Online Supplement

Inflation of Crisis Coverage?

(accepted for publication in *Journal of Communication*)

Stefan Geiß
Christina Viehmann
Conor A. Kelly

Table of contents

Chapter 1

Appendix A: Scanning Real-World Developments That May Help Explain Changes in Crisis Coverage

1.1 Purpose

The purpose of this appendix is to document various indicators of real-world developments that may be able to explain the changes in crisis labelling salience, crisis news wave salience, or crisis news wave wave count.

We also describe some qualitative characterizations of developments for which we did not find any quantitative indicators (or they did not have sufficient temporal coverage or resolution).

This is per se selective. Our rationale for exploring indicators was:

- The development is associated with events and issues that regularly find expression in crisis coverage, mainly
  - Disaster and accident crisis
  - Epidemics
  - Economic crisis, strikes, boycotts
  - Geopolitical crisis
  - Government crisis
- Existence of appropriate indicators, particularly if there is...
  - ...good coverage of the time period (including pre-WWII period)
  - ...high temporal resolution
- Existence of scientific concepts or analyses that posit a plausible trend (example: risk perception, value change)

1.2 Main categories

The explorations are grouped into three main categories which overlap, however.

There are changes in the **rate of critical conditions** that we assume were perceived as highly adverse events during the entire period of study. Examples include... * major natural disasters and accidents with many casualties, * incidents with extremely high monetary costs.

This presumes that the value society puts of human life, money, etc. is approximately stable.

There are changes in how society **handles and responds to critical conditions**. This includes that… * …new information technology and denser bureaucratic supervision create more information and data and allow information and data to flow more easily. * …changes in social organization, economy, and technology enable society to effectively respond to or prevent the critical conditions, opening up new sets of responses.

Finally, there are changes in how society **views itself and imagines itself**. This includes… * …the expectations towards standards of living and the tolerance of grievances change; * …alternatively, the threshold for viewing a critical condition as a crisis changes or linguistic habits of expression change * …the priorities of values in society change (for example because more citizens receive voting rights and their interests start to count for decision-makers and their interests become more visible in the public sphere)

1.3 Note on the Process of Scanning for Real-World Developments

In earlier stages of the project we were very interested in mapping the adequacy of crisis coverage relative to indicators of severity of a variety of social problems. As a starting point, we looked at the main crisis types we identified in our corpus, which were *economic crisis*, *disasters* (natural disasters, major accidents), *geopolitical crisis* and *epidemics*. Therefore, we looked into economic performance of the UK as reconstructed by economic historians Figure 1.3, number of deaths in the UK by epidemics, wars, natural disasters Figure 1.6. None of these indicators shows any increases at face value. This is also reflected in the development of general indicators of the human condition, such as life expectancy Figure 1.5 (see also event timelines in Section 1.4.1.1).

We then explored other arguments that may explain an increase in crisis coverage, like the general population increase, (increasing the likelihood of adverse events and increasing the rate and damage of disasters that do damage to humans), increasing life expectancy (creating challenges in providing healthcare and elderly care at a larger scale) and the increasing degree of social complexity were explored Figure 1.1, Figure 1.2, Figure 1.5. Even if these trends actually exist, they do not seem to manifest in an increase in rate or intensity of the most widely covered types of crisis events (see previous point).

The most likely explanatory factors we found were factors relating to how society itself is structured, observes its environment, and thinks of itself and its values (and to some extent what it views as crisis) (*Prioritized values and systems* and *Sensitivity for critical conditions* in society). For example, more different government top-level departments/ministries were added Figure 1.4, more people become literate, can follow the news, want to follow the news, and do follow the news Section 1.6.2, changes in risk perception Section 1.6.10 and information availability and access Section 1.5.1.

Quantifying all these dimensions with high temporal resolution and back in time for several centuries was not realistic due to lack of indicators, insufficient coverage of the time period, or insufficient temporal resolution. Even if indicators existed, it would prove impossible to put them together into a general index of crisis-strickenness of society, or at least such an index would be challengeable from various directions and hard to defend.

In conjunction with the explanatory factors that are conceptualized in the theoretical approaches we utilize (policy arenas model, policy dynamics model, Thermostat model, Mediatization theory) (Soroka and Wlezien 2010; Hilgartner and Bosk 1988; Baumgartner, Jones, and Wilkerson 2002; Strömbäck 2008), these explorations suggest that, in line with our explanatory model, we would focus on factors involving society's problem solving capacities (intensity of public spending), shifting/diversifying value system (diversity of public spending), and an inclination of the public sphere to focus on broad popular interests and maximize audience size (media penetration and media autonomy), and to cover problems because they are interesting to the audience and not because they are politically instrumental to a certain side (media autonomy) that would promise to explain the changes in salience of crisis coverage.

1.4 Rate of critical conditions

1.4.1 Observed event rates

If we turn to statistics of adverse events Section 1.4.1.1 that are likely to become subject of crisis coverage—e.g. disaster statistics (natural disasters, accidents), war casualties, economic growth statistics—we do not observe clear trajectories in any of these areas. We would probably expect more accidents given the rapid population growth, but improved safety measures and mechanization seem to compensate this. The relatively low number of war casualties after WWII are checked by the potentially devastating consequences of a possible nuclear war. Economic recessions have occurred throughout the period of study, with no clear trend. Famines and epidemics cost fewer human lives today (and much less in terms of percentage of the world population), but the COVID-19 pandemic has shown that undoubtedly dangerous diseases do not necessarily find expression in the statistics of actual deaths—if all 8 Billion human had been infected and case fatality ratio was 2, 1 or 0.5%, respectively, that would mean that 160 Million, 80 Million, or 40 Million deaths would have resulted instead of the 7.01 Million that we have observed until now (July 2024) Worldometer: COVID-19 Coronavirus Pandemic. Storms and floods tend to become more frequent and more devastating (maybe as a consequence of climate change), but not yet to the extent that we would see a substantial change in terms of crisis coverage.

1.4.1.1 Event timelines

1.4.1.1.1 Wars with UK involvement or Wars with more than 1 Million dead

- Napoleonic Wars (1803-1815), 3.5-7.0 Million (Europe),
- Taiping Rebellion (1850-1864), 20-30 Million (China),
- Indian Rebellion (1857), 0.8-1.0 Million (India),
- US Civil War (1861-1865), 0.65-1.0 Million (USA),
- Boxer Rebellion (1899-1901), 0.1 Million (China),
- WWI (1914-1918), 16.0-40.0 Million (Europe),
- Russia Civil War (1917-1922), 5-9 Million (Russia),
- Sino-Japanese War (1937-1945), 20-25 Million (China),
- WWII (1939-1945), 56-85 Million (Worldwide),
- Korean War (1950-1953), 1.5-4.5 Million (Korea),
- Algerian War (1954-62), 0.4-1.5 Million (Algeria),
- Vietnam War (1955-75), 1.3-4.3 Million (Vietnam),
- Nigerian Civil War (1967-70), 1-3 Million (Nigeria),
- Bangladesh Liberation War (1971), 3+ Million (Bangladesh),
- Ethiopian Civil War (1974-91), 0.5-1.5 Million (Ethiopia),
- Falklands War (1982), ~1 Thousand (Falkland Islands),
- Soviet-Afghan War (1979-89), 0.6-2.0 Million (Afghanistan),
- Iran-Iraq War (1980-88), 0.3-1.1 Million , (Iraq),
- Gulf War (1991), 0.2 Million (Kuweit, Iraq),
- Bosnian War (1992-1995) 0.1 Million (Croatia, Bosnia-Herzegovina, Serbia),
- Kosovo War (1998/99), ~13 Thousand (Kosovo), 1 Million refugees,
- Sierra Leone Civil War (1991-2002), 0.05-0.3 Million
- Afghanistan War (2001-today), 0.2 Million
- Iraq War (2003-11), 0.06 Million
- Libyan Civil War (2011), 0.01-0.025 Million

Sources: Compiled 10 October 2021

Wikipedia: List of wars involving the United Kingdom

Wikipedia: List of wars by death toll

1.4.1.1.2 Natural disasters with more than 100 000 deaths

- Cyclone in India (1839, 300,000 dead),
- Yellow River Flood in China (1887, 2,000,000 dead),
- Yangtze River Flood in China (1911, 100,000 dead)
- Haiyuan Earthquake (1920, 273,000 dead),
- Great Kanto Earthquake (1923, 143,000 dead)
- Flooding in China (1931, 4,000,000 dead),
- Bhola Cyclone (1970, 500,000 dead),
- Flooding in Vietman (1971, 100,000 dead),
- Typhoon Nina (1975, 229,000 dead),
- Tangshan Earthquake (1976, over 240,000 dead),
- Bangladesh Cyclone (1991, 138,000 dead),
- Indian Ocean earthquake/tsunami (2004, 227,000 dead),
- Cyclone Nargis (2008, 138,000 dead),
- Haiti earthquake (2010, at least 100 000 dead)
- Epidemics 1812 Russia Typhus epidemic (1812, at least 300,000 dead)
- Ottoman Plague epidemic (1812-1819, at least 300,000 dead)
- First Cholera Pandemic (1817-1824, at least 100,000 dead)
- Second Cholera Pandemic (1826-1837, at least 100,000 dead)
- Third Cholera Pandemic (1846-1860, at least 1,000,000 dead)
- Third Plague Pandemic (1855/1894, at least 12,000,000 dead)
- Fourth Cholera Pandemic (1863-1875, at least 600,000 dead)
- European Smallpox Epidemic (1870-1875, at least 500,000 dead)
- Fifth Cholera Pandemic (1881-1896, at least 298,000 dead)
- Russian Flu (1889-1890, 1,000,000 dead)
- Congo Sleeping Sickness epidemic (1896-1906, at least 500,000 dead)
- Sixth Cholera Pandemic (1899-1923, at least 800,000 dead)
- Uganda Sleeping Sickness epidemic (1900-1920, at least 200,000 dead)
- Encephalitis lethargica pandemic (1915-1926, at least 500,000 dead)
- Spanish Flu (1918-1920, at least 17,000,000 dead)
- Russia Typhus epidemic (1918-1922, at least 2,000,000 dead)
- Asian Flu (1957-1958, at least 1,000,000 dead)
- Hong Kong Flu (1968-1969, at least 1,000,000 dead)
- 1977 Russian Flu (1977-1979, at least 700,000 dead)
- AIDS (1981-present, at least 35,000,000 dead)
- Swine Flu (2009-2010, at least 150,000 dead)
- COVID-19 pandemic (2020-present, 3,200,000 dead)

Sources: Compiled 10 October 2021

Wikipedia: List of natural disasters by death toll

1.4.1.1.3 Shipwrecks with more than 500 deaths

- HMS Ardent (Villefranche-sur-Mer, France, April 1794, 500)
- HMS Queen Charlotte (Cabrera, 17 March 1800, 673)
- HMS Minotaur (off Texel, Netherlands, 22 December 1810, 570)
- Haak Sand on 24 December 1811, 600
- HMS Defence (Jutland. Denmark, 24 December 1811, 583)
- HMS St George (Ringkøbing, Denmark, 24 December 1811, 731)
- San Telmo (Drake Passage, Antarctica, 2 September 1819, 644)
- Tek Sing (1822, 1,600)
- Lefort (Gulf of Finland, 22 September 1857, 826)
- Sultana (1865, 1,168)
- RMS Atlantic (Nova Scotia, 1 April 1873, 546)
- SS Princess Alice and SS Bywell Castle (River Thames, 3 September 1878, 640)
- SS Vaitarna (lost in cyclonic storm of coast of Saurashtra, Gujarat, India, 8 November 1888, 746)

- Ertuğrul (Kushimoto, Japan, 18 September 1890, 533)
- SS Utopia (Bay of Gibraltar, 17 March 1891, 564)
- SS La Bourgogne (Sable Island, Nova Scotia, 4 July 1898, 565)
- HMS Bulwark (1899) (Sheerness, England, 26 November 1914, 736)
- SS Camorta (Irrawaddy Delta, 6 May 1902, 737)
- SS Norge (Rockall, 28 June 1904, 627)
- General Slocum (1904, 1,021)
- Kiche Maru (1912, 1,000)
- Titanic (1912, 1,514)
- Empress of Ireland (1914, 1,024)
- SS Eastland (1915, 844)
- Hsin-Yu (1916, 1,000)
- SS Principe de Asturias (off Brazil, 5 March 1916, 558)
- SS Afrique (Bay of Biscay, 9 January 1920, 575)
- Hong Moh (1921, 1,000)
- Wusung (1927, 900)
- SS Indigirka (Sarufutsu, Japan, 12 December 1939, 741)
- SS Ramdas (Bombay, 17 July 1947, 625)
- Kiangya (1948, 2,750+)
- Taiping (1949, 1,500)
- Toya Maru (1954, 1,159)
- Novorossiysk (Sevastopol, 29 October 1955, 608)
- Indonesian passenger ship Tampomas 2 (caught fire and sank in Java Sea, 27 January 1981, 580)
- Atlas Star (Dhaleswar River, Munshiganj, Bangladesh, 29 January 1986, 500)
- Shamia (Meghna River, Bangladesh, 25 May 1986, 500+)
- Dona Paz (1987, 1,562+)
- Ferry Neptune (16 February 1993, 500+)
- MS Estonia (1994, 852)
- MV Bukoba (1996, 894)
- Cahaya Bahari (off Sulawesi, Indonesia, 29 June 2000, 550)
- Le Joola (2002, 1,863)
- MV Nazreen 1 (Chandpur, Bangladesh, 8 July 2003, 528)
- Al-Salam Boccaccio 98 (2006, 1,012)
- MV Princess of the Stars (2008, 832)
- Spice Islander I (2011, 1,573)

Sources: Compiled 10 October 2021

Wikipedia: List of maritime disasters in the 18th century

Wikipedia: List of maritime disasters in the 19th century

Wikipedia: List of maritime disasters in the 20th century

Wikipedia: List of maritime disasters in the 21st century

1.4.1.1.4 Building collapse with more than 500 deaths

- Puentes Dam collapse, 1802; 608 dead
- Etai Bridge Collapse (1807, at least 500 dead)
- Ponte das Barcas (1809, 4000 dead)
- Johnstown Flood, 1889; 2,209 dead
- Tigra Dam failure, 1917; at least 1,000 dead
- Panshet Dam failure, 1961; 1,000 dead
- Vajont Dam, 1963, 1,910 dead
- Iruka Lake Dam failure, 1968; 941 dead
- Banqiao Dam, 1975; 171,000 dead

- Machchhu Dam, 1979; 5,000 dead
- Sampoong department store collapse, 1995; 502 dead
- Rana Plaza collapse, 2013; 1,129 dead

Sources: Compiled 10 October 2021

Wikipedia: List of accidents and disasters by death toll

Wikipedia: List of building and structure collapses

From this inspection of relevant available data, we draw two tentative conclusions:

- Event rates have not substantially changed beyond changes in reporting.
- Real-world event rate changes do not explain changes in the extent of CL and CNWs.

1.4.1.2 Population increase and population density

A marked upward trend in UK (more than fourfold) Figure 1.1 and world population (almost 20-fold) Figure 1.2. This increases social activity, making human-made crisis (e.g. major accidents) more likely. And this increases population density, leading to a higher damage potential per crisis (e.g. an earthquake of the same strength will on average kill more people if population density is greater).

From this inspection of relevant available data and analyses, we draw three tentative conclusions:

- Population size and density have increased in the UK and worldwide.
- There will be more and more serious critical conditions that give rise to CL and CNWs.
- The more people there are, the more can go wrong, the more do they interfere with each other, and the more people will be affected by adverse events.

1.4.1.3 Increasing wealth, security, and supply

Fast economic growth has been a consistent characteristic of modern societies, going hand in hand with rapid changes in people's living conditions: a major transition of labor from agricultural to industry and service, urbanization, the almost universal employment of computers and powered machinery Figure 1.3. This allowed a major expansion of social security systems Our World in Data: Public social spending as a share of GDP and educational possibilities (Clark 2005; Carpentier 2018), but these also became necessary in the face of the waning potential of families to provide social security, and higher educational requirements (Smelser 1967; Haferkamp and Smelser 1992). This means that some problems were solved or became easier to tackle, but new needs and expectations were created that can give rise to totally new types of crises.

Very much the same argument applies to improved medical treatment: this solved many immediate issues and extended life expectancy tremendously (it approximately doubled from ~40 to ~80 years). But the extended life expectancy means that there are new challenges that can lead up to critical conditions, for instance when pensions, healthcare, and elderly care become increasingly expensive.

From this inspection of relevant available data and analyses, we draw four tentative conclusions:

- Between 1785 and 2020, wealth, medical treatment and supply with basic goods have improved substantially, leading to less poverty and longer life expectancy. Social security systems have been implemented to prevent many serious critical conditions.
- Critical conditions associated with extreme poverty or shortages have declined, leading to less CL and fewer CNWs.
- Critical conditions associated with the functioning of social security systems have grown with the growing importance and complexity of these systems. This leads to more CL and CNWs.
- There will not be fewer critical conditions as wealth grows, but their profile will change from immediately existential crises to systemic and adaptation crises.

1.4.1.4 Increasing interdependency and division of labor.

The rapid economic growth led to increased specialization, interdependence and division of labor. While increasing efficiency, this also increases the susceptibility and consequentiality of disturbances in the economy that spread from sector to sector and can lead up to chain reactions (Hidalgo and Hausmann 2009; Isik 2010; Biggs et al. 2011). This also concerns international trade and globalization, where interdependency and complexity (e.g., of supply chains) increase the potential for chain-reactions.

Similarly, the expansion of the state (in terms of budget and policy fields/responsibilities) led to establishment of more and more institutions Figure 1.4 whose failure to fulfill their functions can also cause chain reactions.

From this inspection of relevant available data and analyses, we draw two tentative conclusions:

- Social and economic complexity and interdependency has increased 1785-2020.
- With increases in social complexity, government size, and interdependency, there will be (a) more and more severe critical conditions, (b) a greater chance for chain reactions that lead to spill-over from critical conditions in one region or sector to other regions or sectors.

1.4.1.5 Increasing life expectancy

In the late 1700s, life expectancy in the UK was below 40 years and has increased tremendously to almost 80 years for children born in 2020 Figure 1.5. This reflects better living conditions and more effective medical service (such that the potential for crises related to living conditions decreased) but raises new problems such as pensions and elderly care.

From this inspection of relevant available data and analyses, we draw one tentative conclusion:

- Critical conditions have shifted to provision of pensions and care for the elderly, and causes of death typical for elderly people (e.g. many types of cancer).

1.4.1.6 War and violence – ambiguous trends

The trends in wars and violence are ambiguous. With fewer people living in existence-threatening conditions and with longer education phases, rates of violence have decreased (e.g. in terms of violent crimes). However, warfare has seen a major industrialization with the two World Wars showing the potential of how much death and devastation mass armies and modern weapons can cause, culminating in the use of atomic bombs that ended World War II. The risk of atomic warfare destroying humanity's livelihood is still real. So even if the period since 1945 has not extreme-scale wars as the early 1900s, potential wars may be just as critical, as e.g. the Cuban Missile Crisis of 1962.

