., 2003).

To assess the dimensionality of opinions, PCA is performed to assess what proportion of variance in the relative differences of opinion can be explained by by each each factor as done by Page (2020). To analyze how this dimensionality evolves over time PCA is performed on the differences in opinion measures for each time period. According to the Kaiser-Gutman criterion, all principal components with an eigenvalue above one should be retained (Kaiser, 1960). Additionally, a likelihood ratio test is performed to test the null hypotheses that one or two factors are sufficient to explain the variation in the differences in opinion at time t (Revelle & Revelle, 2015). To give the unidimensionality hypothesis the best chance, this test is performed at the end of the period because opinions should become closer and more unidimensional over time. Given the uncertainty and crisis conditions during the pandemic, however, it is likely that obstacles to convergence and dimensionality reduction arise. Thus, in addition to the last data in the observation period, 31st of March 20201, the date for which unidimensionality is most likely is also considered. That is, for each opinion measure the date on which the proportion of variance in the relative differences in opinion is maximized is also used.

4.8 Sensitivity Analysis

To assess sensitivity to the chosen sentiment analysis method I repeat the analysis with a sentiment dictionary designed for political texts (Lexicoder). Alternative imputation techniques, such as exponentially weighted moving average, Kalman smoothing, linear interpolation and last observation carried forward can also be used to impute missing values in the time series (Moritz & Bartz-Beielstein, 2017). I assess sensitivity to the chosen method of imputation by comparing the results across moving average imputation, linear imputation and last observation carried forward imputation. Additionally, I assess the sensitivity to the topic labels provided in the guided topic modeling procedure by using less precise labels, consisting of two words for each suggested topic instead of four (see Table 4.3).

Table 4.3. *Less precise topic labels*

Indicator	Name	Suggestion
C1	School Closures	school, children
C2	Workplace Closing	work, remote
C3	Cancel Public Events	eventl, venue
C4	Restrictions on Gatherings	gathering, bubble
C5	Public Transportation	transport, public
C6	Stay at Home Order	stay, home,
C7	Restrictions on Internal Movement	tri, movement,
C8	International Travel Controls	international, travel
E1	Income Support	income, support
E2	Debt/Contract Relief for Households	debt, relief
H1	Public Information Campaigns	information, campaign
H2	Testing Policy	test, result
H3	Contact tracing	contact, trace
H6	Facial coverings	mask, face
H7	Vaccination Policy	vaccine, dose
H8	Protection of Elderly People	elderly, protect

5 Results

In the following Section, I discuss the results of the analysis outlined above. I begin by characterizing the discovered topic structure. Next, I compare the opinions across groups and then compare opinions between experts and politicians within the Netherlands and the UK respectively. Then, I will discuss the dimensionality of opinions over time. Finally, I discuss the sensitivity of the results to the chosen dictionary and emotion model, the method of imputation and the guide topics.

5.1 Topic structure

Using the meeting notes and speech transcripts from the Netherlands and the UK, 16 Topics are constructed. The number of sentences found for each topic by group can be found in Table A.10. Each topic is linked to the 16 policy indicators based on the similarities of the topic label and the guided topics shown in Table 4.1. The topics are ranked based on the number of sentences within each topic cluster when considering all groups. That is, the topic ranked 0 is the most frequently discussed topic overall, and the topic ranked 15 is the topic least discussed across all documents.

Table 5.1. *Topic label, rank, and frequency by group*

	Topic Label	Rank	Number of Sentences					
			WHO	OECD	SAGE	OMT	UK	NL
C1	schools school education children	2	451	2024	193	264	225	829
C2	industry hospitality entrepreneurs economy	7	233	2447	12	23	122	672
C3	half metres distance metre	4	172	106	8	48	92	805
C4	group groups people crowd	9	181	56	25	44	31	375
C5	transport public bus trains	13	145	306	12	24	129	169
C6	home curfew lockdown stay	3	328	474	47	82	425	819
C7	travel visit movements advice	10	292	624	12	45	24	257
C8	quarantine orange flights travel	8	344	410	50	158	51	408
E1	money income support tax	14	64	2075	5	6	41	142
E2	debt aid support loans	15	73	1940	0	0	4	69
H1	hands public wash washing	11	811	1136	90	57	54	205
H2	testing test tests coronavirus	1	1513	1255	156	663	739	682
H3	app contact source tracing	12	252	235	53	34	35	227
H6	mouth masks caps face	5	574	536	86	178	187	409
H7	vaccine vaccination vaccines vaccinated	0	5857	1226	385	428	564	1304
H8	nursing elderly homes care	6	224	342	82	179	119	393
Total Number of Sentences			68978	63304	6395	8144	8039	47058
Total Number of Documents			149	213	81	57	138	91

Note. Rank ranges from 0 to 15, with 0 being the largest group and 15 being the smallest.

The number of sentences in each topic varies across the groups according to how relevant or controversial the topic is in each case. For example, the topic linked to the policy on school closures (C1) is the third largest cluster of sentences overall, it contains 2 024 sentences in the OECD documents and 451 in the WHO documents. In contrast, there are 5 857 sentences in the vaccination policy related topic (H7) for the WHO and only 1 226 for the OECD. Indeed, the most commonly discussed topic is vaccination followed by testing (H2). Word clouds containing the most salient words for each topic can be found in Appendix Section A.2.1.

For most topics, a sufficient number of sentences is found, however, for the UK less than 10 sentences were found for the topic related to debt relief (E2). Economic policies in general (E1 and E2) were not frequently discussed in the OMT or SAGE meetings. Another cluster with less than 10 sentences corresponds to policies on public events (C3) in the SAGE meeting notes. As these observations represent noisy signals of opinion, no inferences can be made regarding these topics for SAGE, OMT and the UK. Indeed, as the topic related to debt relief policy (E2) is empty or contains less than five sentences for three of the fours groups in focus, this topic will be removed from the following analysis.

The heatmap in Figure 5.1 shows a comparison of the 16 topics. Here the topics are identified in the following format: '*r_label*' where r is the rank and label consists of the topic label words separate by an underscore (see Table A.10). The topics are ordered by rank and each box in the 16 by 16 square corresponds to the comparison of the topic listed on the left and the bottom. The comparison uses the cosine similarity between the clusters of sentence embeddings. On the diagonal topics are compared to themselves and thus have a similarity score of 1, as shown by the dark blue boxes. The lighter the colour of the box the lower the similarity between the two groups.

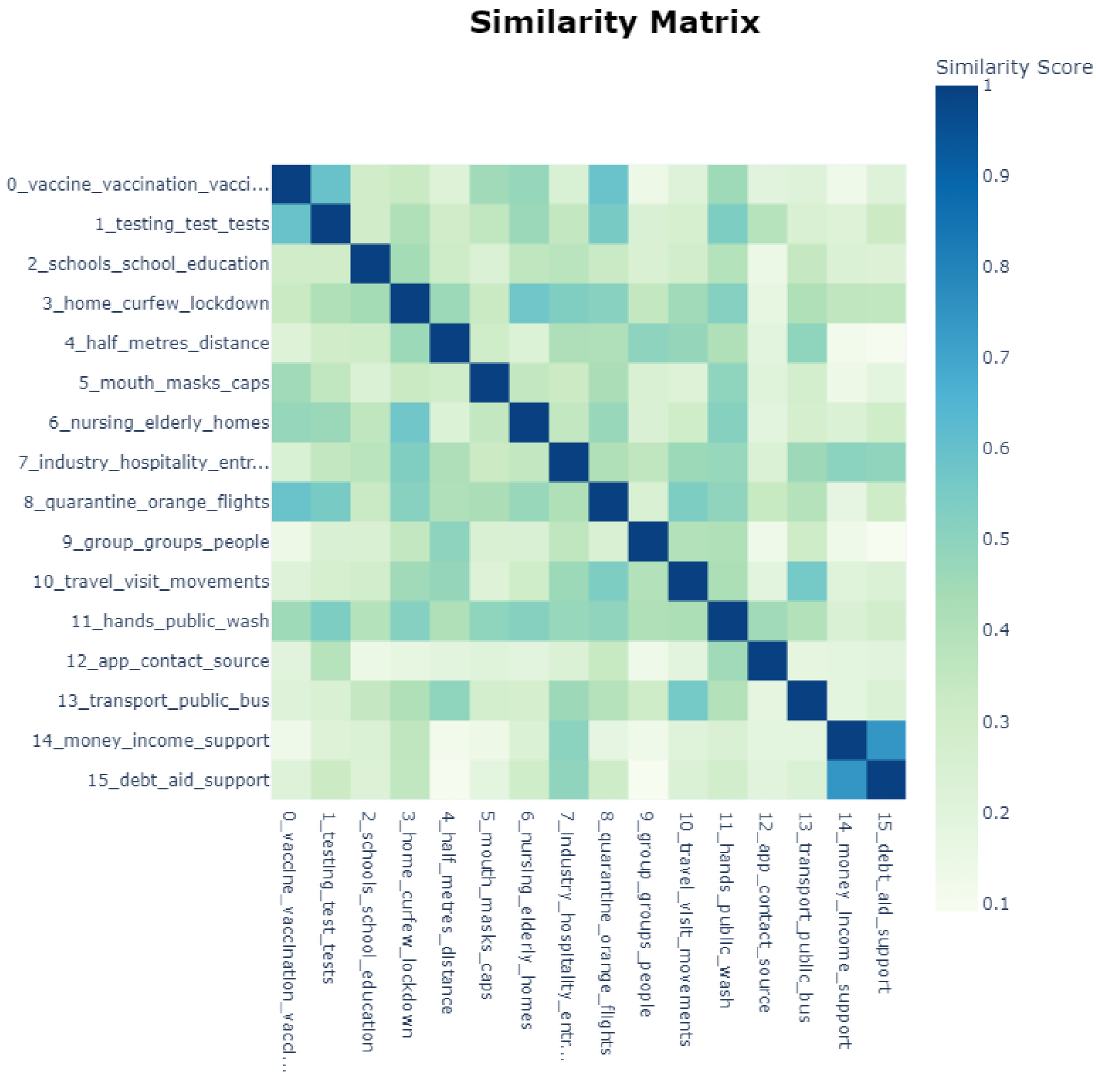

Figure 5.1. Similarity Matrix

When comparing the constructed topics it is clear that topics ranked 14th and 15th corresponding to the policies on income support and debt relief (E1 and E2) respectively are the most similar. Interestingly a similar relationship is absent for the set of containment polices or the set of health policies. Instead, a strong relationship exists for the topics with rank 0 and 1, which focus on vaccination and testing. A similar relationship could be expected between 8th and 10th ranked topics corresponding to internal movement (C7) and international travel (C8) respectively, due to their focus on traveling. In fact, both topic labels contain the word travel. However, these topics receive a similarity score of only around 0.5. Instead, there is a higher similarity score between internal movement and the topic focused on public transport (C5), the 10th and 13th ranked topics respectively. Meanwhile, international travel is closest in similarity to the topics on vaccination and testing, with rank 0 and 1. This exemplifies how the model uses words in context to derive their meaning rather than grouping sentences according to key terms. While the labels are assigned based on the most representative words for each topic, this does not reflect how sentences are classified. This assignment can be illustrated further by two-dimensional representations of the word embeddings of sentences in the documents (see Appendix Section A.2.2).

5.2 Convergence across groups

The estimated convergence rates from the regression used in the log t test are shown in Table 5.2. Positive values are shown in bold to distinguish between estimates that indicate convergence and those that indicate divergence.

Table 5.2. *Estimated convergence rates*

Indicator	Sentiment	Affection	Excitement	Uncertainty	Frustration	Sadness
C1	**0.469**	-0.707	-0.272	-0.297	**0.327**	**0.074**
	(0.493)	(0.245)	(0.277)	(0.312)	(0.241)	(0.168)
C2	-0.490	-0.529	-0.344	-0.119	-0.175	-0.287
	(0.584)	(0.105)***	(0.243)	(0.119)	(0.115)	(0.112)
C3	**2.624**	-0.815	-0.961	-0.385	-0.519	**0.083**
	(0.593)	(0.233)	(0.253)*	(0.322)	(0.317)	(0.576)
C4	-2.911	-0.669	-0.771	-0.724	-0.593	-0.049
	(0.779)*	(0.385)	(0.432)	(0.252)	(0.192)	(0.459)
C5	**0.703**	-0.075	**0.006**	-1.125	-0.030	**0.063**
	(0.632)	(0.459)	(0.312)	(0.189)***	(0.220)	(0.251)
C6	-0.178	-0.528	-1.140	-0.162	-0.359	-0.628
	(0.681)	(0.219)	(0.302)*	(0.273)	(0.212)	(0.229)
C7	-0.337	-0.609	-0.290	-0.124	-0.435	-0.698
	(0.459)	(0.261)	(0.145)	(0.176)	(0.281)	(0.260)
C8	**0.462**	**0.071**	-0.672	-0.360	-0.247	-0.154
	(0.865)	(0.340)	(0.268)	(0.181)	(0.394)	(0.370)
E1	-0.970	-1.241	**0.356**	-0.264	**0.170**	-0.721
	(0.582)	(0.215)***	(0.299)	(0.271)	(0.222)	(0.130)***
H1	-1.454	-0.800	-0.527	-0.571	-0.732	-0.236
	(0.836)	(0.402)	(0.216)	(0.171)	(0.305)	(0.510)
H2	-0.065	**0.276**	-0.546	**0.134**	**0.012**	**0.625**
	(0.610)	(0.286)	(0.141)*	(0.174)	(0.228)	(0.436)
H3	-1.914	-1.270	-0.27	-0.163	-1.612	-1.519
	(0.936)	(0.174)***	(0.634)	(0.241)	(0.476)	(0.411)*
H6	-2.265	-1.155	-0.923	-0.319	-0.376	**0.869**
	(0.836)	(0.349)	(0.267)	(0.169)	(0.185)	(0.717)
H7	**1.003**	**0.430**	**0.120**	**0.482**	**0.454**	**0.079**
	(0.734)	(0.313)	(0.256)	(0.183)	(0.271)	(0.437)
H8	**1.929**	-1.042	-0.008	-0.918	-0.756	-0.579
	(0.734)	(0.333)	(0.358)	(0.235)*	(0.283)	(0.213)

Note. Standard errors are shown in round brackets below the coefficients. Positive estimates are shown in bold.
Note. If a standard error is labeled *, **, or *** its coefficient is significant at a 5%, 1% or 0.1% level respectively, after applying the Bonferroni-Holm correction (Abdi, 2010).

As discussed previously, an estimated convergence rate of at least 2 suggests level convergence and an estimate between 0 and 2 suggests the rate of change in the series is converging. The only evidence of level convergence is found for sentiment towards policies focusing on public events

(C3). The polarity of sentiment regarding this topic over time is shown in Figure 5.2. Here, SAGE and UK are not shown due to a lack of observations. For SAGE only eight sentences are clustered in this topic (see Table A.10) and for the UK it was last discussed on the 20th of October 2020, thus no inference can be made based on the imputed values. Due to the imputation of missing values the polarity has a higher variance in periods where it is often discussed as opposed to periods with a lot of missing values.

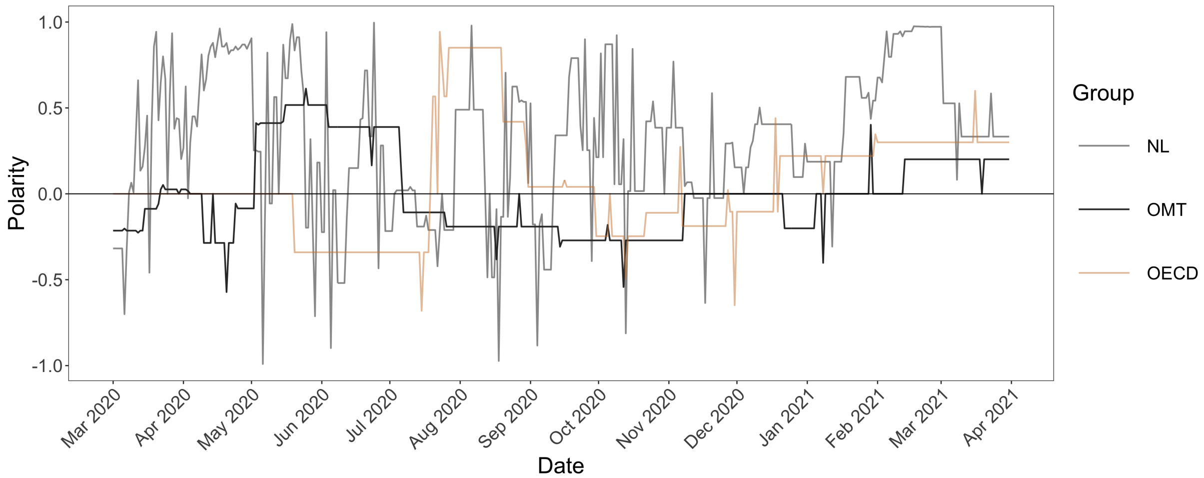

Figure 5.2. Sentiment on the cancellation of public events (C3) in press conferences

Despite initial differences in the curves the series do seem to converge to the OECD opinion over time. The sentiment in NL and OMT begin to follow the OECD sentiment more closely from December and October 2020 onwards respectively. The next largest estimated convergence rates are found for sentiment regarding vaccination policy (H7) and regarding the protection of elderly people (H8). These values do not exceed two thus they suggest conditional convergence, that is rather than a convergence in the level of opinion, groups converge in the rates of change in opinion. The sentiment related to vaccination policy (H7) is shown in Figure**??** for the UK and NL and in **??** for SAGE and OMT.

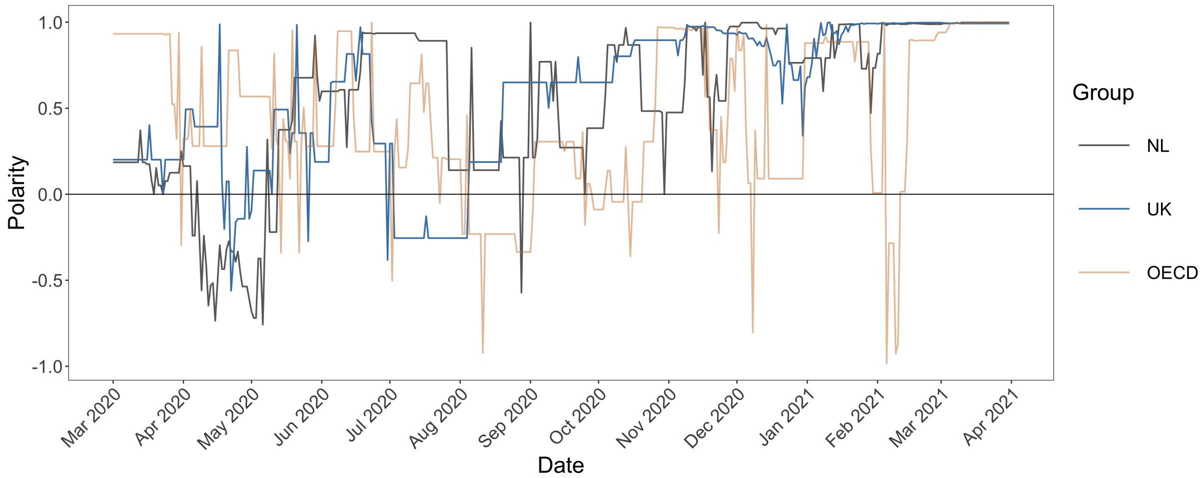

Figure 5.3. Sentiment regarding vaccination policy (H7) in press conferences

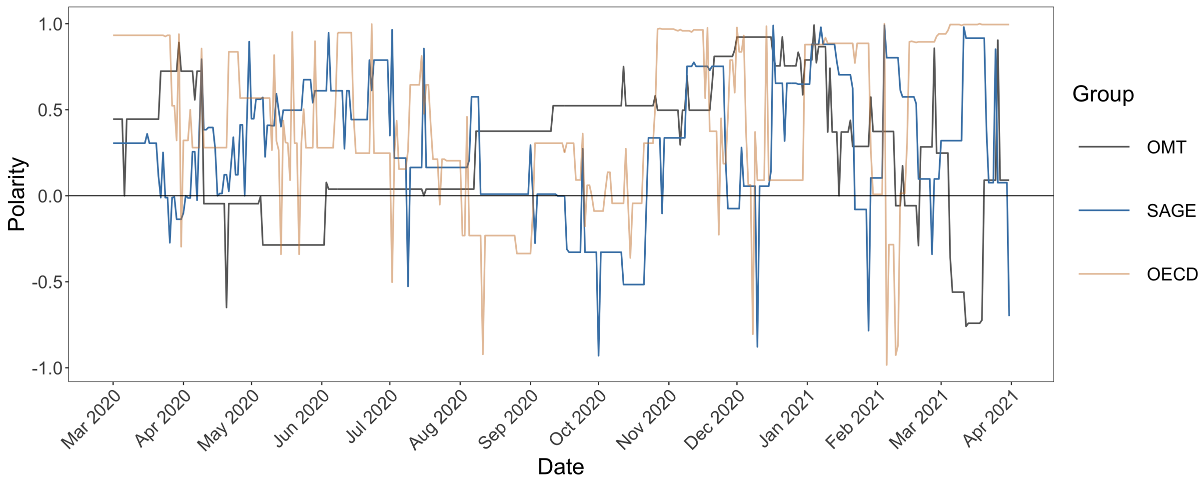

Figure 5.4. Sentiment regarding vaccination policy (H7) in meeting notes

While the groups shown in Figure **??** seem to converge in level to a very positive stance, the sentiment expressed in meeting notes from the OMT and SAGE is much more negative and does not seem to follow a similar rate of change. Infact, there seems to be an increasing trend in the polarity expressed in press conferences as expressed by the OECD from Octover 2020 onwards. A similar trend can be discerned from the meeting notes but in January 2021 sentiment expressed by the OMT and SAGE becomes much more polarizing in that it is either very positive or very negative.

