



The underlying dynamics of the institutional crisis
of the Dutch Tax and Customs Administration:
An unsupervised machine learning approach.

Vera Leijtens

EasyChair preprints are intended for rapid
dissemination of research results and are
integrated with the rest of EasyChair.

June 10, 2019

The underlying dynamics of the institutional crisis of the Dutch Tax and Customs Administration: An unsupervised machine learning approach.

V.L. Leijten

Corresponding author e-mail address: v.l.leijten@tilburguniversity.edu

Tilburg Institute of Governance, Tilburg University, the Netherlands.

Prepared for the 2019 Public Management Research Conference, Chapel Hill, NC, United States, 11- 14 June.

Abstract

In this article, I explore the causes of institutional crisis by analyzing mismatches in news coverage for the Dutch Tax and Customs Organization (DTCA). An institutional crisis is defined in this article as the delegitimation process of a previously well-established public organization. Crucially, institutional crisis indicate a mismatch between the organization and its environment. In this process, the media is considered an important stakeholder. This article offers a longitudinal analysis of news coverage for the DTCA, a public organization that deteriorated from an exemplary authority in the 90s, to a much criticized and scrutinized organization. A combination of unsupervised machine learning and qualitative content analysis shows that two mismatches and a catalyst caused the institutional crisis of the DTCA. Based on this study, I conclude that in contrast to previous research, institutional crisis do not necessarily revolve around performance deficits.

1. Introduction

Public organizations, having a governmental function and performing public tasks, are exposed to judgments and criticism from stakeholders and society. If this criticism endures for a long period or suddenly erupts, the organization may become controversial. Consequently, the organizational legitimacy decreases, causing problems regarding the organizational performance and operational activities (Waeraas, 2019). While previously being a well-established public organization, the future of the organization may now be uncertain. In the literature, these situations are identified as institutional crisis (Boin & 't Hart, 2000).

The aim of this article is to provide an understanding of the causes of institutional crisis. Understanding what dynamics or mechanisms underlie institutional crisis is of great importance. Firstly, institutional crisis portend serious threats to public organizations (Ansell, Boin, & Kuipers, 2016). Second, public organization are financed with taxpayers money, crisis may offend legislators and society (Heath, 2010). Finally, a developed understanding may help in coping and even preventing these crisis from emerging (Ansell et al., 2016; Nohrstedt, Bynander, Parker, & Hart, 2018; Schmidt, Boersma, & Groenewegen, 2018).

In this article, institutional crisis is defined as the delegitimation process of a well-established public organization. Crucially, institutional crisis indicate a mismatch between the organization and their environment, meaning that the organization is not (perceived as) conforming to expectations, norms and values of the environment (Suchman, 1995). Society judges the organization based on their prevailing values and norms. If society perceives that the organization is not conforming to these norms and values, delegitimation may occur.

The media has an important role in this delegitimation process (Deephouse, 2000; Deephouse & Suchman, 2008; Pollock & Rindova, 2003). Media can both be considered as an

indicator of legitimacy and as a source (Carroll & McCombs, 2003; Deephouse & Suchman, 2008). In this article, I use the media as indicator of legitimacy. As illustrated by Dowling and Pfeffer (1975), and Deephouse and Suchman (2008), media represent public opinion, and norms and values prevalent in society. Likewise, media has been used to gauge (de)legitimation by examining the (amount of) negative media attention (Alink, 2006; Alink, Boin, & 't Hart, 2001; Ansell et al., 2016; Boin & 't Hart, 2000; Pollock & Rindova, 2003; Schmidt et al., 2018).

To date, research has mainly focused on the media as quantitative measure for institutional crisis (Alink, 2006; Alink et al., 2001; Boin & 't Hart, 2000). Scant attention has been paid to the content of news coverage and what dynamics can we identified that cause institutional crisis. Researchers have pointed to the importance of understanding the underlying dynamics of institutional crisis (Ansell et al., 2016; Nohrstedt et al., 2018; Schmidt et al., 2018). This article, therefore, fills this gap by offering a longitudinal analysis of the content of news coverage to identify mismatches causing institutional crisis, for the case of Dutch Tax and Customs organization.

The Dutch Tax and Customs organization (DTCA) offers a good opportunity to study institutional crisis. The organization was perceived as a good functioning and stable public administration in the 90s (Stevens, 2007). The Tax Monitor, an instrument used by the DTCA organization to measure what the taxpayer thinks about the DTCA, and how it could perform better (Alink & Kemmer, 2012), supports this. However, in the late 90s, early 00s, the Tax Monitor shows a decrease in legitimacy. Especially in 2007, there is a decline in legitimacy. Furthermore, the state secretary of Finance expressed his concerns regarding the viability and future of the administration (NTFR, 2014). In addition, incidents are reported on a frequent basis, and the

administration receives a lot of criticism and scrutiny from specific stakeholders (recently for example from the Ombudsman and the Council of State).

The research question of this article is as follows: which mismatches in media coverage can be identified that cause institutional crisis? Firstly, news items of 8 different newspapers over the last 24 years are analyzed. To be able to analyze a large number of news items (N = 5451), a state-of-the-art research method was applied in this research, i.e. topic modelling an unsupervised machine learning technique. This technique reveals which topics are being addressed over the years in the media in relation to the DTCA. In addition, a sentiment analysis identifies in which period in time the DTCA has received negative media attention. By combining the sentiment analysis and the topic modelling, the prevalent and dominant topics will be distinguished that are being discussed in the process towards institutional crisis. In the second phase, the data is studied more in-depth to understand what mismatches occurred.

This paper is organized as follows. Firstly, I discuss the theoretical background on institutional crisis. This is followed by an overview of the method used in this paper. Next, I discuss preliminary results. The final sections offers some preliminary conclusions.

2. Theoretical Background

Institutional crisis denote deep destabilizing crises that cannot be dealt with by short term actions (Boin & 't Hart, 2000; Boin, Kuipers, & Otten, 2000; Rosenthal and Muller, 1998). The foundations and basis of the public organization are subject of intense criticism. In other words, policies, performance, regulations, routines i.e. the building blocks of the public organization are being challenged. Institutional crisis, thus, do not revolve around single incidents or scandals, rather they entail broad-based challenges for the entire organization. The continuity and viability

of the organization are at stake. Examples of institutional crisis in the public sphere are: the Dutch law enforcement, diagnosed with a three folded crisis (crisis of norms, crisis of authority and an organizational crisis) following a row between two large police forces (Rosenthal and Muller, 1998; Boin and 't Hart, 2000). The Washington Department of Corrections in 2015, after releasing a large number of inmates before expiration of their sentence (Useem, Pacholke, & Mullins, 2016). Following Hurricane Katrina, the U.S. federal disaster response system (Ansell et al., 2016). Finally, earthquakes caused by gas drilling, causes an institutional crisis of the gas production policy sector in the Netherlands (Schmidt et al., 2018).

Often, institutional crisis emerge following a trigger event. A trigger event is an event that draws public and political attention to an organization (Baumgartner & Jones, 2010). The public backlash and criticism following a trigger event, generally, are not about the event itself, instead they revolve around the failures and deficiencies of the entire public organization (Ansell et al, 2016). Rarely, a trigger event emerges out of nowhere, rather it is the result of an accumulation of problems, deficits, inefficiencies and failures, which if remain unresolved cumulate in institutional crisis. Examples of trigger events are: Hurricane Katrina, exposing the defects of the once leading light of public administration the Federal Emergency Management Agency (FEMA), the influx of asylum seekers in the 20th century in Europe or more recently the earthquakes in the province of Groningen in the Netherlands, both revealing questionable policies and interests (Alink et al, 2001; Schmidt et al, 2018).