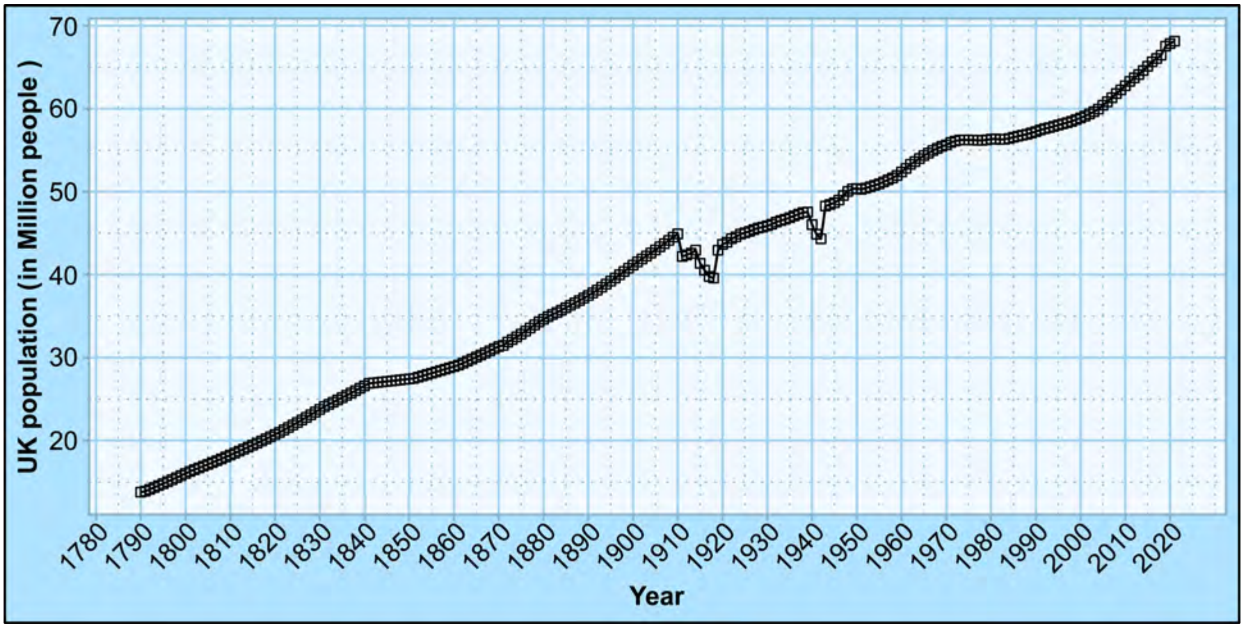
From this inspection of relevant available data and analyses, we draw one tentative conclusion:

- Critical conditions tied to wars, civil wars, and criminal violence have developed nonlinearly; they tended to be low in the entire period after WWII, but the threat of a nuclear war can be considered a critical condition that does not find expression in number of deaths.

Overall, the idea that we live (1) in an era of unprecedented wealth (2) with only few critical conditions that could legitimately become subject of crisis news coverage proves porous. You could also call this the *spoilt brats theory* of crisis: we are just "spoilt brats" that have never seen any really bad conditions and therefore complain and cry "crisis" without a legitimate reason.

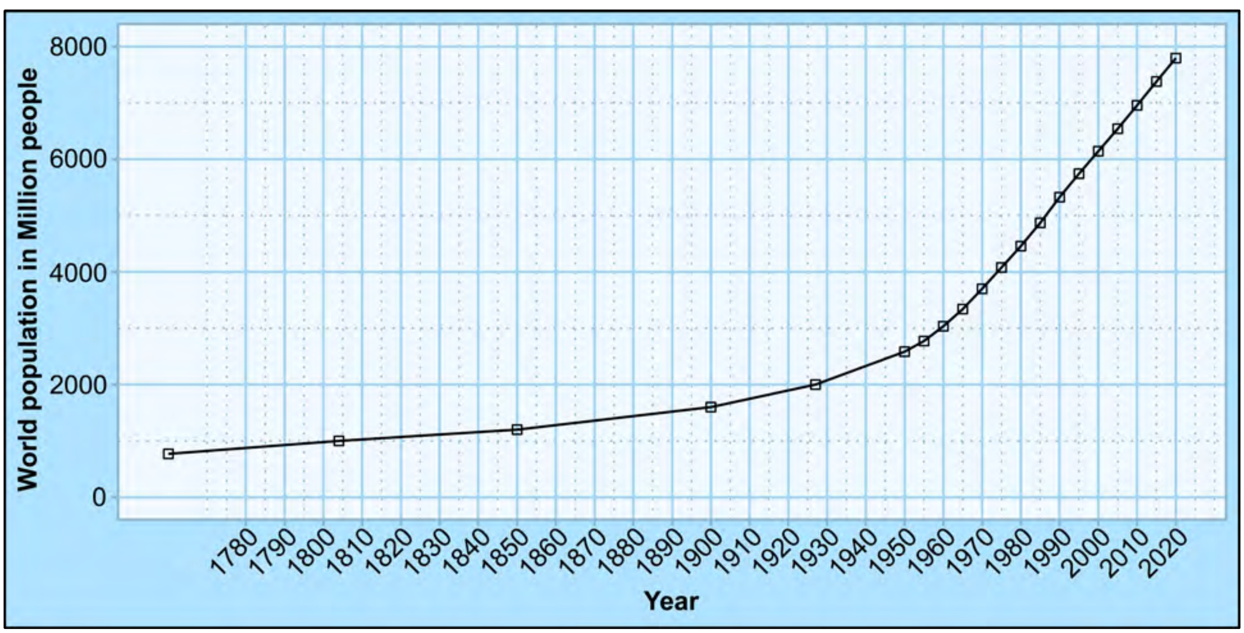
While wealth (and better medical treatment) removes or cushions some of the most immediate and existential critical conditions, it has historically been associated with rising life expectancy, population growth and greater social complexity that produce additional critical conditions. The state expands and creates new institutions and new areas it regulates each of which can fail or whose function can be disturbed. These failures and disturbances can also become subject of crisis coverage. No consistent upward or downward trends of CL or CNW are expected across domains.

Figure 1.1: Population of the UK in Million people



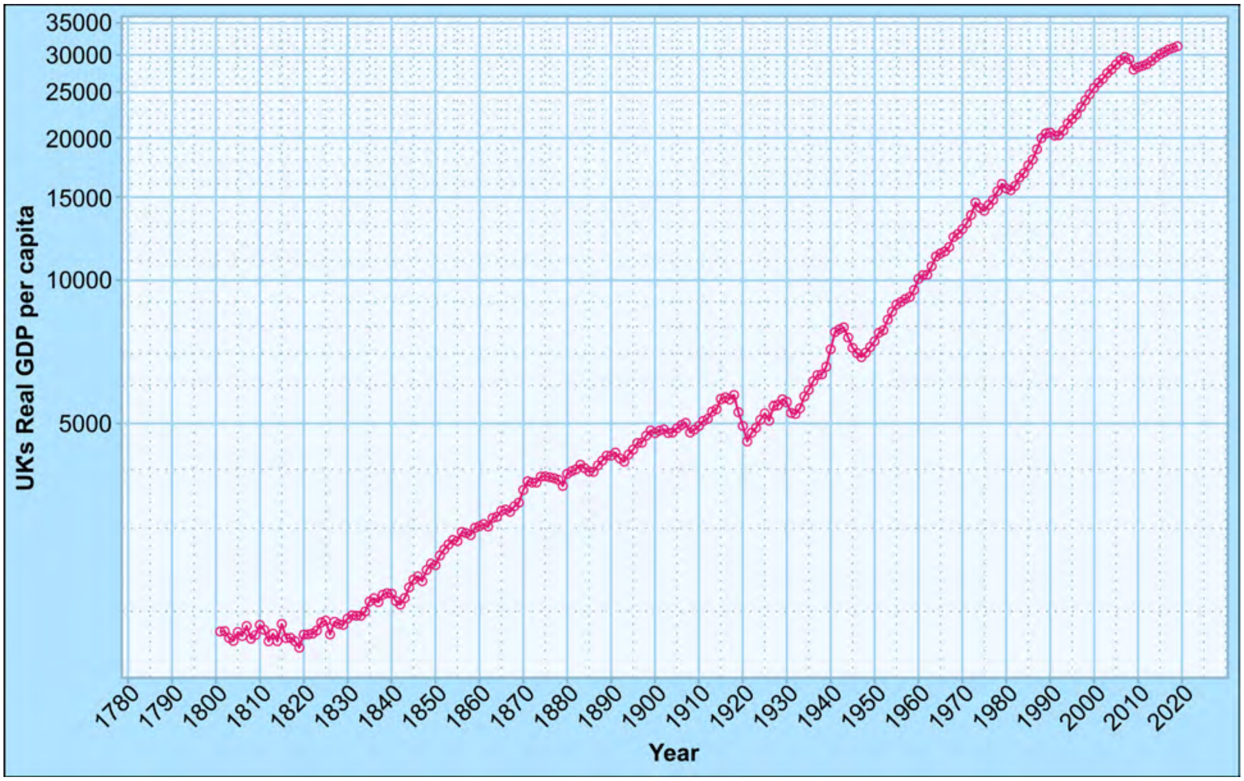
Source: Chantrill (2023); Thomas and Dimsdale (2017)

Figure 1.2: Worldwide human population, in Million people



Source: Worldometer. 1950 to current: United Nations: World Population Prospects

Figure 1.3: Real GDP per capita, in 2016 pounds



Source: Ryland and Williamson (2021)

1.4.1.7 Natural Disasters

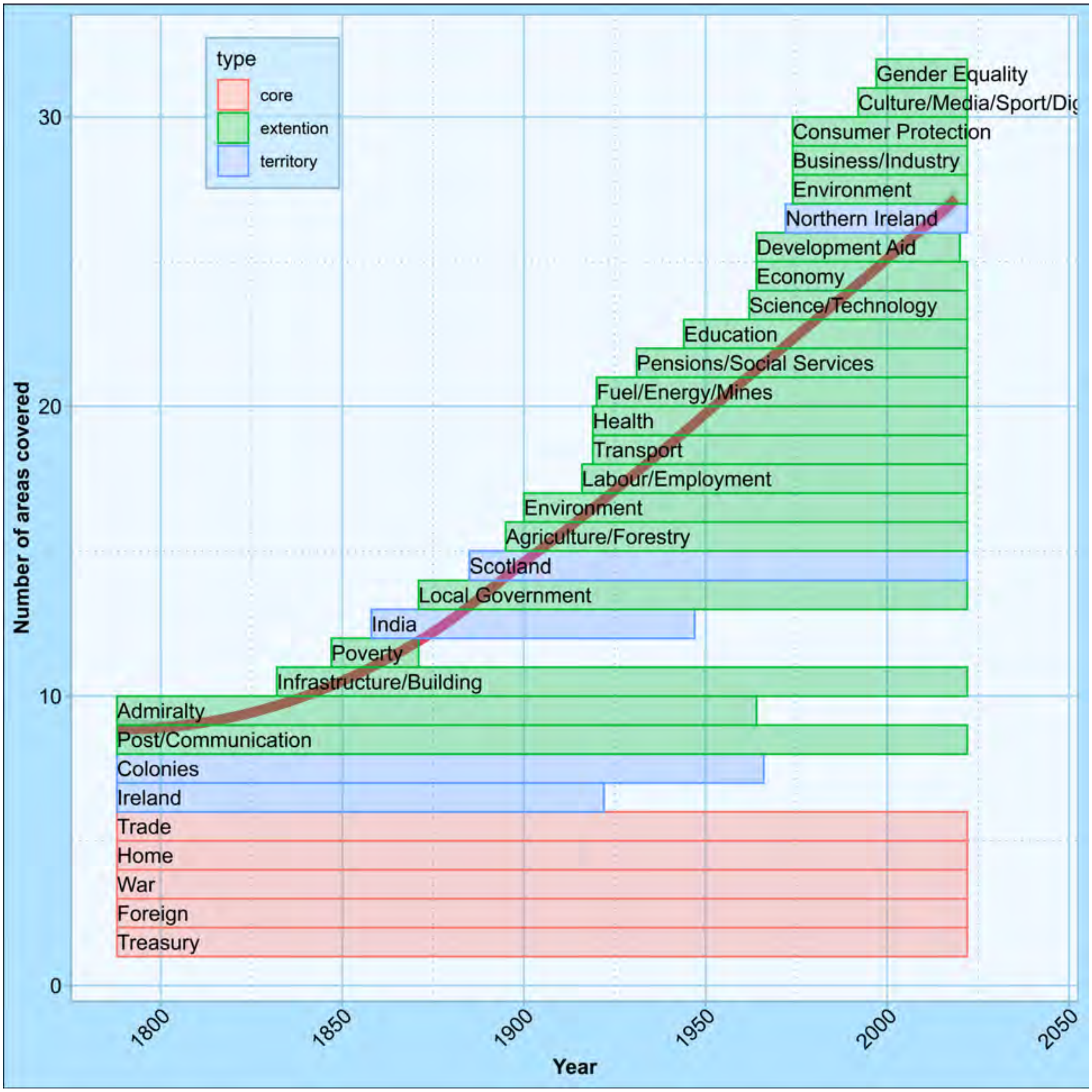
While one would assume that the occurrence of natural disasters is not systematically changing, there are ways in which human activity can affect the likelihood and extent of disasters, for example if human activity influences the climate (which can change the risks of storm, floods, etc.), and by the density and distribution of human settlement and other facilities: A flood in a largely uninhabited area would not cause severe damage in terms of news value, but a flood in a major city will. On the other hand, humans can improve disaster prevention and management to reduce the damage in case a natural disaster strikes (e.g., by building dams, flood barriers).

Therefore, we roughly estimated the news value of the recorded natural disasters in the EMDAT database (Centre for research on the Epidemiology of Disasters 2021). The underlying equations are massive simplifications but should allow getting a rough sense of the news value of disaster events. Based on the literature on news selection, the most significant sources of news value in disaster coverage are "damage" and "proximity", i.e. the most media attention is to be expected if massive damage occurs in the country (in case of a national news outlet, such as *The Times*) that the news outlet is based in and caters to.

The amount of damage was calculated based on the EMDAT database entries for deaths (1 point), injuries (0.1 points), homelessnesses (0.01 points), and damages in 1000 US$ [inflation-adjusted] (0.0001 points). This was computed separately for the total damage that occurred in a natural disaster category in a given year (more useful to see less-intense disasters that spread out over a longer period, e.g. most famines), and the maximum daily damage (more useful to see intense disasters with a lot of damage in a short time).

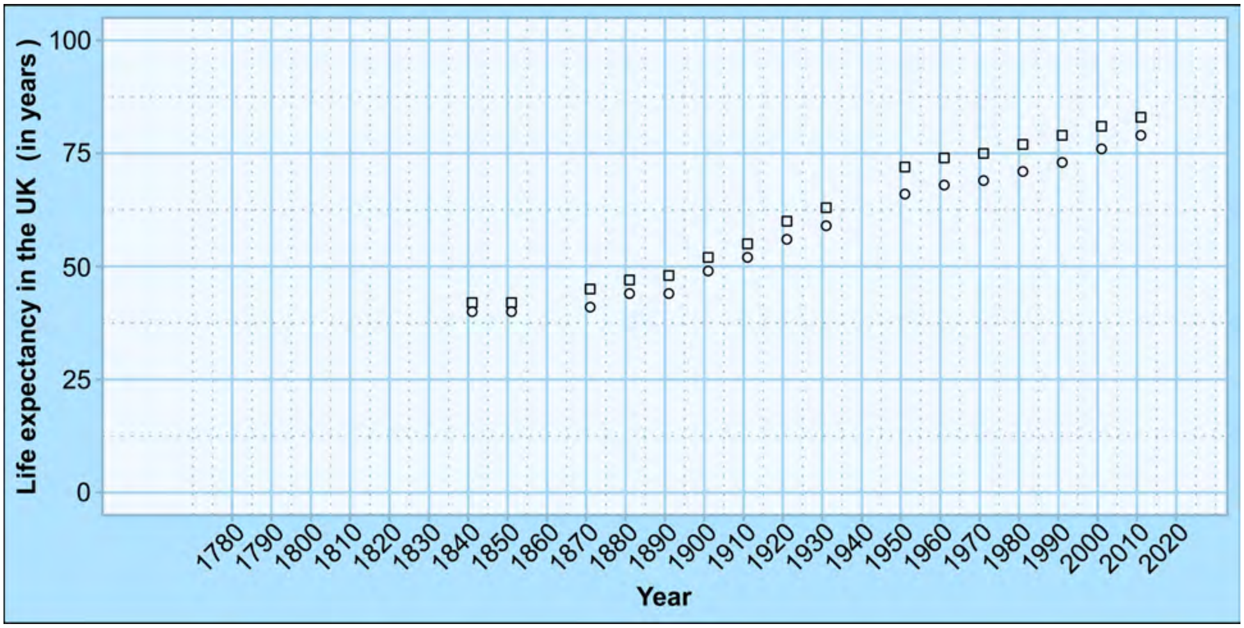
The relation was calculated using distance of the countries to the UK on five dimensions: colonial relation (all former colonies), political relation ([fellow] EU members), geographical relation (distance between

Figure 1.4: Top-level ministries in the British government



Source: https://en.wikipedia.org/wiki/List_of_British_governments and the articles about each ministry/administration. These list the portfolios of the ministers in the ministry. The accuracy of the lists was spot-checked against the following sources: "Cabinets and Administrations" (2005); "Chronology of Major Cabinet Changes" (2005); Cook and Stevenson (1980); Cook and Keith (1984); Butler and Butler (1994).

Figure 1.5: Life expectancy for females (squares) and males (circles) in the UK.



Source: Office for National Statistics (2015)

London and country capital), military relation (list of former war allies, list of former war enemies, fellow NATO members), trade relation (trade volume with the top 30 trading partners; only current data were used). These relation indicators ranged between 0 and 1 and were weighted equally into a relationship value ranging from 0 to 100.

The resulting indicators of amount of damage and relationship proximity were multiplied to arrive at a measure of maximum daily news value Figure 1.7 or total news value Figure 1.8.

1.5 Capabilities for detecting, handling, responding to critical conditions

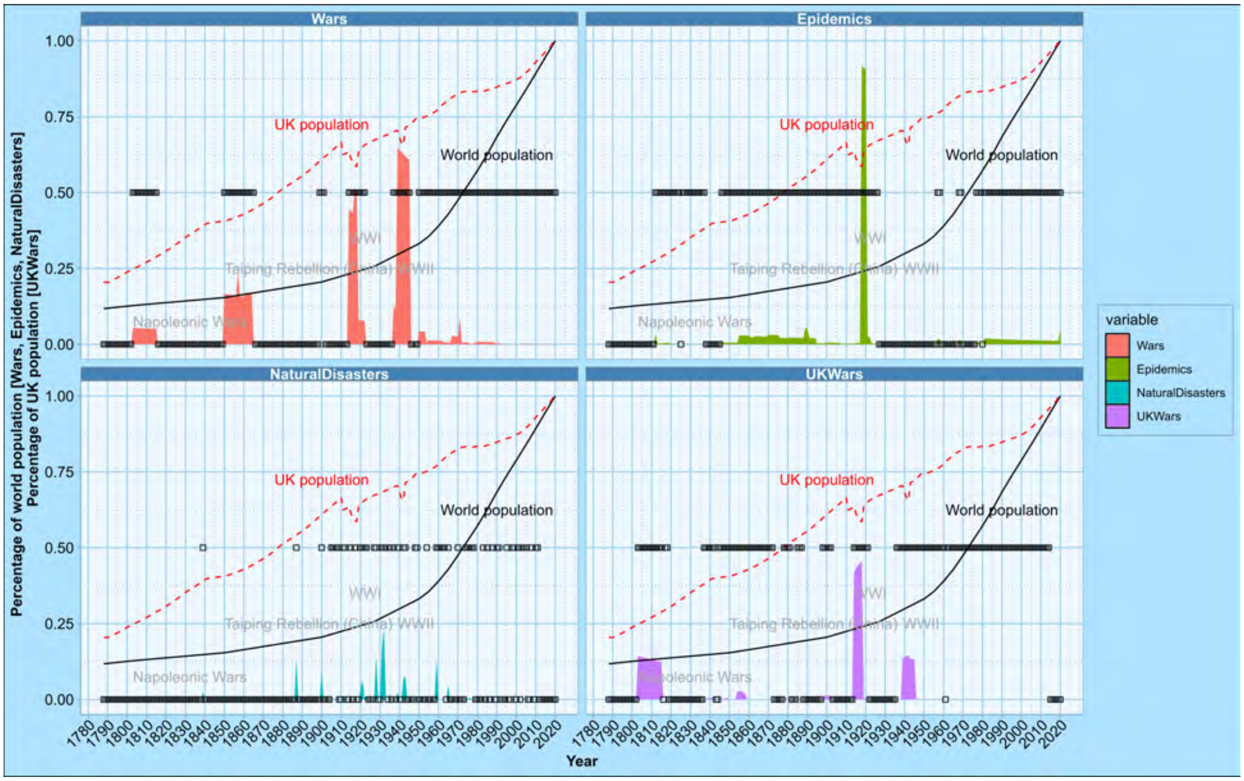
1.5.1 Improved event reporting and information flow

Obtaining information, keeping records and transmitting information was a very resource-intensive undertaking in the past which has become much cheaper and easier with modern information and communication technology (Kaukiainen 2001). The state (and its administration) and major corporations probably were the only sources that had enough resources to gather reliable information (but nowhere near the reliability and the coverage we expect today), and they were often in control which information would become public and which would not. Independent information gathering by neutral third parties probably was a rare exception before the 20th century. Imagine a mining accident in one of the British colonies. Certainly this was recognized and recorded, but it probably never was made public such that journalists could cover it—also to avert a discussion of safety standards and colonial management. Today, it would be very likely that journalists receive news of the incident through other channels. In effect, there will be a much greater number of critical conditions that are principally known and could be covered by news media, and less control for the state and for corporations to hinder public attention; attempts to keep obviously critical conditions a secret are likely to lead up to a scandal. The "news net" (Tuchman 1978) has become much denser.