Note, even when level convergence is indicated for one measure of opinion on an aspect, the sign of the coefficient often differs across measures. For instance, despite finding evidence of level convergence for sentiment on policy regarding public events, the null hypothesis of convergence is rejected for excitement expressed toward this policy. Indeed, the only policy for which the estimated convergence rates are consistently positive is vaccination policy (H7). Policies for which divergence is indicated through consistently negative estimates are regarding workplace closures (C2), restrictions on public gatherings (C4), stay at home orders (C6), restrictions on internal movement (C7), public information campaigns (H1), and contact and tracing (H3). While many of the estimates are negative the null hypothesis of convergence is only rejected for 11 combinations of aspect and opinion measures. In cases where the estimated convergence rate is significantly less than zero, an attempt to find a subgroup for which opinions converge is made. The resulting convergent clubs and diverging groups are shown in Table 5.3.

Table 5.3. *Convergence Clubs*

Indicator	Measure	Club 1	Club 2	Divergent Groups
C2	Affection	OMT, NL, OECD, UK		SAGE
C3	Excitement	SAGE, OMT, UK, OECD		NL
C4	Sentiment	UK, SAGE, NL		OECD, OMT
C5	Uncertainty	UK, SAGE, NL		OMT, OECD
C6	Excitement	UK, SAGE, OECD		OMT, NL
E1	Affection	SAGE, OECD		UK, OMT, NL
	Sadness	OECD,SAGE, UK		OMT, NL
H2	Excitement	UK, NL, OMT	OECD, SAGE	
H3	Affection	NL, UK	SAGE, OECD, OMT	
	Sadness	SAGE, OMT , NL, OECD		UK
H8	Uncertainty	OMT, NL, SAGE, OECD, UK		

In most cases the convergence clubs identify one to two groups as diverging from the rest. In one exception three groups are identified as divergent. Policymakers and experts can be grouped together, for instance for excitement regarding the stay at home order (C6) or sadness regarding income support policy (E1) where the UK, SAGE and OECD opinion converge. For Affection regarding contact and trace policy (H3) the groups are split into two converging clubs, with opinions from press conferences in one group and the opinions from meeting notes in the other. For uncertainty regarding policy on protecting the elderly (H8), convergence is rejected for the full sample but no groups are removed from the convergence club.

Two interesting cases to highlight from this Table are the clubs found for excitement regarding the cancellation of public events (C3) (see Figure **??**) and affection regarding contact and trace policy (H3) (see Figure **??**). For excitement regarding the cancellation of public events, NL is identified as a diverging group while the rest form a convergence club. The two spikes in the probability of excitement being expressed in Dutch press conferences coincide with the strengthening of lockdown measures in November 2020 and the lifting of restrictions in April 2021.

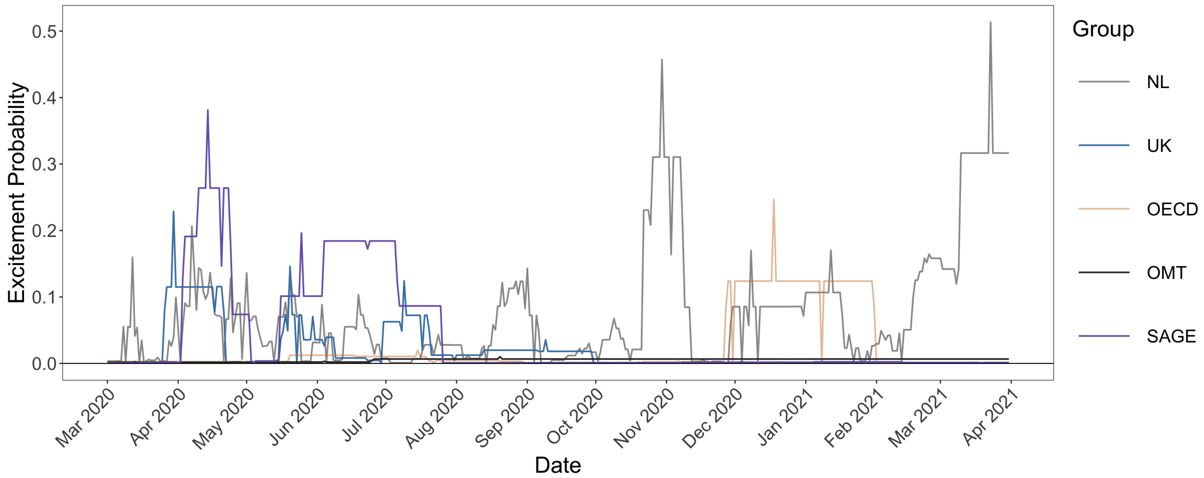

Figure 5.5. Excitement regarding the cancellation of public events (C3)

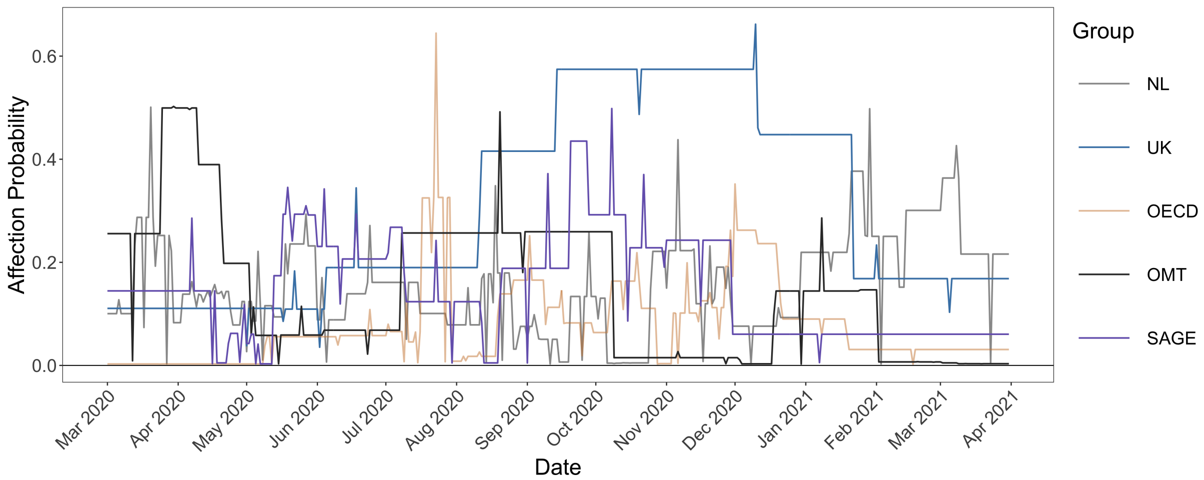

Figure 5.6. Affection regarding contact and trace policy (C3)

For affection regarding contact and trace policy two convergence clubs are formed. The first contains NL and UK, that is affection expressed in press conferences, the second contains SAGE, OECD, and OMT, that is affection expressed in meeting notes or by the OECD. In the figure one can observe that towards the end of the period the probability that affection is expressed in NL and UK converges to a higher level than that in the other convergence club. Thus this would suggest that affection was expressed more in press conferences than in meeting notes.

Finally, the most negative estimates of convergence rates are found for sentiment regarding restrictions on public gatherings (C4). In this case, as shown in Table 5.3, OMT and OECD are identified as diverging groups.

To summarise, level convergence is found for sentiment regarding policy on public events. Conditional convergence is found consistently across all dimensions of opinion for vaccination

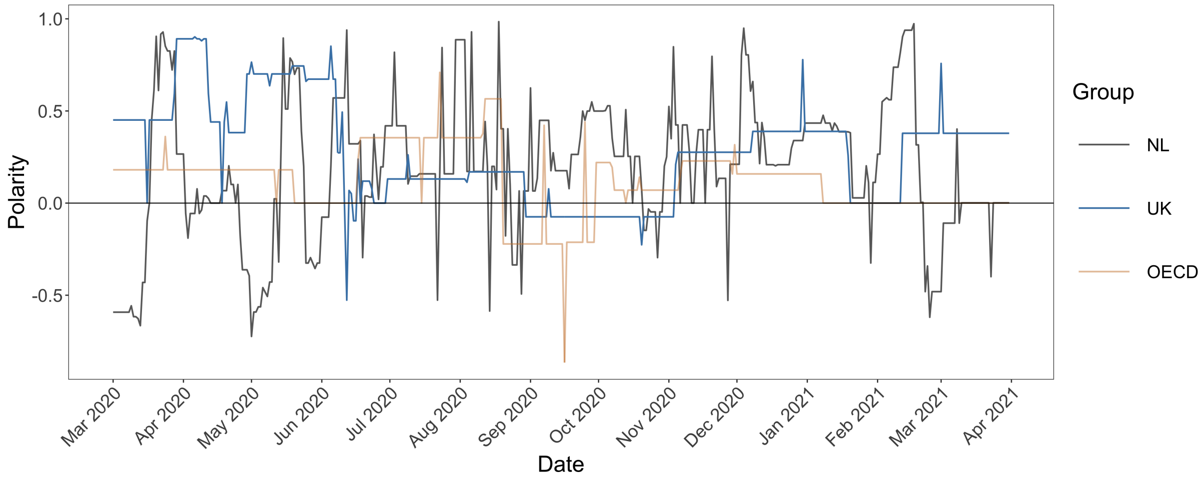

Figure 5.7. Sentiment regarding restriction on public gatherings (C4) in press conferences

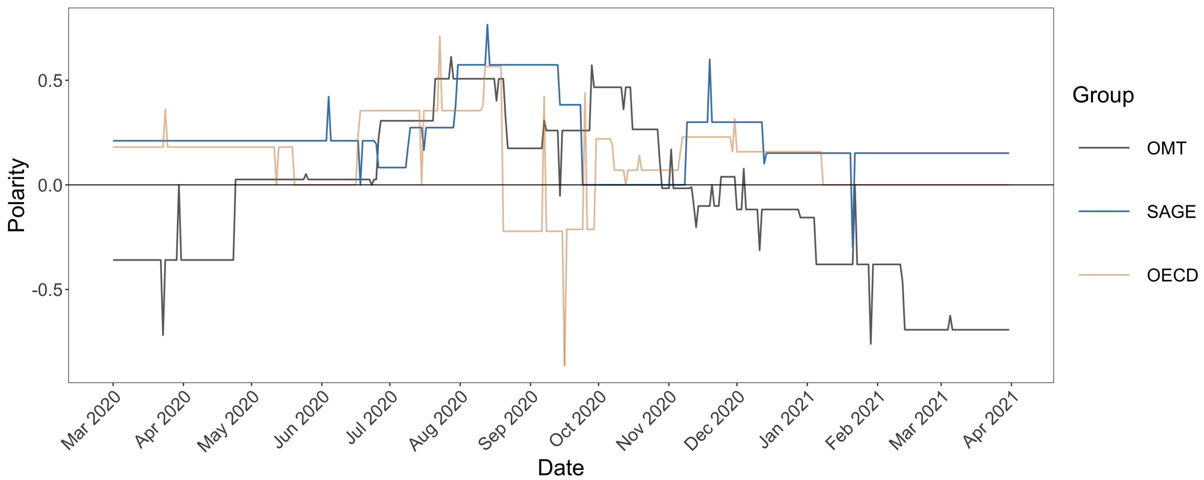

Figure 5.8. Sentiment regarding restriction on public gatherings (C4) in meeting notes

policy. Meanwhile, more divisive aspects of pandemic policy are regarding workplace closures (C2), restrictions on public gatherings (C4), stay at home orders (C6), restrictions on internal movement (C7), public information campaigns (H1), and contact and tracing (H3). The allocation of convergence clubs indicate that SAGE affection for policies regarding workplace closures diverges from the other groups. NL excitement is identified as diverging from the other group regarding policies for public events and UK sadness diverges form the rest on contact and trace policies. Convergence clubs distinguish between affection expressed in meeting notes and press conferences for contact and trace policy.

5.3 Convergence between experts and politicians

To further assess opinion formation, the convergence between policymakers and experts is assessed as well. The resulting estimated convergence rates from the log t test applied to the subgroups NL and OMT and UK and SAGE are shown in Table 5.4, for the full values and significance levels see Table A.12. Here I simplify the table to show only the sign and magnitude of the estimated convergence rates. Positive estimates are shown as a plus and negative ones as a minus. When the absolute value of the estimate exceeds one two symbols are displayed and when it exceeds 2 three are shown. Thus all estimates that suggest level convergence are shown by three plus signs.

Table 5.4. *Sign of the estimated convergence rates*

Indicator	Subgroup	Sentiment	Affection	Excitement	Uncertainty	Frustration	Sadness
C1	NL, OMT	+++	-	-	-	+	- -
	UK, SAGE	**- - -**	- -	- -	-	+	- -
C2	NL , OMT	++	-	- -	-	-	-
	UK, SAGE	- -	**-**	+	- -	**- - -**	- -
C3	NL , OMT	-	- -	- -	-	-	**- -**
	UK , SAGE	+++	- -	-	-	**- -**	+++
C4	NL , OMT	- - -	-	- -	- -	+	+
	UK , SAGE	-	+	+	++	+	+++
C5	NL , OMT	+	+	+++	- -	-	+
	UK , SAGE	++	+++	-	+++	**- -**	-
C6	NL , OMT	-	+	- -	++	-	-
	UK , SAGE	+	**- -**	+	-	-	-
C7	NL , OMT	+	+	-	++	+	-
	UK , SAGE	-	++	**- -**	-	**- -**	**- -**
C8	NL , OMT	-	++	-	-	-	++
	UK , SAGE	++	+	-	- -	**- - -**	-
E1	NL , OMT	-	- - -	-	+	+	**-**
	UK , SAGE	-	**- - -**	+++	+	-	+
H1	NL , OMT	+	- -	- -	-	-	+
	UK , SAGE	**- - -**	- -	+	+	-	-
H2	NL , OMT	++	-	-	+	+	+
	UK , SAGE	-	++	-	+	++	++
H3	NL , OMT	+	- -	+	-	+++	+++
	UK , SAGE	- - -	**- - -**	- -	**- - -**	+++	**- - -**
H6	NL , OMT	-	-	-	-	- -	++
	UK , SAGE	**- -**	- -	**- - -**	-	**- - -**	+
H7	NL , OMT	++	++	-	++	+	-
	UK , SAGE	++	+	-	+	-	-
H8	NL , OMT	+++	+	+	+	+	- -
	UK , SAGE	+++	- - -	+	- -	-	+

Note. + stands for a positive estimate, - stands for a negative estimate. Positive estimates above 1 are shown as ++, positive estimates above 2 as +++. The same applies to negative estimates below 1 or below 2. Estimates that are significantly less than 0 are shown in bold.

Level convergence is found for 13 cases. For instance, in the sentiment regarding the protection of elderly people (H8) for both comparisons. For sentiment on vaccination policy (H7), only conditional convergence is found. Interestingly, the only opinion for which the convergence rate estimates are consistently negative are for policies on public events when comparing NL and OMT. This is in contrast to the level convergence found for this policy for sentiment in Section 5.2. Convergence is rejected for affection regarding contact and trace policy for UK and

SAGE, confirming the convergent clubs for this indicator. Similarly, the estimated coefficients confirm the convergence clubs found previously for policies on contact and trace (H3), income support (E1), public transportation (C5) and the stay at home order (C6).

In fact, the null hypothesis of convergence is rejected for 20 cases, only 2 of which are for the comparison between NL and OMT. Thus, there is more evidence of diverging opinions between UK policymakers and their advisors than in NL.

When looking at the estimated signs across comparisons, it is shown that policy regarding face masks (H6), is the only one for which the sign matches across all measures of opinion. This indicates that there are similar listening structures in place between experts and policymakers when it comes to policy on face masks.

5.4 Dimensionality of opinions

The result of the test for whether one or two factors are sufficient to explain the variation in the constructed relative differences of opinions is shown in Table 5.5.

Table 5.5. *Sufficient number of Factors*

	Chi-Sqaure test statistic		First principal component	
	One Factor Sufficient	Two Factors Sufficient	Eigenvalue	Prop. Var. Explained
Sentiment				
31-03-21	5.145	0.942	6.786	45.239
13-04-20	8.367	2.713	8.809	58.726
Affection				
31-03-21	5.427	0.345	6.127	40.845
11-04-20	11.737	0.0126	8.545	56.966
Excitement				
31-03-21	2.549	0.578	6.637	44.250
30-09-20	6.669	1.537	10.306	68.704
Uncertainty				
31-03-21	4.569	1.0160	7.379	49.195
06-05-20	10.260	0.295	8.631	57.543
Frustration				
31-03-21	0.584	0.009	5.390	35.930
02-06-20	1.051	0.0773	10.600	70.669
Sadness				
31-03-21	39.265***	11.307	6.044	40.295
17-03-20	3.022	0.346	9.680	64.532

Note. Prop. Var. Explained stands for Proportion of Variance Explained.
Note. If a test statistic is labeled *, **, or *** its coefficient is significant at a 5%, 1% or 0.1% level respectively, after applying the Bonferroni-Holm correction (Abdi, 2010).

When the first test statistic is not rejected it is an indication that one factor could be sufficient to explain the variance. However, given that the proportion of variance explained by the first principal component for all measures of opinion at the end of the period is less than 50%, the factor does not adequately capture the variation in the relative differences. Interestingly, the proportion of variance explained by the first principal component is highest for frustration in June of 2020, followed by excitement in September of 2020 and sadness, after only 17 days in March 2020.

The eigenvalues for the principal components derived from the relative differences in the six opinion measures are shown over time in Figures 5.9-5.14. The plots include a horizontal line indicating when the eigenvalues drop below one. Using the Kaiser-Guttman criterion, drop in the eigenvalue below one indicates a reduction in the required number of factors, and thus a reduction in dimensionality (Kaiser, 1960).

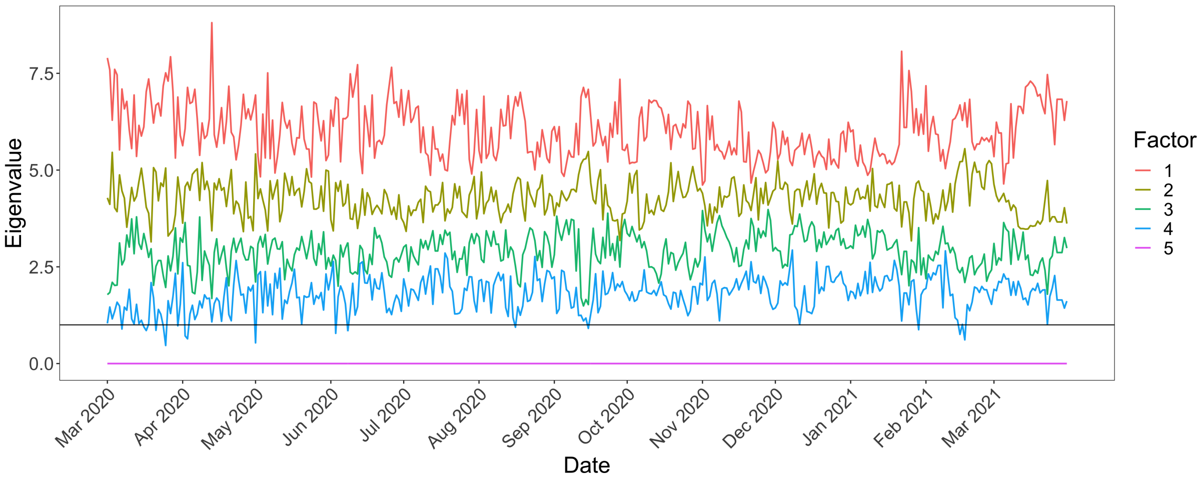

Figure 5.9. Eigenvalues of Principal Components over Time for relative differences in Sentiment

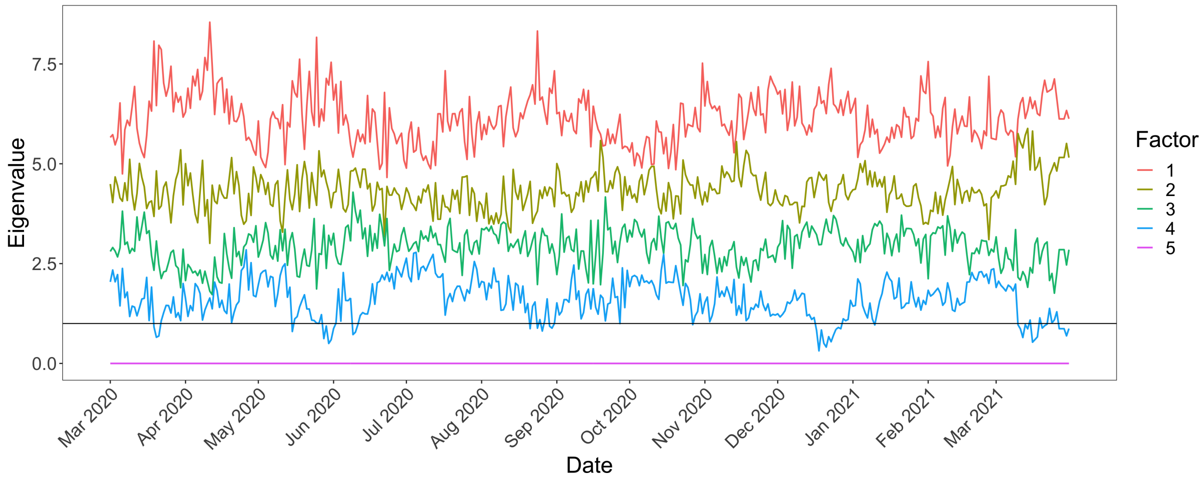

Figure 5.10. Eigenvalues of Principal Components over Time for relative differences in Affection