In times of institutional crisis, legitimacy declines (Boin & 't Hart, 2000). Hence, institutional crisis in this paper is defined as the delegitimation process of a well-established public organization. The concept of legitimacy is ambiguous. Although the concept exists for a very long time, there is no consensus regarding its definition (Anderson, 2015; Haldenwang, 2016; Suchman,

1995). To illustrate, legitimacy is often described as a normative concept referring to 'justness', 'appropriateness' and 'the right to rule', meaning that an organization rules in accordance with the accepted standards of moral and legal behavior. Along these lines, an organization is legitimate because it does what is best for society, seen from a public perspective (Gilley, 2006). At the same time, scholars also refer to the perception of legitimacy, addressing the degree of acceptance of an institute by society and/or politics (Anderson, 2015; Wallner, 2008).

Following the latter, in this paper, legitimacy is understood as *'the generalized perception or assumption that the actions of an entity are desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs, and definitions'* (Suchman, 1995, p. 574). Legitimacy is important for an organization: it signifies support for the organization, promotes its existence, practices and enables the organization to conduct its operations (Waeraas, 2019) (Bitektine, 2011; Deephouse & Suchman, 2008; Suchman, 1995; Waeraas, 2019; Wallner, 2008). Correspondingly, a lack of legitimacy may be problematic, affecting the organizations' performance and in the long term threaten the survival chances of the organization (Dowling and Pfeffer, 1975).

Organization are judged by their environment based on the prevailing norms and values. Public organizations, performing a governmental function and financially supported with taxpayers money, are highly subjected to these judgements (Waeraas, 2019). Consequently, organizations try to conform to social values, norms and expectations from the environment in order to deem legitimate (Ashforth & Gibbs, 1990; Deephouse & Suchman, 2008; Hirsch and Andrews, 1984; Suchman, 1995; Waeraas, 2019).

If the organization is not successful in (creating the perception of) conforming to the expectations, norms and values, a mismatch occurs between the organization and its environment.

Subsequently, the organizations legitimacy may be challenged and questioned by increased criticism and scrutiny, possibly resulting in institutional crisis.

To date, we do not know what mismatches underlie institutional crisis. In general, it is assumed that a gap between performance and expectations cause institutional crisis. However, performance can only explain institutional crisis to some extent. As illustrated by Ansell et al (2016), the Dutch prison sector experienced an institutional crisis in the 90s after a series of escapes, while at the same time their performance had increased as the number of prison escapes had never been so low. Thus far, little is known about the causes of institutional crisis. Specifically, research to date has not addressed what mismatches underlie institutional crisis. In this article, the media is used to address this question.

The media is an important platform where legitimacy struggles are being articulated (Deephouse, 2000; Deephouse & Suchman, 2008; Schmidtke, 2018). Researchers have been working with media to measure legitimacy for decades (Deephouse & Suchman, 2008). For example by counting negative and positive news articles (Pollock & Rindova, 2003; Alink, 2001, 2006; Resodihardjo, 2009), or with content analysis (Grafström, Windell, & Karlberg, 2015; Jonsson, Greve, & Fujiwara-Greve, 2009; Joutsenvirta & Vaara, 2009; Schmidtke, 2018).

There is, however, no consensus in research on the role of media in legitimacy (Pollock & Rindova, 2003). Media has a dual role in legitimacy research. On the one hand, media is an indicator of legitimacy because it represents public opinion. On the other hand, media can also be a source of legitimacy, as it affects public opinion (Deephouse & Suchman, 2008; Pollock & Rindova, 2003). Ivengar & Simon(1993), for example, have showed the latter role of media in their research to news coverage of the Gulf Crisis. They argue that the more prominent issues and events are covered in the media, the greater its weights in judgements. Additionally, in describing

social and political problems, people refer to events and issues as covered in the media (Ivengar & Simon, 1993). Furthermore, media may also influence public opinion, by focusing attention to specific issues, making sense of issues, or by framing actions of organization positively or negatively (Deephouse & Suchman, 2008; Grafström et al., 2015; Mccombs & Shaw, 1972; Pollock & Rindova, 2003).

However, in this article the media is used as indicator of legitimacy to explore what mismatches underlie institutional crisis. As illustrated by Dowling and Pfeffer (1975), and Deephouse and Suchman (2008), media represent public opinion, and norms and values prevalent in society (Dowling and Pfeffer, 1975; Deephouse and Suchman, 2008). Likewise, media has been used to gauge (de)legitimation by examining the (amount of) negative media attention (Boin & 't Hart, 2000; Alink, 2001; Alink 2006; Ansell & Boin, 2016; Pollock & Rindova, 2003; Schmidt, 2018).

3. Methods and Data Collection

In the first phase of the study, I used an unsupervised machine learning approach topic modeling with Latent Dirichlet allocation (LDA). Machine learning techniques are increasingly being used in social sciences. These techniques make it possible to analyze large amounts of data (Anastasopoulos & Whitford, 2018; Hollibaugh, 2018). Specifically, topic modeling is a method that analyzes latent themes in text data (Blei, 2012). For example, Hollibaugh (2018) used topic modeling to study changes in topic discussed by the Centers for Medicare & Medicaid Services and its predecessor agency, are affected by changes in the political environment. Roberts et al (2014) show in their article how structural topic modeling can be helpful in analyze open-ended responses.

LDA is a simplistic and powerful topic model (Blei, 2012), which is generally and widely used (Hollibaugh, 2018; Zhu, Witko, & Meier, 2018). It is a computational content analysis technique that can identify hidden topics in large amounts of texts without manual coding (Blei, Ng, & Edu, 2003). No prior annotations or labeling of the data is required; LDA combines an inductive approach with quantitative measurements. Topics emerge from the analysis of the text (Blei, 2012; Meier, 2018). The idea behind LDA is that documents exhibit multiple latent topics, with which each topic is characterized by a distribution over words (Blei, 2003).

LDA suggests that words carry semantic information, and that documents discussing similar topics, will use the same words. LDA finds these topics from recurring patterns of word occurrence in documents (Hollibrough, 2018). Simplistically said, LDA generates a list of the probability of words per topic and attempts to find the set of topics that best capture the documents (Törnberg & Törnberg, 2016). To illustrate, the current article might be represented by a mixture over topics as crisis, public organizations, and media. Each of these topics is a distribution over words. For example, the topic crisis might be defined by a list of high probability words as legitimacy, criticism, challenges et cetera.

More technically speaking, the collection of texts is called the corpus, items within the corpus are referred to as documents, and the words in the documents are terms. LDA suggests that the documents, in a corpus are probability distributions over latent topics. Analogously, topics are probability distributions over words (Blei, 2003; Meier, 2018). For each document and word in that document, LDA samples a topic from the respective topic distribution in the document, and a word from the respective topic's word distribution (Meier, 2018). For a more technical description of LDA see Blei et al (2003) and Blei (2012).

Furthermore, after revealing the latent topics in the corpus, a sentiment analysis is performed to identify positive/neutral and negative media articles. Sentiment analysis can be done by corpus based techniques, e.g. supervised machine learning or lexicon based techniques. In this article, the latter is used. Lexicon based techniques use sentiment dictionaries, and matches these to the data to determine the polarity (Kreutz & Daelemans, 2018). The sentiment analysis in this article will be performed by the Pattern package for Python¹. The sentiment function in the package returns a polarity value between -1.0 and +1.0. The value is based on the adjectives for the given sentences. In general, the pattern package shows that an article is positive if it returns a polarity score of +0.1. As articles may also be categorized as neutral, I categorized news articles with a <0.0 score as negative.

Following the literature on institutional crisis, increased negative media attention signifies a possible emergent institutional crisis (Boin & 't Hart, 2000; Alink, 2001; Alink, 2006; Resodihardjo, 2009; Schmidt et al, 2018). Thus, by using a sentiment analysis, possible periods of institutional crisis can be identified. Once this has been done, the sentiment analysis is combined with the topic modeling to show what topics were dominantly negatively discussed during these periods.

Finally, in the second phase of the study, the dominant negative topics in the institutional crisis period are explored more in-depth with qualitative content analysis. The textual data is manually coded, which is defined as *'the operations by which data are broken down, conceptualized, and put back together in new ways'* (Strauss & Corbin, 1990, p. 57). I start with open coding by asking questions as 'what is this news article about?', 'what are possible reasons

¹ See <https://www.clips.uantwerpen.be/pattern>

for the negative sentiment?', 'what is the role of the Dutch tax and customs administration?', 'what negative/positive consequences are reported?'