From this inspection of relevant available data and analyses, we draw two tentative conclusions:

- In the UK and worldwide, more and better data on social affairs (and potentially critical conditions)

Figure 1.6: Percentage of Worldwide or UK population dying from wars, epidemics, and natural disasters, per year

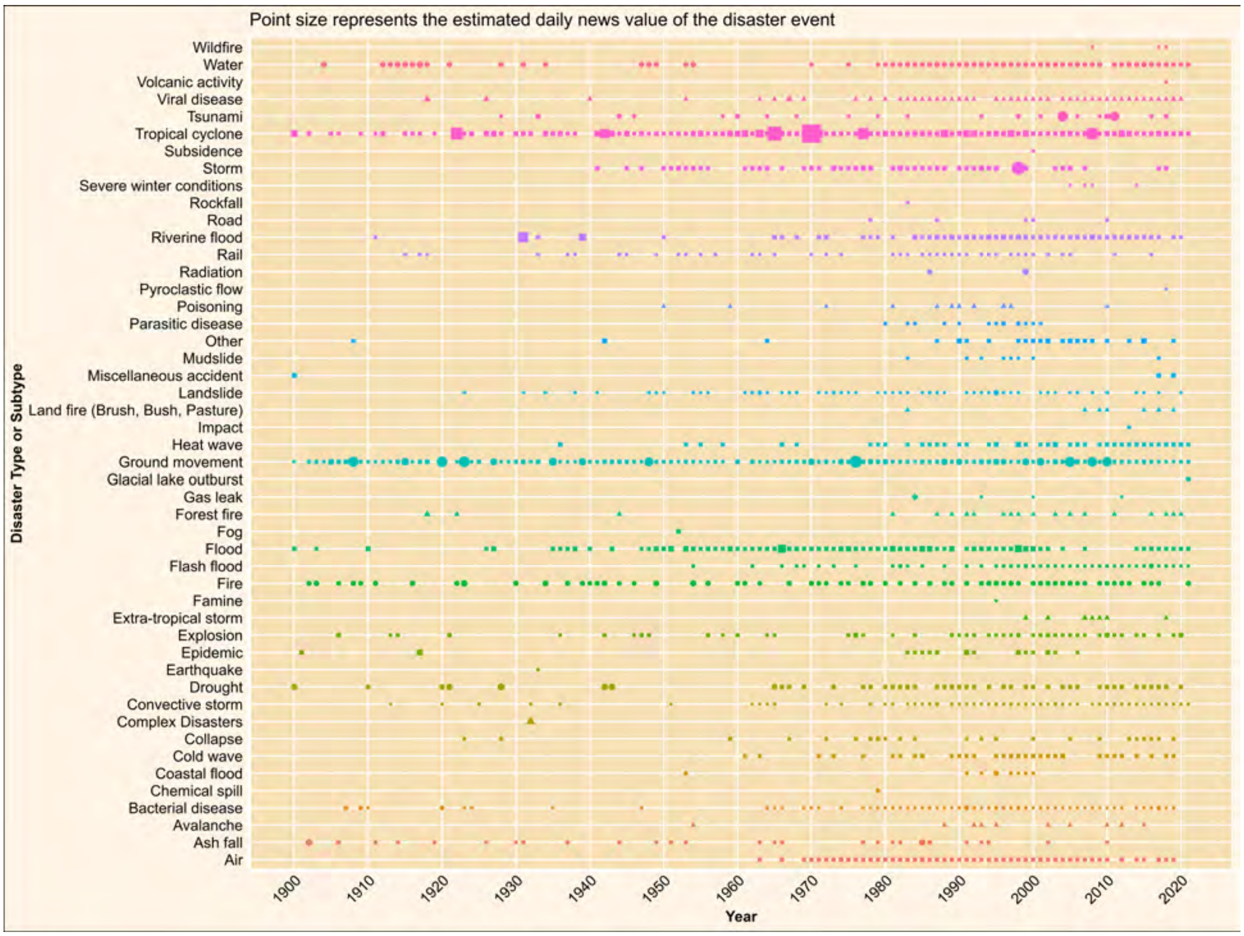


Note: The EM-DAT database (Centre for research on the Epidemiology of Disasters 2021) is more reliable since 1950 but does not go further back than 1900 and is obviously incomplete 1900-1950, it was therefore complemented by floods, earthquakes, epidemics, and famines listed in Wikipedia lists that rely on a wide range of sources some of which may be problematic. We include them nevertheless and always use the upper margin of estimates of fatalities available because ambiguity regarding the number of fatalities will usually be exploited by the news media to draw more attention, like in "up to 5'000 fatalities in major flooding".

Sources:

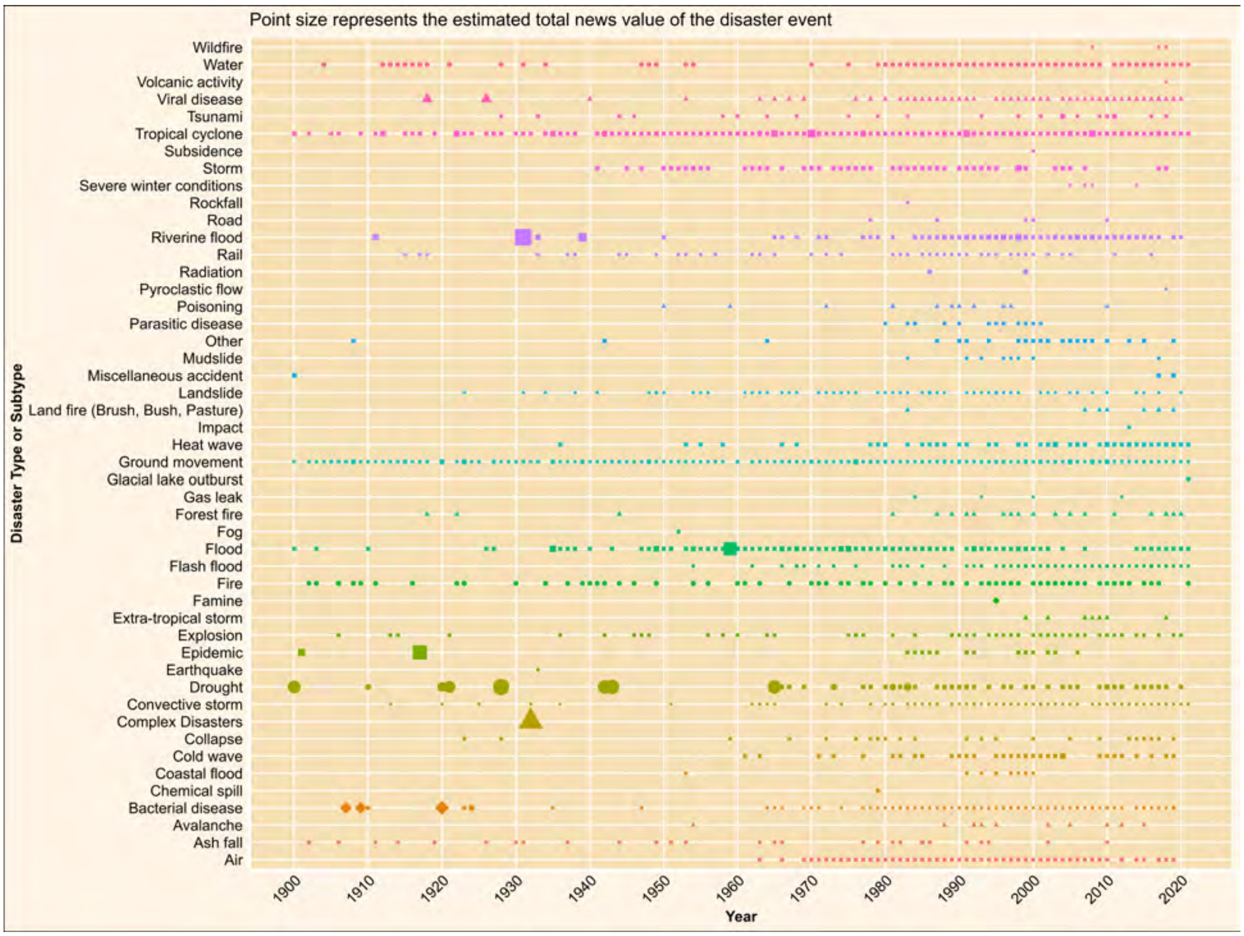
1. The international disaster database (EM-DAT) [epidemics, famines, floods, earthquakes, 1900-2020] (Centre for research on the Epidemiology of Disasters 2021);
2. Conflict Catalogue [wars, 1788-2020] (Brecke 2021)
3. Wikipedia lists 1788-2020 (compared to the EM-DAT data): Floods [all listed with >1000 fatalities; all reported in the British Isles]; Earthquakes, pre-1900 [all listed with >1000 fatalities; all reported in the British Isles]; Earthquakes, 20th [all listed with >1000 fatalities; all reported in the British Isles]; Earthquakes, 21st century [all listed with >1000 fatalities; all reported in the British Isles]; Epidemics [all listed with >10'000 fatalities; all reported in the British Isles] Famines [all listed > 10'000 fatalities worldwide; all reported in the British Isles]

Figure 1.7: Estimated maximum daily news value of different types of natural disasters over time (emphasizes temporal concentration of damage)



Source: (Centre for research on the Epidemiology of Disasters 2021)

Figure 1.8: Estimated total news value of different types of natural disasters over time (emphasizes total amount of damage)



Source: (Centre for research on the Epidemiology of Disasters 2021)

are collected and become more easily available, also by independent observers.
- Better available data greater pool of critical conditions the news media can choose from (and probably with higher news value). This will contribute to more CL and CNWs.

1.5.2 Increased public spending and government activity

Constant and rapid economic growth both necessitated and enabled an increase in the tax revenue, an expansion of the state's responsibilities beyond its basic provisions of administration, international affairs as well as internal (law and order) and external security (military). This means that state activity has increased tremendously, and society increasingly presupposes and relies that the state takes on many different responsibilities – e.g. in education, infrastructure and transport, welfare, housing, pensions, and healthcare. All these systems and institutions can produce critical conditions because their environment is constantly changing and they need to adapt to these changes (e.g. demographic change, changes in funding). This widens the pool of critical conditions that could become subject of CL and CNWs.

From this inspection of relevant available data and analyses, we draw two tentative conclusions:

- The state has expanded in quantity of spending and topics and responsibilities it covers.
- With state expansion, more critical conditions in state institutions and systems can become subject of CL and CNWs.

1.5.3 Effective medical treatment and sufficient nutrition

Many diseases that are well-treatable today were major causes of death in the past. In addition, hygiene has improved tremendously and has restricted the spread of many diseases, as has better knowledge of transmission of the various diseases. Working conditions have improved in terms of ergonomics, working hours and physical strain, with rights to pauses and vacation. Finally, pregnancy and birth complications as a major cause of death for women (1850: more than 40 deaths per 1000 total births) and children (1840: over 150 deaths per 1000 total births) have been declined to 0.07 deaths per 1000 life births (mothers) and 4 per 1000 life births (infants), respectively (Davenport 2021). Despite the enormous population growth around the world, famine and hunger have also decreased heavily [Figure 1.2; Our World in Data: Famine].

From this inspection of relevant available data and analyses, we draw one tentative conclusion:

- The frequency and extent of critical conditions in (formerly) frequent sources of death has decreased with more effective medical treatment.

1.6 Sensitivity, values, and identity

Even if critical conditions in society did not change at all, their appearance in newspaper coverage may change nonetheless when society changes. New goals come into sight, others fade into the background, others get transformed. Some of humanity's most deep-cutting upheavals have occurred in the past 240 years this study covers, and forced societies to also mentally adapt to major changes in their environment: The Industrial Revolution, the dissolution of the British Empire, both World Wars, and the Digital Revolution (Haferkamp and Smelser 1992; Inglehart and Flanagan 1987; Majima and Savage 2007; Duch and Taylor 1993; O'Brien 1996). Family structure got transformed, work life shifted from rural to industrial rhythms, many village-dwellers moved to the major towns, voting rights got expanded to all classes and to both sexes (Hudson 2014). It is hard to imagine that such fundamental transformations would not lead to a major shift in society's values and goals. And newspapers would accommodate to these changing patterns of expectation, perception, and sense-making. Five main trajectories of changes can be extracted:

1. changes in values and their relative importance (or priority);

2. changes in self-conception or identity of society, and its sense of control over its environment and ability to tackle challenges;
3. expectations and aspirations of what society wants to achieve, who is in charge of what, and who is accountable for what;
4. risk perception, in how we wage and accept risks in terms of their potential damage and the likelihood of that damage to occur; also, the relative importance of damage potential on the one hand and likelihood of occurrence on the other hand can be subject to change;
5. higher standards and broader coverage of data and reporting, coupled with higher speed and capacity of communication and information storage/retrieval.

1.6.1 Widening crisis definition

In the general history of the concept of crisis, a widening scope of application has been diagnosed. The term, originally referring to the decisive stage in the timeline of a sickness (Habermas 1973), has seen application to the decisive moments in a drama (Habermas 1973), to emergency situations (Schneider 2007-06-05/2007-06-07), to decisive moments in politics and diplomacy (James 1987), in the history of ideas (Ossewaarde 2018), and the cyclical development of the economy (Quiring et al., 2013), up until exis-tential threats of abstract social institutions and social systems (Estes 1983; Hay 2016; McBeth et al. 2013). The underlying process is tied to the metaphorical power of the term: if something can be called a crisis, this in turn signals that there is an existential external threat, that a decisive intervention is necessary, that there is time-pressure, that there is a demand for consensus and for working together (rally round the flag) (Oneal and Bryan 1995)—a very powerful set of meanings that can exert very powerful effects on the media agenda, the public agenda, and the policy agenda (Bennett, Lawrence, and Livingston 2007). Interest groups and politicians will therefore tend to call an event or condition a "crisis" and bit by bit contribute to expanding the meaning of the term further and further. The basic process we assume to occur is the widening of the conceptual field of crisis as it is commonly understood (and as it can be used by newspaper journalists).

From this inspection of relevant analyses, we draw three tentative conclusions:

- The conceptual meaning of "crisis" has widened 1785-2020.
- This would contribute to more CLs and more CNWs in newspaper coverage.
- More events and conditions fulfill the "stretched" definition of crisis.

1.6.2 Widening of the public sphere, composition of the reading public

The size of the readership of newspapers expanded tremendously in the 19th century, and the "Times" was a pioneer in developing and implementing printing technology innovations. This has also shifted the composition of the audience of the "Times" and of the part of the population that is considered as "the public". Along the same lines, more and more UK citizens became eligible to vote in the general elections, slowly but surely, with major breakthroughs in 1918 and 1928. This also came with a major transformation of the party landscape because workers—now eligible to vote—pushed the Labour Party to become the other major political party besides the Conservatives, while the Liberals were marginalized. This gave traction to new political topics and areas of politicization that were promoted by the worker's movement and the Labour Party. Also, issues that were important to the general masses, like unemployment, welfare, pensions, child care, would become more salient in political discourse; this could result in more media salience for them as well. They would also enter the realm of "crisis coverage". This would be true for the "Times" even though its readership tends to be upper-class because the whole structure of the public sphere changed with universal suffrage and universal access to public information and public debate through means of mass communication.

From this inspection of relevant available data and analyses, we draw three tentative conclusion:

- The public sphere in the UK has widened 1785-2020.
- This would contribute to more CL and more CNWs

- Topical areas interesting for lower-class people such as pensions, welfare, healthcare become more relevant and add to the spectrum of topics where crises can potentially arise.

1.6.3 Prioritized values and systems

How far one is willing to go to save a human life, to prevent an injury and trauma, to avert poverty and homelessness, to avert damages and loss of money? In contrast to these more material damages, how far are we willing to go allow freedom of expression of opinion, voting rights, personality rights? And how do these potential damages compare in serious to systems and institutions and their functions: international balance of power, government operability, the system of unemployment benefits, the healthcare system, and other abstract entities?

Crises are about a situation in which something that society places value on might be lost. Different sorts of critical conditions touch on different types of values: natural disasters, accidents, wars, and civil wars lead to loss of human lives, injury and physical damages; economic crises involve the loss of money, shortages of goods or services, the loss of employment and livelihood; government crises threaten the values of leadership and authority, and maybe also of democratic values; geopolitical crises revolve around values such as sovereignty of states, the "glory" of nations, the power of the involved states, and balance in the international system of power; social "systems" and institutions like healthcare, elderly care, child care, pensions, welfare and education and their functioning are a means of preventing and managing critical conditions to even arise and system functioning itself thereby emerged as a potential topic of crisis coverage (e.g. a "crisis of the pension system").

1.6.4 Ongoing politicization of society

When the state is relatively weak in terms of its size and budget, many social problems are not considered a matter the state has any responsibility for solving. For instance, the livelihood of older people, the handling of child care, and the maintenance of the environment and the infrastructure was mostly handled within local communities and families. With increased individualization, social mobility, urbanization, industrialization, these established structures became less and less effective. At the same time, economic growth allowed the state to expand and take on more and more responsibilities. The government apparatus grew in coverage and in power. This process of increasing politicization of many aspects of society that used to be taken for granted in a less mobile agricultural society opened up the opportunity to cover these aspects of societies in terms of "crisis" with a call for public attention and political reactions. Thereby, new topics entered the scope of "crisis coverage" (or significantly gained in attention): child care, elderly care, healthcare, welfare, pensions, the environment, education.

From this inspection of relevant available data and analyses, we draw two tentative conclusions:

- Social issues such as child care, elderly care, healthcare, welfare, pensions, the environment, and education have become increasingly politicized in the last centuries.
- Social issues have become more prominent topics of CL and CNW. The pool of potential critical conditions has expanded, which should lead to more CL and CNWs.

1.6.5 Extended expectations towards standards of living

On the one hand, higher standards of living and constant growth would lead to expectations of ever more growth and improvement; and not meeting these expectations may be experienced as a crisis. Colloquially put, expanding wealth may turn us into "spoilt brats" who take everything for granted and want ever more and more.

But there is another side to economic growth: The unprecedented economic growth that the UK (and many countries in the world) has transformed social life (and disrupted established practices) such that many new problems and demands arose and previous life styles were no longer possible. And satisfying these demands is necessary to take part in society and contribute to economic growth; life gets more expensive. For instance, owning a smartphone was a luxury 10 years ago and has become a requirement to participate in social life and even on the job. In a similar vein, industrialization and urbanization meant that the established ways of dealing with older people and younger people within the family and the community were no longer applicable; and after long struggles, the social security and healthcare systems that we know today were implemented.

From this inspection of relevant available data and analyses, we draw two tentative conclusions:

- The expected minimum standard of living has grown more expensive between 1785-2020. Economic growth is necessary to satisfy the growing expectations.
- The value of economic development has increased, lowering the threshold for economic topics to become object of CL and CNWs.

1.6.6 Extended expectations towards politics and science

With a feeling of controllability of many critical conditions, and a stronger (and more costly) state, would come the expectation that the state effectively combats critical conditions, or even effectively prevents them from happening. The same goes for science and technology, who have contributed many solutions to problems that humankind has faced helplessly in a long time. In the same vein, some critical conditions may receive more attention because controlling, preventing, containing them is now viewed as a realistic possibility:

From this inspection of relevant available data and analyses, we draw two tentative conclusions:

- Society's capacity to prevent and solve critical conditions has improved 1785-2020. Expectations that critical conditions can be prevented or solved has also increased 1785-2020.
- Critical conditions are more likely to provoke CL or CNWs because solving the conditions and/or stimulating political activity is viewed as realistic in more cases.

1.6.7 The increasing value of systemic functioning

Wealth and economic growth have led to new social problems that have largely been solved by expanding the state and creating social security and healthcare systems. Funding and maintaining these systems has therefore emerged as a major topic of political debates because a threat to these systems would indirectly lead to a re-emergence of many of the problems these systems were designed to solve or limit. Threats to the functioning of these systems has in turn become a major set of critical conditions.

From this inspection of relevant available data and analyses, we draw two tentative conclusions:

- The functioning of social systems has become a major value for modern societies.
- New critical conditions that threaten the functioning of social systems have become relevant occasions for CL and CNWs.