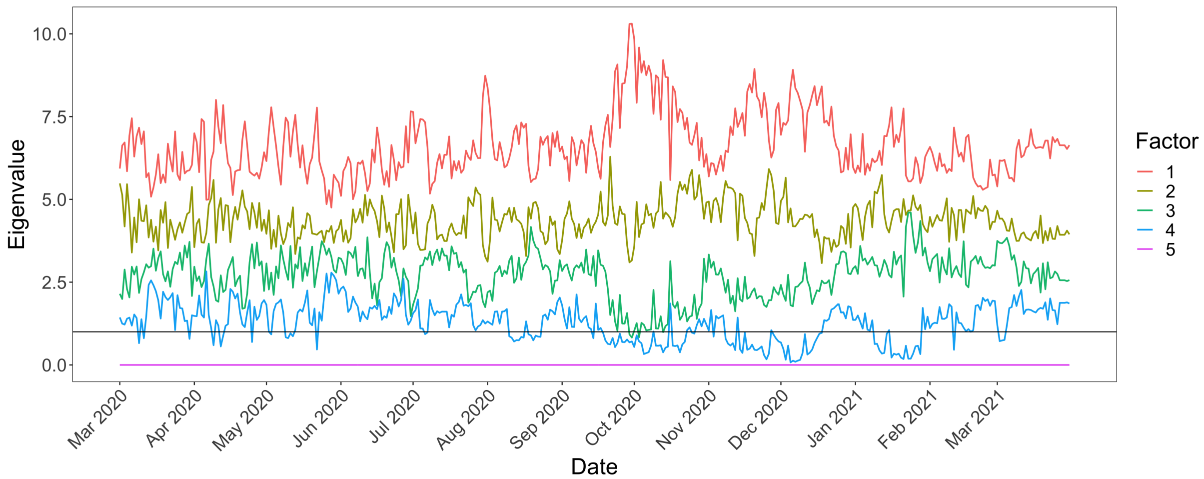

Figure 5.11. Eigenvalues of Principal Components over Time for relative differences in Excitement

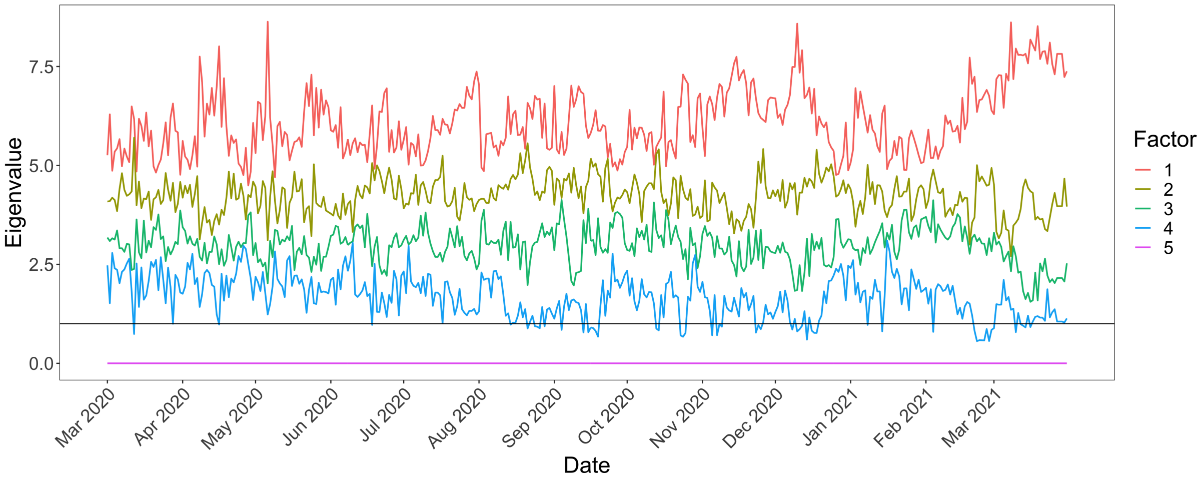

Figure 5.12. Eigenvalues of Principal Components over Time for relative differences in Uncertainty

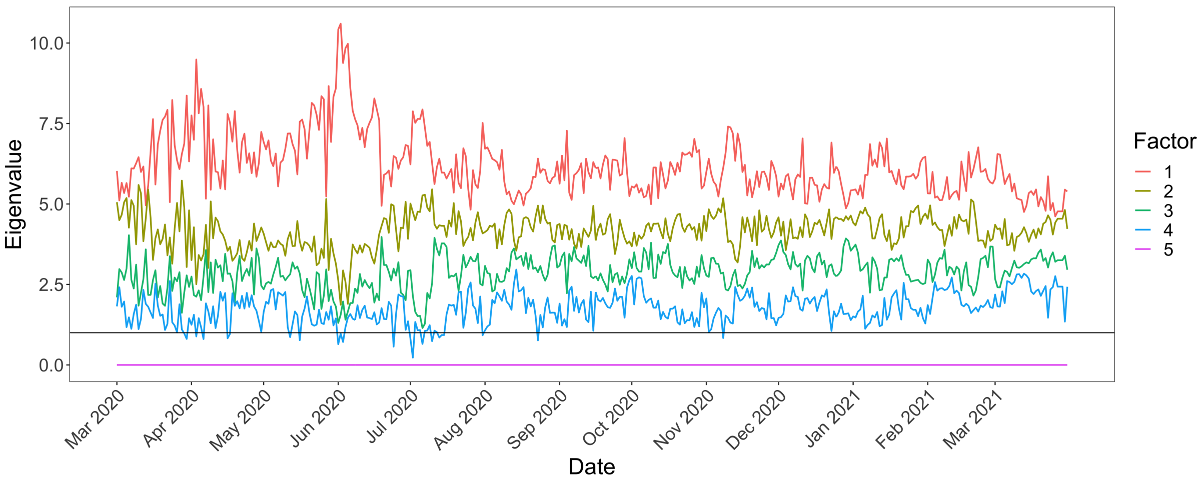

Figure 5.13. Eigenvalues of Principal Components over Time for relative differences in Frustration

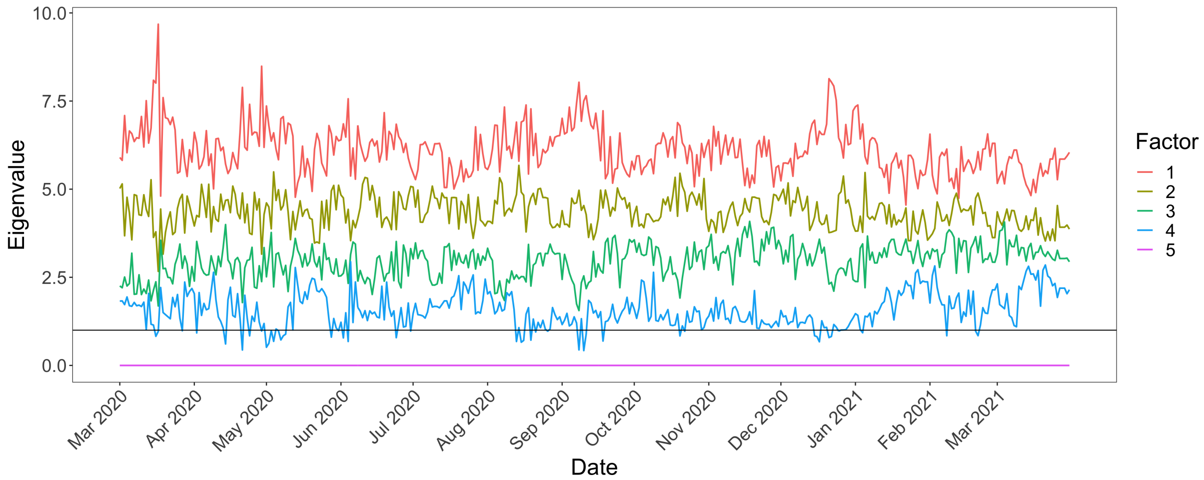

Figure 5.14. Eigenvalues of Principal Components over Time for relative differences in Sadness

Although the Kaiser-Guttman criterion is susceptible to selecting too few components when the number of variables is small, as in this case, it still suggests that 3-4 principal components are sufficient to explain the variation in the relative differences in opinions (Streiner, 1998). For excitement, there is one exception where only 2 components have eigenvalues larger than one. Indeed, the eigenvalues for excitement suggest that 2 principal component would have been sufficient to explain the variation in relative differences in excitement at the end of September, which is shortly before the reintroduction of lockdown measures. Similarly, the dips in the eigenvalues for affection coincide with the introduction and lifting of lockdowns or containment measures. However neither figure suggests any trend towards unidimensionality, that is a decreasing trend in the eigenvalues of principal component 2-5 and an increasing trend in the eigenvalue of the first. Such a trend can be observed in Sentiment in March 2021 and in Uncertainty from February 2021 onwards.

Increased unidimensionality can also be seen in the proportion of relative differences that fall above or below zero. The proportion of relative differences that is positive for each group and opinion measure is shown in Table 5.6. Here the further the proportion is from 0.5 the more unidimensional opinions are. A proportion lower than 0.5 indicates that the majority of relative differences fall below the OECD opinion for that measure and a proportion higher than 0.5 indicates the majority falls above. For sentiment this reflects a more positive stance across topics, while for the rest of the emotion measures it reflects a higher intensity in the emotion.

Table 5.6. *Proportion of positive relative differences at the end of the observation period by group*

	NL	UK	OMT	SAGE	WHO
Sentiment	0.688	0.563	0.333	0.467	0.500
Affection	0.563	0.733	0.533	0.625	0.875
Excitement	0.800	0.800	0.688	0.563	0.813
Uncertainty	0.533	0.375	0.688	0.563	0.600
Frustration	0.500	0.250	0.438	0.533	0.267
Sadness	0.250	0.250	0.333	0.333	0.500

Note. Prop. Var. Explained stands for Proportion of Variance Explained.
Note. If a test statistic is labeled *, **, or *** its coefficient is significant at a 5%, 1% or 0.1% level respectively, after applying the Bonferroni-Holm correction (Abdi, 2010).

Taking into account the absolute deviation from 0.5 in these proportion across all groups suggests that unidimensionality is most evident in excitement. However, this is likely due to the fact that excitement for OECD is very low across all topics. The relative opinions are shown in the scatterplot in Figures 5.15-5.15 against the OECD opinion at the end of the observation period for sentiment, affection and uncertainty. The unidimensionality hypothesis predicts that all observations will fall above or below the x-axis for each group, however this is not consistently the case for any group.

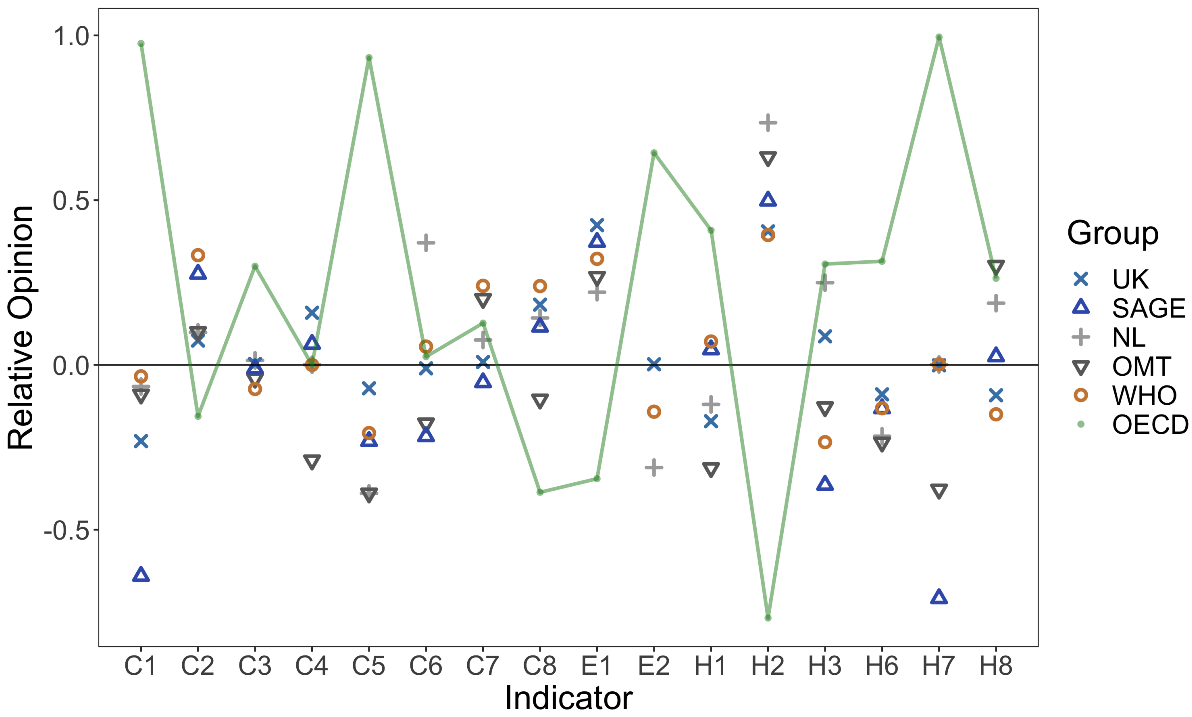

Figure 5.15. Relative Differences in Sentiment across Indicators compared to the OECD opinion at the end of March 2021

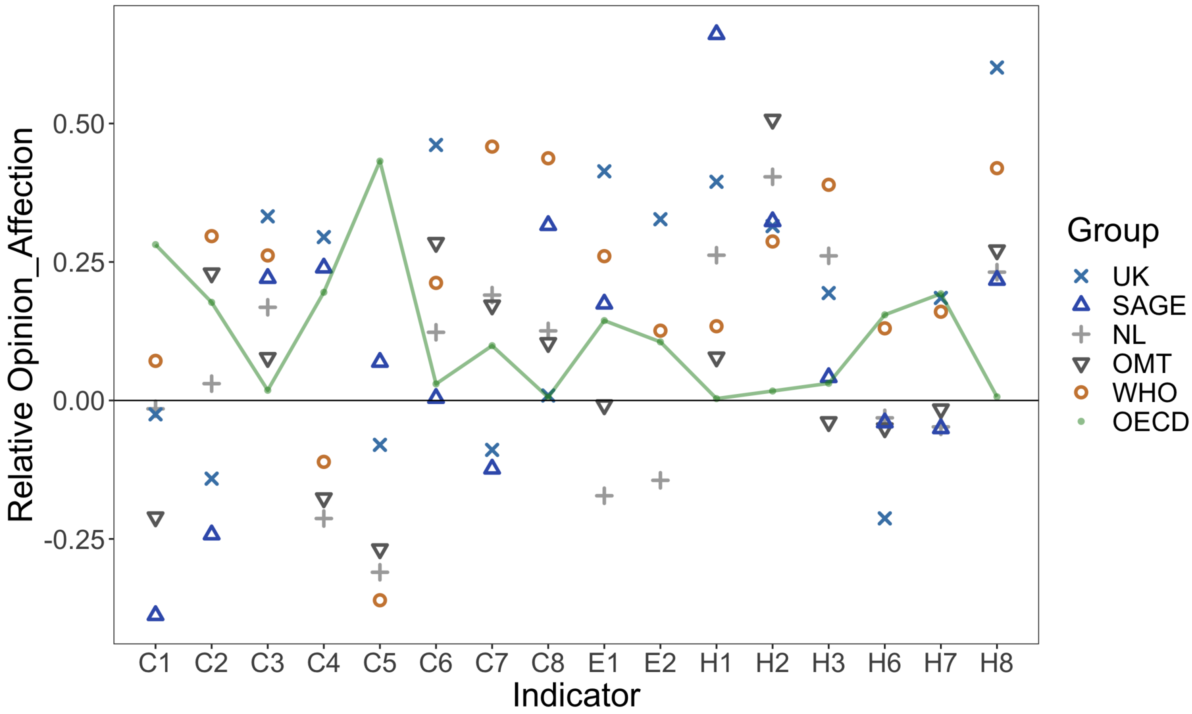

Figure 5.16. Relative Differences in Affection across Indicators compared to the OECD opinion at the end of March 2021

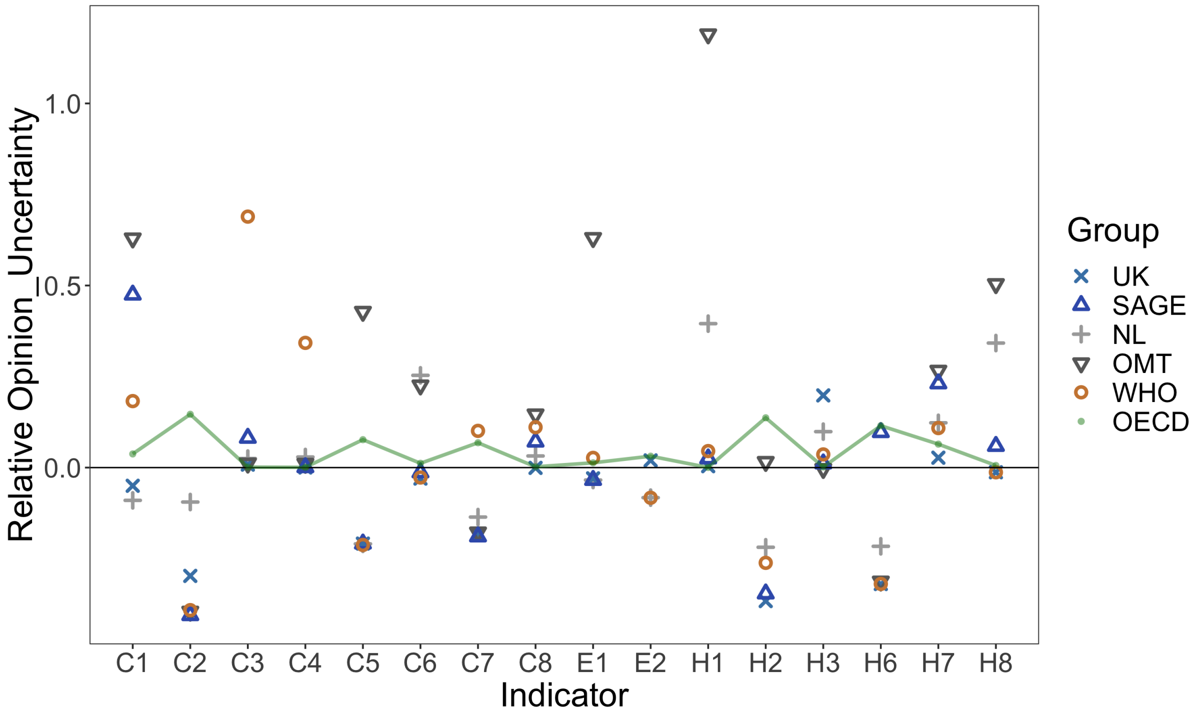

Figure 5.17. Relative Differences in Uncertainty across Indicators compared to the OECD opinion at the end of March 2021

For affection, some clusters of policies where unidimensionality could be present can be identified. For instance, the WHO affection is more intense than the OECD on all topics with the exception of restrictions on public gathering and public transportation. Similarly, SAGE, UK and WHO affection is unidimensional for economic policies. For school and workplace closures (C2 and C1 respectively), there is unidimensionality in uncertainty for NL and UK and in affection for SAGE and UK. The latter highlights another pattern in that the relative difference in opinion of the advisors is often close to that of the policymakers, especially for sentiment. Relative differences in sentiment are closest for UK and SAGE for all policies except relating to face masks and vaccination (H6 and H7 respectively). In contrast, relative differences in opinion for NL and OMT are the most different for sentiment, especially in C1, C7 and C8.

5.5 Sensitivity Analysis

Sensitivity of the estimates of convergence rates and principal component analysis is assessed towards the chosen method of imputation, Lexicon and emotion model. The results under these alternative methdos are shown in Section A.1, In most cases when comparing the series of 5 groups to each other, indications of level convergence are robust to the method of imputation. As expected the linearity imposed by linear imputation and last observation carried forward imputation induce more convergence than MA imputation in most cases. These estimates are less robust when considering only two groups. This is likely because the OLS estimates are not consistent when N is not sufficently large. Similarly, the Convergence clubs found in Table 5.3 are not robust to either model changes or changes in the imputation method.

The distribution of sentences across topics for the less precise guide topics is shown in Table 4.3. Topic ranks and correlation are sensitive to these labels. The most difference is observed in the topic related to public events as there was no distinction made between the closing of venues and industries closing down due to lockdowns. As a the topic related to the stay at home order, which includes lockdowns, is a much bigger topic here. Figure 5.18 shows the similarity structure of the alternative topic model.