Data collection

Relevant newspaper articles were searched in Lexis Nexis Academic using the keywords 'Belastingdienst' (Tax authority) or 'Fiscus' (Tax authority) for the headline in the period 1st January 1995 – 31st December 2018. I choose to analyze newspaper articles as this for a long period was the prevailing source for the public (Vergne, 2011), and newspaper content and public opinion are closely aligned (Dowling & Pfeffer, 1975).

News articles were selected from eight national newspapers (Financieel Dagblad, NRC, Volkskrant, Telegraaf, Trouw, Algemeen Dagblad, Parool, Reformatisch dagblad). In total, the search generated 5860 news articles. After manually removing news articles concerning foreign tax agencies (N = 329), 5531 news articles remained.

Before the thematic structure of the text can be uncovered, the text collection must be made available for analysis (Hollibaugh, 2018). The first step is data cleaning, meaning that irrelevant items are removed. Information on Copyright, Project ID, journalists, cities et cetera were removed. Second, the texts needs to be further preprocessed. After data cleaning, the documents are tokenized, meaning that the documents are divided into word units. Then, all capital letters are converted to lowercase for term unification. This is followed by removal of punctuations and special characters. In the next step, stop-words are removed. Stop-words are words that appear frequent and are not necessary for document representation (see Meier, 2018). Subsequently, I used lemmatization to convert words to their lemma (e.g. the lemma for collection is collect). Finally, duplicates are removed (N = 80). This resulted in a final set of N = 5451 news articles.

4. The case of the Dutch Tax and Customs Organization

The Dutch Tax and Customs Administration (DTCA) is responsible for levying and collecting state taxes, and national social insurance contributions, detecting tax fraud, paying income related benefits for childcare, rent and health care, control and supervision of import, export and transit of goods, and supervision of compliance with laws and regulations (Alink & Kommer, 2012). The State Secretary of Finance is accountable for the entire tax domain (Borstlap & Joustra, 2017).

In the 1990s, the DTCA was known as one of the best-organized public organizations in the Netherlands (Stevens, 2007). Before the 1990s, the administration was known for its slow and bureaucratic methods (Alink & Kommer, 2012). However, the DTCA went through a major reorganization between 1987 and 1992 to better meet the demands of a changed environment. To illustrate, tax avoidance and tax fraud became a bigger issue, and there was a large increase in (Alink & Kommer, 2012). The administration changed its primary processes of levying and collection of taxes towards a more joined process, and the hierarchical bureaucratic management structure was adapted to an entrepreneurial orientation (Alink & Kommer, 2012).

Unfortunately, despite or because of this reorganization, the DTCA deteriorated from an example public organization to a problematic organization (Stevens, 2007). Crucially, in 2007 taxpayers' satisfaction with the overall performance of the DTCA dropped (see fig x), satisfaction with regard to the helpdesk by phone declined, and taxpayers increasingly believed that the expertise of the DTCA was declining (Alink & Kommer, 2012). Moreover, in 2007 the State Secretary of Finance expressed his concerns regarding the viability and the future of the DTCA, and the National Ombudsman questioned the performance of the DTCA (Stevens, 2007).

Recently, the situations remains unchanged. The DTCA experiences incident after incident (Borstlap & Joustra, 2018; Ombudsman, 2017), and the Ombudsman, Council of State, and The Court of Audit scrutinize and criticize the administration increasingly.

5. Preliminary results

Topic Modeling

Choosing the right value of N topics is an important task. A low N may represent too few or too broad topics, and a high N results in uninterpretable topics, or topics that overlap. To evaluate the quality of topics, I use topic coherence measure Cv. The coherence is determined by measuring how strong the top 15 words support each other in the corpus. According to Syed and Spruit (2017), Cv achieves a high correlation with human interpretation. Cv has four steps: 1. Divide the data in pairs of words, 2) calculating the probabilities of words and word pairs, 3) calculation of a confirmative measure, which indicates how strongly a word set supports another word set, and 4) aggregation of the confirmative measure into an overall coherence score (Syed & Spruit, 2017).

To determine the right value of N topics, I created 30 LDA models, by varying the number of topics, from 10 to 30 with an interval of 1, and 30 – 70, with an interval of 5. The LDA models were created with the Python library Gensim. Figure 2 and 3 show the obtained Cv coherence scores for all the LDA models created. A visual inspection of the figures shows that the LDA models with N = 30 – 70 achieve lower mean coherence scores in general than the LDA models with N = 10 – 30. Only the LDA model with 55 topics shows a higher mean coherence score. However, a visual inspection of this model showed a lot of overlap between topics. In figure 2 the highest topic coherence is for N = 27 topics. However, these topics as well show a lot of overlap. After inspecting the remaining highest mean coherence scores for the topics, I choose to ran the

LDA model again for $N = 16, 17$ and 22 topics. A visual inspection of the topics resulted in $N = 16$ topics. Figure 4 shows the top 10 relevant words per topic.

Topics are represented by series of words. LDA does not assign a topic name to the words, it is up to the researcher to interpret what overall topic can be assigned to the words. Figure 4 shows the top 10 relevant words per topic. For some topics, the overall topic can be easily derived from the relevant words. For example, topic 2, 3, 5, 9, 10, 11, and 13. The words for these topics are evident (for example topic 2 has words as suspicion, criminal, justice, lawyers, and suspect. Topic 3 contains Supreme Court, objection, verdict, and judge. Both topic 9 and 13 include words referring to Dutch soccer clubs, and topic 9 in particular to transfers).

To interpret topic 4, 6, 7, 11, 12, an understanding of the Dutch context is needed. Topic 4 contains words as parliament, State Secretary, problems, employees, and authority. In addition, the topic has individual names of ministers and State Secretaries (Wiebes, Jager, Weekers). Further inspection of this topic shows that this topic also includes words as 'mistakes', 'reorganization', and 'National Audit Committee'. This implies that the topic revolves around problems with the DTCA. For that reason, topic 4 is labeled accordingly. Next, I labeled topic 6 as 'Tax rates', because it includes words as boxes (referring to the boxes or categories in which income for taxation can be divided), deduction, taxation, rates and interest. Topic 7 contains names of different banks (e.g. ABN AMRO, ING), investments, and shares, and is therefore categorized as 'Banks'. To determine what topic 11 is referring to, the most relevant words ('multinationals', 'Panama Papers', 'tax avoidance', and 'rulings') have to be seen in combination with the other words. The other words in this topic are 'the Netherlands', 'Ireland', and 'Starbucks'. This implies that this topic is does not merely refers to tax avoidance, but to 'Tax Havens'. Next, the words in topic 12 mostly refer to certain incidents of fraud in the Netherlands. To illustrate, some words refer to

fraud in the construction industry (construction companies, construction fraud, and off shore accounts). Other words refer to illegal tax deals people from trailer parks made with the DTCA (Vinkenslag (the name of a former trailer park), trailer park, and haven).

The remaining five topics are more difficult to interpret. For all topics, a further inspection of the full topic list is required. Firstly, the words in topic 1 do not seem to be related. The topic includes words as China, advertisement, villages, hospitals, and individual names of researchers. Further inspection of other words included in this topic did not lead to a clear coherent picture of what topic is represented. Hence, this topic reflects a residual category, and is labeled as 'other'. Topic 8 includes words as Energy Company, Ahold, AEX, concert hall, Nuon, products. The words imply that this topic revolves around companies. Words in topic 14 are bankruptcy, creditors, inventory, and insolvency practitioner, and is therefore labeled as 'Bankruptcy'. For topic 15, the most relevant words hint that the topic might revolve around money. However, other words in this topic describe an issue between the DTCA and parents who received childcare allowance (day care, parents, and defaulter), therefore I chose to label the topic as 'Childcare allowance'. Finally, topic 16 includes words as church, public benefit organization, foundation, and donation, implying that this topic refers to 'Funds'.