1.6.8 The increasing importance of postmaterialist and individualist values

Wealth contributes to a change in values that gives more weight to postmaterialist values such as "self-expression" and less weight to materialist values; the more basic needs to sustain one's livelihood (represented by materialist values) are pushed into the background in wealthy societies where many people simply take them for granted. However, this shift to postmaterialist values is not continuing in a linear fashion (Majima and Savage 2007); there is a floor effect for materialist values below which they do not decline, e.g. because there is still a high number of persons who struggle economically (e.g. those who are unemployed) or even among well-off citizens there is a substantial number of people who aspire to ever more income and wealth. Postmaterialist values more or less immediately respond to changes of the economic situation of the individual, giving way to materialist values if the individual experiences economic struggle (Duch and Taylor 1993).

From this inspection of relevant available data and analyses, we draw two tentative conclusions:

- Despite the growing importance of postmaterialist values, materialist values remain activatable and important. If lives are threatened or wealth is threatened, this will still cause widespread concern and push back other values.
- Disasters, wars, and economic recessions – if they happen – still have a high potential to cause CL and CNWs.

1.6.9 The development of imperialism and nationalism

The 19th century is often viewed as the "age of nationalism" and the "age of imperialism", with the UK as its prime example. The "glory of the nation" was an important driver of both foreign and domestic policy, and the struggle for outcompeting other nations is often viewed as a major reason behind the two world wars. Despite nationalism's persistence today, it is likely that policymaking for the "glory of the nation" has decreased in importance in the last decades: the dissolution of the British Empire and its transformation into a loose confederation (the "Commonwealth of Nations"), Britain's joining (and later leaving) of the European Union, Globalization trends, and the struggle to keep the different nations inside the UK (prominently the English, Scottish, Welsh, (Northern) Irish) within the union shows how the UK struggles with its national identity (Goulbourne 1991).

1.6.10 From everday risks to dread risks

There are no longitudinal studies on structural changes in the way individuals assess risks. Assuming that the general mechanisms that are observed today have not changed fundamentally, we can try to reconstruct risk perception as it occurred in the past based on changes in the environment. Exposure to a risk in terms of direct experience or indirect experience (e.g. through media coverage) raises awareness of that risk (for indirect experience, geographical proximity moderates the effect) and tends to lead to higher ratings of that risk. This effect would be stronger for low-probability/high-damage events (e.g. airplane disasters) than for high-probability/low-damage events (e.g. car accidents) (Lichtenstein et al. 1978; Slovic, Fischhoff, and Lichtenstein 1986). Today, better reporting and communication infrastructure may increase the salience of low-probability/high-damage events, mainly through news media. In contrast, the experience of serious everyday risks occurs at a much lower rate with the marginalization of many existential threats (e.g. deadly infections, poverty, undernourishment). In conjunction, this would lead to a risk perception and a risk management that concentrates on dread risks and unknown risks (and even more dread/unknown risks) (which have increased in salience) than on known/non-dread risks (which have decreased in salience)—and the level of risk management moves from individual responsibility (to manage non-dread/known risks) to public/government/science responsibility (to manage dread/unknown risks).

Also, the way risks are supposed to be managed will change in this environment. According to Prospect Theory (Tversky and Kahneman 1981), a shift of attention towards "dread risks" (high potential damage, low probability) would lead to very risk-averse decision-making (because few heavy losses rather than many small losses are feared) compared to the behavior that would be expected under perfect rationality: people confronted with dread risks would invest strongly into averting the risk.

From this inspection of relevant available data and analyses, we draw two tentative conclusions:

- Societal risk management changes to a focus on averting/preventing high-damage/low-probability ("dread") events; so information about such events become more valuable (and is more readily available as well).
- The share of total coverage devoted to crises (CC and CNW) will increase

1.6.10.1 Self-conception of control

The last 230 years have seen an incredible expansion of knowledge and technology that increased the leverage that humankind has for combating critical conditions. Problems that seemed unsolvable in the past now have begun to appear solvable. For instance, famine and starvation have been part of the human condition for a long period and continue to be, but the ability to reliably produce enough food for a population of 8 Billion people (despite a major shrinkage of the agricultural workforce) suggest that famine and starvation are not a matter of fate but something we can effectively reduce or even eradicate.

With a feeling of controllability of many critical conditions (also through technology), and a stronger state, would come the expectation that the state effectively combats critical conditions, or even effectively prevents them from happening. The much increased government apparatus would take the form of specialized institutions that keep "systems" running that serve to prevent critical situations or cope with them if they arise. This again means that the "systems" or institutions themselves can come to the center of attention when their functioning is threatened. For instance, underfunded pension systems can be treated as a critical condition in and of itself before any pensioner has lost anything. Possible mismanagement of a financial regulation agency can pose a critical condition without any economic crisis being in sight.

In the same vein, some critical conditions may receive more attention because controlling, preventing, containing them is now viewed as a realistic possibility:

From this inspection of relevant available data and analyses, we draw one tentative conclusion:

- The topical spectrum of crises and the types of crises will shift towards "systemic" crises that are detached from acute critical events: education, health (non-epidemic), welfare, pensions.

Chapter 2

Appendix B: External data sources for indicators for potential drivers of crisis coverage

2.1 Purpose

This appendix provides visualizations of the data used as predictors of crisis coverage in the study: public spending intensity, public spending diversity, suffrage in the UK, media penetration, voter de-alignment (as a measure of media autonomy), and VDEM indicators of media autonomy.

2.2 Voter Dealignment in the UK: Comparing possible indicators

The data used for estimating the degree of voter dealignment in the UK were the cross-sectional data sets of the British Election Study (2019, 2017, 2015, 2010, 2005a, 2005b, 2001, 1997, 1992, 1987, 1983, 1979, 1974a, 1974b, 1970, 1966, 1964, 1963) (British Election Studies, cross-sectional data).

We considered two different possible indicators of voter dealignment in the UK (as a proxy to media autonomy in the UK):

- The proportion of voters surveyed that reported no party identification (**no party ID indicator**) [0/1].
- The average strength of party identification that voters reported (**strength of party ID indicator**) [0/1/2/3/4].

The former measure dichotomizes the latter measure with a split below 1 [$0 \rightarrow 0$; $1/2/3/4 \rightarrow 4$] (Figure 2.1).

The correlation between the two measures at the individual level is $R = -.825; t(48437) = -321.8; p < .001$. At the aggregate level (per survey), the correlation between the survey-level averages is $R = -.950; t(16) = -9.65; p < .001$.

Given the very slight differences the higher-resolution measure makes at the aggregate level, we decided to use the simpler dichotomized measure which is easier to describe and to comprehend, with similar informational content. Also, we saw some issues with the *1—weak party ID* mapping on an equidistant scale; we wanted to avoid these problems unless necessary.

2.3 Data Visualizations

Figure 2.1: Question Model for Party ID in the British Election Study

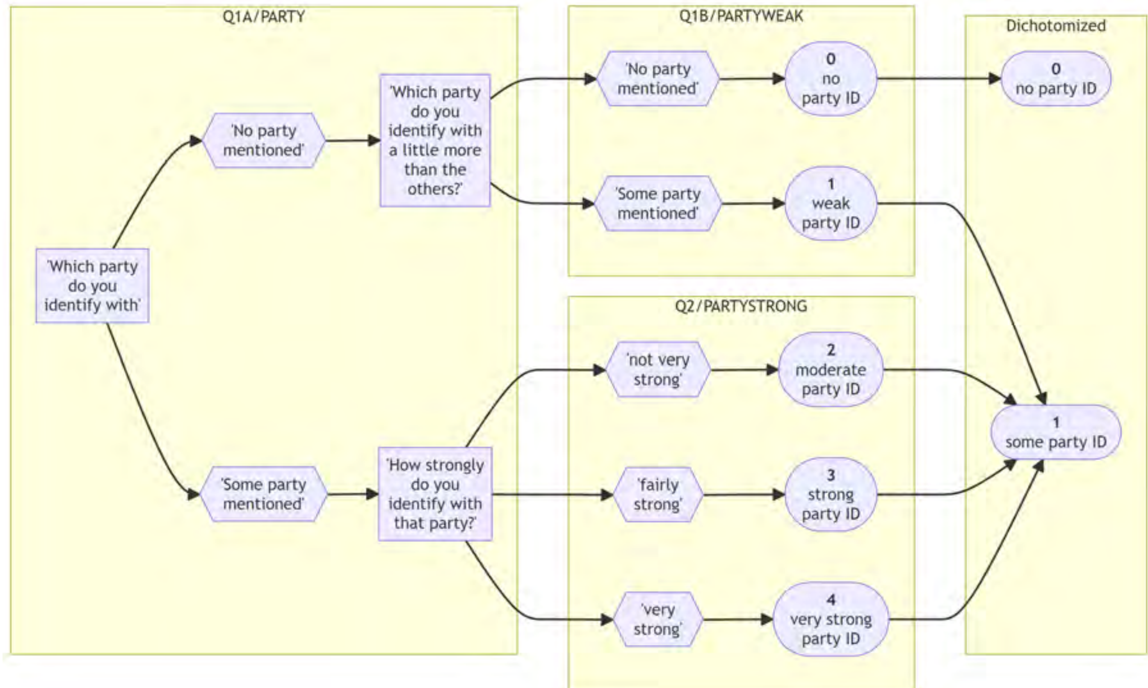
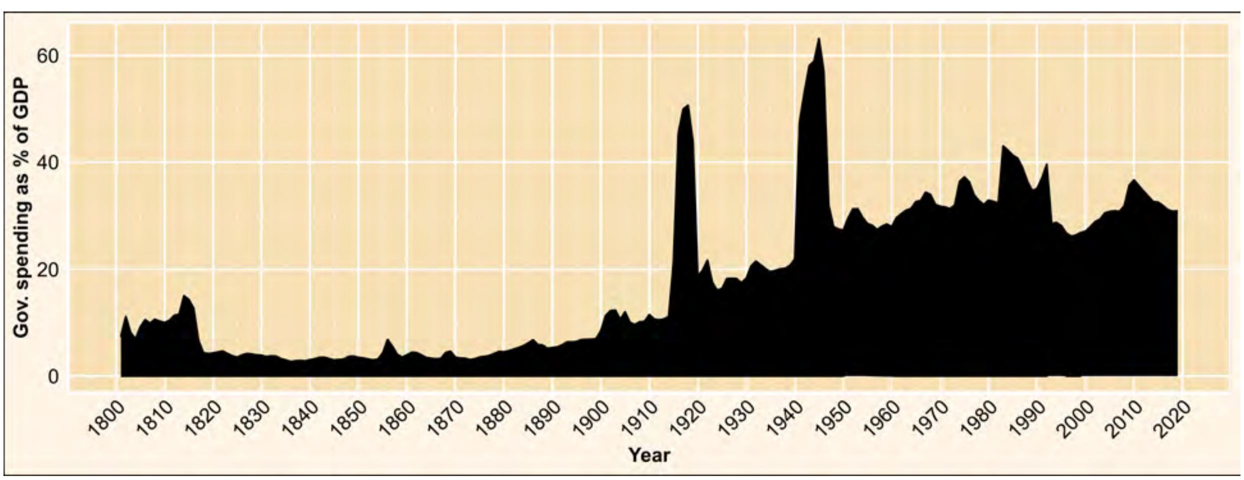
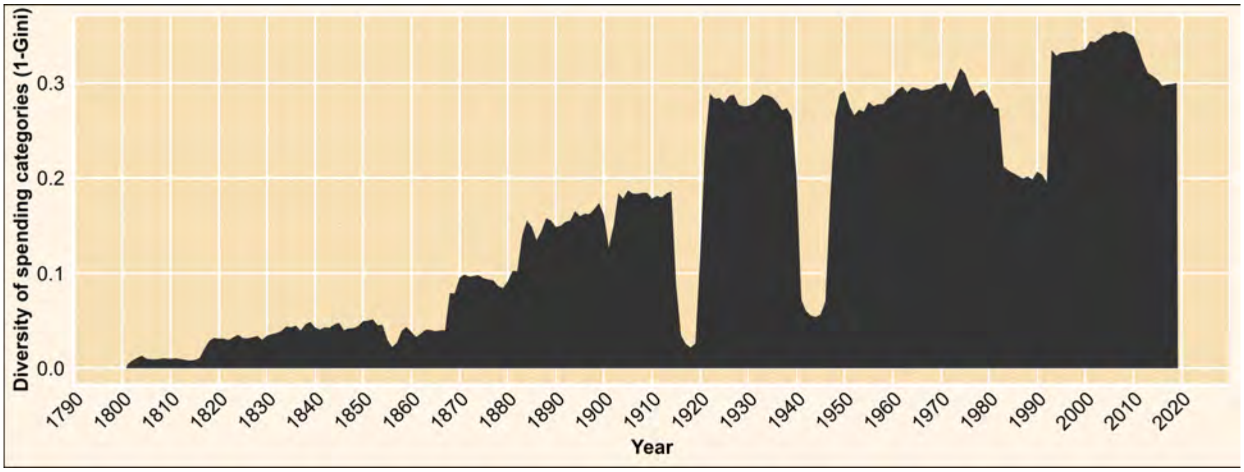


Figure 2.2: Public spending intensity in the UK 1800–2019



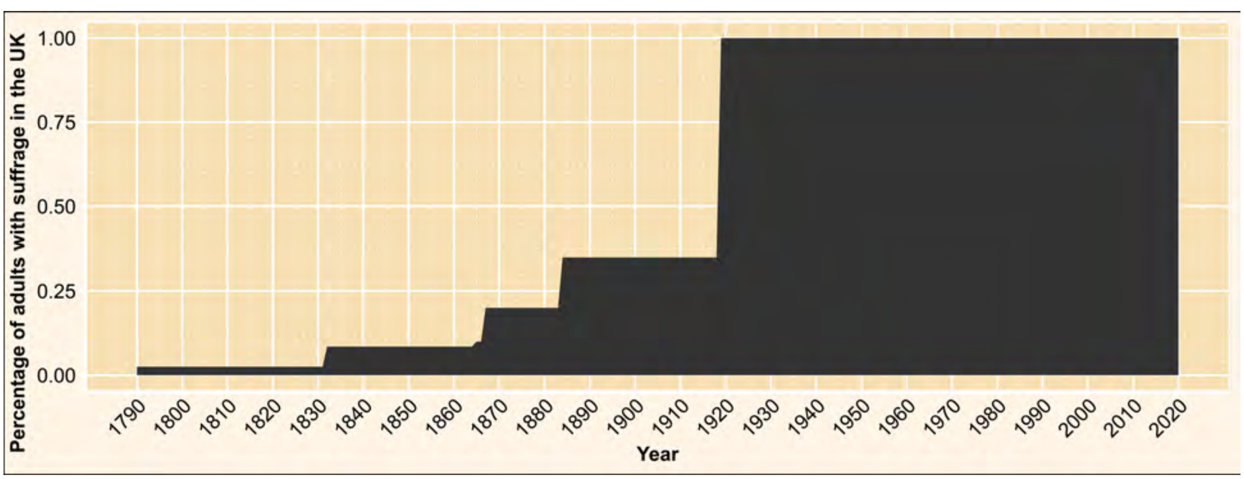
Source: Data on government spending are available from official sources (Central Statistical Office, Public Expenditure Statistical Analysis) and from a historical statistics compilation (Mitchell 2011). The data were compiled by and retrieved from a private website (Chantrill 2023) whose data we spot-checked against the official data repositories and found to be accurate. The GDP data was retrieved from (Ryland and Williamson 2021).

Figure 2.3: Public spending diversity in the UK 1800–2019



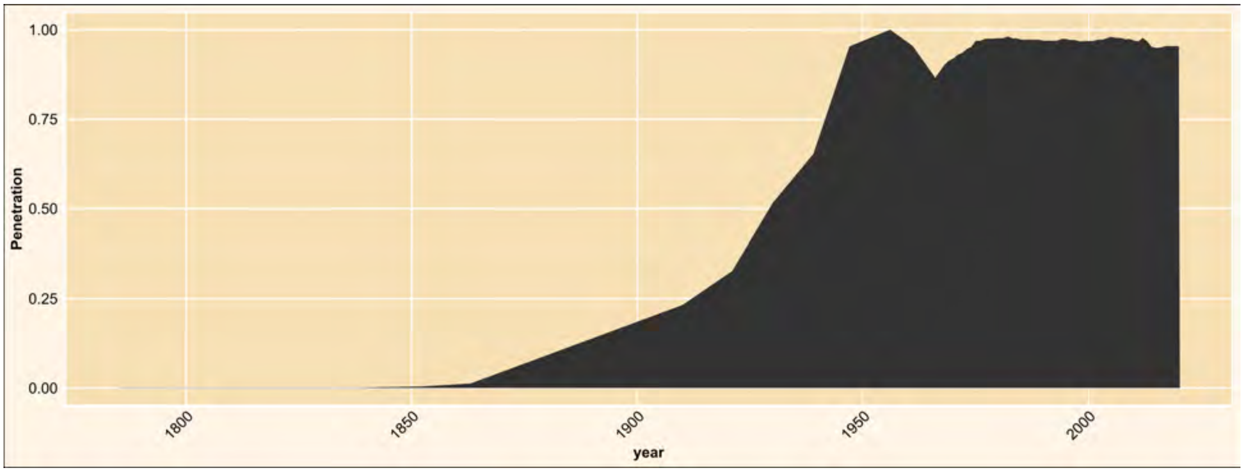
Source: Data on government spending are available from official sources (Central Statistical Office, Public Expenditure Statistical Analysis) and from a historical statistics compilation (Mitchell 2011). The data were compiled by and retrieved from a private website (Chantrill 2023) whose data we spot-checked against the official data repositories and found to be accurate. The GDP data was retrieved from (Ryland and Williamson 2021).

Figure 2.4: Percent of adults with suffrage in the UK



Source: (V-Dem Institute 2022).