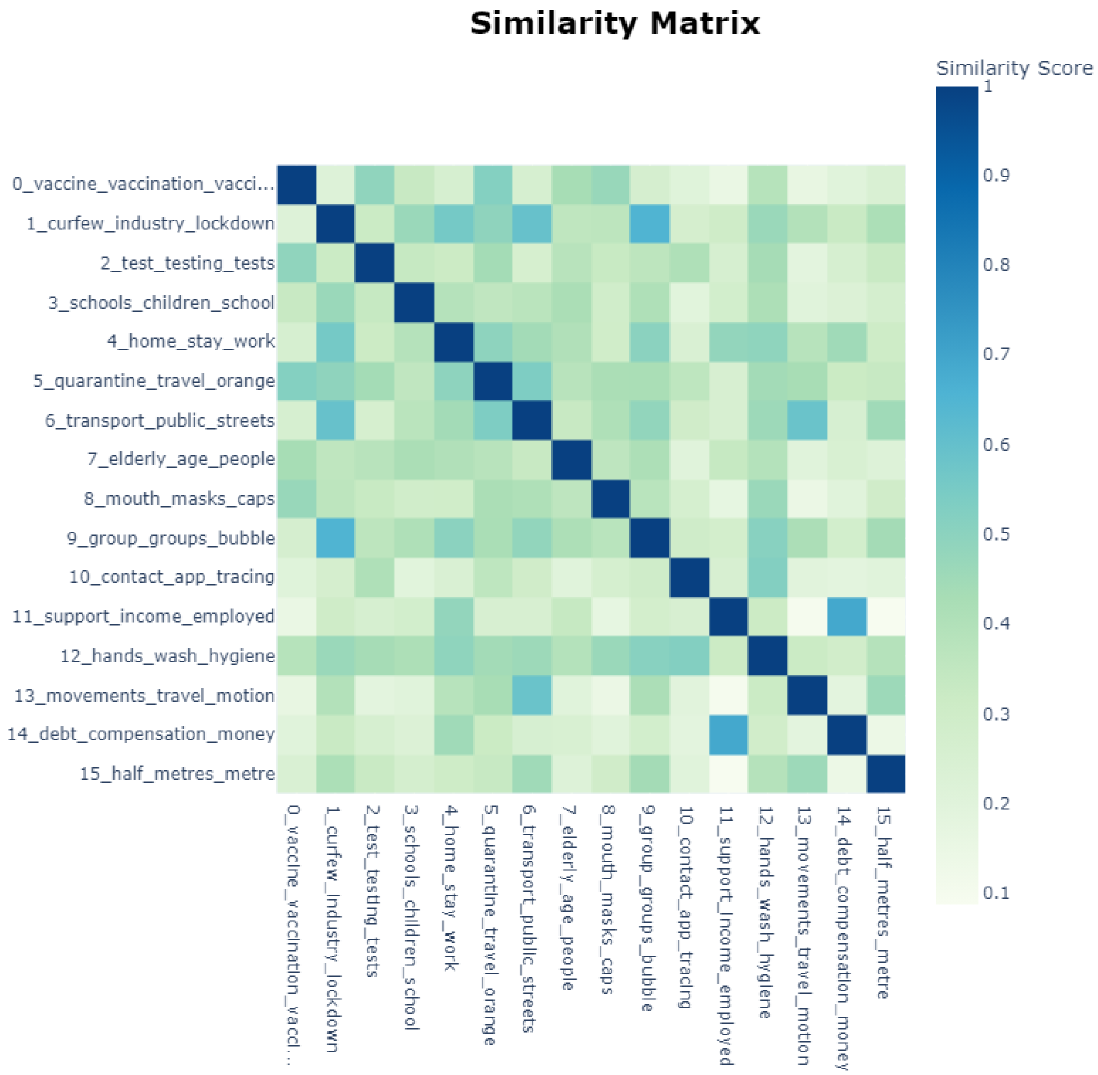

Figure 5.18. Similarity Matrix

The similarity between vaccination polcies and testing is much lower here. While, economic policies are still highly related, the relationships distinguishing between internal movement and international travel disappear. Nevertheless, an interesting similarity between the stay-at-home indicator, workplace closure, public transport and public gatherings is revealed. Overall, topics on containment policies are much more sensitive to the guide topics than health policy topics and economic policy topics. This is likely due to the ambiguity and overlap in the policies related to each containment indicator.

6 Conclusion

This study investigated to what extent opinions in COVID-19 related press conferences and meeting notes converge in the Netherlands and the UK over the course of one year, from the start of the pandemic in March 2020 to the third wave of cases in March 2021. The results show that opinions remain multidimensional but a tendency towards unidimensionality can be found for some combinations of aspects.

Indeed, policy relating to workplace closures and stay at home orders are found to be particularly divisive. This result is robust to the chosen lexicon used in sentiment analysis and the chosen emotion model. While converge to the OECD average opinion was found in levels of sentiment for policies on public events, this result is not robust when considering an alternative lexicon or imputation method.

When investigating convergence between policymakers' and their advisors' opinions there is more evidence of convergence in levels of opinion measures. In particular, sentiment on policies regarding the protection of the elderly converges in level for the UK and the Netherlands and is robust to different methods of imputation. Convergence is rejected more frequently when comparing UK policymakers' opinion to their advisors opinions than for the Netherlands, however relative differences of opinions at the end of the observation period are closer for the UK and SAGE then for NL and OMT. This suggests that policymakers and advisors exchange ideas more effectively than in the UK throughout the pandemic but the UK and SAGE are more alligned on their opinions relative to the OECD in the long run. Furthermore, matching diverging and converging opinions on policies regarding face masks suggest that the listening structure between policymakers and advisors is similar in the UK and the Netherlands for this aspect of COVID-19 policy. This result is robust to the considered imputation method.

While, convergence is not reached in the period considered, principal component analysis suggests a reduction in dimensionality of sentiment and uncertainty towards the end of the observation period. Changes in dimensionality, especially for affection, also coincide with the dates of COVID-19 lockdown measures and easing measures. Finally, dimensionality is not minimised at the end of the observation period. Instead, the most evidence of unidimensionality can be found at the during the first lockdown between March and May of 2020.

Sensitivity of the estimates of convergence rates and principal component analysis is assessed towards the chosen method of imputation, Lexicon and emotion model. In most cases when comparing the series of 5 groups to each other, indications of level convergence are robust to the method of imputation. However, tests of convergence in smaller groups are very sensitive to the applied methods. Namely, convergence is found more often when using linear imputation and even more so when using last observation carried forward imputation. Additionally, sensitivity in the guided topic modelling approach is assessed. It shows that topics on containment policies are much more sensitive to the guide topics than health policy topics and economic policy topics. This is likely due to the ambiguity and overlap in the policies related to each containment indicator.

Thus this research contributes to the literature on topic modelling by highlighting the benefits of additional guidance when topics are ambiguous or overlap. It also contributes to the literature on opinion formation by providing a framework to assess aspect-based opinion through tone and sentiment over time. The findings regarding unidimensionality at the end of the observation period provide further evidence for what conditions create obstacles to convergence, namely uncertainty and economic hardship.

The main limitations of this is the limited time period for which press conference transcripts are frequently available. If a longer time frame would be considered document level opinions could be aggregated by weeks or months to filter out some of the noisy signals in opinions. This would also allow an assessment of the persuasion bias hypothesis over a sufficiently long period of time. Future research could also focus on expanding the group of countries considered in this analysis. This would provide more consistent OLS estimates of convergence rates. Additionally, it would remove the need to use a proxy for the average opinion. Alternatively, more fine-grained comparisons could be made on the opinions of politicians or exper advisors indiviually. These opinons could be extracted from meeting notes or press conferences using Part of Speech Tagging or Entity Recognition methods. Finally, the topic modelling approach can be improved by performing domain specific fine-tuning after task specific fine-tuning. The lack of robustness to the chosen lexicon in sensitivity analysis further suggest that a domain specific word-embedding generated dictionary may be superior.

7 References

Abdi, H. (2010). Holm's sequential Bonferroni procedure. *Encyclopedia of Research Design*, *1*(8), 1–8. https://personal.utdallas.edu/~Herve/abdi-Holm2010-pretty.pdf

Allcott, H., Boxell, L., Conway, J., Gentzkow, M., Thaler, M., & Yang, D. (2020). Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic. *Journal of Public Economics*, *191*, 104254. https://doi.org/10.1016/j.jpubeco.2020.104254

Bertsch, C., Hull, I., & Zhang, X. (2021). Narrative fragmentation and the business cycle. *Economics Letters*, *201*, 109783. https://doi.org/10.1016/j.econlet.2021.109783

Brandts, J., Giritligil, A. E., & Weber, R. A. (2015). An experimental study of persuasion bias and social influence in networks. *European Economic Review*, *80*, 214–229. https://doi.org/10.1016/j.euroecorev.2015.07.007

Bursztyn, L., Rao, A., Roth, C. P., & Yanagizawa-Drott, D. H. (2020). *Misinformation during a pandemic* (tech. rep.). National Bureau of Economic Research. https://doi.org/10.3386/w27417

Cochrane, C., Rheault, L., Godbout, J.-F., Whyte, T., Wong, M. W.-C., & Borwein, S. (2022). The automatic analysis of emotion in political speech based on transcripts. *Political Communication*, *39*(1), 98–121. https://doi.org/10.1080/10584609.2021.1952497

Corazzini, L., Pavesi, F., Petrovich, B., & Stanca, L. (2012). Influential listeners: An experiment on persuasion bias in social networks. *European Economic Review*, *56*(6), 1276–1288. https://doi.org/10.1016/j.euroecorev.2012.05.005

Coronavirus tijdlijn. (n.d.). https://www.rijksoverheid.nl/onderwerpen/coronavirus-tijdlijn

DeMarzo, P. M., Vayanos, D., & Zwiebel, J. (2003). Persuasion bias, social influence, and unidimensional opinions. *The Quarterly Journal of Economics*, *118*(3), 909–968. https://doi.org/10.1162/00335530360698469

Dreier, S. K., Gade, E. K., Card, D., & Smith, N. A. (2022). Patterns of bias: How mainstream media operationalize links between mass shootings and terrorism. *Political Communication*, *39*(6), 755–778. https://doi.org/10.1080/10584609.2022.2111484

Du, K. (2017). Econometric convergence test and club clustering using stata. *The Stata Journal*, *17*(4), 882–900. https://doi.org/10.1177/1536867X1801700407

Foerster, M. (2018). Finite languages, persuasion bias, and opinion fluctuations. *Journal of Economic Behavior & Organization*, *149*, 46–57. https://doi.org/10.1016/j.jebo.2018.03.001

Georgarakos, D., & Kenny, G. (2022). Household spending and fiscal support during the COVID-19 pandemic: Insights from a new consumer survey. *Journal of Monetary Economics*, *129*, S1–S14. https://doi.org/10.1016/j.jmoneco.2022.02.007

Golub, B., & Jackson, M. O. (2010). Naive learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, *2*(1), 112–149. https://doi.org/10.1257/mic.2.1.112

GOV.UK. (2022). Slides, datasets and transcripts to accompany coronavirus press conferences. https://www.gov.uk/government/collections/slides-and-datasets-to-accompany-coronavirus-press-conferences#transcripts

Goyal, N., & Howlett, M. (2021). "Measuring the Mix" of policy responses to COVID-19: Comparative policy analysis using topic modelling. *Journal of Comparative Policy Analysis: Research and Practice*, *23*(2), 250–261. https://doi.org/10.1080/13876988.2021.1880872

Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint arXiv:2203.05794*. https://doi.org/10.48550/arXiv.2203.05794

Grootendorst, M. P. (n.d.). Guided Topic Modeling - BERTopic. https://maartengr.github.io/BERTopic/getting_started/guided/guided.html

Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., Webster, S., Cameron-Blake, E., Hallas, L., Majumdar, S., et al. (2021). A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature Human Behaviour*, *5*(4), 529–538. https://doi.org/10.1038/s41562-021-01079-8

Han, X., Zhu, Y., Zhang, Y., & Chen, Y. (2022). Policy effectiveness on the global COVID-19 pandemic and unemployment outcomes: A large mixed frequency spatial approach. *Available at SSRN 4049509*. https://doi.org/10.2139/ssrn.4049509

Hidalgo-Ternero, C. M. (2020). Google Translate vs. DeepL: Analysing neural machine translation performance under the challenge of phraseological variation. *MonTI. Monografías De Traducción E Interpretación*, 154–177. https://doi.org/10.6035/MonTI.2020.ne6.5

Hutto, C., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), 216–225. https://doi.org/10.1609/icwsm.v8i1.14550

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*(1), 141–151. https://doi.org/10.1177/001316446002000116

Kamath, R., Ghoshal, A., Eswaran, S., & Honnavalli, P. B. (2022). EmoRoBERTa: An enhanced emotion detection model using RoBERTa. *IEEE International Conference on Electronics, Computing and Communication Technologies.*

Madrigal, G. (2023). Community matters: Content analysis of children in immigration media coverage, 1990-2020. *Political Communication*, 1–27. https://doi.org/10.1080/10584609.2023.2172117

McInnes, L., Healy, J., & Astels, S. (2017). HDBSCAN: Hierarchical density based clustering. *Journal of Open Source Software*, *2*(11), 205. https://doi.org/10.21105/joss.00205

McInnes, L., Healy, J., & Melville, J. (2018). UMAP: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*. https://doi.org/10.48550/arXiv.1802.03426

Meyer-Gutbrod, J., & Woolley, J. (2021). New conflicts in the briefing room: Using sentiment analysis to evaluate administration-press relations from Clinton through Trump. *Political Communication*, *38*(3), 241–259. https://doi.org/10.1080/10584609.2020.1763527

Ministerie van Algemene Zaken. (2022). Documenten. https://www.rijksoverheid.nl/documenten?trefwoord=persconferentie&type=Mediatekst&pagina=1

Moritz, S., & Bartz-Beielstein, T. (2017). imputeTS: Time series missing value imputation in R. *The R Journal*, *9*(1), 207–218. https://doi.org/10.32614/RJ-2017-009

Motta, M., & Stecula, D. (2023). The effects of partisan media in the face of global pandemic: How news shaped COVID-19 vaccine hesitancy. *Political Communication*, 1–22. https://doi.org/10.1080/10584609.2023.2187496

Müller, M., Salathé, M., & Kummervold, P. E. (2023). COVID-Twitter-BERT: A natural language processing model to analyse COVID-19 content on Twitter. *Frontiers in Artifical Intelligence*. https://doi.org/10.3389/frai.2023.1023281

Müller, S. (2020). Media coverage of campaign promises throughout the electoral cycle. *Political communication*, *37*(5), 696–718. https://doi.org/10.1080/10584609.2020.1744779

OECD. (2021). OECD policy responses to coronavirus (COVID-19). https://www.oecd.org/coronavirus/en/policy-responses?submodel=statement

Organization, W. H. (n.d.). Coronavirus disease (COVID-2019) previous press briefings. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/media-resources/press-briefings/previous

Page, L. (2020). Assessing the unidimensionality of political opinions. An indirect test of the persuasion bias. *Journal of Economic Behavior & Organization*, *174*, 469–477. https://doi.org/10.1016/j.jebo.2019.11.003

Phillips, P. C. B., & Sul, D. (2007). Transition modeling and econometric convergence tests. *Econometrica*, *75*(6), 1771–1855. https://doi.org/10.1111/j.1468-0262.2007.00811.x

Phillips, P. C. B., & Sul, D. (2009). Economic transition and growth. *Journal of Applied Econometrics*, *24*(7), 1153–1185. https://doi.org/10.1002/jae.1080

Ramaekers, M. J., Verbakel, E., & Kraaykamp, G. (2023). Decline in informal helping during the first COVID-19 lockdown: A longitudinal analysis of dutch data. *European Societies*, 1–28. https://doi.org/10.1080/14616696.2023.2200503

Revelle, W., & Revelle, M. W. (2015). Package 'psych'. *The comprehensive R archive network*, *337*(338). https://cran.r-project.org/web/packages/psych/psych.pdf

RIVM. (n.d.). Outbreak management team (OMT). https://www.rivm.nl/coronavirus-covid-19/omt#adviezen_omt

Sichera, R., Pizzuto, P., et al. (2019). ConvergenceClubs: A package for performing the phillips and sul's club convergence clustering procedure. *The R Journal*, *11*(2), 142. https://journal.r-project.org/archive/2019/RJ-2019-021/RJ-2019-021.pdf

Streiner, D. L. (1998). Factors affecting reliability of interpretations of scree plots. *Psychological reports*, *83*(2), 687–694. https://doi.org/10.2466/pr0.1998.83.2.687

Timeline of UK government coronavirus lockdowns and restrictions. (2022). https://www.instituteforgovernment.org.uk/data-visualisation/timeline-coronavirus-lockdowns

Timeline: WHO's COVID-19 response. (n.d.). https://www.who.int/emergencies/diseases/novel-coronavirus-2019/interactive-timeline#event-232

Todd, I., Bulder, C., McCauley, D., & Burns, M.-K. (2022). Assessing the post-COVID prospects for the energy transition in the Netherlands and the UK, using a policy barriers approach. *Local Environment*, *27*(6), 712–727. https://doi.org/10.1080/13549839.2022.2068142

Viehmann, C., Beck, T., Maurer, M., Quiring, O., & Gurevych, I. (2022). Investigating opinions on public policies in digital media: Setting up a supervised machine learning tool for stance classification. *Communication Methods and Measures*, 1–35. https://doi.org/10.1080/19312458.2022.2151579

Yarchi, M., Baden, C., & Kligler-Vilenchik, N. (2021). Political polarization on the digital sphere: A cross-platform, over-time analysis of interactional, positional, and affective polarization on social media. *Political Communication*, *38*(1-2), 98–139. https://doi.org/10.1080/10584609.2020.1785067

Yulianto, A., & Supriatnaningsih, R. (2021). Google Translate vs. DeepL: A quantitative evaluation of close-language pair translation (French to English). *AJELP: Asian Journal of English Language and Pedagogy*, *9*(2), 109–127. https://doi.org/10.37134/ajelp.vol9.2.9.2021

Zhou, A., Liu, W., & Yang, A. (2023). Politicization of science in COVID-19 vaccine communication: Comparing US politicians, medical experts, and government agencies. *Political Communication*, 1–23. https://doi.org/10.1080/10584609.2023.2201184

A Appendix

Table A.1. *Policy indicators*

Indicator	Name	Differentiated	Scale
C1	School Closures	Y	0-3
C2	Workplace Closing	Y	0-3
C3	Cancel Public Events	Y	0-2
C4	Restrictions on Gatherings	Y	0-4
C5	Public Transportation	Y	0-2
C6	Stay at Home Order	Y	0-3
C7	Restrictions on Internal Movement	Y	0-2
C8	International Travel Controls	Y	0-4
E1	Income Support	N	0-2
E2	Debt/Contract Relief for Households	N	0-2
H1	Public Information Campaigns	N	0-2
H2	Testing Policy	N	0-3
H3	Contact tracing	N	0-2
H6	Facial coverings	Y	0-4
H7	Vaccination Policy	N	0-5
H8	Protection of Elderly People	Y	0-3

Note: Y stands for Yes and N stands for No. A differentiated policy indicator refers to one where different policy is recorded for vaccinated and non-vaccinated individuals.

A.1 Sensitivity Analysis

A.1.1 Lexicon

Table A.2. *Estimated convergence rates when using Lexicoder and EmoBerta*

Indicator	Sentiment	Joy	Surprise	Anger	Disgust	Sadness	Fear
C1	-0.229	-1.017	-1.604	-1.127	-0.856	-0.586	-0.766
	(0.465)	(0.469)	(0.667)	(0.442)	(0.315)	(0.410)	(0.393)
C2	-3.572	-0.607	**1.865**	**0.488**	-1.670	-0.338	-0.156
	(0.866)**	(0.256)	(0.686)	(0.364)	(0.203)***	(0.249)	(0.452)
C3	**0.817**	-0.166	**0.077**	**1.934**	-0.032	-0.464	-1.598
	(0.721)	(0.315)	(0.469)	(0.485)	(0.421)	(0.102)***	(0.257)***
C4	-2.198	-1.375	-1.713	-1.800	-0.702	-0.758	-0.307
	(0.887)	(0.707)	(0.708)	(0.974)	(0.572)	(0.477)	(0.348)
C5	-2.416	**0.124**	**1.930**	**1.898**	**1.599**	-1.667	-0.531
	(0.743)	(0.594)	(0.663)	(0.836)	(0.581)	(0.448)*	(0.170)
C6	-0.865	-0.855	-1.532	-0.789	**0.229**	-1.043	-0.114
	(0.730)	(0.298)	(0.352)**	(0.538)	(0.469)	(0.223)***	(0.169)
C7	-1.539	-0.791	**0.191**	-0.996	-1.211	-0.719	-0.824
	(0.616)	(0.679)	(0.975)	(0.429)	(0.296)**	(0.309)	(0.386)
C8	-0.252	**0.068**	-0.038	-0.255	-0.973	-0.212	**0.377**
	(0.737)	(0.497)	(0.246)	(0.388)	(0.505)	(0.243)	(0.414)
E1	-0.968	-0.257	-1.024	**0.300**	-0.330	-0.082	**0.599**
	(0.605)	(1.183)	(0.100)***	(0.353)	(1.188)	(0.716)	(0.906)
H1	-2.552	-0.767	**0.165**	-0.113	**0.127**	-0.522	-0.225
	(0.661)*	(0.574)	(0.664)	(0.444)	(0.689)	(0.562)	(0.759)
H2	-0.708	-0.464	-0.226	-0.168	-0.045	-0.859	-0.020
	(0.701)	(0.375)	(0.429)	(0.306)	(0.460)	(0.351)	(0.302)
H3	-1.634	-0.896	**0.168**	-0.118	-0.017	-0.834	-0.077
	(0.817)	(0.397)	(0.643)	(0.484)	(0.703)	(0.565)	(0.303)
H6	-1.448	-1.495	-1.523	-1.082	-1.628	-0.569	-0.563
	(0.703)	(0.437)	(0.417)*	(0.546)	(0.553)	(0.425)	(0.420)
H7	-0.511	**0.388**	-1.953	**0.031**	**0.894**	-0.438	-0.556
	(0.868)	(0.422)	(0.495)*	(0.533)	(0.363)	(0.458)	(0.283)
H8	**0.953**	**1.178**	-0.387	**0.257**	**0.081**	**0.489**	**0.155**
	(0.915)	(0.669)	(0.525)	(0.309)	(0.318)	(0.813)	(0.450)

Note. Standard errors are shown in round brackets below the coefficients. Positive estimates are shown in bold.
Note. If a standard error is labeled *, **, or *** its coefficient is significant at a 5%, 1% or 0.1% level respectively, after applying the Bonferroni-Holm correction (Abdi, 2010).