After labeling the topics, I aggregated the topic shares per topic per year and compared each topic with the total each year. The topics Funds, Bankruptcy, Soccer, Soccer Transfer, and Companies were removed because their average shares accounted for 1% or 2%. Finally, the topic 'other' has an overall share of 8%, with a high of 11% in 2016 and a low of 3% in 2001. However, because it is not clear what this topic is about, this topic is removed as well.

The remaining topics are displayed in figure 6. The topics Tax Haven and Fraud have an average share of 2% - 3%, but both have an evident peak in share. The topic Fraud in 2004 with a

share of 12% and the topic Tax Haven in 2017 with 10%. The topic Criminals, and Banks have both an average share of 7%. Criminals has a low of 4% in 1999 and a high of 13% in 2012. Banks has a low of 1% in 1999 and a high of 24% in 2000. The average shares of the topic Court, Savings Abroad, and Childcare Allowance is 8%. Court has a low of 6% in several years, and a high of 13% in 1996. Savings abroad has a low of 4% in 1998 and a high of 17% in 2003, and Childcare Allowance has a low of 3% in 1999 and a high of 14% in 2013. The topics Tax Rates, Tax Return and Problems have the highest shares. Tax rates has an average share of 12% with a low of 7% in several years and a high of 24%. The topic Tax Return has a low of 7% in 2017 and a high of 25% in 2005. The topic Problems has the highest average share of 18%, with a low of 8% in 1995 and a high of 35% in 2017.

Sentiment Analysis

After topic modeling, I performed a sentiment analysis on the full text of all DTCA news articles from 1995 to 2018. Following the literature on institutional crisis, an institutional crisis can be identified if there is an increase in negative media attention (Boin & 't Hart, 2000). The sentiment analysis shows that in 2007 there is a steep increase of negative media attention, comprising over 40% of total news coverage. Together with the concerns of the State Secretary of Finance expressed in 2007, regarding the future and viability of the DTCA, it can be assumed that in 2007 the DTCA experienced an institutional crisis.

Next, I combined the sentiment analysis with the topic modeling. Based on this I could identify three prominent negative topics: Problems, Tax return and Court. Figure 8 displays the number of articles that contain the topics Problems, Tax return and Court. Noteworthy, news

articles can exhibit multiple topics. Hence, it is possible that a single news article can represent the topics 'Problems', 'Court', as well as 'Tax Return'.

Qualitative Analysis

Based on the topic modeling and the sentiment analysis we know that the prominent topics discussed during the institutional crisis in 2007 are: Problems, Court and Tax Return. However, to understand what caused the institutional crisis, more context is needed. Therefore, I performed an additional qualitative analysis on the negative news articles exhibiting the topic 'Problems' between the period 2004 and 2007 (N = 267). I chose this period, because the first increase in negative news articles began in 2004, leading up to the institutional crisis in 2007. Second, I chose to focus on the topic 'Problems' for now, because this is the most discussed topic in these years.

From this analysis, two mismatches and one catalyst emerged as the cause of the institutional crisis. I will discuss these below.

Two mismatches one catalyst

A mismatch between (government) ambitions and organizational capacities

One of the major causes of the problems with the DTCA is the introduction of the 'Whale law' (in Dutch: de wet Walvis) in 2006. The goal of this law was to simplify the laws and regulations of the employee insurance. To achieve this goal several changes in (tax) laws had to be made. Secondly, an allowance authority was created within the DTCA, responsible for the allowance for rent, healthcare and childcare. Besides, tasks had to be transferred from the UWV (Employee Insurance Agency) to the DTCA.

While these changes had the aim to simplify regulations, and a cost reduction of 200 million per year for businesses, it actually became “a nightmare”. A year and a half after the introduction of the Whale law, it became apparent that the systems had major flaws. Consequently, tax returns were lost, businesses had to resubmit employee's data, and taxpayers ran into financial difficulties because of delays in payments. The minister of Social Affairs concluded in 2007 that the implementation of the law Whale with hindsight was irresponsible.

Another cause of the problems is a reorganization of the DTCA in the early 00s. The Director- General wished to end the hierarchical structure of the DTCA. Against the wishes of the employees, cutbacks were made, experts left, and the DTCA was reorganized into self-steering collectives. Consequently, the DTCA became less consumer-friendly, deadlines could not be met, employees had to overwork, and audits became less sufficient, causing many mistakes. The Netherlands Court of Audit indicated that the DTCA lost € 400 million tax revenues, because of the insufficient audits.

Overall, the introduction of new laws, regulations and a reorganization, albeit well intentioned, seemed to achieve little. The increase of tasks, the cutbacks on personnel, and the problems with IT systems, led to numbers of mistakes. Furthermore, the DTCA was no longer able to perform their core tasks: monitoring and enforcing taxation. As a result, taxpayers were at the receiving end, and the DTCA had a significant loss of tax income.

A mismatch in the asymmetric relationship

The second cause of the institutional crisis has its roots in the asymmetric relationship between DTCA and taxpayers. In general, the asymmetric relationship is both assumed and accepted: the DTCA has more (financial) resources and expertise, and taxpayers in turn, more information

regarding their tax position. Hence, both parties are interdependent (Gribnau, 2013). However, for the period 2004-2007, an increased gap in the asymmetry between taxpayers and DTCA caused a mismatch.

To illustrate, media report on a frequent basis that taxpayers were at the receiving end of mistakes made by the DTCA. Over the years, taxpayers repeatedly experienced financial difficulties because of delays and mistakes on the side of the DTCA. In addition, a number of times after the DTCA paid too much tax refund and allowances, because of errors, taxpayers had to refund the money with interest. An extreme example of the latter is the situation that the DTCA continued to pay allowances to a deceased person. While the family had for several times informed the DTCA about the death, the allowances did not stop. Eventually the family had to pay the money back with interest.

Another example of the gap in asymmetry is the lack of ownership of the DTCA for their mistakes. To illustrate, the DTCA collected a wrong amount of taxes, due to a misalignment between computer systems. However, instead of fixing the problem, taxpayers had to object to the tax bill, in order to reimburse their money. Furthermore, in 2006 the DTCA lost between 35.000 and 100.000 tax returns. Instead of being honest, the DTCA tried to conceal it, and requested taxpayers to resubmit their tax return, because the DTCA had never received any.

Furthermore, in 2007, several media report that tax advisors and taxpayers complained about the unreasonable gap in asymmetry between taxpayers and DTCA. It was perceived that the DTCA frequently used its means of power. The DTCA extended its own deadlines, while penalizing taxpayers for exceeded deadlines. Moreover, the DTCA time and again (wrongly) threatened taxpayers with penalties.

Following the analysis, taxpayers perceived that the DTCA could make any mistake, without any consequence, while mistakes on taxpayers' side lead to immediate penalties. Together with the lack of ownership for mistakes, the asymmetry was no longer perceived to be within reasonable limits, resulting in a mismatch between taxpayers and DTCA.

Culture as catalyst

The culture of both government and management created an environment in which the mismatches could rise without hindrance. For starters, both government and management ignored expert warnings on emerging problems. To illustrate, warnings on the implementation of the law Whale were ignored. For example, in 2005 the Netherlands Court of Audit warned the government for taking too little time on transferring tasks from the UWV to the DTCA. Following the Court of Audit, the schedule was too tight, endangering the reliability of the administration. Secondly, the Court of Audit concluded that insufficient agreements were made between authorities and the responsible ministries (Ministry of Social Affairs and Ministry of Finance) on exchanging and monitoring data. In addition, in 2003 computer experts warned the DTCA that bad project management, lack of knowledge, and miscommunication, endangered the implantation. Indeedn, years later, Minister Donner (Social welfare), had to admit that the implementation of the law Whale, was irresponsible.