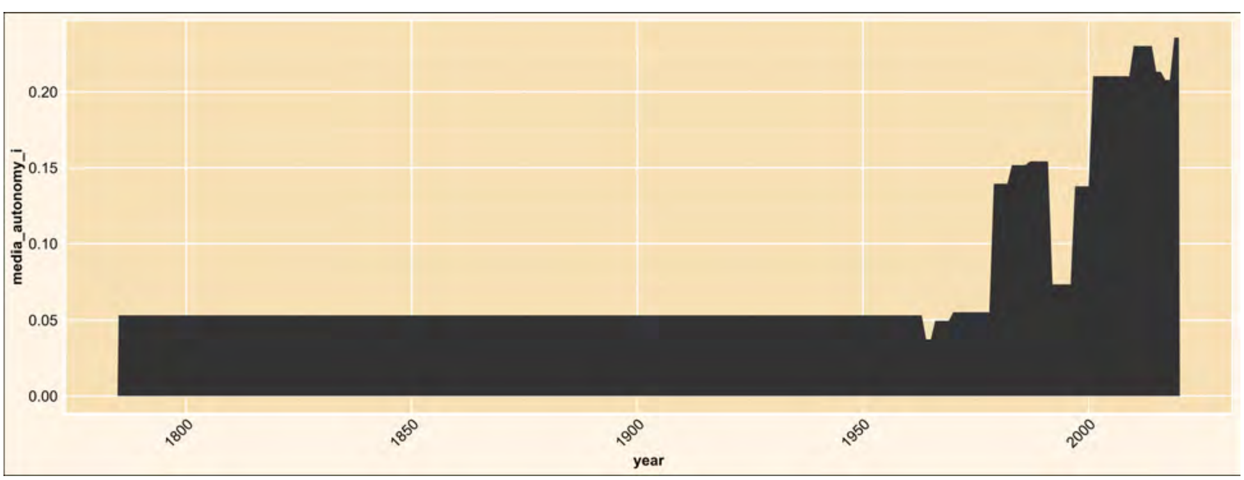
Figure 2.5: News media (newspapers, TV, online) reach in the UK



Source:

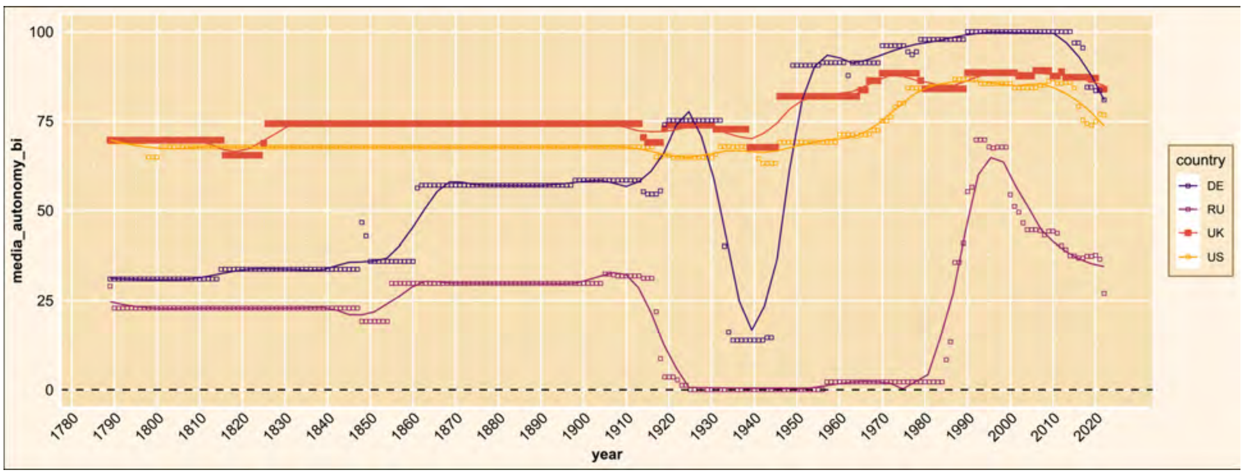
- Newspapers: After 1950, the Audit Bureau of Circulation (ABC) provides definitive figures on newspaper circulation; before that, we rely on a mixture of ABC data, data collected by advertisers, and data collected for collecting stamp tax. The years covered are 1838, 1852, 1863, 1910, 1921, 1930, 1939, 1947, 1956, 1961, 1966, 1976, 1980, 1987, 1992, 1997, 2000-2020. For the years not covered, we used linear interpolation. The data from these sources was compiled and we collected it via the Wikipedia page List of newspaper in the United Kingdom by circulation. We checked the original sources cited on the Wikipedia page for a small subset of the numbers.
- Television: The share of UK households with TV sets (1956-2019) was documented by Closer.ac.uk, a longitudinal data hosting service hosted by the University College London's Social Research Institute; it is a collaboration involving UK Data Service and The British Library.
- Internet: We used the World Bank's data repository to retrieve the International Telecommunication Union's data on Internet usership in the UK, given as percentage of the population. We used these proportions (maximum value: 2020: 0.95) as an indicator of Internet penetration.
- Aggregation to media penetration index: According to our model, the vital part of media penetration is what share of the population can be expected to receive regular updates on political matters. A conservative estimate of that is the maximum of newspaper penetration, TV penetration, or Internet penetration in the respective year. This estimate is conservative because it presumes that there is full overlap between the higher-penetration sources and all the lower-penetration sources such that the full reach equals "only" the reach of the highest-penetration source.

Figure 2.6: Voter dealignment in the UK as indicator of media autonomy



Source: British Election Studies, cross-sectional data.

Figure 2.7: Index from VDEM indicators of media autonomy



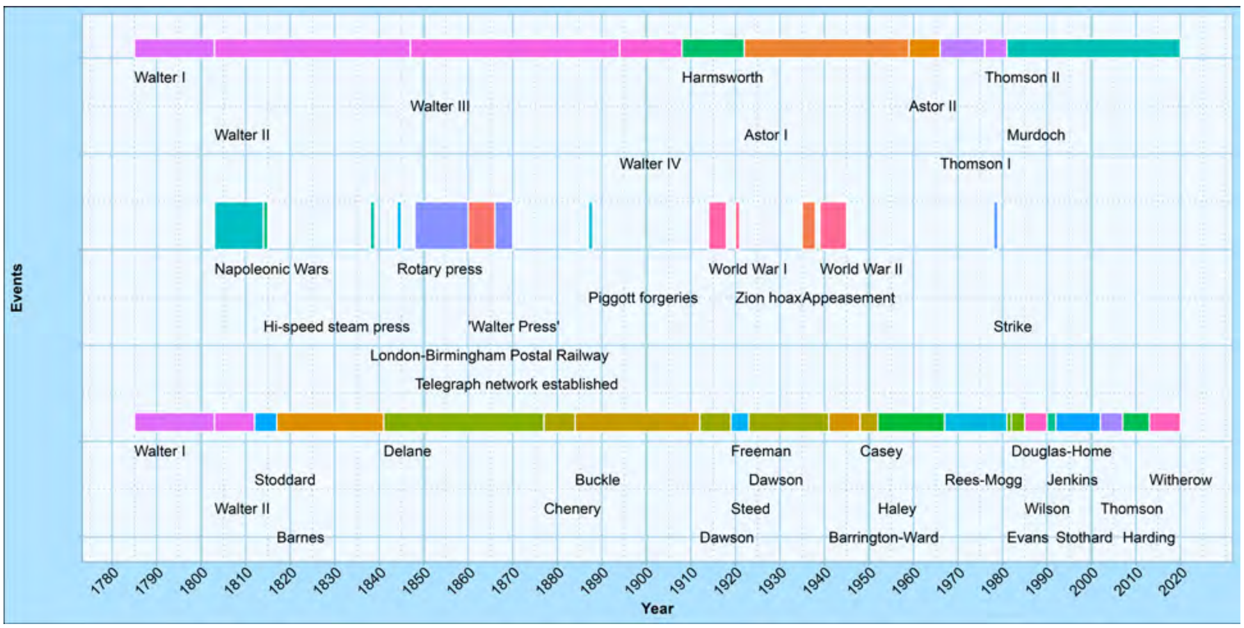
Note. Index based on VDEM (V-Dem Institute 2022) indicators *Harassment of journalists* (0-4), *Media self-censorship* (0-3), *Media bias against opposition* (0-4), *Print/broadcast media perspectives (broadness)* (0-3), *Print/broadcast media critical towards government* (0-3). Additive index where higher values represent greater media autonomy. Standardized against the highest observed value (maximum: Germany 1990–2014) and the lowest observed value (minimum: Russia 1925–1956) in the following countries: China, Switzerland, Germany, Netherlands, Norway, Russia, UK, USA.

Chapter 3

Appendix C: *The Times* in Context

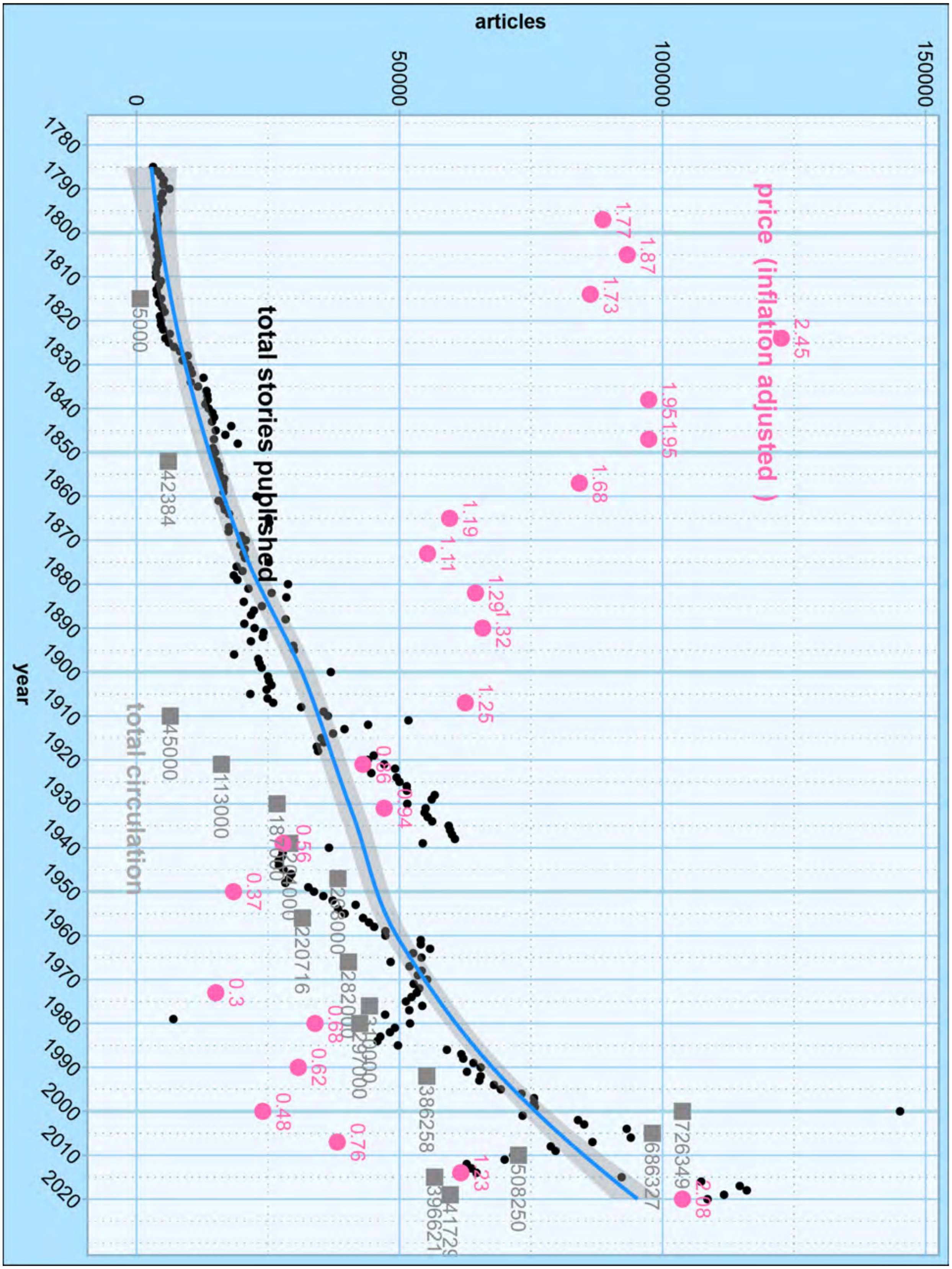
This appendix provides background information about the history of The Times that may be relevant for interpreting the results.

Figure 3.1: Timeline of the history of *The Times* 1785–2020



Note. **Top timeline**: Ownership regimes. **Center timeline**: Historical events (Times specific, newspaper specific, and general). **Bottom timeline** Editor regimes. **Sources**: Editor regimes (editors, including the articles about each editor). Ownership regimes (owners, including the articles about each owner)

Figure 3.2: Timeline of the price, circulation, and volume of *The Times* 1785–2020



Note. **Total stories published:** The number of total news stories was retrieved by searching with a blank search term, year by year (Gale Digital Scholar lab). **Price.** Raw price was recorded based on the price listed on the scanned front page for one randomly selected edition in each decade; inflation adjustment based on MeasuringWorth (2023). **Total circulation.** Audit Bureau of Circulations. Reid, Alanah (2020). A History of the Times Newspaper. Historic Newspapers. https://www.historic-newspapers.co.uk/blog/the-times-newspaper-history/; Encyclopaedia Britannica (2021). The Times. British Newspaper. https://www.britannica.com/topic/The-Times for circulation estimates. Ellegård, Alvar (1971). The readership of the periodical press in Mid-Victorian Britain: II. Directory. Victorian Periodicals Newsletter, 4(3), 3-22.

Chapter 4

Appendix D: Comparing the Long-Term Development of Crisis Labelling Salience in *The Times* and *The Guardian*

4.1 Purpose

The purpose of this appendix is to illustrate the potential for generalization of the findings we obtained in an in-depth analysis of crisis coverage in *The Times*. Our approach is to collect comparable data for *The Guardian*, another long-running UK newspaper-of-record that is considered center-left rather than center-right (as *The Times*). The search string was simplified to "crisis OR crises" due to the limitations of the search function in the archive. The search prompted 153,526 newspaper pages with hits (CL pages).

4.2 Data and Method

Data source 1: We rely on the data for *The Times* from Gale and Factiva full-text archives that we also use for our main study. *Articles* are the unit of analysis (so we distinguish between articles with and without crisis coverage).

Data source 2: We rely on keyword searches for the keyword *crisis* (as well as a dummy search for the generic term *today* to measure the total volume of coverage) in a searchable (but not downloadable) archive of *The Guardian*. These data serve only for validation and exploration purposes and, other than the analyses based on *The Times*, we have not and cannot explore the quality of the data these results are based on (and also we cannot obtain results regarding crisis news waves given the lack of downloadable full-text data).

4.3 Results

The explanatory model for *The Times* and *The Guardian* together matches very well the explanatory model for *The Times* alone that is reported in the paper (Table 4.1). The time series of crisis labelling salience in *The Times* and *The Guardian* are strongly correlated ($r = 0.93$, $p < .001$ for number of crisis labelling articles/pages; and $r = 0.70$, $p < .001$ for crisis labelling salience) and the time series to a large degree run in parallel (Figure 4.1). Mediation analyses show that the impact of *media autonomy* Figure 4.2 and *government spending diversity* Figure 4.3 is only partially mediated by *crisis frame sponsor activity*.

4.4 Discussion

Since the Guardian is considered a center-left newspaper-of-record in the UK, we take the correspondence of the time series and the robustness of the regression results as evidence that our results reflect a broader pattern observable in the public sphere in the UK more generally; whether this affects other press segments such as regional and tabloid press is not certain, however. Still, these newspapers should be considered pivotal in shaping elite and public views of current events and potential crises, and hence our results hold strong general national-level implications regarding the construction of crisis in the UK.

Figure 4.1: CL salience—The Times and Guardian

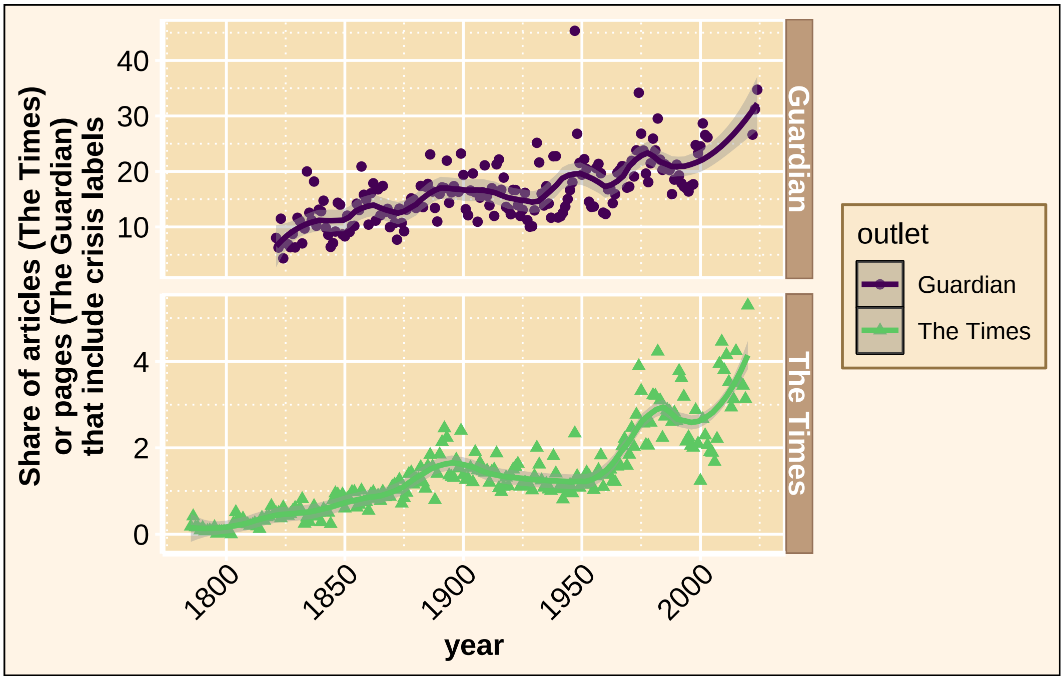


Table 4.1: Impact of time, pervasiveness of media logic, and public spending on CL salience—The Times and The Guardian

	(1)	(2)	(3)	(4)
(Intercept)	0.018 *	0.030 *	0.019 +	0.009
	(0.008)	(0.015)	(0.011)	(0.011)
Decades since 1785	0.011 ***			0.004 **
	(0.000)			(0.001)
Media Penetration		-0.013	-0.024	-0.029
		(0.025)	(0.019)	(0.019)
Media Autonomy		0.760 ***	0.234 ***	0.196 **
		(0.070)	(0.060)	(0.061)

	(1)	(2)	(3)	(4)
Public Spending Intensity		0.101 *	0.113 ***	0.054
		(0.045)	(0.034)	(0.039)
Public Spending Diversity		0.336 ***	0.165 ***	0.062
		(0.053)	(0.041)	(0.054)
Crisis Frame Sponsor Activity			0.024 ***	0.022 ***
			(0.001)	(0.002)
Num.Obs.	422	419	419	419
R2 Marg.	0.618	0.582	0.762	0.769
R2 Cond.	0.628	0.625	0.785	0.788
AIC	-1259.8	-1229.9	-1456.8	-1451.7
BIC	-1243.6	-1201.6	-1424.4	-1415.3
ICC	0.0	0.1	0.1	0.1
RMSE	0.05	0.05	0.04	0.04

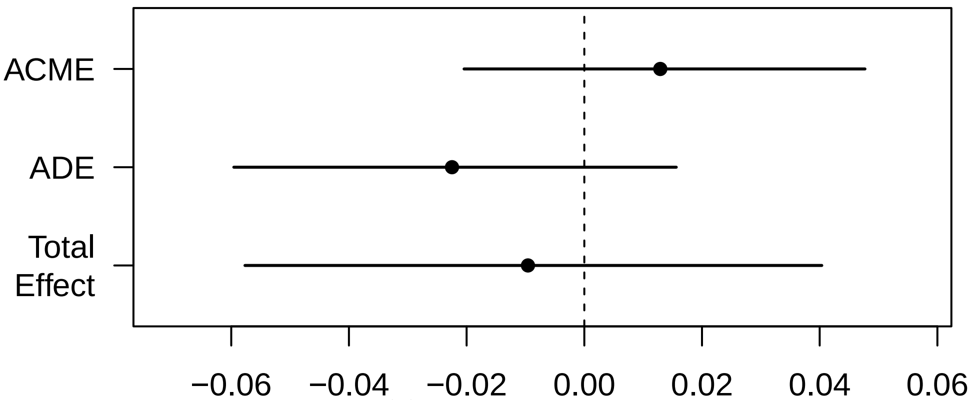
Table 4.2: Impact of pervasiveness of media logic, and public spending on crisis frame sponsor activity—The Times and The Guardian

	(1)
(Intercept)	0.419 +
	(0.233)
Media Penetration	0.466
	(0.683)
Media Autonomy	21.602 ***
	(1.924)
Public Spending Intensity	-0.496
	(1.239)

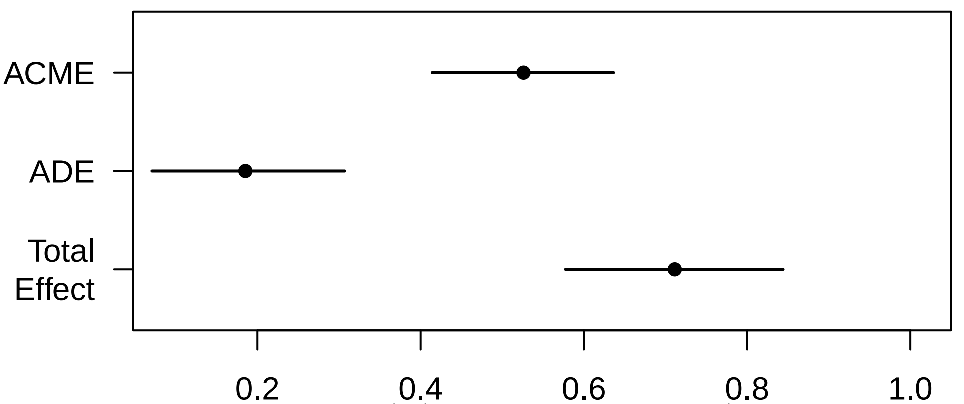
	(1)
Public Spending Diversity	7.093 ***
	(1.459)
Num.Obs.	419
R2 Marg.	0.549
R2 Cond.	0.555
AIC	1509.4
BIC	1537.6
ICC	0.0
RMSE	1.44

Figure 4.2: Mediation analysis of media logic-related drivers' impact on CL salience with crisis frame sponsor activity as mediator—The Times and The Guardian