Table A.3. *Convergence Clubs when using Lexicoder and EmoBerta*

Indicator	Measure	Club 1	Club 2	Divergent Groups
C2	Sentiment	NL,SAGE,OMT,OECD		UK
C2	Disgust	NL, OECD		UK, SAGE, OMT
C3	Sadness	UK, NL, OECD, SAGE		OMT
C3	Fear	SAGE, NL, OMT, OECD		UK
C6	Surprise	NL, SAGE	OMT, OECD	UK
C6	Sadness	OMT, UK	OECD, SAGE	NL
C7	Disgust	OECD, NL	SAGE, OMT	UK
E1	Surprise	OECD, UK, SAGE, NL		OMT
H1	Sentiment			
H6	Surprise	OMT, UK, OECD		NL, SAGE
H7	Surprise	NL, OECD, OMT, SAGE		UK

Table A.4. *Sufficient number of Factors using Lexicoder and EmoBerta*

	Chi-Sqaure test statistic		First principal component	
	One Factor Sufficient	Two Factors Sufficient	Eigenvalue	Prop. Var. Explained
Sentiment				
31-03-21	3.33	0.48	5.93	39.56
01-05-21	5.79	1.34	8.69	57.91
Joy				
31-03-21	4.79	0.82	7.18	47.85
14-03-21	2.89	0.37	10.95	72.98
Surprise				
31-03-21	7.60	2.31	6.50	43.32
29-11-21	4.44	0.17	9.51	63.37
Anger				
31-03-21	11.80	1.98	5.77	38.46
15-08-21	5.49	1.34	9.26	61.76
Disgust				
31-03-21	1.35	0.20	5.76	38.42
18-03-21	9.82	4.57	10.56	70.40
Sadness				
31-03-21	5.88	0.08	6.44	42.90
09-05-20	5.27	0.70	11.49	76.60
Fear				
31-03-21	2.23	0.27	5.22	34.77
07-05-20	10.15	2.77	10.49	69.95

Note. Prop. Var. Explained stands for Proportion of Variance Explained.

Note. If a test statistic is labeled *, **, or *** its coefficient is significant at a 5%, 1% or 0.1% level respectively, after applying the Bonferroni-Holm correction (Abdi, 2010).

A.1.2 Linear imputation

Table A.5. *Estimated convergence rates when using linear imputation*

Indicator	Sentiment	Affection	Excitement	Uncertainty	Frustration	Sadness
C1	**1.082**	-0.448	-0.109	-0.267	**0.228**	**0.200**
	(0.703)	(0.300)	(0.227)	(0.278)	(0.241)	(0.221)
C2	-2.335	-1.016	-0.294	-0.085	-0.348	-0.246
	(0.862)	(0.124)***	(0.213)	(0.124)	(0.178)	(0.159)
C3	**2.394**	-1.361	-0.905	-0.743	-0.317	**0.835**
	(0.529)	(0.379)*	(0.214)**	(0.463)	(0.385)	(0.362)
C4	-2.902	-0.904	-0.596	-0.942	-0.703	**0.643**
	(0.972)	(0.562)	(0.409)	(0.234)**	(0.183)*	(0.490)
C5	**0.487**	-0.041	-0.966	-0.901	**1.166**	**1.509**
	(0.497)	(0.359)	(0.329)	(0.220)**	(0.608)	(0.614)
C6	**0.408**	-0.434	-1.175	-0.138	-0.618	-0.646
	(0.700)	(0.295)	(0.298)*	(0.339)	(0.266)	(0.293)
C7	-1.300	-0.131	-0.203	-0.155	-0.547	-0.754
	(0.394)	(0.121)	(0.222)	(0.255)	(0.301)	(0.182)**
C8	**0.481**	-0.443	-0.464	-0.825	-0.118	-0.595
	(0.705)	(0.262)	(0.351)	(0.238)	(0.326)	(0.368)
E1	-0.711	-1.076	-0.506	-0.414	0.887	-0.395
	(0.456)	(0.117)***	(0.222)	(0.282)	(0.495)	(0.589)
H1	-1.114	-0.693	-0.916	-0.278	-0.897	-0.659
	(0.845)	(0.369)	(0.329)	(0.194)	(0.264)	(0.567)
H2	-0.309	**0.324**	-0.370	**0.167**	-0.061	**0.531**
	(0.759)	(0.287)	(0.170)	(0.160)	(0.182)	(0.415)
H3	-3.875	-1.443	-0.594	-0.335	-0.076	-1.190
	(1.379)	(0.173)***	(0.469)	(0.157)	(0.673)	(0.708)
H6	-1.811	-1.293	-0.511	-0.589	-0.527	**0.754**
	(0.863)	(0.290)***	(0.285)	(0.228)	(0.156)	(0.810)
H7	**1.021**	**0.686**	**0.199**	**0.438**	**0.295**	**0.049**
	(1.004)	(0.391)	(0.327)	(0.245)	(0.267)	(0.493)
H8	**1.634**	-0.924	**0.174**	-0.725	-0.604	-0.524
	(1.015)	(0.267)	(0.322)	(0.219)	(0.287)	(0.273)

Note. Standard errors are shown in round brackets below the coefficients. Positive estimates are shown in bold.
Note. If a standard error is labeled *, **, or *** its coefficient is significant at a 5%, 1% or 0.1% level respectively, after applying the Bonferroni-Holm correction (Abdi, 2010).

Table A.6. *Convergence Clubs when using linear imputation*

Indicator	Measure	Club 1	Club 2	Divergent Groups
C2	Affection	OMT, NL, OECD,	UK, SAGE	
C3	Affection	SAGE, NL	OECD, UK, OMT	
C3	Excitement	OMT, UK, OECD		NL, SAGE
C4	Uncertainty	UK, SAGE, OECD, OMT		NL
C4	Frustration	OMT, NL, UK;	OECD, SAGE	
C5	Uncertainty	OECD, NL	UK, SAGE,	OMT
C6	Excitement	OMT, NL	UK, SAGE, OECD	
C7	Sadness	NL, OECD, SAGE, OMT		UK
E1	Affection	UK, OECD	OMT, NL	SAGE
H2	Affection	UK, NL	SAGE, OMT, OECD	
H6	Affection	OMT, SAGE	NL,OECD	UK

Table A.7. *Sufficient number of Factors using linear imputation*

	Chi-Sqaure test statistic		First principal component	
	One Factor Sufficient	Two Factors Sufficient	Eigenvalue	Prop. Var. Explained
Sentiment				
31-03-21	4.41	0.47	5.71	38.10
27-05-20	10.51	2.54	8.21	54.73
Affection				
31-03-21	2.13	0.76	6.08	40.50
01-02-21	5.18	1.02	8.42	56.10
Excitement				
31-03-21	2.92	0.89	6.01	40.07
29-09-20	5.26	1.83	9.27	61.78
Uncertainty				
31-03-21	2.63	0.13	6.12	40.78
07-12-20	13.10	4.71	9.81	65.39
Frustration				
31-03-21	8.80	3.95	5.99	39.93
05-06-20	1.27	0.19	9.66	64.42
Sadness				
31-03-21	32.66**	25.66***	7.19	47.93
31-01-21	4.35	0.36	8.86	59.09

Note. Prop. Var. Explained stands for Proportion of Variance Explained.

Note. If a test statistic is labeled *, **, or *** its coefficient is significant at a 5%, 1% or 0.1% level respectively, after applying the Bonferroni-Holm correction (Abdi, 2010).

A.1.3 Last observation carried forward imputation

Table A.8. *Estimated convergence rates when using last observation carried forward imputation*

Indicator	Sentiment	Affection	Excitement	Uncertainty	Frustration	Sadness
C1	**1.071**	-0.306	-0.082	-0.283	**0.107**	-0.090
	(0.768)	(0.299)	(0.197)	(0.216)	(0.219)	(0.207)
C2	-1.478	-0.988	-0.314	-0.196	-0.129	-0.334
	(0.868)	(0.191)***	(0.202)	(0.120)	(0.192)	(0.216)
C3	**1.271**	-0.862	-0.495	-0.243	-0.252	**0.471**
	(0.592)	(0.404)	(0.187)	(0.245)	(0.223)	(0.712)
C4	-2.451	-0.486	-0.294	-0.263	-0.384	**1.586**
	(0.809)	(0.344)	(0.375)	(0.218)	(0.426)	(0.716)
C5	**0.496**	**0.370**	-0.719	-0.404	**0.977**	**1.231**
	(0.556)	(0.340)	(0.819)	(0.246)	(0.646)	(0.770)
C6	**0.472**	-0.189	-1.277	-0.462	-1.186	-0.777
	(0.665)	(0.266)	(0.305)**	(0.184)	(0.293)**	(0.383)
C7	-2.000	-0.274	**0.120**	-0.586	-1.240	-1.881
	(0.537)*	(0.205)	(0.290)	(0.239)	(0.273)***	(1.078)
C8	**0.767**	-0.110	-0.056	-0.655	-0.326	**0.065**
	(0.597)	(0.271)	(0.272)	(0.251)	(0.257)	(0.339)
E1	-0.737	-1.621	-0.672	-1.333	**1.838**	**0.569**
	(0.435)	(0.239)***	(0.230)	(0.183)***	(0.756)	(0.469)
H1		-0.790	-0.290	-0.263	-0.549	-0.802
-0.343						
	(0.560)	(0.291)	(0.318)	(0.170)	(0.374)	(0.517)
H2	**0.391**	**0.465**	-0.313	-0.004	-0.847	**0.646**
	(0.717)	(0.287)	(0.165)	(0.135)	(0.283)	(0.399)
H3	-3.188	-0.927	**0.063**	-0.332	-0.421	-0.508
	(1.116)	(0.344)	(0.557)	(0.148)	(0.866)	(0.555)
H6	-1.379	-1.210	-0.817	-0.710	-0.591	**0.075**
	(1.054)	(0.357)	(0.376)	(0.169)**	(0.179)	(0.818)
H7	**1.595**	**0.637**	**0.143**	**0.342**	**0.220**	**0.089**
	(1.293)	(0.244)	(0.260)	(0.316)	(0.290)	(0.397)
H8	**1.557**	-0.351	**0.226**	-0.434	-0.118	-0.572
	(0.724)	(0.194)	(0.354)	(0.162)	(0.303)	(0.248)

Note. Standard errors are shown in round brackets below the coefficients. Positive estimates are shown in bold.
Note. If a standard error is labeled *, **, or *** its coefficient is significant at a 5%, 1% or 0.1% level respectively, after applying the Bonferroni-Holm correction (Abdi, 2010).

Table A.9. *Sufficient number of Factors using last observation carried forward imputation*

	Chi-Sqaure test statistic		First principal component	
	One Factor Sufficient	Two Factors Sufficient	Eigenvalue	Prop. Var. Explained
Sentiment				
31-03-21	4.41	0.47	5.71	38.10
10-04-20	3.85	0.09	7.85	52.33
Affection				
31-03-21	2.13	0.76	6.08	40.50
29-05-20	6.95	0.33	8.81	58.70
Excitement				
31-03-21	2.92	0.89	6.01	40.07
15-10-20	3.51	1.52	8.93	59.52
Uncertainty				
31-03-21	2.63	0.13	6.12	40.78
22-03-21	4.19	0.53	9.47	63.14
Frustration				
31-03-21	8.80	3.95	5.99	39.93
30-04-20	0.82	0.10	9.27	61.80
Sadness				
31-03-21	32.66**	25.66***	7.19	47.93
17-09-20	1.49	0.08	9.54	63.61

Note. Prop. Var. Explained stands for Proportion of Variance Explained.

Note. If a test statistic is labeled *, **, or *** its coefficient is significant at a 5%, 1% or 0.1% level respectively, after applying the Bonferroni-Holm correction (Abdi, 2010).

A.1.4 Guide Topics

Table A.10. *Indicator, Corresponding topic label, and frequency by group*

Indicator	Topic Label	Rank	Number of Sentences					
			WHO	OECD	SAGE	OMT	UK T	NL T
C1	schools children school education	3	719	2501	249	353	233	1354
C2	home stay work entrepreneurs	4	5896	1935	13	49	229	1096
C3	half metres metre distance	15	17	11	2	13	11	392
C4	group groups bubble bubbles	9	**85**	39	44	34	35	357
C5	transport public streets traffic	6	250	556	22	69	225	392
C6	curfew industry lockdown hospitality	1	797	1273	91	175	297	2056
C7	movements travel motion trip	13	177	**14**	**3**	**12**	14	228
C8	quarantine travel orange countries	5	651	1005	72	146	56	520
E1	support income employed self	11	119	2470	16	12	103	179
E2	debt compensation money tax	14	**122**	**4550**	**2**	**9**	**18**	**246**
H1	hands wash hygiene washing	12	806	**723**	**25**	40	57	228
H2	test testing tests positive	2	**1016**	1191	238	842	**362**	727
H3	contact app tracing source	10	298	212	76	34	43	194
H6	mouth masks caps mask	8	475	410	437	149	118	306
H7	vaccine vaccination vaccines vaccinated	0	5856	1241	388	471	567	1369
H8	elderly age people nursing	7	259	535	83	206	**47**	314

A.2 Topic model

Table A.11. *Indicator, Corresponding topic label, and frequency*

Indicator	Topic Label	Number of Sentences
C1 School Closures	2 schools school education children	1733
C2 Workplace Closing	7 industry hospitality entrepreneurs economy	834
C3 Cancel Public Events	4 half metres distance metre	1043
C4 Restrictions on Gatherings	9 group groups people crowd	621
C5 Public Transportation	13 transport public bus trains	363
C6 Stay at Home Order	3 home curfew lockdown stay	1449
C7 Restrictions on Internal Movement	10 travel visit movements advice	515
C8 International Travel Controls	8 quarantine orange flights travel	666
E1 Income Support	14 money income support tax	268
E2 Debt/Contract Relief for Households	15 debt aid support loans	148
H1 Public Information Campaigns	11 hands public wash washing	475
H2 Testing Policy	1 testing test tests coronavirus	2304
H3 Contact tracing	12 app contact source tracing	406
H6 Facial coverings	5 mouth masks caps face	971
H7 Vaccination Policy	0 vaccine vaccination vaccines vaccinated	2706
H8 Protection of Elderly People	6 nursing elderly homes care	853

Note: The Number of Sentences is based on the set of meeting notes and speech transcripts from the Netherlands and the UK consisting of 72504 sentences.