Furthermore, despite signals from tax advisors, the labor union, and the interest group of senior tax officials regarding the problems caused by the reorganization, both government, and the director-general did not take measures to solve the problems. On the contrary, it took four years before government ended the hiring freeze, because they realized that the disturbances and work pressures were too high.

Another surprising observation that emerged from the data is that government in the period 2004-2007 had the tendency to minimize, deny or pass problems. To illustrate, a theme that caused a lot of public outcry in 2004 was the illegal agreements DTCA officials made with certain population groups in the Netherlands. In short, certain population groups (for example people living in trailer parks), refused to pay tax. Besides, DTCA officials were threatened if they tried to collect taxes. Consequently, DTCA officials made agreements with these people that either they paid a low share of their taxes or even nothing at all. Most striking is that the State Secretary of Finance firstly stated that these agreements were very incidental. However, a few weeks later, the State Secretary of Finance had to admit that this actually was a widespread practice, already known since 1996 and even founded on an internal instruction of the Ministry of Finance.

Unfortunately, it became apparent that the DTCA had more unresolved known issues. For example, the DTCA did not pursue a computer error regarding inheritance tax for six years. As a result, the DTCA lost € 6 million on tax revenue. More strikingly, when this became known in media, the DTCA referred to this loss as an inevitable accident.

In addition, as mentioned above warnings from the Netherlands Court of Audits with regard to the Law Whale were denied, as well as problems due to system failures, and responsibilities for problems in transferring tasks from UWV to the DTCA were passed.

6. Preliminary conclusion

In this article, I studied the causes of institutional crisis for the case of the Dutch Tax and Custom Administration. An institutional crisis is defined as the delegitimation process of a well-established public organization. Specifically, an institutional crisis indicates a mismatch between the

organization and its environment. Delegitimation can be gauged by examining (the amount of) negative news articles.

Through a combination of unsupervised machine learning and qualitative content analysis, I longitudinally analyzed news articles on the DTCA for the period 1995 – 2018. The unsupervised machine learning approach, topic modeling revealed that the prevalent topic in the period up to the institutional crisis in 2007 was 'Problems'. A content analysis of the negative news articles in the period between up to 2007 (starting from 2004), showed that two mismatches and a catalyst caused the institutional crisis. Contrary to previous literature, the roots of the institutional crisis did not lie in performance deficits, but rather in political ambitions, and operational activities.

Firstly, the mismatch between political ambitions and organizational capacities, created a situation in which the DTCA had to implement new tasks, with less capacity. As a result, the DTCA was no longer able to perform their core tasks: monitoring and enforcing taxation.

Secondly, the increased gap in the asymmetric relationship between taxpayers and DTCA was the second cause of the institutional crisis. The DTCA operated in a way that was not perceived to be within reasonable limits. Taxpayers were at the receiving end of mistakes of the DTCA, and the DTCA could make mistakes without any consequence, while taxpayers did get penalties for their mistakes.

Finally, the culture of the DTCA and government acted as a catalyst for the institutional crisis. It became apparent that on a structural basis, problems were minimalized, denied and passed. Strikingly, a number of problems were already known for years, but not solved. In addition, warnings from experts were ignored.

This study answers recent calls to research underlying dynamics of institutional crisis (Ansell et al, 2016; Norhstedt, 2018). Contrary to existing research, I found that institutional crisis

do not necessarily revolve around performance deficits (cf. Boin & 't Hart, 2001; Ansell et al, 2016), but factors as political ambitions, organizational capacity and operational activities are also a breeding ground for institutional crisis. Furthermore, the culture of an organization or government may act as a catalyst if problems are ignored, denied and not solved.

TABLES AND FIGURES

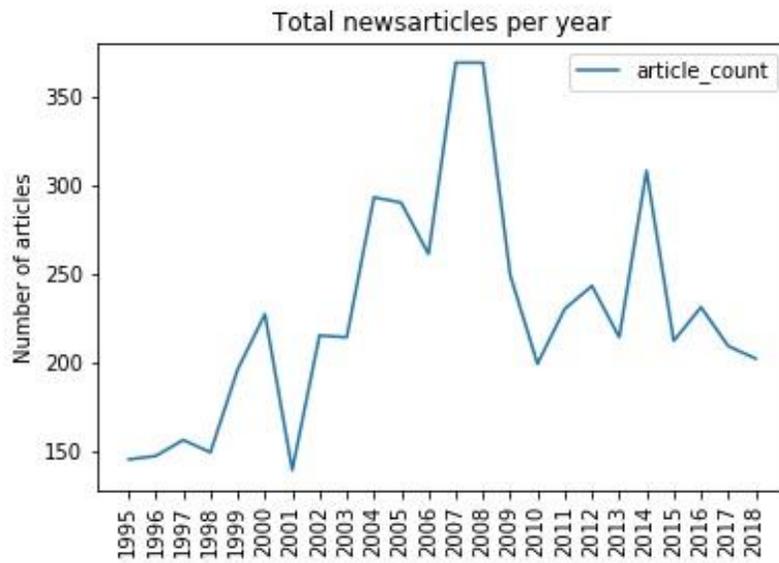


Fig. 1. Total number of news articles per year

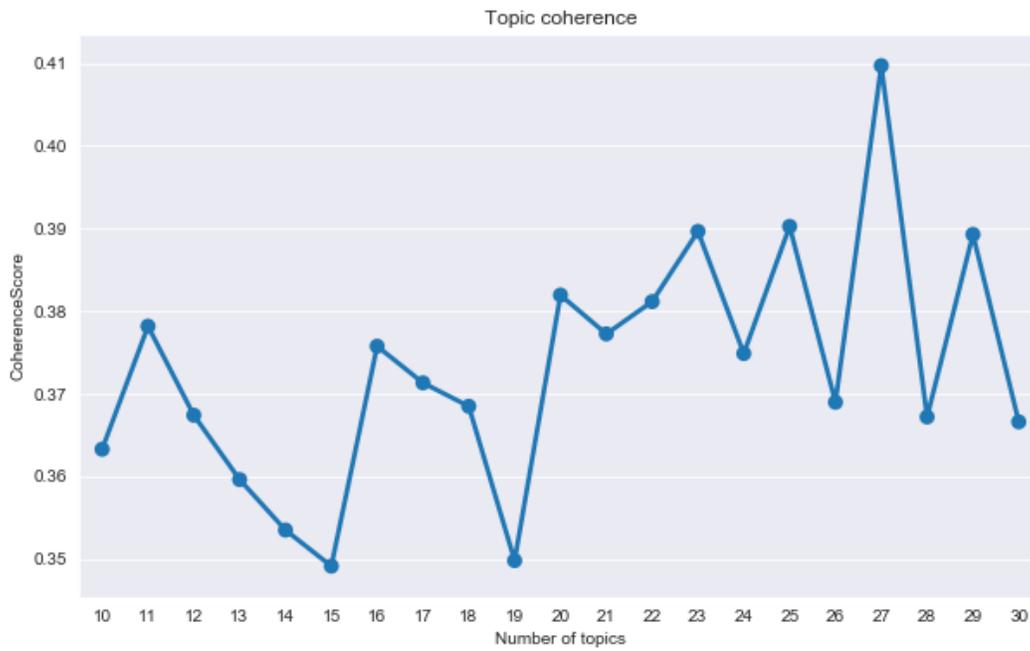


Fig. 2. Calculated Cv topic coherence score for LDA models with $K = \{10, \dots, 30\}$.