(a) Independent variable—Media penetration



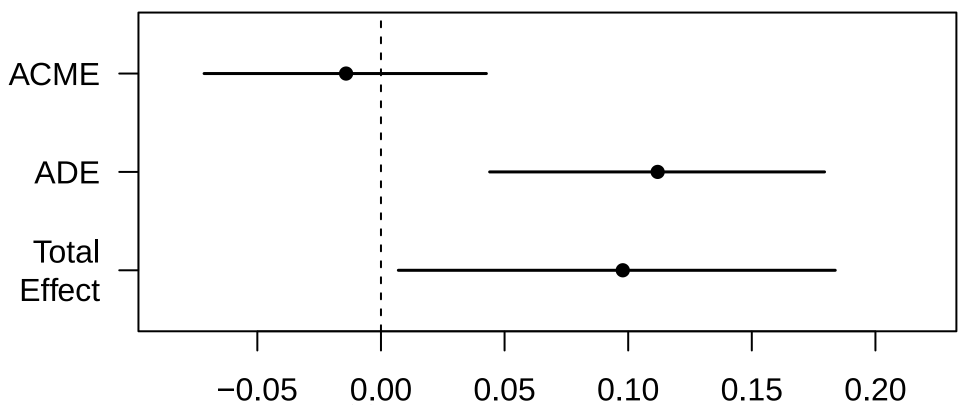
(b) Independent variable—Media autonomy



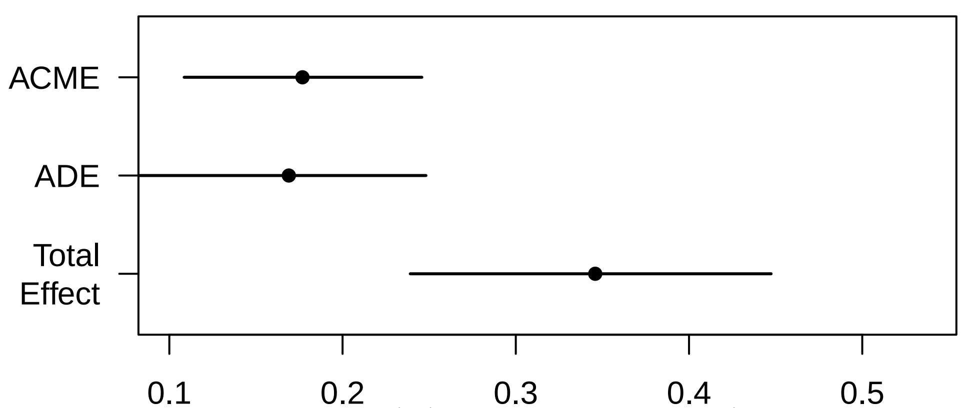
Note. ACME: Average causal mediation effect; ADE: Average direct effect

Figure 4.3: Mediation analysis of public spending-related drivers' impact on CL salience with crisis frame sponsor activity as mediator—The Times and The Guardian

(a) Independent variable—Public spending intensity



(b) Independent variable—Public spending diversity



Note. ACME: Average causal mediation effect; ADE: Average direct effect

Chapter 5

Appendix E: Comparative Analysis in Times, Economist, NZZ, and Washington Evening Star

5.1 Purpose

This appendix presents results for models equivalent to those used in the main study, but with an extended set of four outlets that are highly diverse in being from three different countries (USA, UK, Switzerland) and two different periodicities (daily, weekly). In order to consider the inherent between-outlet differences, we rely on mixed models with random intercepts for outlets. Given the unavailability of budget and media data for USA and Switzerland, we can only investigate H1.1–H1.3.

5.2 Method

We used the same search string we used for *The Times* (for *NZZ*: A German translation thereof) to collect full texts of crisis coverage from *NZZ*, *The Economist*, and *Washington Evening Star*. (*NZZ*: 162,101 CL articles, *The Economist*: 104,910 CL articles, *Washington Evening Star*: 321,988 CL pages).

“cris!s” OR “disast” *OR “catastrophe*” OR “pandemic” *OR “epidemic*” OR “recession” *OR “breakdown*” OR “collapse” *OR “debacle*” OR “emergency” OR “emergencies”.

“Krise” *OR “Desaster*” OR “Katastrophe” *OR “Pandemie*” OR “Epidemie” *OR “Rezession*” OR “Zusammenbruch” *OR Einsturz*” OR “Debakel” *OR “Notfall*” OR “Notfäll*.

5.3 Visualizations

5.4 Regression Results

Model G1.1 shows that across a diverse set of media from different countries, there is an upward trend in CL salience Table 5.1, in line with H1.1.

Model G1.2 shows that across a diverse set of media from different countries, there are no upward or downward trends in CL salience Table 5.1, which does not support H1.2.

Model G1.3 shows that across a diverse set of media from different countries, there is an upward trend in CNW count Table 5.1, in line with H1.3.

Figure 5.1: Total number of articles published per year in four newspapers

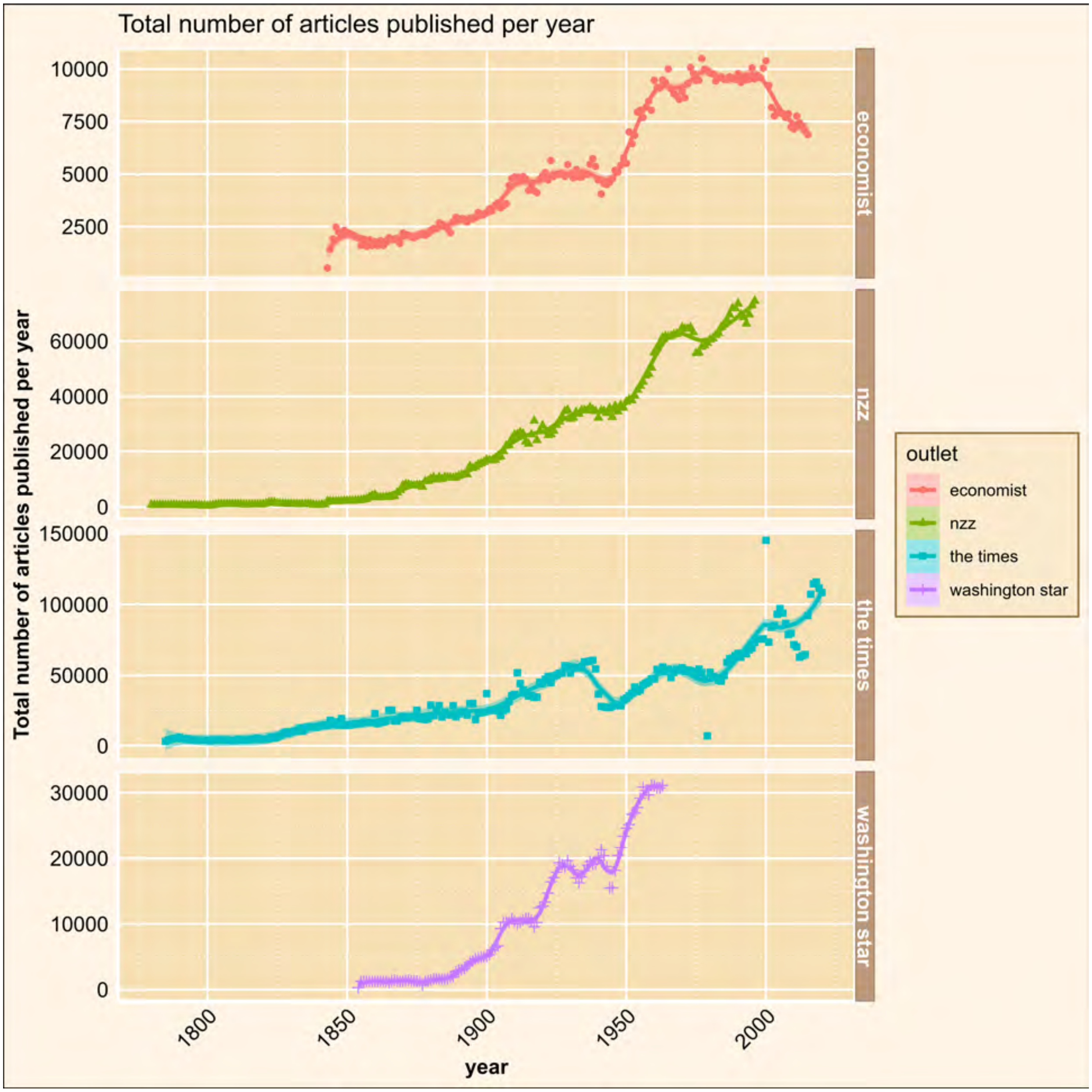


Figure 5.2: Count of articles with crisis labelling published per year in four newspapers

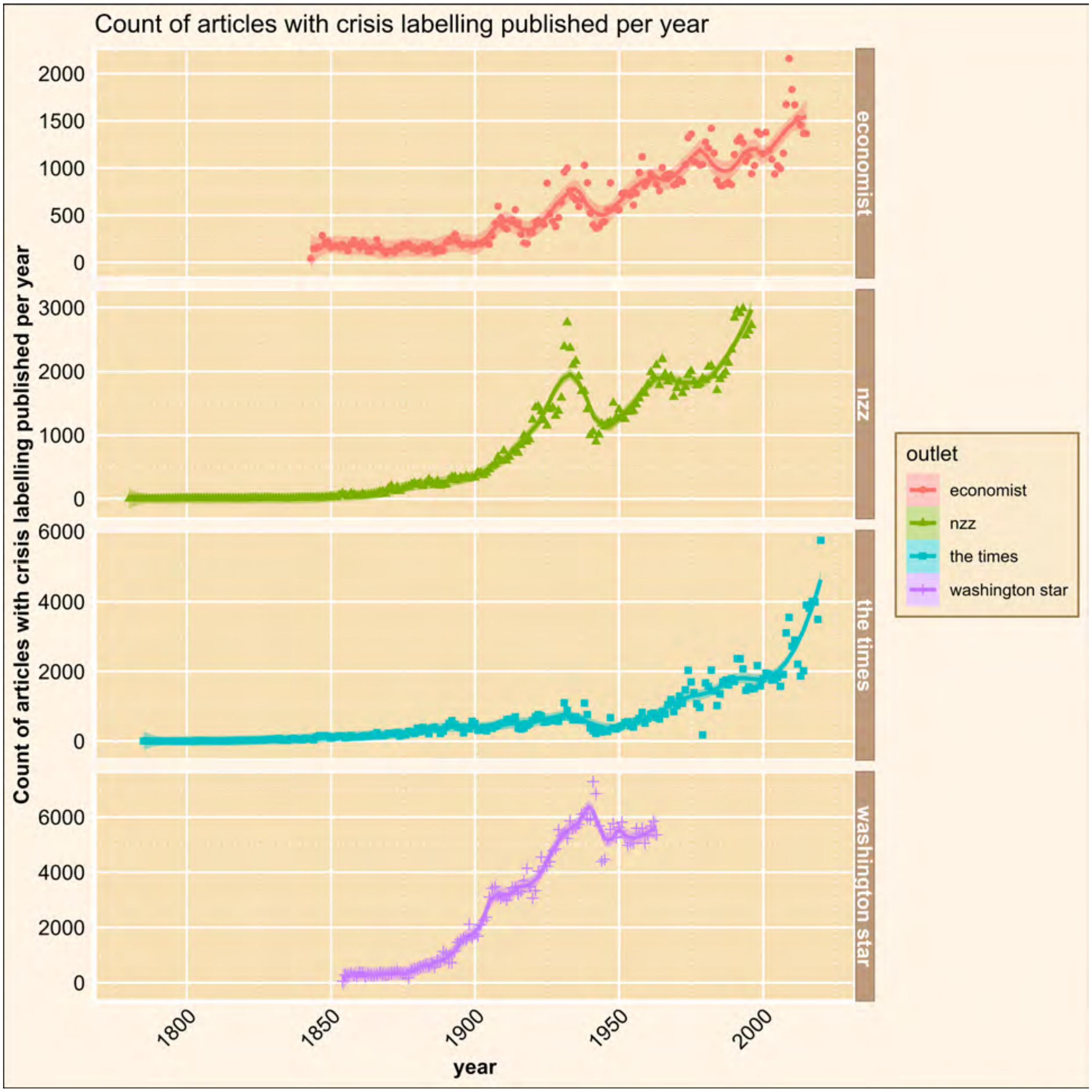


Figure 5.3: Share of articles with crisis labelling published per year in four newspapers

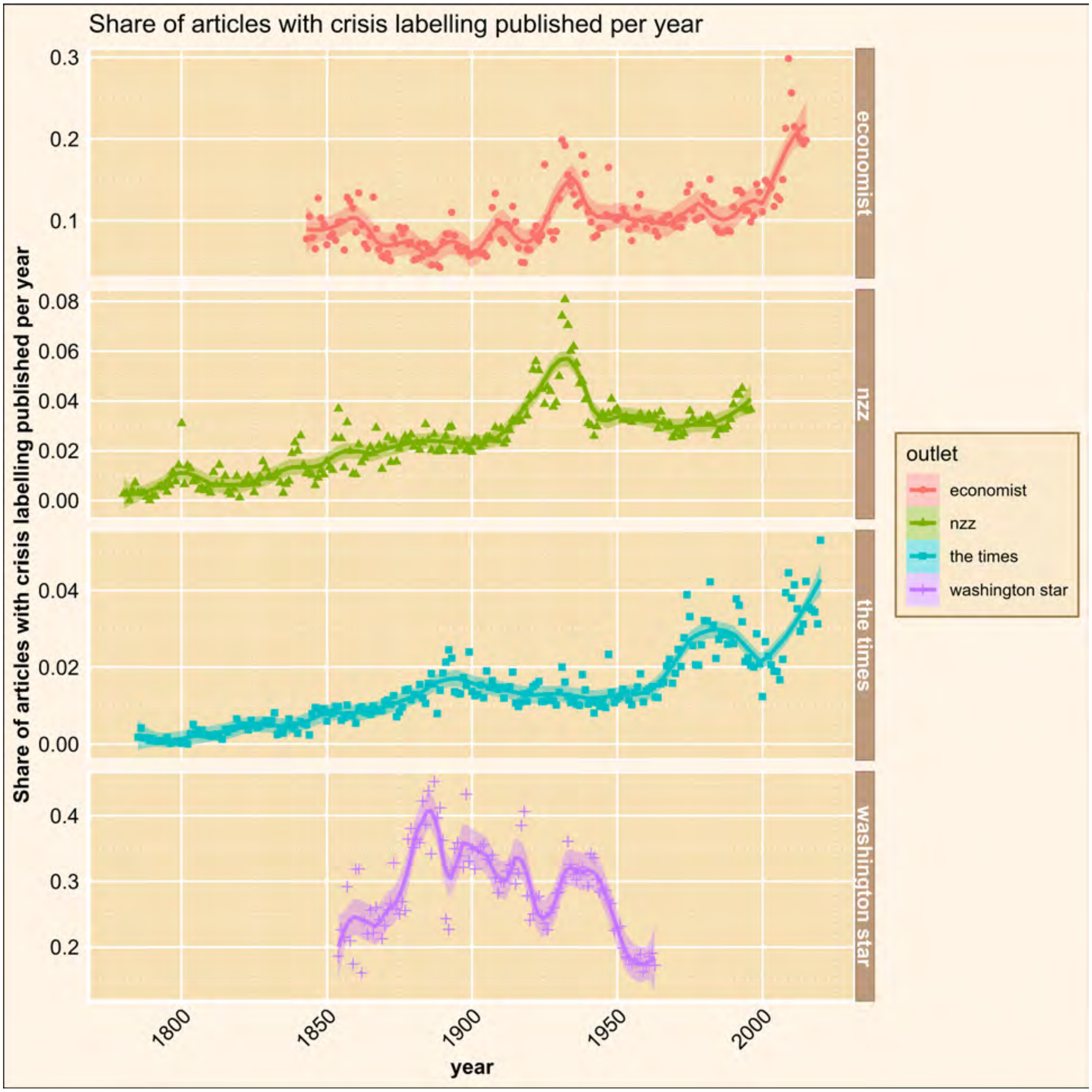


Figure 5.4: Count of articles that belong to crisis news waves per year in four newspapers

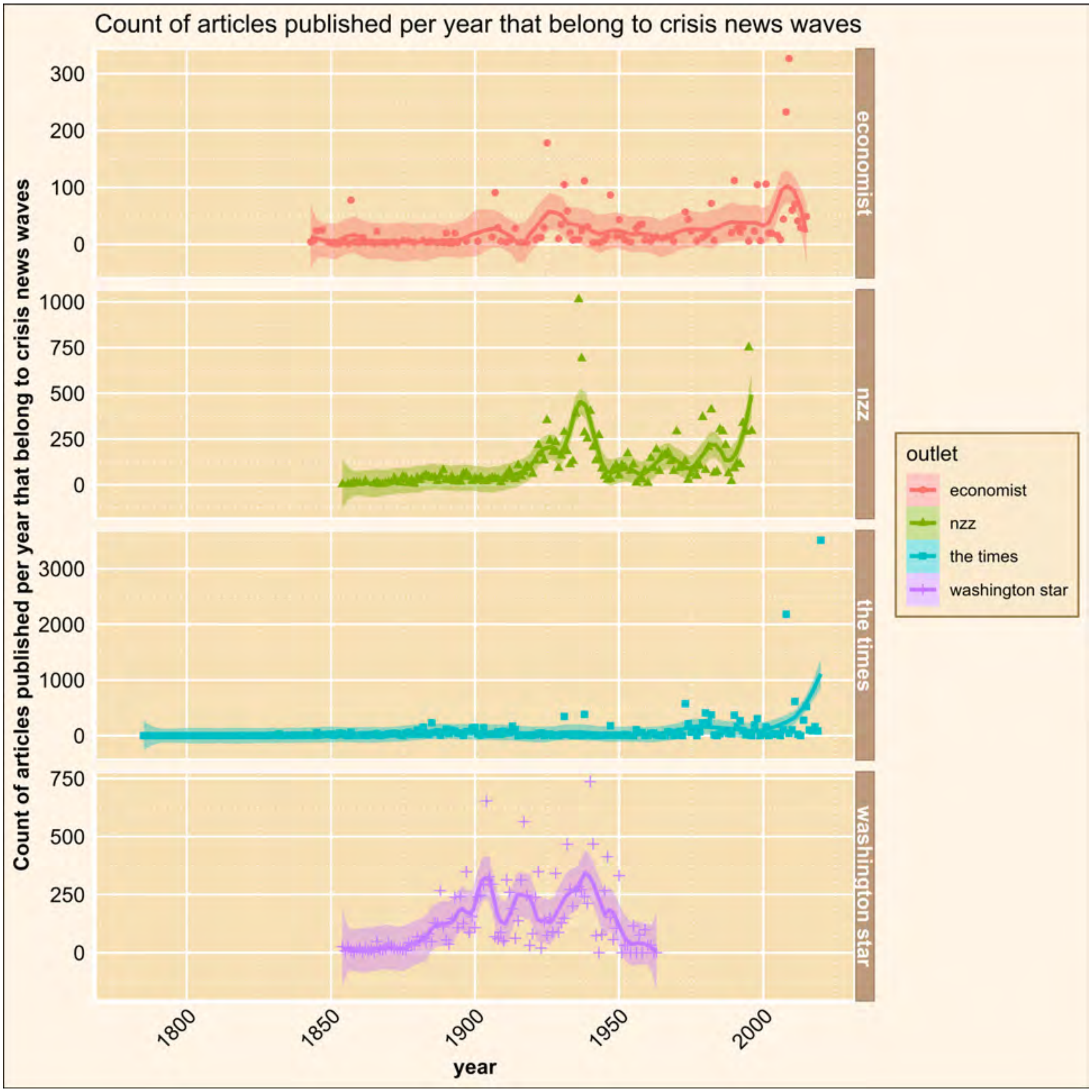


Figure 5.5: Share of articles that belong to crisis news waves per year in four newspapers