A.2.1 Wordclouds

Figure A.1. School Closures

Figure A.2. Workplace Closing

Figure A.3. Cancel Public Events

Figure A.4. Restrictions on Gatherings

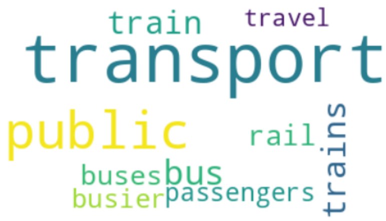

Figure A.5. Public Transportation

Figure A.6. Stay at Home Order

Figure A.7. Restrictions on Internal Movement

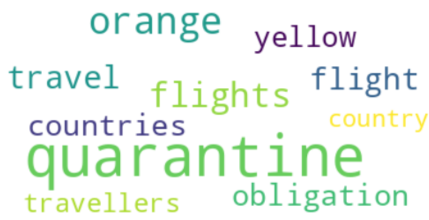

Figure A.8. International Travel Controls

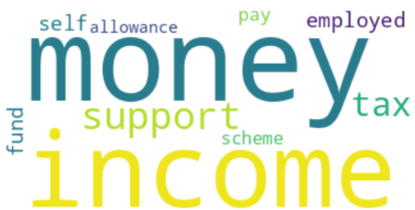

Figure A.9. Income Support

Figure A.10. Debt/Contract Relief for Households

Figure A.11. Public Information Campaigns

Figure A.12. Testing Policy

Figure A.13. Contact tracing

Figure A.14. Facial coverings

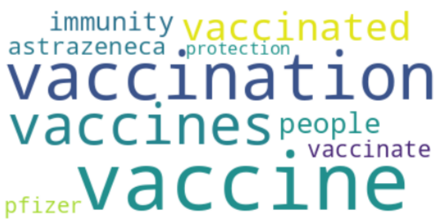

Figure A.15. Vaccination Policy

Figure A.16. Protection of Elderly People

A.2.2 Word embeddings

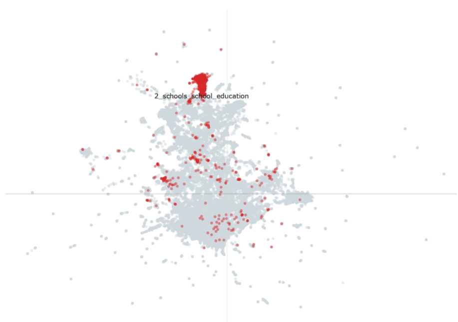

Figure A.17. School Closures

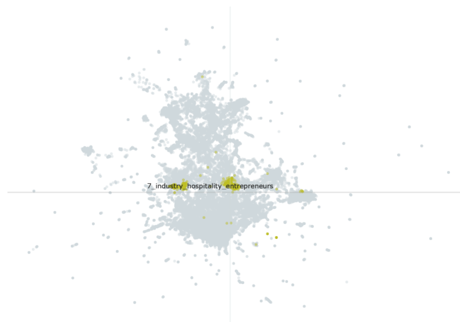

Figure A.18. Workplace Closing

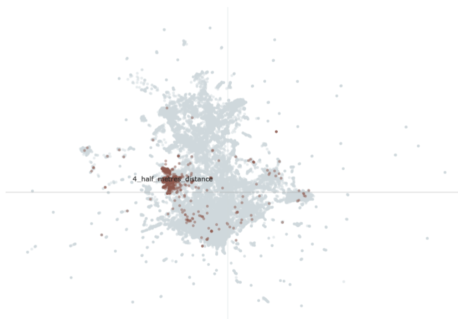

Figure A.19. Cancel Public Events

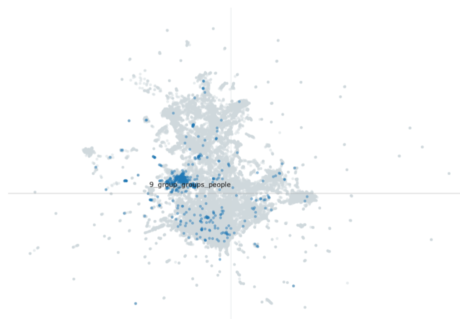

Figure A.20. Restrictions on Gatherings

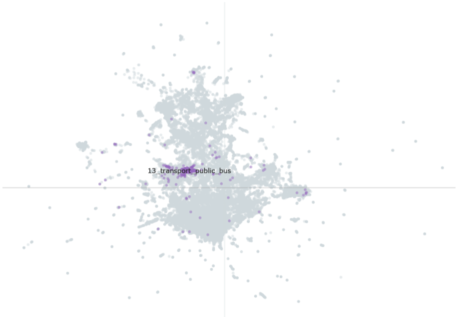

Figure A.21. Public Transportation

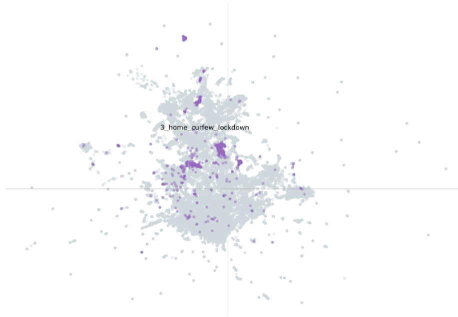

Figure A.22. Stay at Home Order

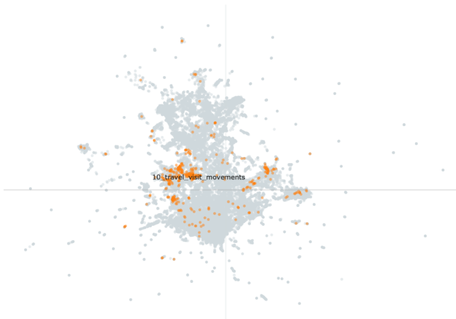

Figure A.23. Restrictions on Internal Movement

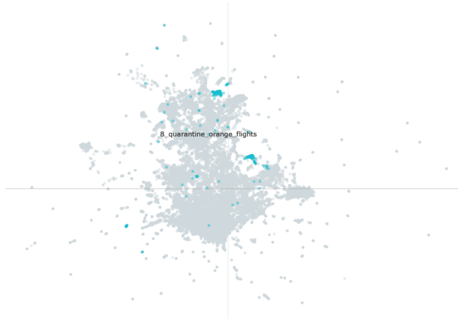

Figure A.24. International Travel Controls

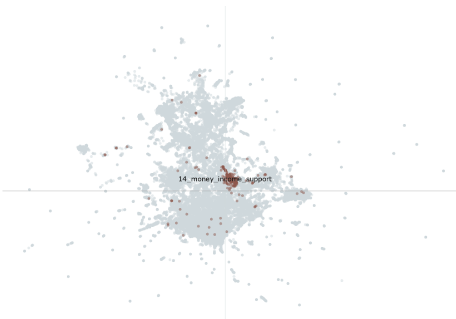

Figure A.25. Income Support

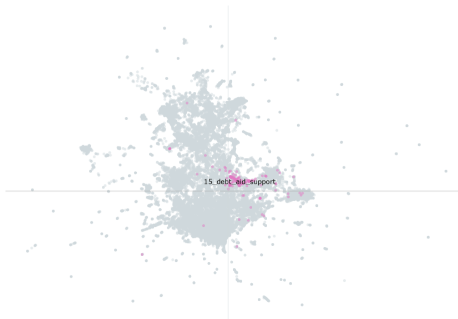

Figure A.26. Debt/Contract Relief for Households

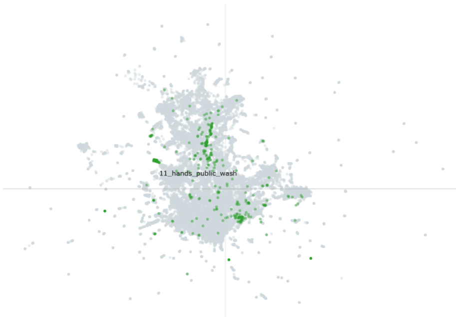

Figure A.27. Public Information Campaigns

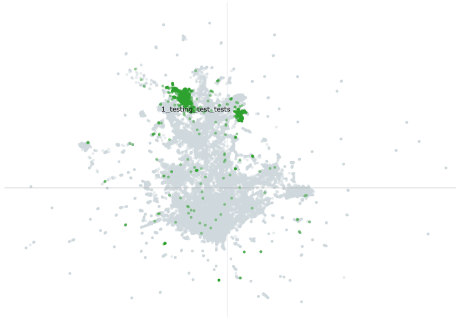

Figure A.28. Testing Policy

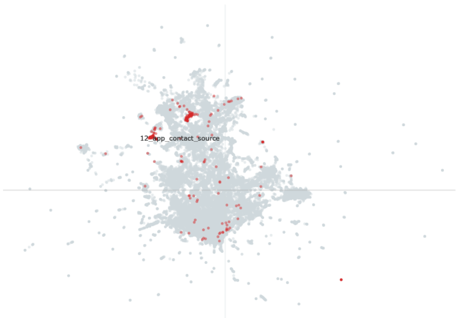

Figure A.29. Contact tracing

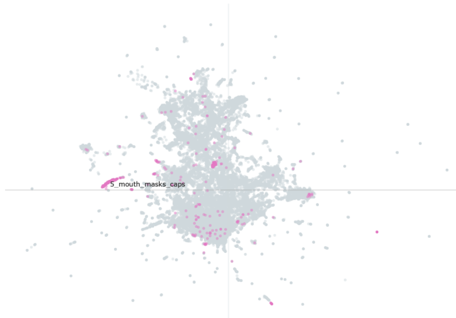

Figure A.30. Facial coverings

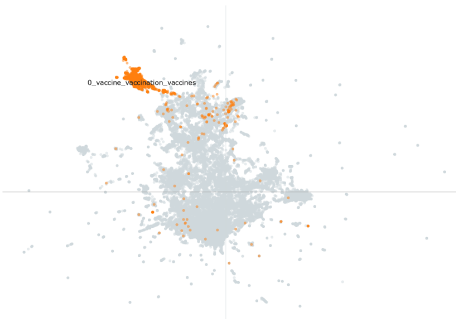

Figure A.31. Vaccination Policy

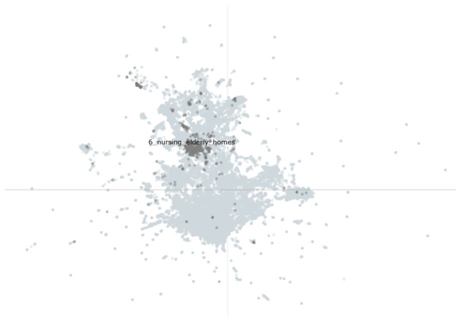

Figure A.32. Protection of Elderly People

A.3 Table

A.4 Plots

Table A.12. *Estimated convergence rates*

Indicator	Subgroup	Sentiment	SE	Affection	SE	Excite.	SE	Uncert.	SE	Frust.	SE	Sadness	SE
C1	NL, OMT	**4.819**	(1.534)	-0.878	(0.537)	-0.353	(0.726)	-0.737	(0.461)	**0.637**	(0.746)	-1.023	(0.645)
	UK, SAGE	-3.795	(0.659)***	-1.462	(0.595)	-1.457	(0.686)	-0.003	(0.390)	**0.697**	(0.824)	-1.823	(0.969)
C2	NL , OMT	**1.704**	(1.067)	-0.075	(0.693)	-1.349	(0.386)	-0.736	(0.545)	-0.099	(0.463)	-0.455	(0.268)
	UK, SAGE	-1.836	(0.926)	-0.740	(0.061)***	**0.730**	(0.157)	-1.102	(1.697)	-2.895	(0.567)***	-1.961	(3.296)
C3	NL , OMT	-0.573	(0.774)	-1.369	(0.587)	-1.146	(0.325)	-0.077	(0.512)	-0.006	(0.445)	-1.818	(0.488)*
	UK , SAGE	**3.668**	(0.399)	-1.305	(0.526)	-0.748	(0.261)	-0.657	(0.590)	-1.128	(0.145)***	**3.779**	(1.222)
C4	NL , OMT	-2.692	(1.284)	-0.103	(0.799)	-1.022	(0.345)	-1.305	(0.620)	**0.859**	(0.794)	**0.272**	(0.558)
	UK , SAGE	-0.022	(1.096)	**0.471**	(1.072)	**0.761**	(0.717)	**1.654**	(1.342)	**0.773**	(1.501)	**2.767**	(1.083)
C5	NL , OMT	**0.773**	(0.975)	**0.080**	(1.925)	**2.133**	(1.489)	-1.148	(1.116)	-0.401	(0.385)	**0.360**	(0.845)
	UK , SAGE	**1.067**	(0.225)	**2.016**	(0.435)	-0.048	(0.028)	**7.387**	(1.474)	-1.863	(0.510)*	-0.934	(0.294)
C6	NL , OMT	-0.104	(1.139)	**0.987**	(0.465)	-1.221	(0.703)	**1.557**	(0.585)	-0.899	(0.465)	-0.244	(0.662)
	UK , SAGE	**0.527**	(1.002)	-1.818	(0.490)*	**0.369**	(0.764)	-0.824	(0.452)	-0.585	(0.516)	-0.669	(0.557)
C7	NL , OMT	**0.179**	(0.975)	**0.797**	(0.982)	-0.377	(0.722)	**1.844**	(0.595)	**0.155**	(0.935)	-0.886	(0.485)
	UK , SAGE	-0.942	(1.940)	**1.026**	(1.349)	-1.434	(0.200)***	-0.365	(0.305)	-1.938	(0.273)***	-1.700	(0.392)**
C8	NL , OMT	-0.325	(0.842)	**1.208**	(0.589)	-0.859	(0.506)	-0.134	(0.422)	-0.118	(0.731)	**1.110**	(0.410)
	UK , SAGE	**1.283**	(1.373)	**0.844**	(0.747)	-0.096	(0.951)	-1.309	(0.681)	-3.952	(0.620)***	-0.355	(0.687)
E1	NL , OMT	-0.116	(1.847)	-2.806	(0.856)	-0.188	(0.932)	**0.898**	(2.376)	**0.558**	(1.152)	-0.343	(0.068)***
	UK , SAGE	-0.215	(0.693)	-4.670	(0.362)***	**3.089**	(0.781)	**0.074**	(1.029)	-0.459	(1.992)	**0.579**	(0.598)
H1	NL , OMT	**0.474**	(1.127)	-1.154	(0.481)	-1.355	(0.570)	-0.365	(0.424)	-0.256	(0.762)	**0.438**	(0.828)
	UK , SAGE	-3.186	(0.497)***	-1.359	(1.159)	**0.667**	(0.767)	**0.295**	(1.471)	-0.603	(0.547)	-0.876	(0.653)
H2	NL , OMT	**1.738**	(0.767)	-0.898	(0.515)	-0.322	(0.451)	**0.618**	(0.412)	**0.368**	(0.518)	**0.987**	(0.588)
	UK , SAGE	-0.779	(0.621)	**1.422**	(0.989)	-0.965	(0.317)	**0.290**	(0.348)	**1.513**	(1.722)	**1.068**	(0.567)
H3	NL , OMT	**0.739**	(1.083)	-1.334	(0.394)	**0.477**	(0.676)	-0.366	(0.265)	**2.526**	(0.928)	**3.034**	(0.739)
	UK , SAGE	-6.571	(2.692)	-2.482	(0.569)**	-1.780	(1.030)	-2.722	(0.726)*	**2.629**	(1.619)	-2.363	(0.464)***
H6	NL , OMT	-0.292	(0.898)	-0.845	(0.814)	-0.897	(0.450)	-0.814	(0.620)	-1.016	(0.913)	**1.516**	(0.816)
	UK , SAGE	-1.357	(0.343)*	-1.940	(0.813)	-2.123	(0.450)***	-0.564	(0.271)	-2.916	(0.454)***	**0.225**	(0.764)
H7	NL , OMT	**1.731**	(1.501)	**1.136**	(0.535)	-0.836	(0.336)	**1.767**	(0.813)	**0.120**	(0.714)	-0.727	(0.552)
	UK , SAGE	**1.388**	(1.405)	**0.096**	(0.273)	-0.316	(0.397)	**0.082**	(0.493)	-0.606	(0.544)	-0.916	(0.961)
H8	NL , OMT	**2.573**	(0.712)	**0.197**	(0.654)	**0.352**	(0.554)	**0.167**	(0.581)	**0.435**	(0.498)	-1.282	(0.532)
	UK , SAGE	**2.791**	(1.240)	-2.333	(0.984)	**0.577**	(1.598)	-1.124	(0.526)	-0.872	(1.54)	**0.150**	(0.738)

Note. Standard errors are shown in round brackets below the coefficients. Positive estimates are shown in bold.

Note. If a standard error is labeled *, **, or *** its coefficient is significant at a 5%, 1% or 0.1% level respectively, after applying the Bonferroni-Holm correction (Abdi, 2010).