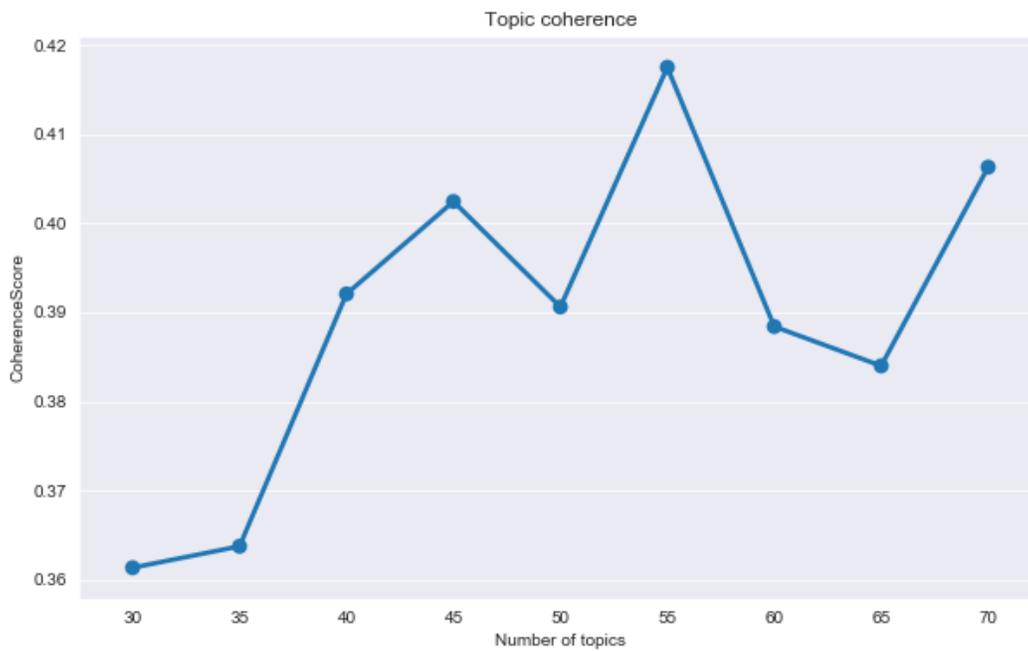


Fig. 3. Calculated Cv topic coherence score for LDA models with $K = \{30, \dots, 70\}$.

Topic	words with Relevance
0 Topic1	{stevensen, pearleen, baren, pomphoudersen, alberen, chineseen, beuvingen, reclameen, dorpen, joustraan}
3 Topic12	{vinkenslaggen, schaduwboekhoudingen, piramidespelen, airbnben, bouwbedrijven, woonwagenkampen, bouwfrauen, auen, woonwagenbewoners, wijnen}
1 Topic10	{invullen, werkgeversen, teruggaven, aangiffen, heffingskortingen, mensen, voorlopigeeen, formulieren, premiezen, socialeen}
10 Topic4	{staatssecretarissen, wiebessen, kameren, problemen, jageren, financiënen, cdaan, medewerkeren, weekersen, diensen}
12 Topic6	{rennen, waren, tarieven, optiezen, aftrekbaren, belasteen, kpnen, woningen, boxxen, aftrekken}
11 Topic5	{belgischeen, luxemburgen, zwarten, nederlandseen, buitenland, zwitserlanden, buitenlandseen, nederlandersen, europeseen, zwartspardersen}
6 Topic15	{miljarden, mrdn, kwartalen, miljoen, euroen, wanbetalersen, hfien, mlnen, opvangen, oudersen}
8 Topic2	{verdenken, fioden, valseen, crimineleen, frauen, advocaten, justitieen, tipgeveren, moszkowiczen, verdachen}
9 Topic3	{inspecteuren, hogeen, bezwaren, rechteren, corporatiezen, uitspraken, gemenen, autoen, raden, huurdersen}
7 Topic16	{beogennen, kerken, instellingen, giftenaftrekken, anbissen, anbien, scientologyen, giften, nutten, predikannen}
2 Topic11	{papersen, rullingen, rullingsen, horizontalen, trussen, panamaan, belastingontwijkingen, paarlbergen, oesoen, multinationalen}
13 Topic7	{aandelen, banken, ingen, fondsen, beleggingsfondsen, bodden, beleggersen, amroen, abnen, pvpen}
4 Topic13	{twennen, voetballen, dueen, feyenoorden, fortissen, ften, spelersen, clubben, clubsen, knvben}
5 Topic14	{postnl, fokkeren, delaan, mollen, dgaan, schuldeiseren, knpen, fien, curatoren, faillieen}
15 Topic9	{laudruppen, nsen, vliegeren, transfersen, naccen, nmaan, ajaxxen, shotaan, arveladzeen, spartaan}
14 Topic8	{energiebedrijven, escapeen, kammen, hummelen, producennen, aexen, aholden, jansen, concertgebouwen, griekseen}

Fig. 4. Top 10 relevant words per topic

Topic number	Label
Topic 1	Other
Topic 2	Criminals
Topic 3	Court
Topic 4	Problems
Topic 5	Savings Abroad
Topic 6	Tax rate
Topic 7	Banks
Topic 8	Companies
Topic 9	Soccer Transfers
Topic 10	Tax Return
Topic 11	Tax haven
Topic 12	Fraud
Topic 13	Soccer
Topic 14	Bankruptcy
Topic 15	Childcare Allowance
Topic 16	Funds

Fig. 5. Topic categories

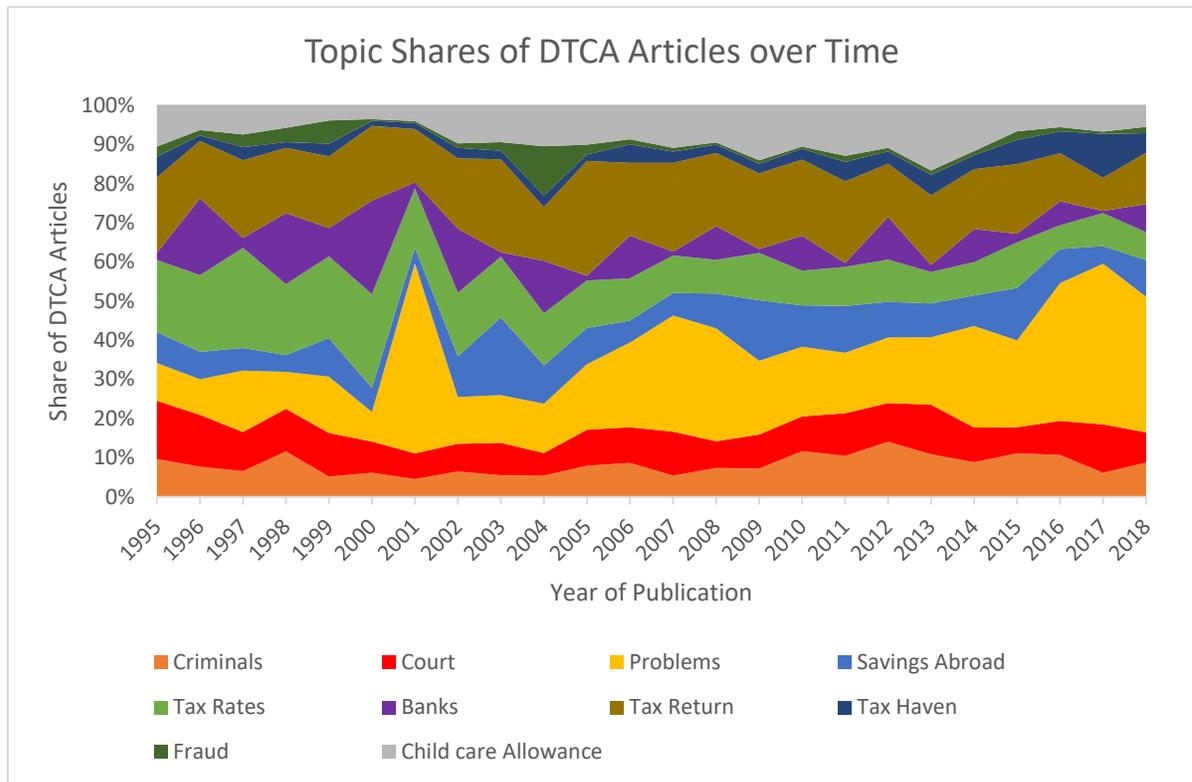


Fig. 6. Area graph of most prominent topics over time of all topics within each topic category per year.