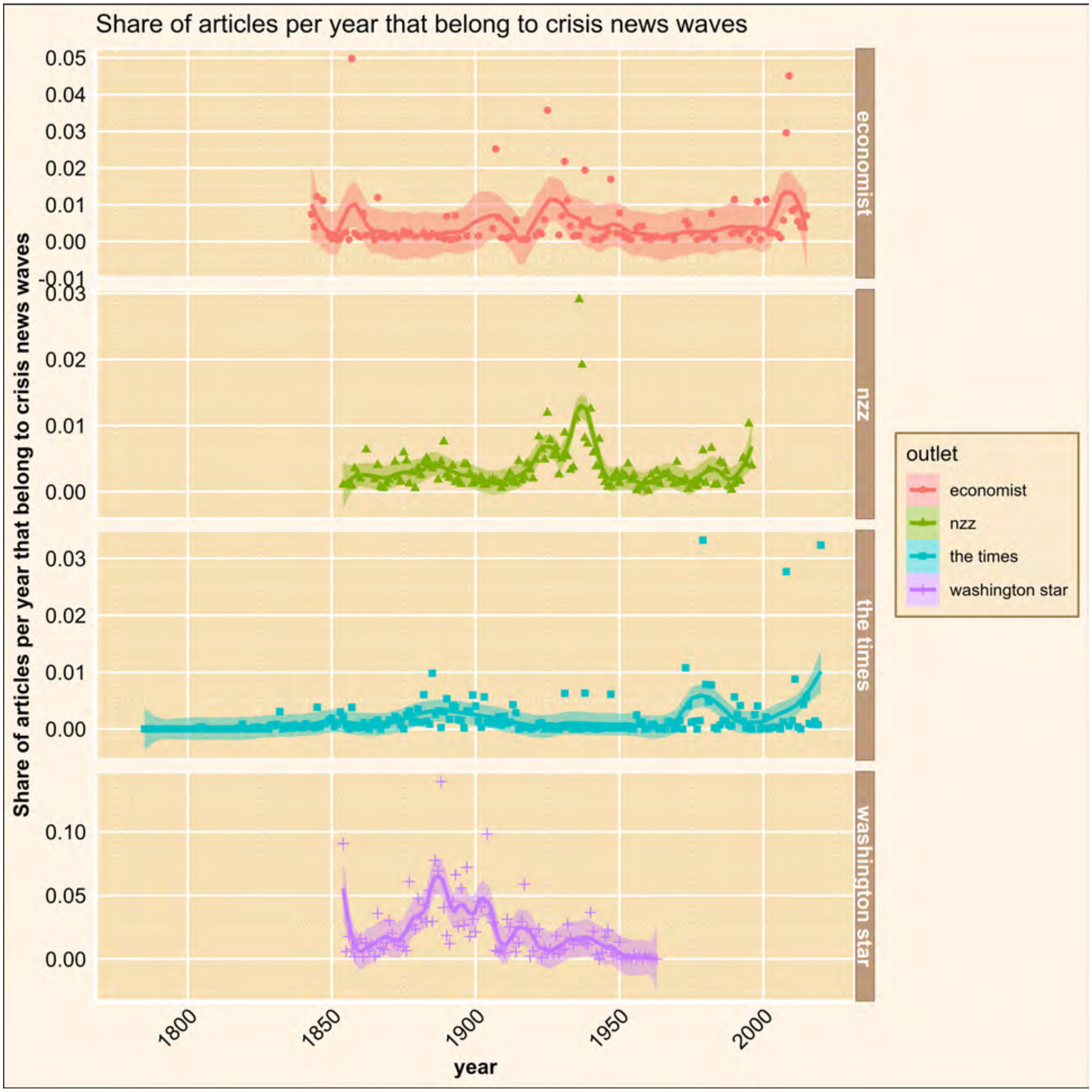


Figure 5.6: Count of crisis news waves per year in four newspapers

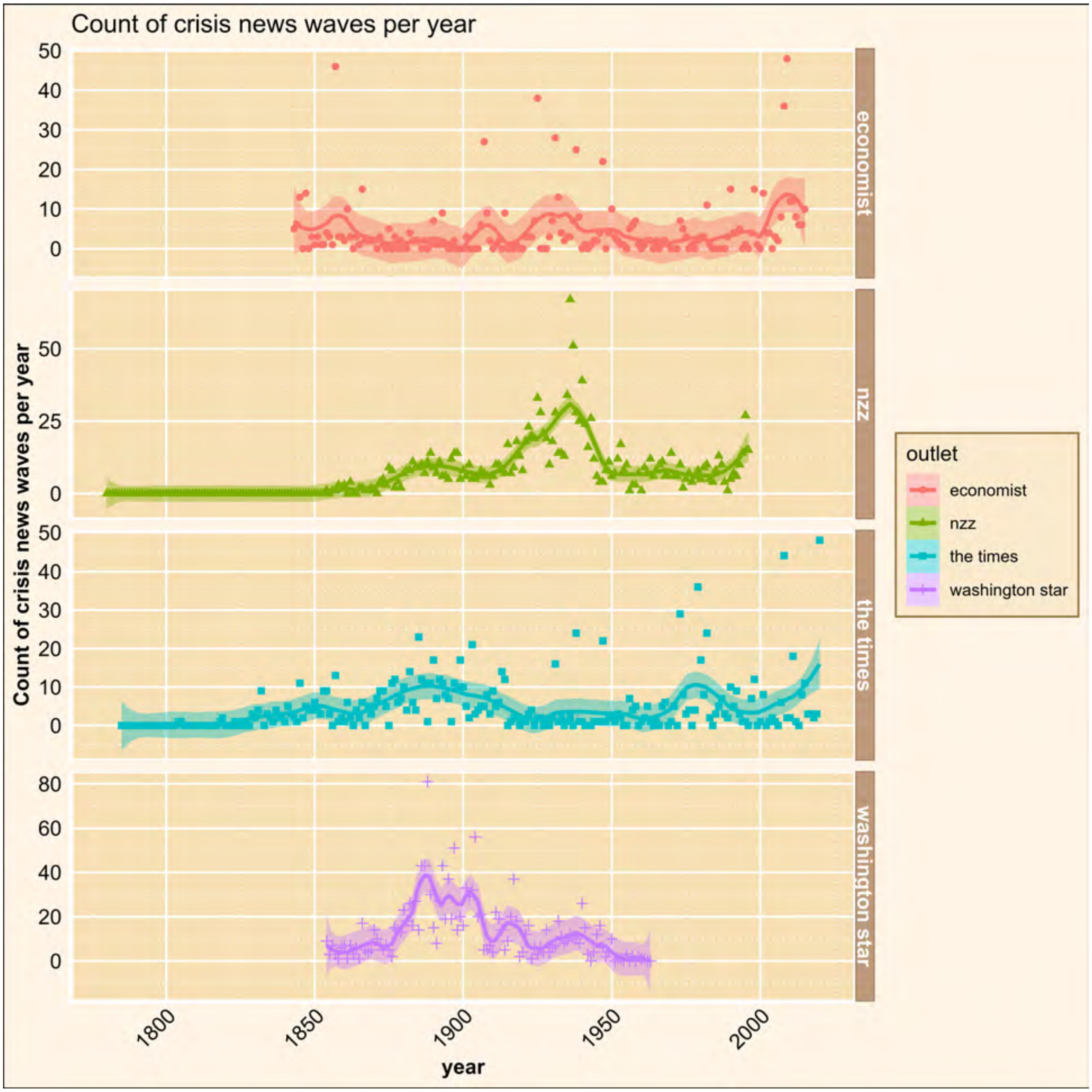


Table 5.1: Mixed linear models across four newspapers from three countries

Predictor	CL salience (G1.1) B(SE)	CNW salience (G2.1) B(SE)	CNW count (G3.1) B(SE)
(Intercept)	8.570 (7.418)	0.830 (1.647)	2.669 (10.662)
Time (years)	0.017*** (0.002)	-0.000 (0.000)	0.036*** (0.005)
Variance Components			
σ^2_{outlet}	157.72	0.807	15.90
$\sigma^2_{residual}$	10.66	1.206	75.14
Explained variance			
$R^2_{marginal}$	.006	.000	.048
$R^2_{conditional}$	.937	.401	.214

Chapter 6

Appendix F: Details on Search String Generation and Validation

6.1 Purpose

This appendix details the procedure for identifying the search string for identifying crisis coverage (or: building the crisis corpus), and tests of the classification performance relative to human coders' impressions.

6.2 Search string identification

This search string was built in several stages and using a variety of methods. We started with a single term, "crisis".

6.2.1 First iteration: Identify candidate words

Three procedures for finding candidate words were used:

First, we retrieved a selection of 50 hits (10 per period: 1785-1850, 1851-1900, 1901-1950, 1951-2000, 2001-2020). Headlines and the first 50 words of the texts were examined for synonyms or alternative terms that unambiguously signaled a "crisis" in line with out working definition.

Second, we used Latent Semantic Scaling (LSS) (Watanabe 2021) to create a collection of words that co-occurred with the keyword "crisis" (or in later stages, the extended keyword list). We scanned the top-1000 correlated words for each of the five time periods.

Thirdly, we automatically identified context words that belonged to the term "crisis" with confidence >0.95 (with `LSX::char_context` function; a window of 30 words) (Watanabe 2021) with a 10% sample of the entire material.

6.2.2 Check candidate words' hit rates

As candidate words, we mainly considered substantives that semantically signal a damage or threat, are unambiguous, and might be a substitute (not an attribute or qualifier) for the term "crisis". Crisis signifiers (which we look for) need to be distinguished from crisis qualifiers (which we do not look for), terms that describe a crisis once it has been signified.

For instance, articles that talk about a "recession" may not mention the word "crisis" in addition but still conform to our working definition, such that we added the word to the list. In contrast, the term "financial" often co-occurs with crisis signifiers but will does not independently and unambiguously signal a crisis, such that we did not add the word to the list. [1]

To that end, we inspected 20 example text per candidate word and inspected each mention of the candidate word ("hit") in its context (50 words before and after the word itself) (using `quanteda::kwic`).

Manual classification categories (the coder starts with the highest value and then checks if the lower-code descriptions can be applied; if yes, the next-lower description is tried; this means that in cases of doubt, the lower of the possible codes is chosen, because auf the "maybe" in the formulation):

4 – clearly a major societal crisis with substantial and broad threat, reverberations across societal domains

3 – maybe a sector-specific societal crisis with moderate reach and/or moderate threat, reverberations in few specific societal domains

2 – maybe a very limited crisis with limited reach and/or limited threat, reverberations in a closely delimited, specific organization, group or person.

1 – maybe only use crisis vocabulary without any threat, reverberations; person-specific issues without noticeable societal threat or reverberations.

0 – maybe not a crisis is meant.

Mentions of keywords that were coded 0 or 1 were possible *false hits*, which is used to calculate the false hit rate.

Terms whose hit rate was 80% or more were included in the search term for the next iterations.

6.2.3 Progression of iterations

We then ran the search with the expanded search term and repeated the procedure. We stopped after the third iteration as no new terms were discovered according to the procedure.

Round 1: «Crisis». Hits: 84537 (=articles). Estimated false hit rate: 0%

Round 2: «Crisis», «Catastrophe», «Disaster», «Collapse», «Breakdown», «Debacle». Hits: 147397 (=articles). Estimated false Hit Rate: 8%

Round 3: «Crisis», «Catastrophe», «Disaster», «Collapse», «Breakdown», «Debacle», «Recession», «Pandemic», «Epidemic». Hits: 187573 (=articles). Estimated false Hit Rate: 6%.

Round 4: No new terms discovered.

Keywords	n	Implied severity of crisis in text no crisis (0–1)	limited crisis (2)	sectorial crisis (2)	full-scale crisis (4)
Crisis %	34	0	12	35	53
Disaster %	24	0	8	42	50
Catastrophe %	19	10	0	38	52
Debacle %	20	5	9	72	14
Collapse %	18	18	23	45	14

[1] More examples: the word "aftermath" may often refer to a critical condition (that what happens before the "aftermath") but would usually not occur independently of more clear-cut crisis signifiers. One would not use the word "aftermath" to talk about the crisis itself but to qualify and extend the picture one draws of the crisis. Hence it would not be included even though the term is correlated with several other crisis terms. Also words that typically occur in crisis-related phrases ("facing" as in "Facing the intensifying crisis...") are not crisis keywords but crisis qualifiers.

Keywords	n	Implied severity of crisis in text no crisis (0–1)	limited crisis (2)	sectorial crisis (2)	full-scale crisis (4)
Breakdown %	20	17	8	67	8
Recession %	27	0	0	63	37
Epidemic %	29	6	4	55	35
Pandemic %	27	0	0	4	96

- Not included: Crisis qualifiers/crisis descriptors: gravest, deepening, aftermath, precipitated, threatens, engulfing, looming, facing, existential, full-blown, shocks, severe, instability, worsening, unprecedented, unpredictable, averted, failures, wake, devastating, strikes, unmitigated, panic, preparedness
- Not included: Borderline cases: Meltdown, Troubles, Tragedies, Tsunami, Earthquake, Calamity, Cyclone, Eruptions, Hurricane, Suffering, Flood, Draught, Famine, Devastation, etc.

6.3 Validation of search string performance against human coders

6.3.1 Method and Data

The keyword classification performance has been tested against two human coders who coded the intensity of crisis rhetoric (on a scale from 1–9) in 50 news stories (25 from our *crisis* corpus and 25 from a *stratified random sample* (stratified by year) of news stories from *The Times*). The coding was blind, i.e. the coders did not know if the article they coded was from the *crisis* or the *random* corpus. The agreement between the two coders was $\alpha_{\text{Krippendorff}} = 0.814[0.691; 0.936]$ and $\kappa_{\text{Brennan\&Prediger}} = 0.800[0.648; 0.952]$.

6.3.2 Results

The keywords performed extremely well in distinguishing news coverage with no or mild crisis rhetoric from news coverage with moderate and intense crisis rhetoric. Defining human coder performance as gold standard, the keyword search corresponds to it with F1-scores between .841 and .889, depending on where one defines the threshold value between crisis and noncrisis coverage on the 1–9 scale Table 6.2. The performance is constant throughout the period of study, with similar classification performance for articles from 1785–1849, 1850–1899, 1900-1949, 1950–1999, and 2000–2020 Table 6.3 (comment 16b).

Table 6.2: Manual crisis classification compared to crisis keyword classification

Cutoff on manual scale	Occurrances	Sensitivity/ Recall	Specificity	Precision	F1
> 1	68	.797	1.000	1.000	.887
> 2	61	.839	.917	.945	.889
> 3	56	.860	.854	.891	.875
> 4	51	.865	.783	.818	.841
> 5	37	.842	.617	.582	.688
> 6	34	.829	.587	.527	.644
> 7	9	.778	.461	.127	.219

Table 6.3: Crisis classification performance in five time periods

Period	Occurrances	Sensitivity/ Recall	Specificity	Precision	F1
1785–1849	10	.800	0.750	0.800	.800
1850–1899	11	.818	1.000	1.000	.900
1900–1949	16	.875	.857	.933	.903
1950–1999	10	.800	1.000	1.000	.889
2000–2020	15	.867	1.000	1.000	.929

Chapter 7

Appendix G: Distinctiveness of the crisis corpus

7.1 Purpose

This appendix is designed to show three things:

1. **Observation**: It shows that the crisis keywords used to construct a crisis corpus lead to massive differences between the so-constructed *crisis corpus* and a random selection of articles from *The Times*, the so-called *noncrisis corpus*. **Conclusion**: Our method of compiling a crisis corpus creates a corpus that is very distinct from a random selection of news coverage and includes a high density of crisis coverage of different intensities.
2. **Observation**: It shows that actors in the crisis corpus with high probability participate in crisis discourse (or have a crisis-related function in the author's crisis narrative). **Conclusion**: It is therefore permissible to conclude from a higher number and diversity of actors that are being mentioned that a more diverse set of societal interests participates in crisis discourse.
3. The two observations are stable over time.

7.2 Study 1: Crisis and noncrisis corpus contrast

7.2.1 Method and Data

We conducted a small manual coder agreement test [1] that contrasts a stratified random selection of 25 articles from our crisis corpus with 25 articles that are a stratified random sample from the full *Times* archive. The coders were blind as to whether an article came from the crisis corpus or the random corpus. They were then asked to classify the articles according to various criteria:

- Which crisis keywords did occur in the texts?
  - Coding agreement (categorical): $\alpha_{Krippendorff} = 0.925[0.852; 0.998]; \kappa_{Brennan\&Prediger} = 0.987[0.953; 1.000]$
- To what degree does the article's portrayal of the situation conform to full-blown alarming crisis coverage?
  - coding agreement (interval): $\alpha_{Krippendorff} = 0.814[0.691; 0.936]; \kappa_{Brennan\&Prediger} = 0.800[0.648; 0.952]$
- Which attributes are attributed to the situation that are elements of widespread crisis conceptualizations?

[1] between two coders, each coding 50 articles based on the first 1000 characters of the articles, including the headline

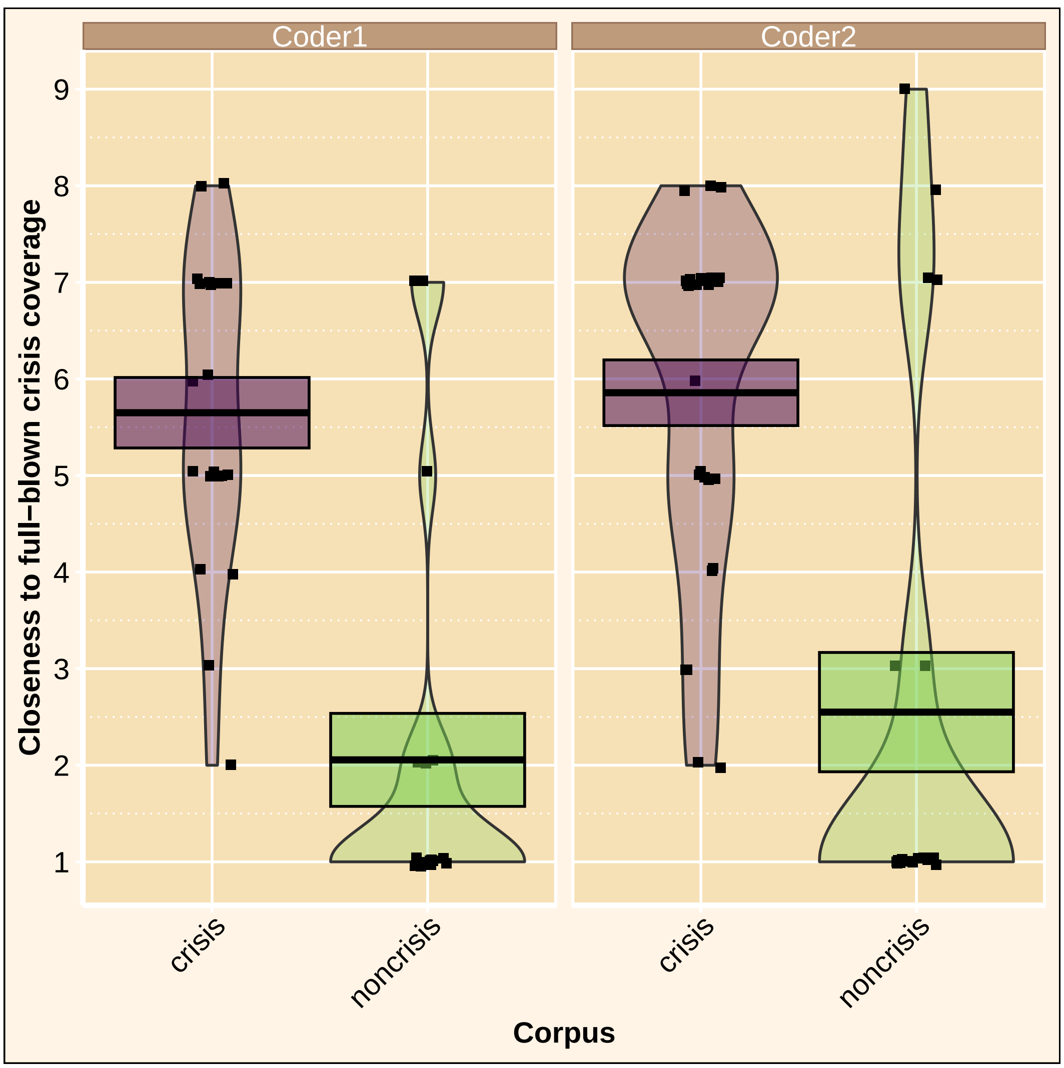
- coding agreement (categorical): $\alpha_{Krippendorff} = 0.610[0.440; 0.781]; \kappa_{Brennan\&Prediger} = 0.622[0.449; 0.795]$
- Which actors/speakers/protagonists appear in coverage, and do they contribute to crisis discourse?
  - [coding agreement (interval): $\alpha_{Krippendorff} = 0.592[0.411; 0.774]; \kappa_{Brennan\&Prediger} = 0.586[0.396; 0.776]$]

7.2.2 Results

Our results show that there is a stark contrast in how closely coverage resembles our conceptualization of crisis. While not all articles with crisis keywords fully conform to all crisis criteria, as a group they strongly differ from the randomly selected *Times* articles. There are clearly varying intensities of crisis coverage, but there is very little overlap in crisis coverage intensity between the crisis and the noncrisis corpora Figure 7.1.