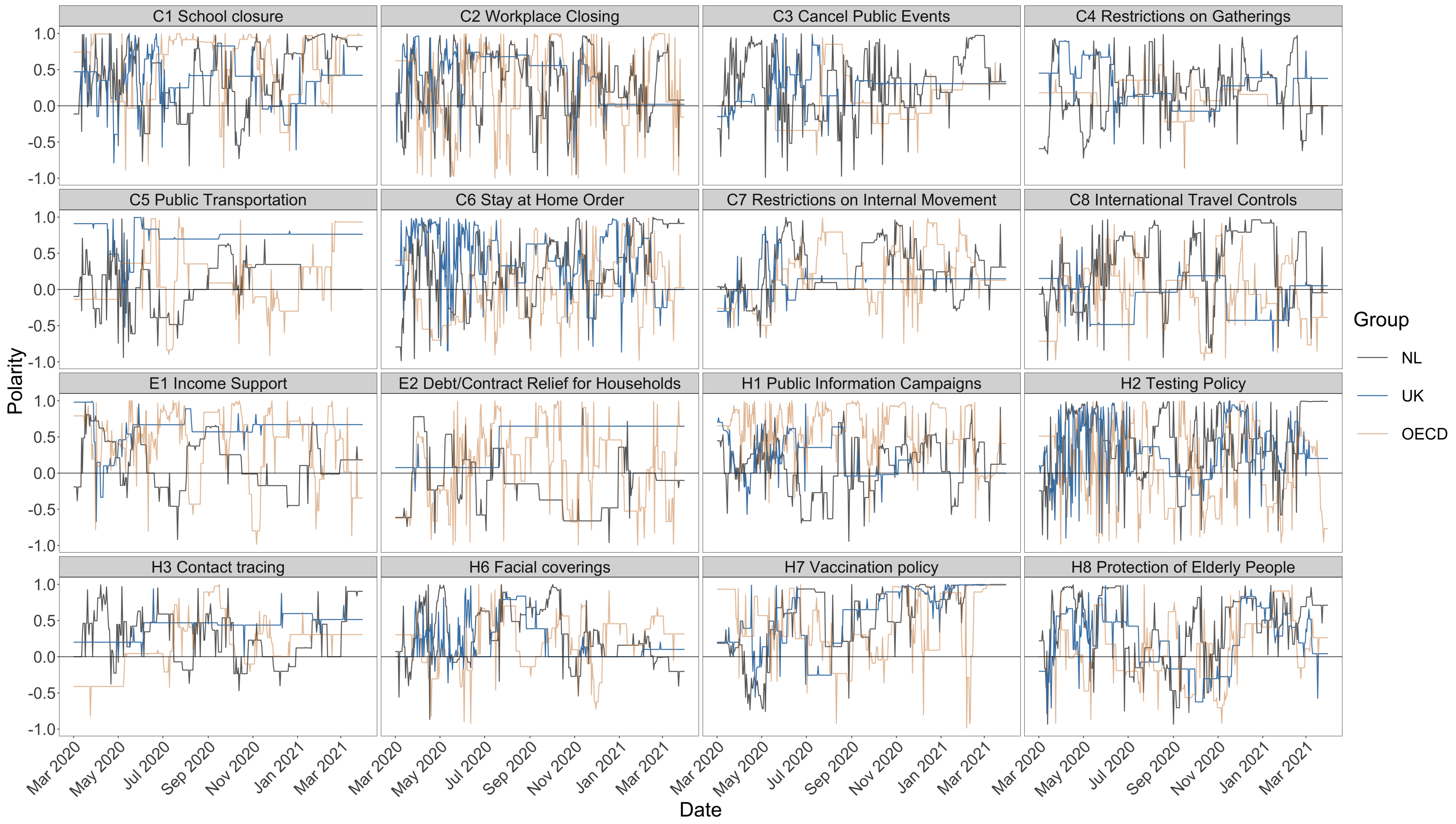

Figure A.33. Sentiment in press conferences

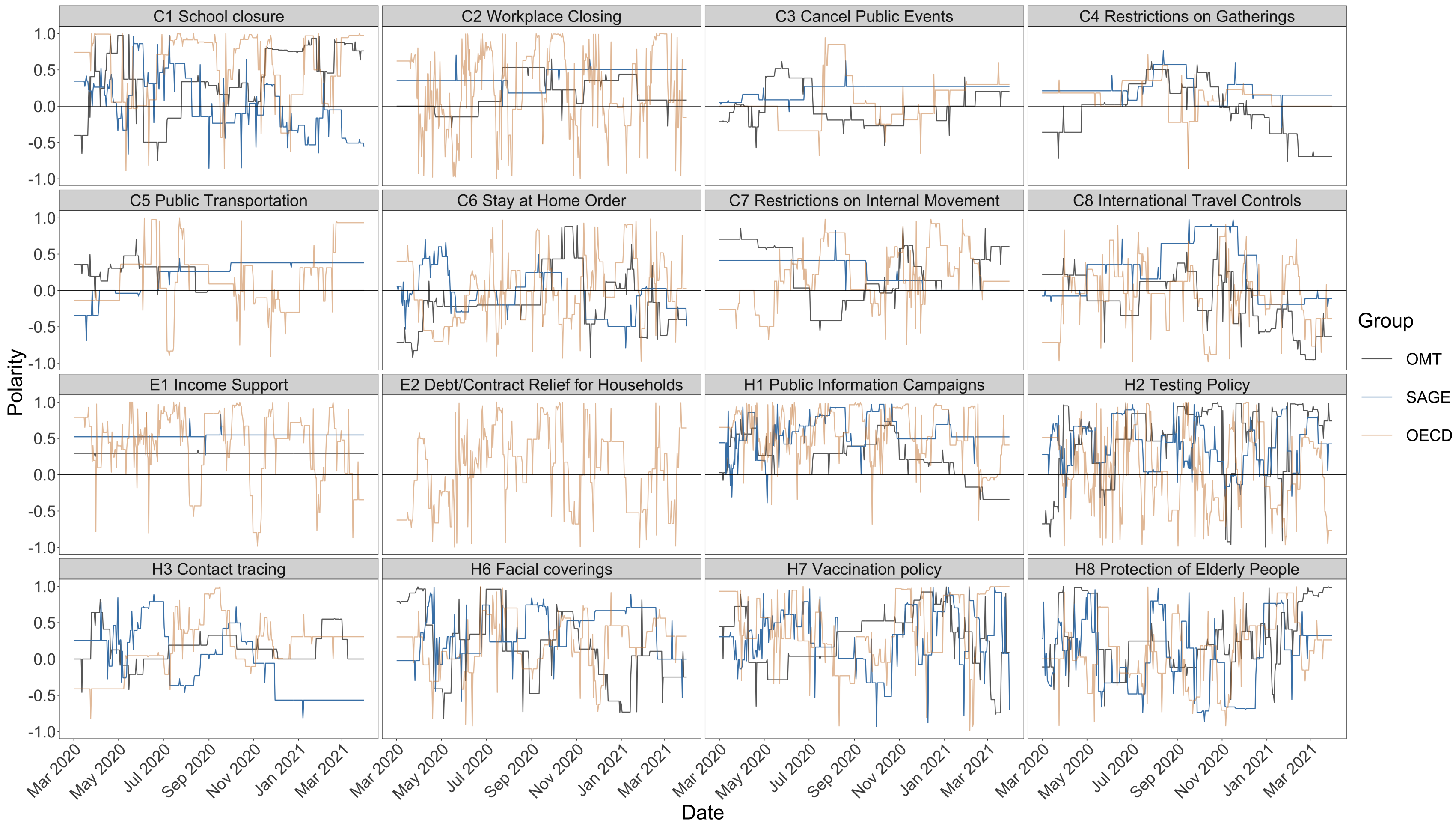

Figure A.34. Sentiment in meeting notes

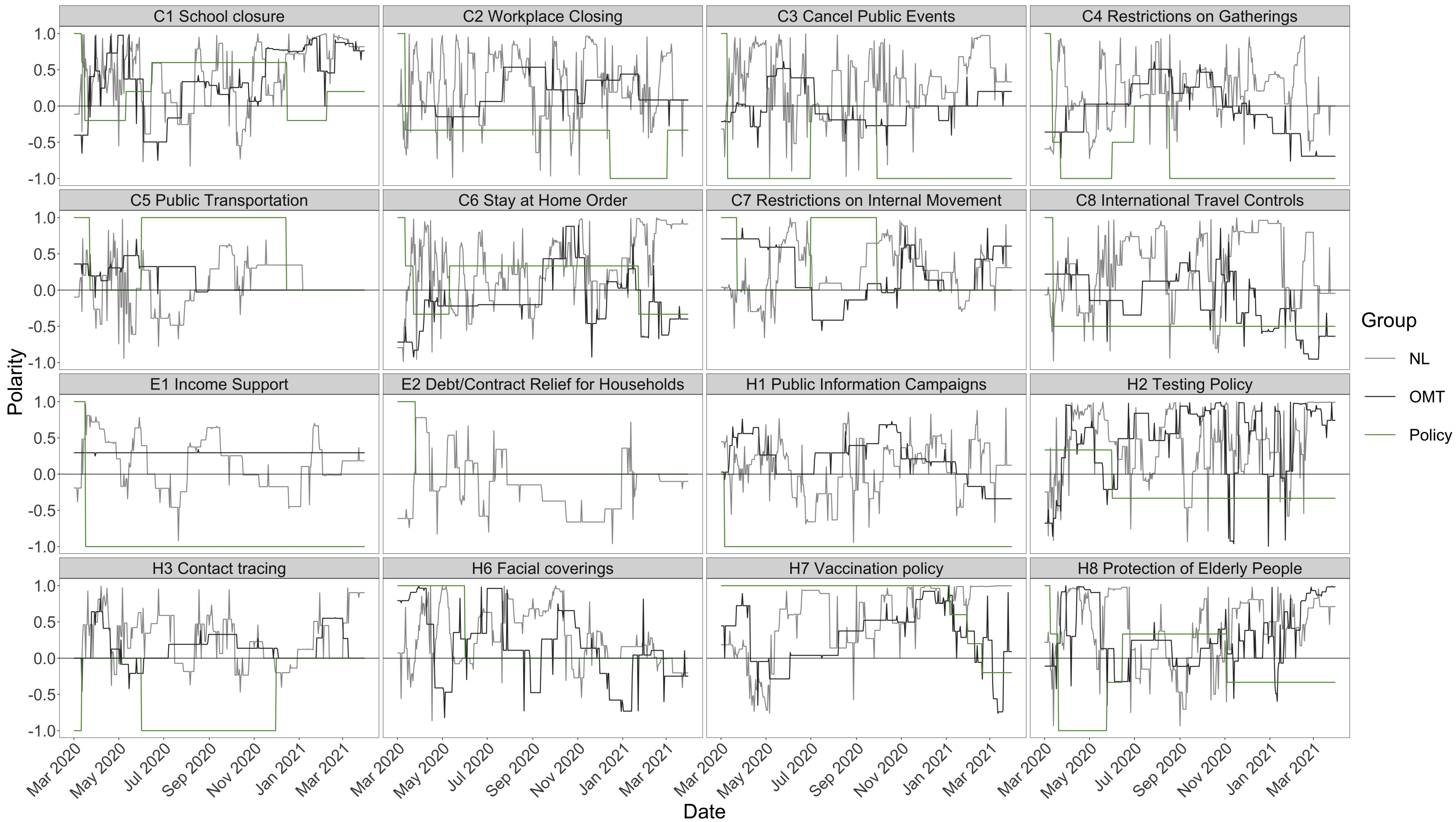

Figure A.35. Sentiment in press conferences

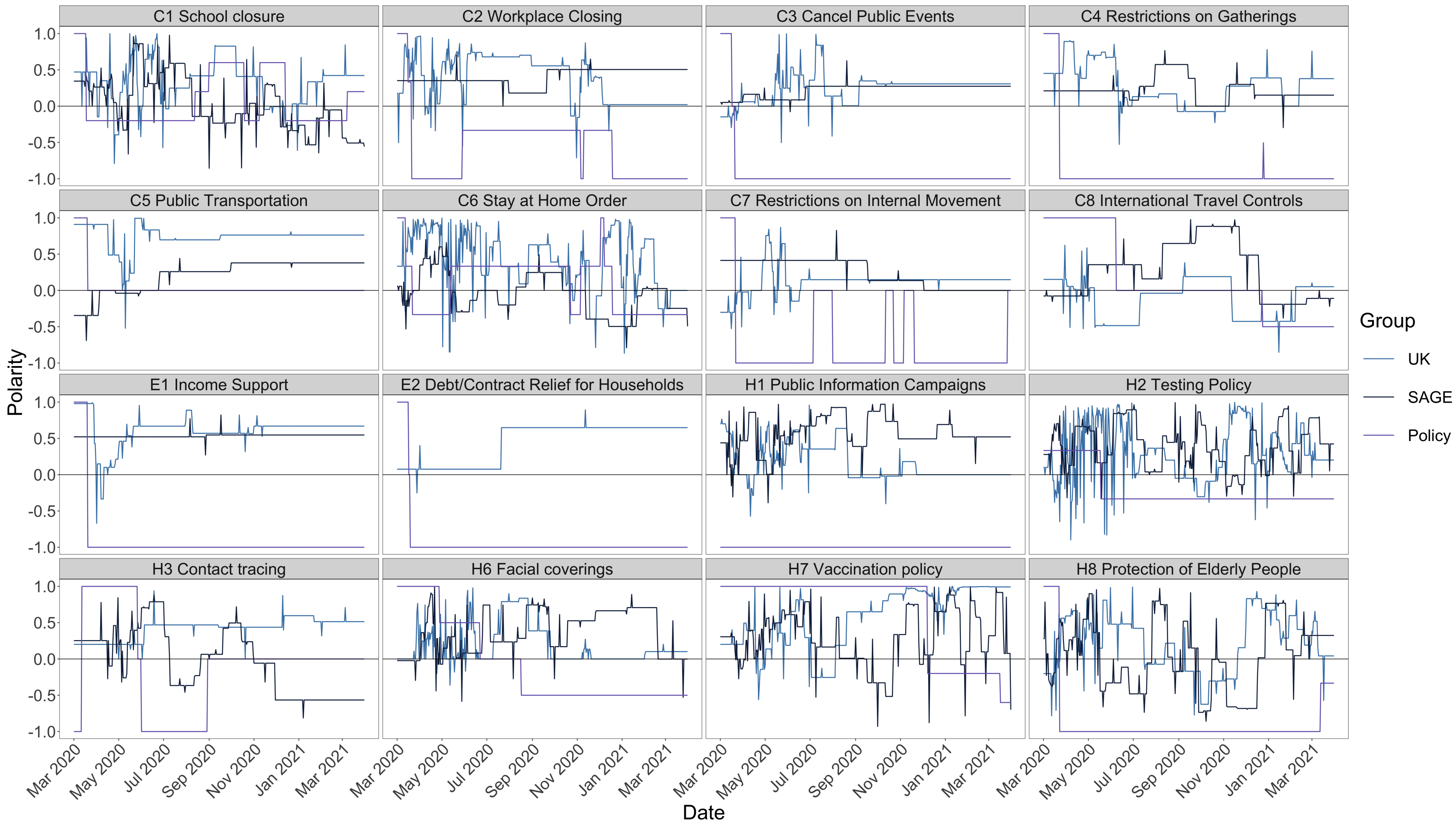

Figure A.36. Sentiment in meeting notes

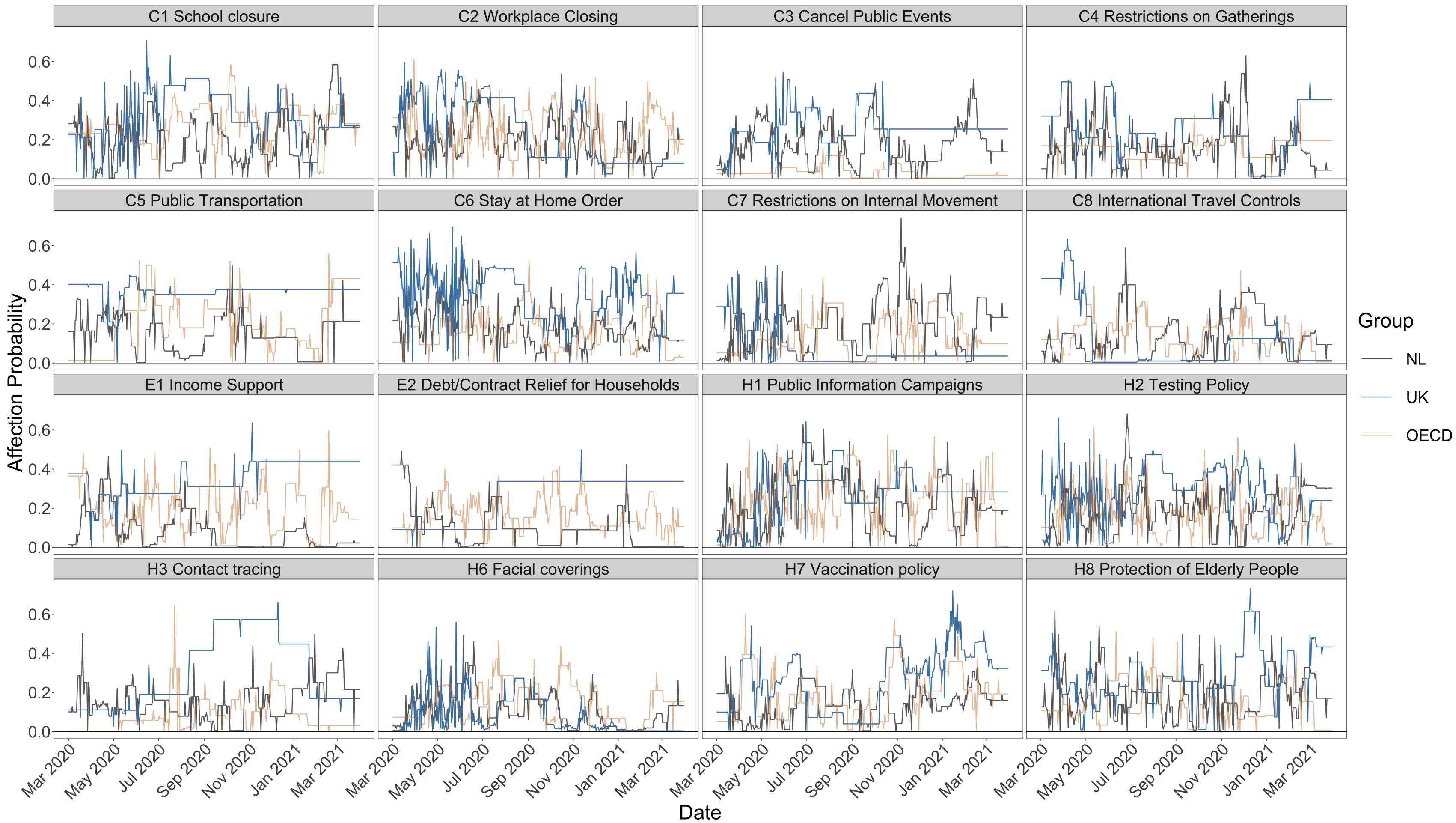

Figure A.37. Sentiment in press conferences

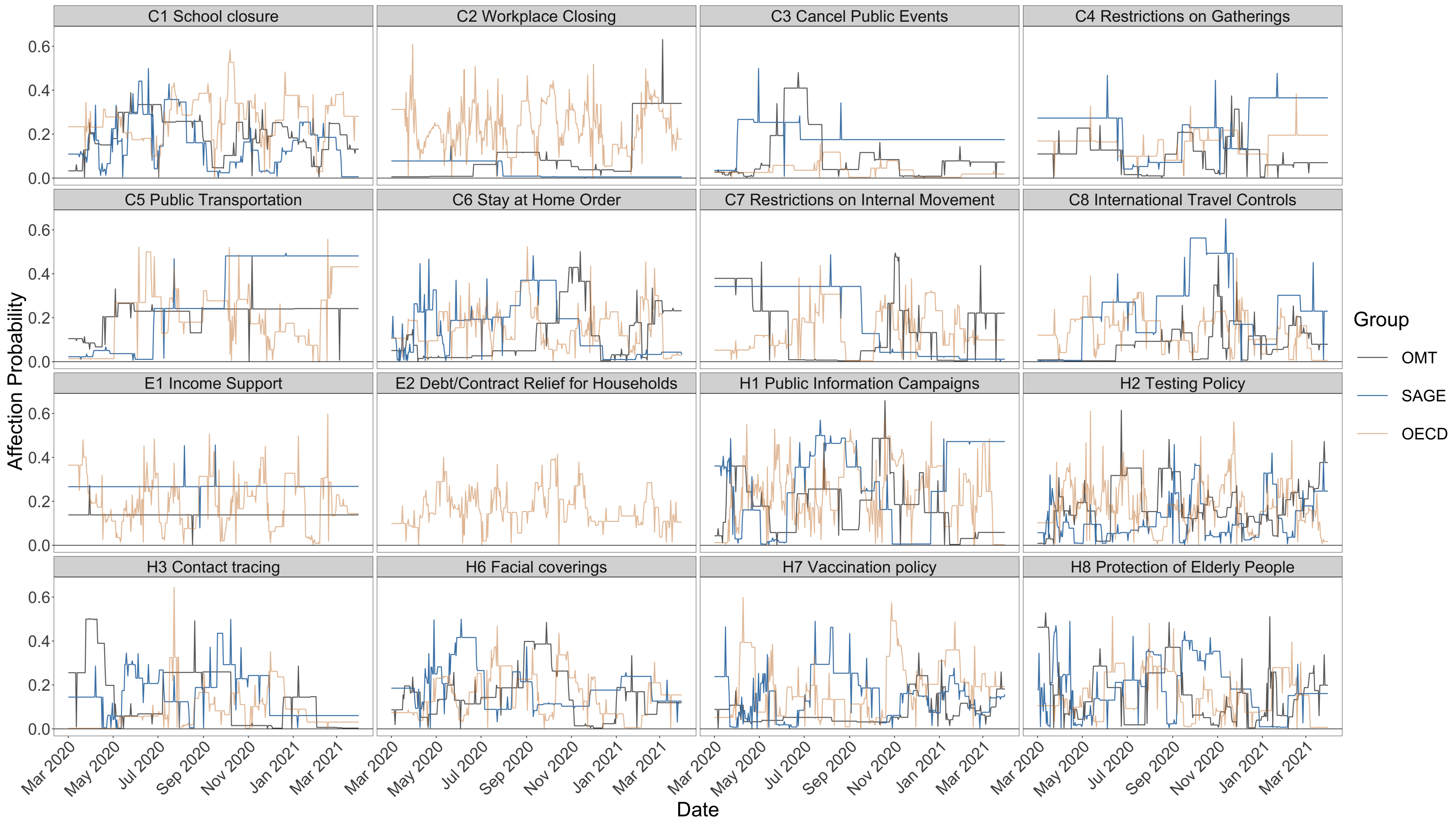

Figure A.38. Sentiment in meeting notes

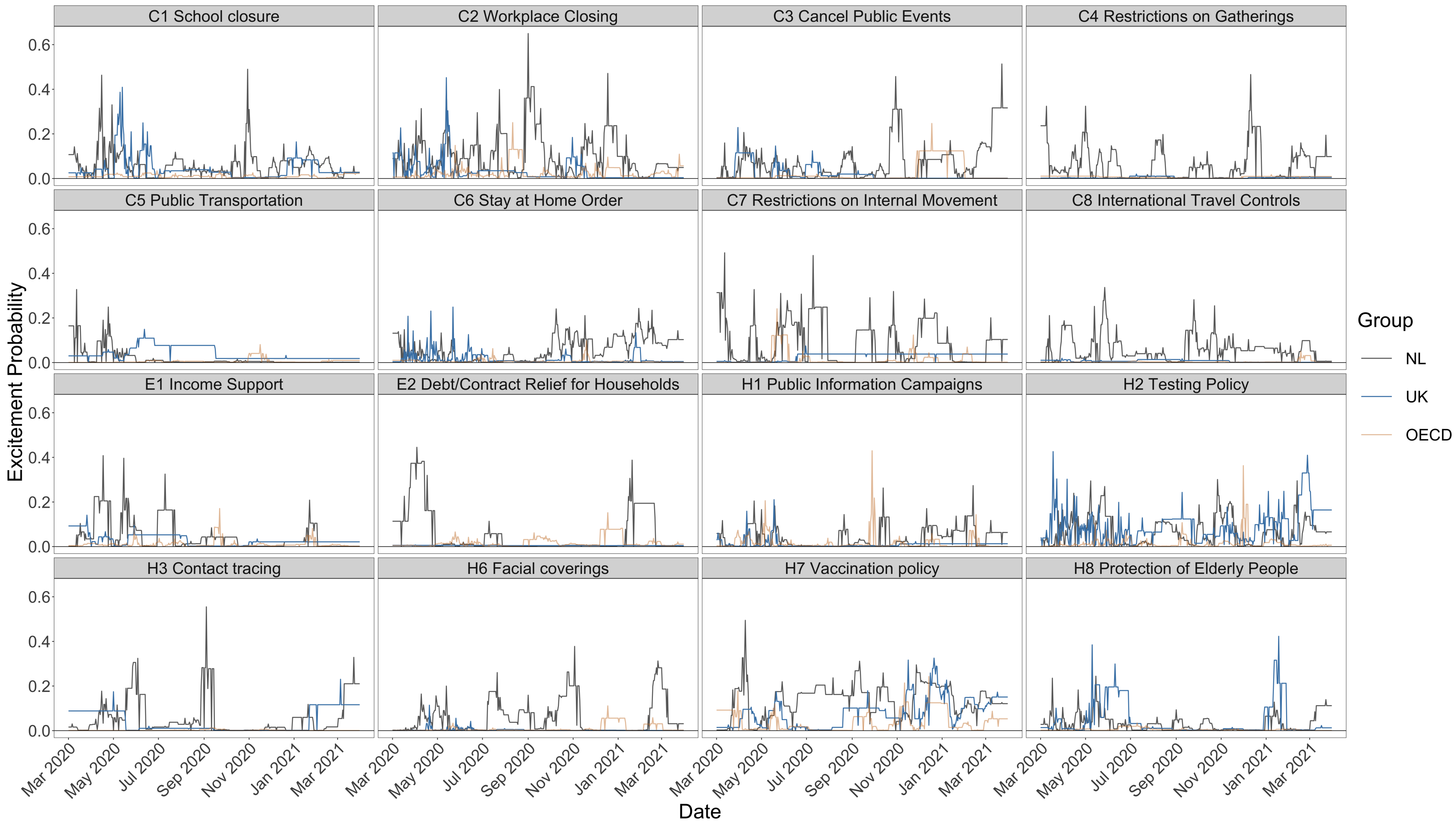

Figure A.39. Sentiment in press conferences