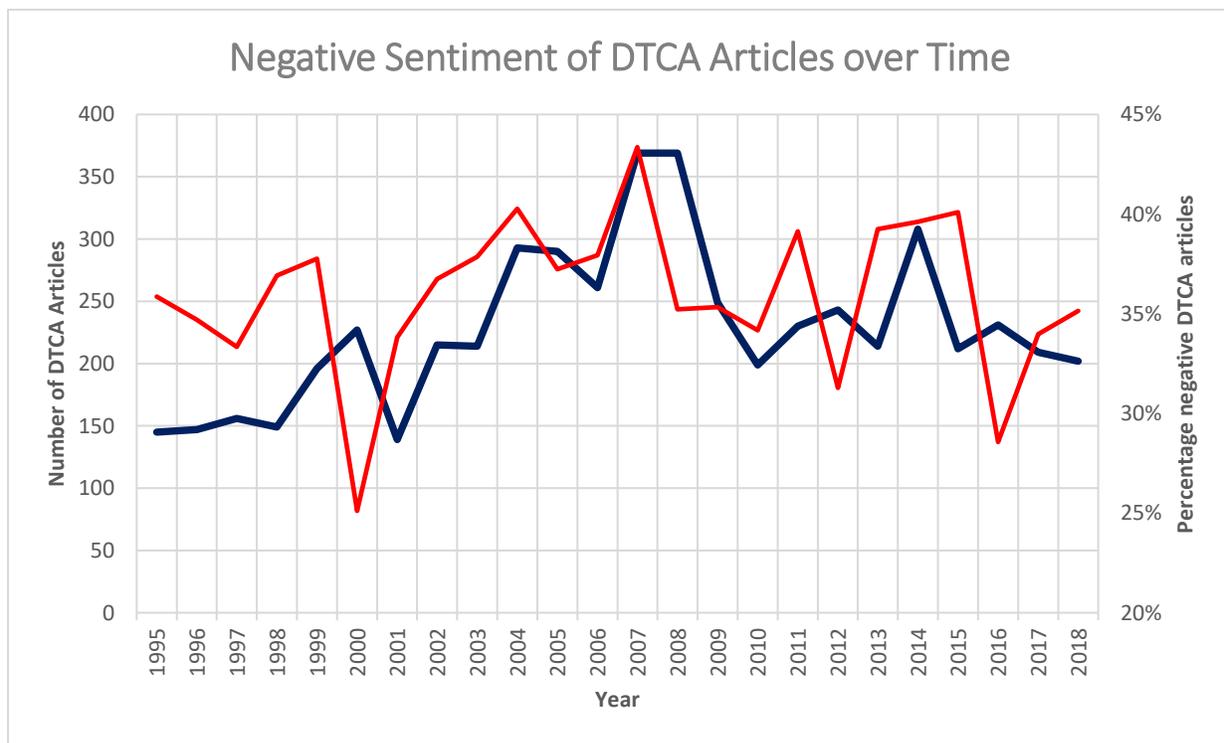


Fig. 7. Development of negative sentiment compared to the total number of news articles.

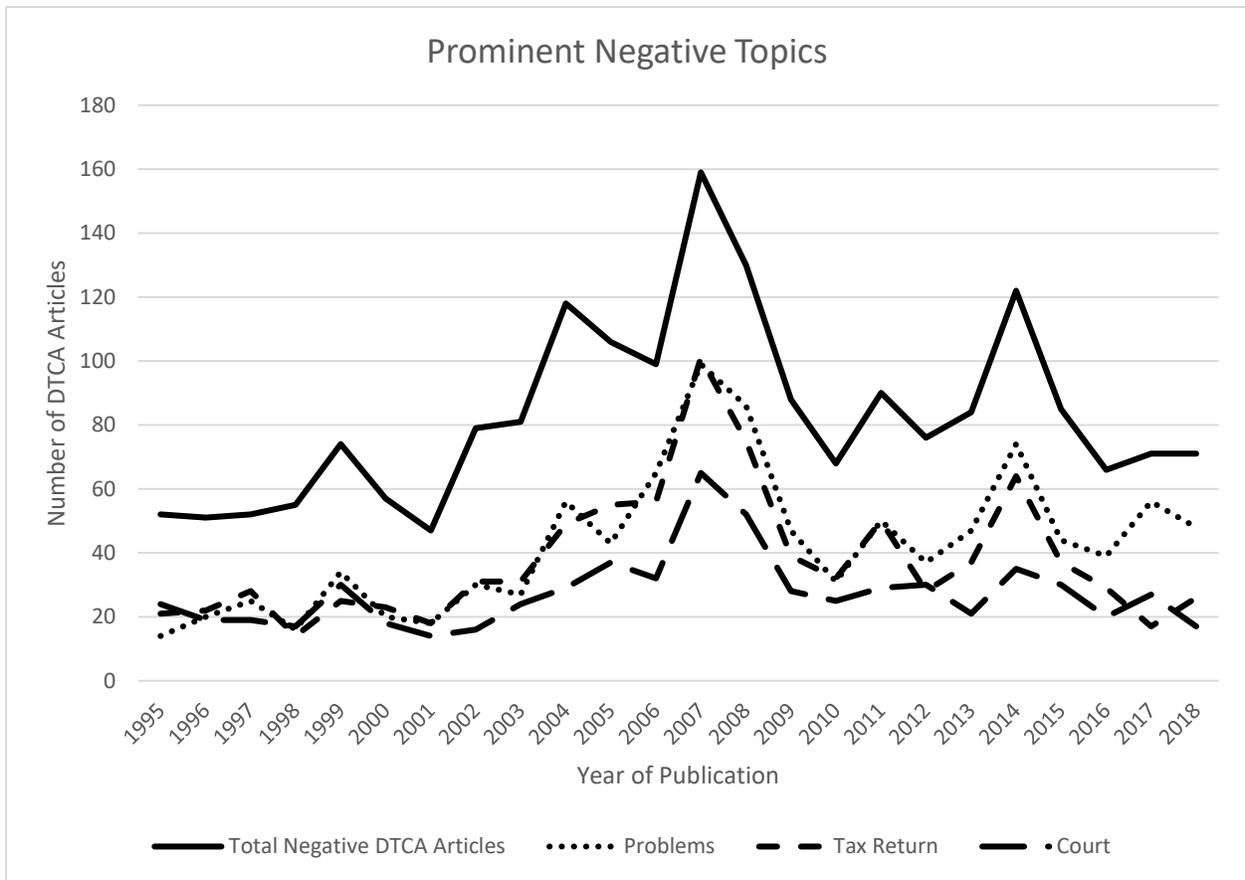


Fig. 8. Development of negative prominent topics negative over the years.

References

- Alink, F. (2006). *Crisis als kans*. Amsterdam University Press.
- Alink, F., Boin, A., & 't Hart, P. (2001). Institutional crises and reforms in policy sectors: the case of asylum policy in Europe. *Journal of European Public Policy*, 8(2), 286–306. <https://doi.org/10.1080/13501760151146487>
- Alink, M. & van Kommer, V. (2012). *The Dutch Approach. Descriptions of the Dutch Tax and Customs Administration*. IBFB.
- Anastasopoulos, L. J., & Whitford, A. B. (2018). Machine Learning for Public Administration Research, with Application to Organizational Reputation. *Journal of Public Administration Research and Theory*, *muy060*. <https://doi.org/https://doi.org/10.1093/jopart/muy060>
- Anderson, B. (2015). Measuring Public Perceptions of International Organizations' Legitimacy. *ECPR Joint Sessions Workshop 2015*, 1–29.
- Ansell, C., Boin, A., & Kuipers, S. (2016). Institutional crisis and the policy agenda. In N. Zahariadis (Ed.), *Handbook of Public Policy agenda Setting*. Cheltenham: Edward Elgar publishing.
- Ashforth, B. E., & Gibbs, B. W. (1990). The Double-Edge of Organizational Legitimation. *Organization Science*, 1(2), 177–194.
- Baumgartner, F. R., & Jones, B. D. (2010). *Agendas and instability in American politics*. University of Chicago Press.
- Bitektine, A. (2011). Toward a theory of social judgments of organizations: The case of legitimacy, reputation and status. *Academy of Management Review*, 36(1), 151–179.
- Blei, D. M. (2012). Surveying a suite of algorithms that offer a solution to managing large document archives. Probabilistic topic models. *Communications of the Acm*, 44(4), 77–84. <https://doi.org/10.1145/2133806.2133826>
- Blei, D. M., Ng, A. Y., & Edu, J. B. (2003). Latent Dirichlet Allocation Michael I. Jordan. In *Journal of Machine Learning Research* (Vol. 3).
- Boin, A., & 't Hart, P. (2000). Institutional Crises and Reforms in Policy Sectors. In H. Wagenaar (Ed.), *Government institutions: Effects, Changes and Normative Foundations* (p. 257). <https://doi.org/10.1007/978-94-010-0963-8>
- Boin, A., Kuipers, S., & Otten, M. (2000). *Institutionele crisis, Boin.pdf*. Alphen aan de Rijn: Samsom.
- Borstlap, H., & Joustra, T. H. J. (2017). *Onderzoek naar de besluitvormingsprocedures binnen de*