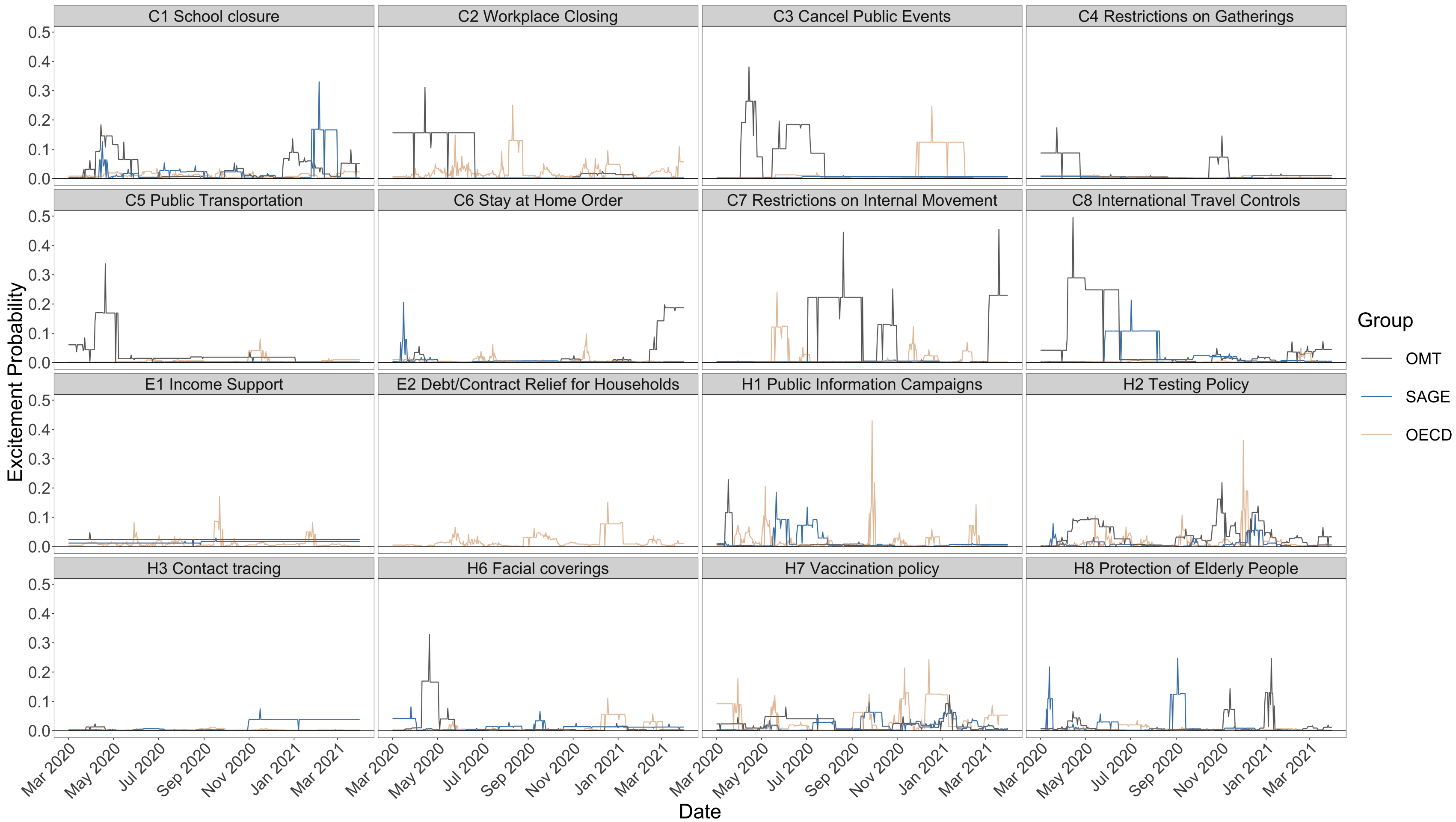

Figure A.40. Sentiment in meeting notes

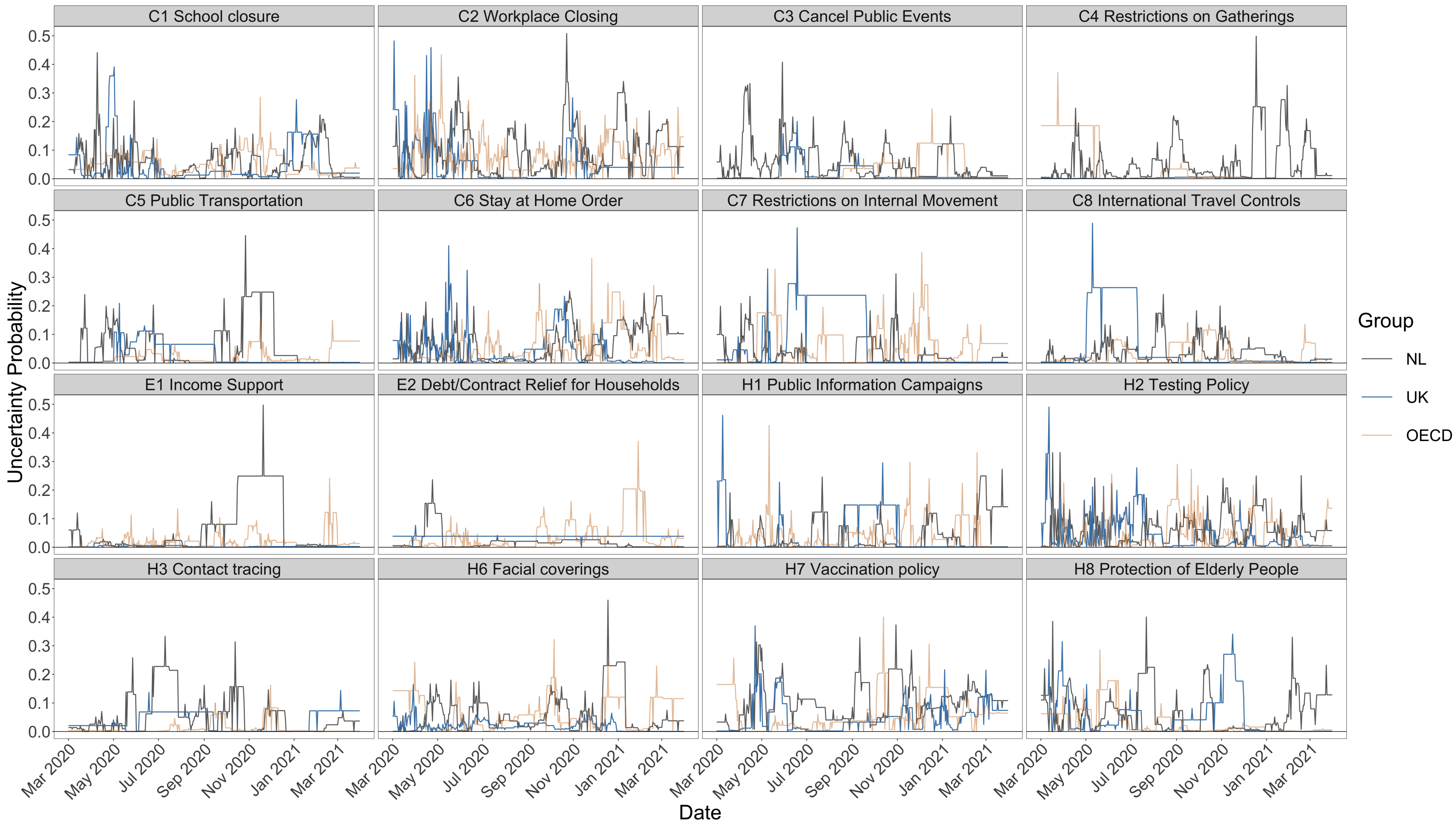

Figure A.41. Sentiment in press conferences

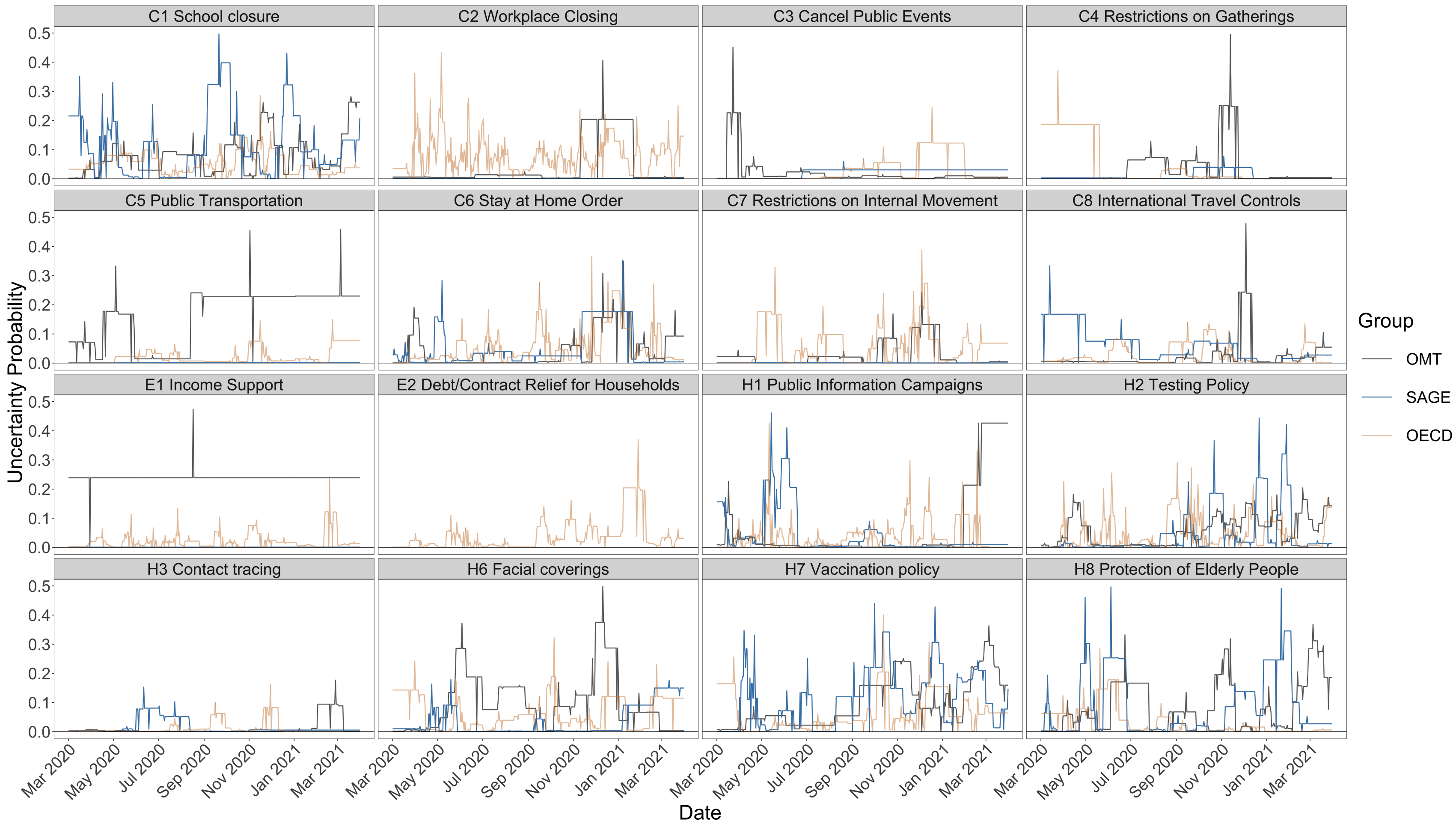

Figure A.42. Sentiment in meeting notes

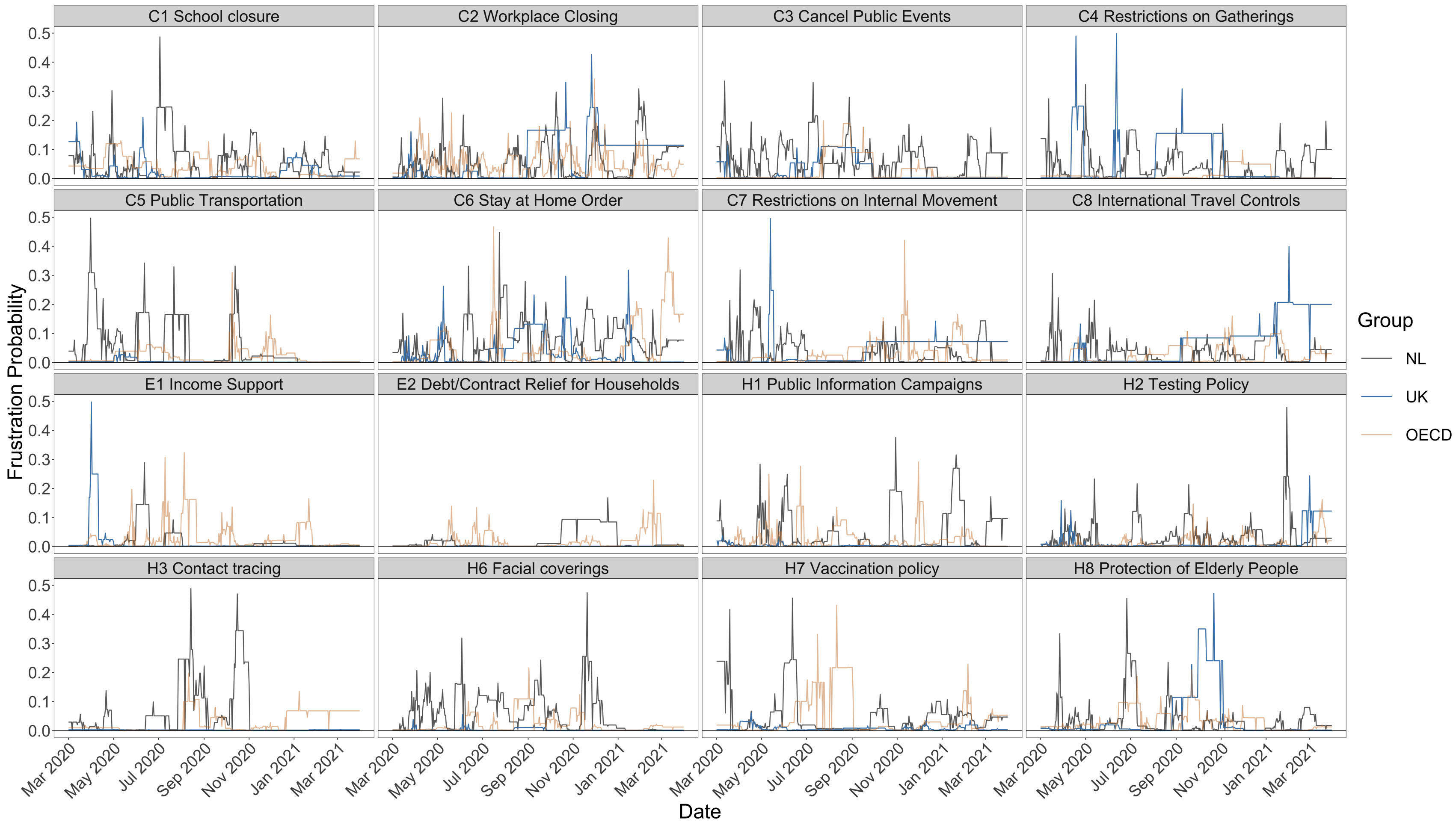

Figure A.43. Sentiment in press conferences

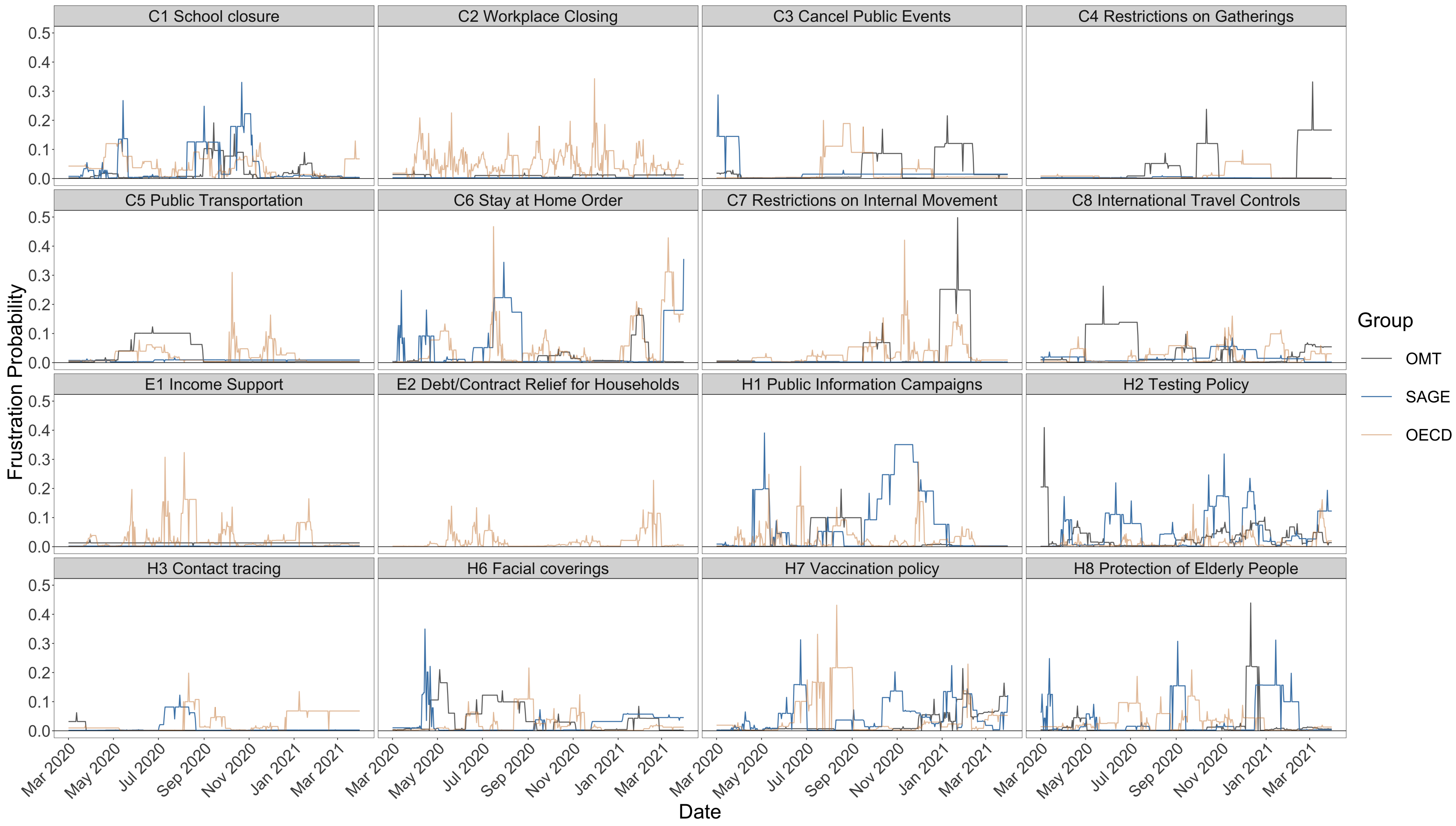

Figure A.44. Sentiment in meeting notes

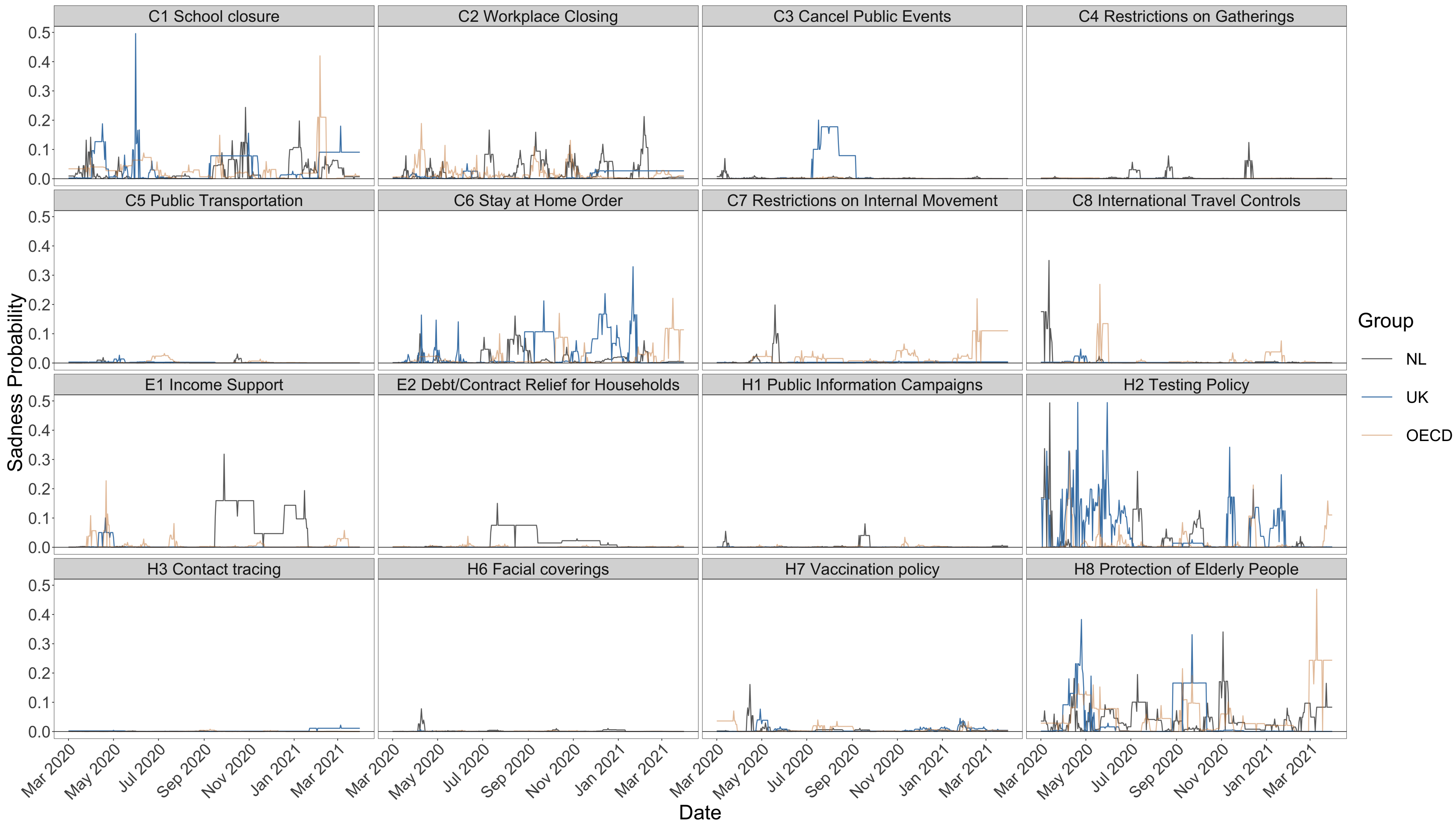

Figure A.45. Sentiment in press conferences

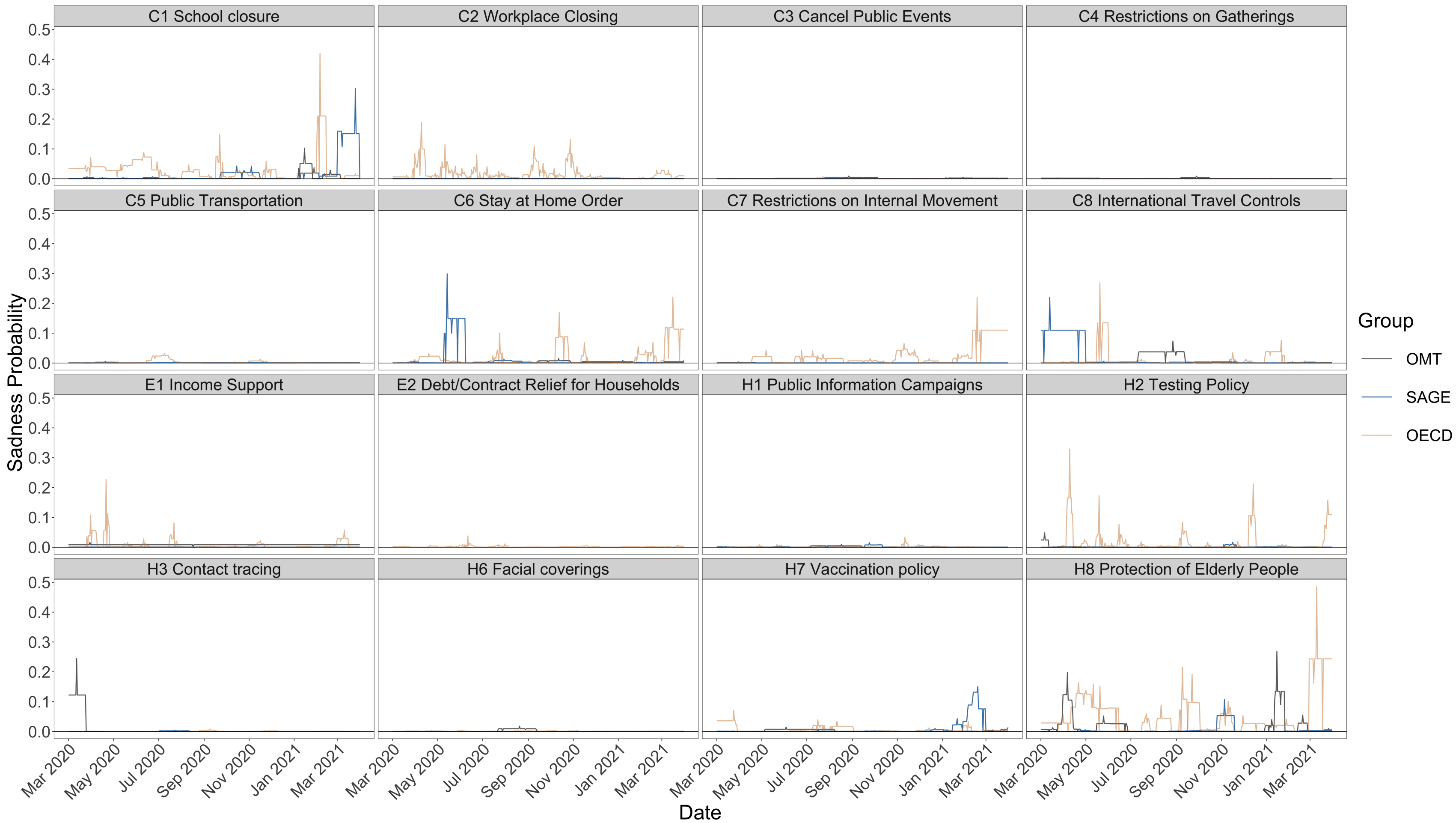

Figure A.46. Sentiment in meeting notes