- Belastingdienst*. 176. Retrieved from <https://www.rijksoverheid.nl/documenten/rapporten/2017/01/27/rapport-commissie-onderzoek-belastingdienst>
- Carroll, C. E., & McCombs, M. (2003). Agenda Setting Effects of Business News on the Public's Images and Opinions About Major Corporations. *Corporate Reputation Review*, 6(1), 36–44. <https://doi.org/10.1057/palgrave.crr.1540188>
- Deephouse, D. L. (2000). *Media Reputation as a Strategic Resource: An Integration of Mass Communication and Resource-Based Theories*.
- Deephouse, D. L., & Suchman, M. (2008). Legitimacy in Organizational Institutionalism. *The Sage Handbook of Organizational Institutionalism*, (April), 49–77. <https://doi.org/10.1080/00221686.2011.557258>
- Dowling, J., & Pfeffer, J. (1975). Organizational Legitimacy: Social Values and Organizational Behavior Social Values and Organizational Behavior. *Source: The Pacific Sociological Review*, 18(1), 122–136.
- Gilley, B. (2006). The meaning and measure of state legitimacy: Results for 72 countries. *European Journal of Political Research*, 45(3), 499–525.
- Grafström, M., Windell, K., & Karlberg, P. P. (2015). Mediatization of Civil Society Organizations: (De)legitimation of the Swedish Red Cross. *Journal of Civil Society*, 11(3), 227–241. <https://doi.org/10.1080/17448689.2015.1057415>
- Gribnau, J. L. M. (2013). *Belastingen als moreel fenomeen: Vertrouwen en legitimiteit in de praktijk*. Den Haag: Boom fiscale uitgevers.
- Haldenwang, C. von. (2016). Measuring legitimacy – new trends , old shortcomings? In *Measuring Legitimacy – New Trends, Old Shortcomings?* <https://doi.org/10.13140/RG.2.2.21332.81286>
- Heath, R.L. (2010). Crisis Communication: Defining the Beast and De-marginalizing Key Publics. In Coombs & Holladay (Ed.), *The Handbook of Crisis Communication*, (1 – 14). Blackwell Publishing
- Hirsch, P. M., & Andrews, J. (1984). Administrators' Response to Performance and Value Challenges - - Or, Stance, Symbols and Behavior in a World of Changing Frames. In T. Sergiovanni, & J. Corbally (Eds.), *Leadership and Organizational Culture: New Perspectives on Administrative Theory and Practice* (pp. 170-185)
- Hollibaugh, G. E. (2018). The Use of Text as Data Methods in Public Administration: A Review and an Application to Agency Priorities. *Journal of Public Administration Research And Theory*, 1–17. <https://doi.org/10.1093/jopart/muy045>
- Ivengar, S., & Simon, A. (1993). News Coverage of the Gulf Crisis and Public Opinion.

Communication Research, 20(3), 365–383.

- Jonsson, S., Greve, H. R., & Fujiwara-Greve, T. (2009). Undeserved Loss: The Spread of Legitimacy Loss to Innocent Organizations in Response to Reported Corporate Deviance. *Administrative Science Quarterly*, 54, 195–228.
- Joutsenvirta, M., & Vaara, E. (2009). Discursive (de)legitimation of a contested Finnish greenfield investment project in Latin America. In *Scandinavian Journal of Management* (Vol. 25).
- Kreutz, T., & Daelemans, W. (2018). Enhancing General Sentiment Lexicons for Domain-Specific Use. *Proceedings of the 27th International Conference on Computational Linguistic*, 1056–1064. Retrieved from <http://www.aclweb.org/anthology/C18-1090>
- Mccombs, M. E., & Shaw, D. L. (1972). The Agenda-Setting Function of Mass Media. In *The Public Opinion Quarterly* (Vol. 36).
- Nohrstedt, D., Bynander, F., Parker, C., & Hart, P. (2018). *Managing Crises Collaboratively: Prospects and Problems — A Systematic Literature Review*. (February). <https://doi.org/10.1093/ppmgov/gvx018>
- Pollock, T. G., & Rindova, V. (2003). Media legitimation effects in the market for initial public offerings Symbolic value creation View project. *Article in The Academy of Management Journal*. <https://doi.org/10.2307/30040654>
- Resodihardjo, S.L. (2009). *Crisis and Change in the British and Dutch Prison Services: Understanding crisis reform processes*, Ashgate Publishing: Farnham
- Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., ... Rand, D. G. (2014). Structural Topic Models for Open-Ended Survey Responses. *Source: American Journal of Political Science*, 58(4), 1064–1082. <https://doi.org/10.1111/ajps>
- Schmidt, A., Boersma, K., & Groenewegen, P. (2018). *Management strategies in response to an institutional crisis: The case of earthquakes in the Netherlands*. 1–15. <https://doi.org/10.1111/padm.12516>
- Schmidtke, H. (2018). Elite legitimation and delegitimation of international organizations in the media: Patterns and explanations. *The Review of International Organizations*. <https://doi.org/10.1007/s11558-018-9320-9>
- Stevens, L. G. M. (2007). De Belastingdienst op drift? *WFR*, 399.
- Strauss, A., & Corbin, J. (1990). *Basics of qualitative research*. Sage publications.
- Suchman, M. C. (1995). Managing Legitimacy: Strategic and Institutional Approaches. *The Academy of Management Review*, 20(3), 571–610.

- Syed, S., & Spruit, M. (2017). Full-Text or Abstract? Examining Topic Coherence Scores Using Latent Dirichlet Allocation. *The 4th IEEE International Conference on Data Science and Advanced Analytics*.
- Törnberg, A., & Törnberg, P. (2016). *Muslims in social media discourse: Combining topic modeling and critical discourse analysis*. <https://doi.org/10.1016/j.dcm.2016.04.003>
- Useem, B., Pacholke, D., & Mullins, S. F. (2016). Case Study - The Making of an Institutional Crisis: The Mass Release of Inmates by a Correctional Agency. *Journal of Contingencies and Crisis Management*. <https://doi.org/10.1111/1468-5973.12146>
- Waeraas, A. (2019). Public Sector Communication and Organizational Legitimacy. In V. Luoma-aho & M.-J. Canel (Eds.), *Wiley Handbook of Public Sector Communication*. Retrieved from <https://www.researchgate.net/publication/323827237>
- Wallner, J. (2008). Legitimacy and Public Policy: Seeing Beyond Effectiveness, Efficiency and Performance. *Policy Studies Journal*, 36(3), 421–443. <https://doi.org/10.1111/j.1541-0072.2008.00275.x>
- Zhu, L., Witko, C., & Meier, K. J. (2018). The Public Administration Manifesto II: Matching Methods to Theory and Substance. *Journal of Public Administration Research and Theory*. <https://doi.org/10.1093/jopart/muy079>