This contrast shows clearly also in the main crisis criteria of *existential threat* and *non-normalcy* that is depicted or implied by the coverage Figure 7.5. Also, the contrast between the crisis corpus and the random corpus remains stable and consistent over time Figure 7.2.

Figure 7.1: Degree of crisis rhetoric in crisis vs non-crisis corpus articles



A similar picture emerges for the actors that are being mentioned, and whether they contribute to crisis

Figure 7.2: Degree of crisis rhetoric in crisis vs non-crisis corpus articles, by time period

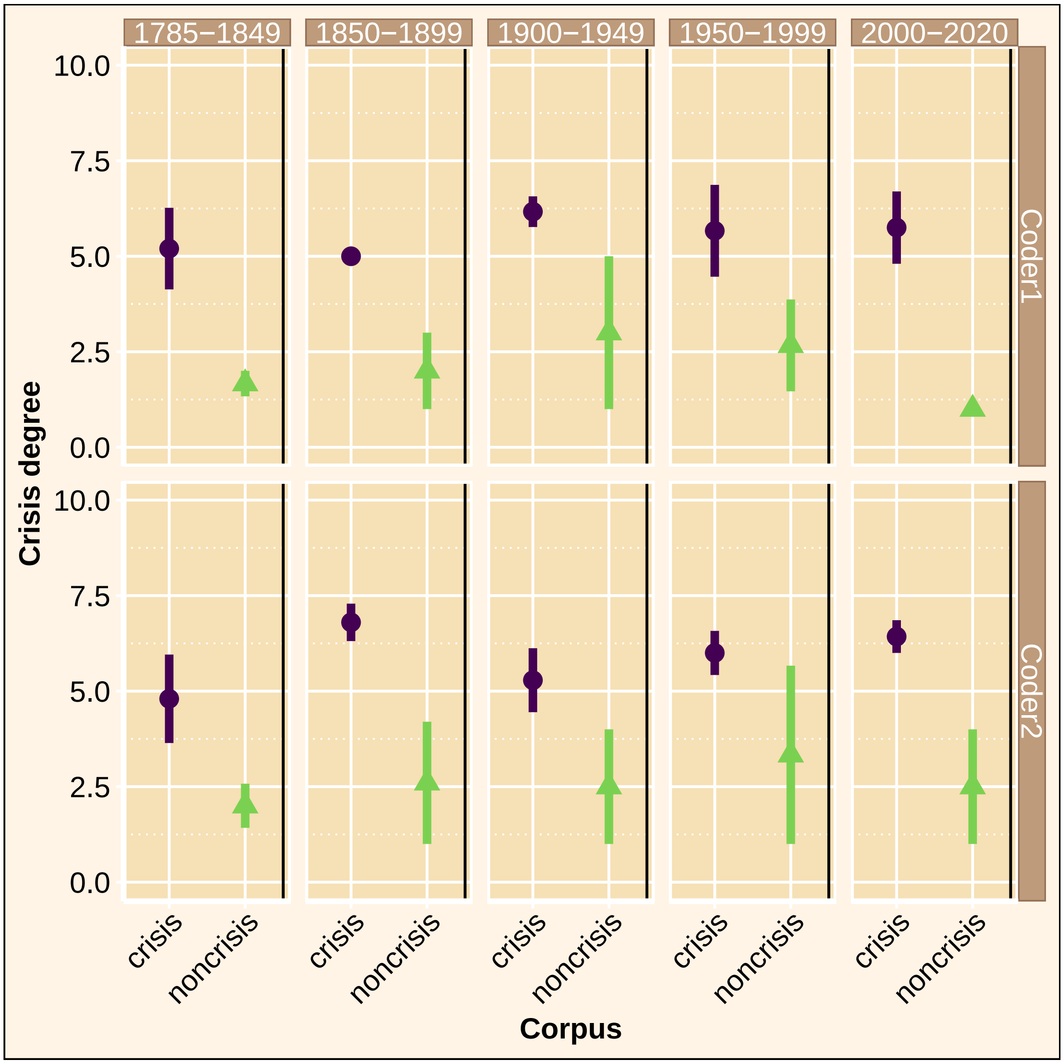


Figure 7.3: Degree to which actors contribute to crisis discourse in crisis corpus vs non-crisis corpus articles

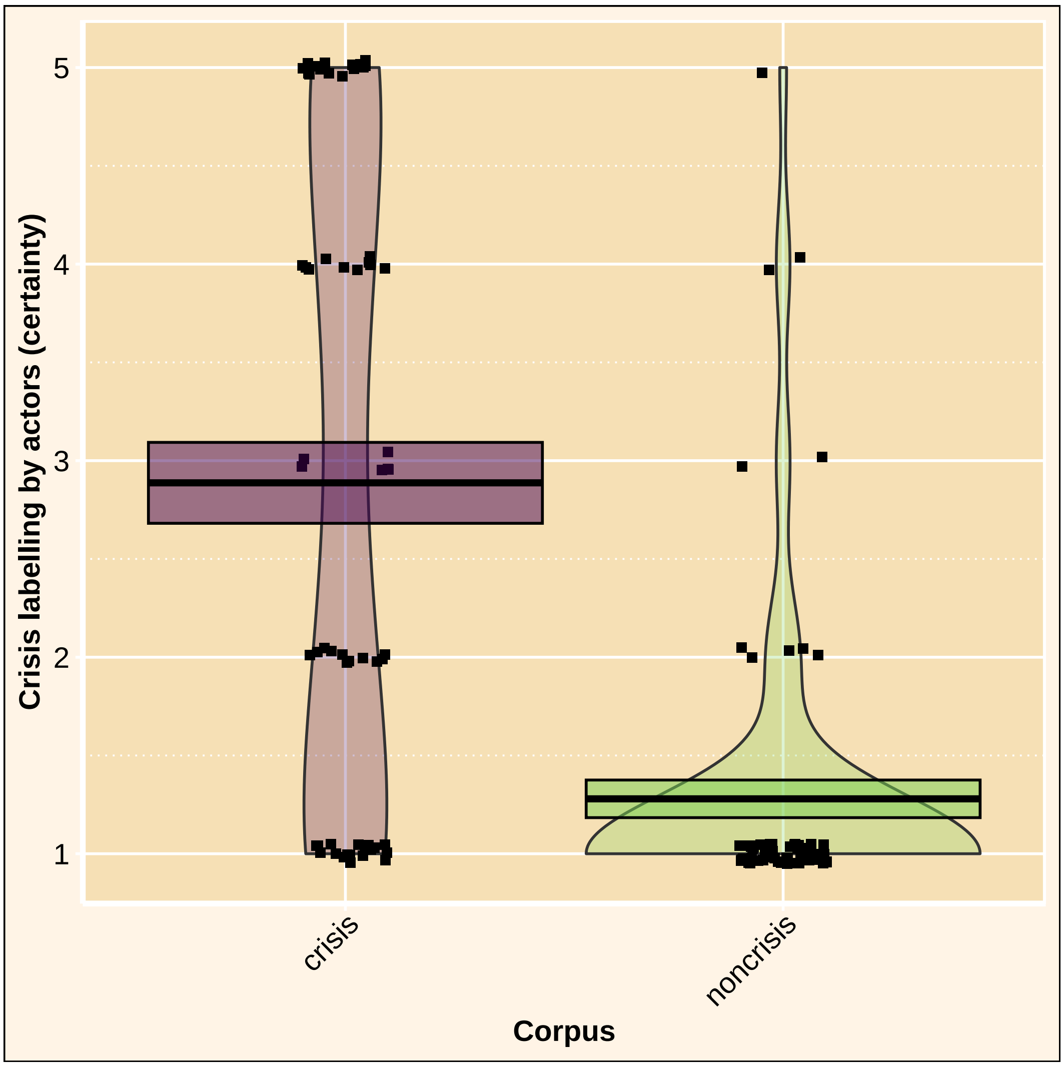


Figure 7.4: Degree to which actors contribute to crisis discourse in crisis corpus vs non-crisis corpus articles, by time period

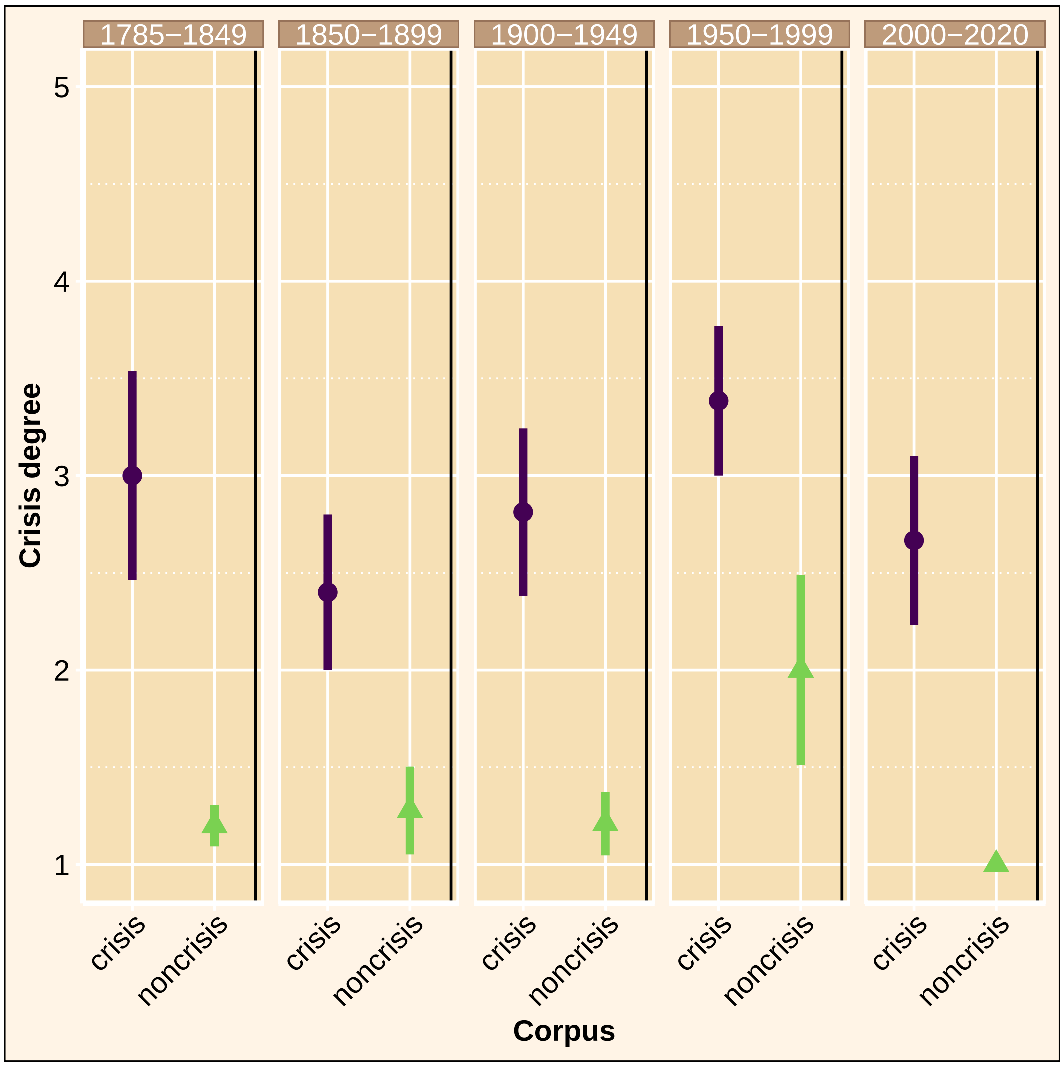


Figure 7.5: Presence of crisis criteria in crisis corpus vs non-crisis corpus articles

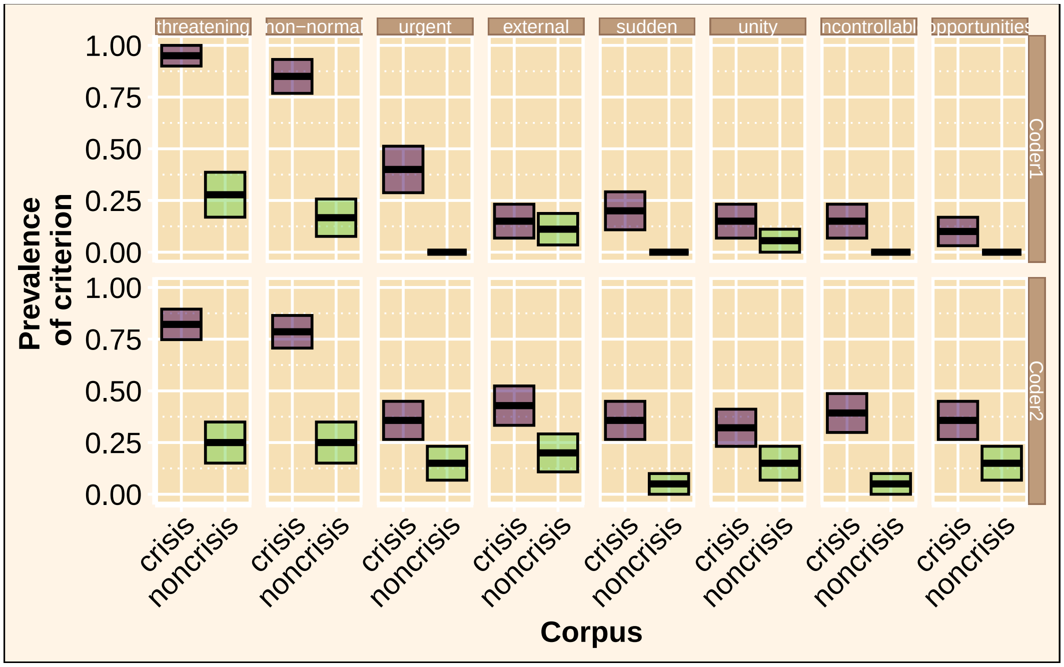
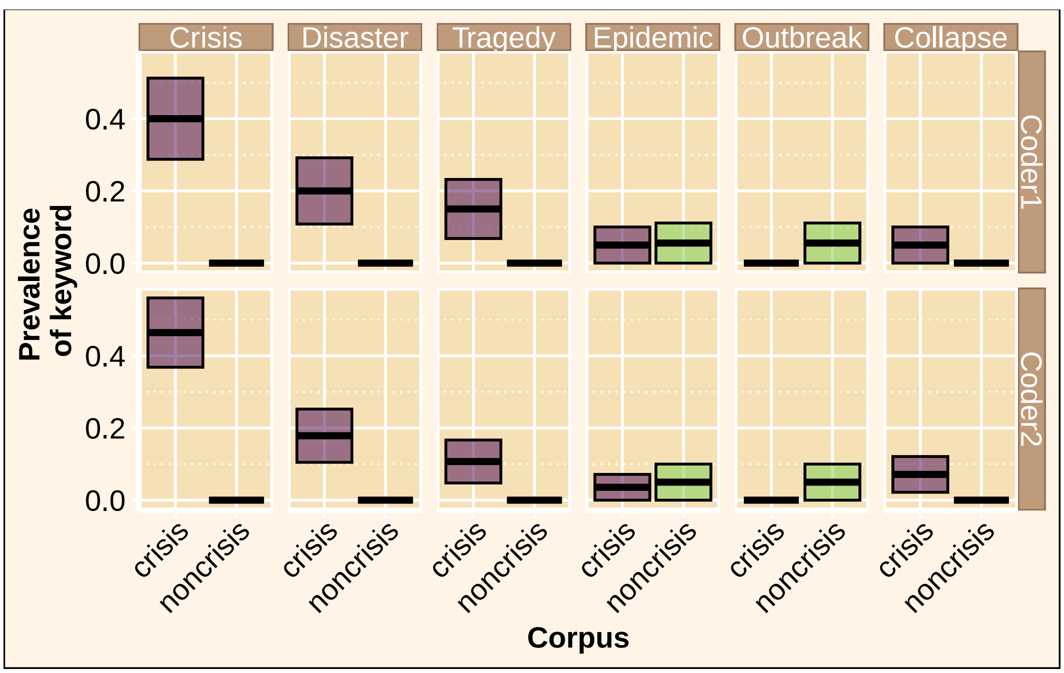


Figure 7.6: Presence of crisis keywords in crisis corpus vs non-crisis corpus articles



discourse: Not all actors that are being mentioned in an article in the crisis corpus definitively contribute to crisis discourse—there is a lot of uncertainty in many cases; 42% of the actors that appear are classified as contributing to crisis discourse with "high" or "very high" certainty. However, almost no actors appearing in the random corpus were classified as contributing to crisis discourse with "high" or "very high" certainty (3.5%) (Figure 7.3, Figure 7.4).

7.2.3 Discussion and Conclusion

The results demonstrate that the crisis keyword search procedure produces texts that are dense with features characteristic of crisis coverage, as intended.

Also, the results indicate that, aggregated to a whole year of crisis coverage, we can assume that a greater diversity of actors that appear in crisis coverage will translate into a greater diversity of actors that participate in crisis discourse (roughly 40% of actors that are being mentioned can be expected to participate in crisis discourse).

7.3 Study 2: Analysis of the crisis corpus

7.3.1 Methods and Data

We conducted a manual coding of 246 randomly selected articles from the crisis corpus as training and validation data for building a classifier.

The manifold categories that were coded mostly had good (above .80) or satisfactory (.67-.80) inter-coder agreement, with few categories falling below .67. We mark all categories where results may be less than satisfactory with double exclamation marks.

7.3.2 Results and Discussion

7.3.2.1 Can actors in crisis labelling coverage be presumed to engage in crisis labelling?

Most actors that are mentioned in crisis labelling coverage also participate in crisis discourse in some way. 68% of articles include some actors, and out of these 168 articles, 78% some actor engages in crisis discourse. Out of all 350 actors recorded, 210 (60%) engaged in crisis discourse. [2] We observed a substantial upward trend in crisis discourse participation of actors that were mentioned. 1785-1899, we 56% of actors recorded in the crisis corpus participated in crisis discourse; 1900-1949, the share was at 50%; then, it rose to 65% (1950-1999), and stood at 76% 2000-2020 (Figure 7.7). We know from the contrast analyis of the two corpora that the share in the random corpus is close to zero.

At the same time, more different actor categories appear in news coverage during the period of study, with 9% of actor categories appearing in the 1785-1899 sample; 11% of actor categories appeared in the 1900-1949 period; 23% of actor categories appeared in the 1950–1999 period; and 24% of actor categories appeared in the 2000-2020 period. And that is despite the fact that the number of actors that appeared has been n=61 in the 1785–1899 subcorpus, n=109 in the 1900–1949 subcorpus, n=97 in the 1950–1999 subcorpus, and only 49 in the 2000–2020 subcorpus. Actor diversity has increased even though the number of actors has decreased. This is also reflected in a decrease in the (1-Gini) coefficient of equal distribution across categories: 0.202 (1785–1899)— 0.214 (1900–1949)— 0.247 (1950–1999)— 0.238 (2000–2020). Interestingly, the diversity of actors increases a lot in crisis labelling and crisis diagnosis while it is reduced for crisis treatment in the last phase.

[2] This corresponds to the findings from the corpus contrast study (Section 7.2) where roughly half the actors mentioned in the crisis corpus engaged in crisis discourse; in the random corpus (Section 7.2), almost no actors engaged in crisis discourse.

7.3.2.2 Can crisis keywords create a crisis corpus rich in crisis coverage?

The crisis degree of coverage was 4.82 on average (95%CI=[4.58–5.07]) on a scale from 1 to 9. The share of articles that show no traces a crisis narrative in line with our definition is low, depending on how we define the threshold: 4.5% of articles received the lowest score (1 on a score from 1-9); 10% received the lowest or second-lowest rating; 25% received one of the lowest three scores. From study 1 (Section 7.2) we know that randomly selected articles from *The Times* rarely have high crisis degree scores.

The average crisis degree 1785–1899 was 5.36, 95%CI [4.80; 5.93]; for 1900–1949 it was 4.68, 95%CI [4.23; 5.13]; for 1950–1999 it was 4.52, 95%CI [4.11; 4.93]; for 2000-2020 it was 5.16, 95%CI [4.53; 5.79].

7.3.3 Discussion and Conclusion

The findings regarding the actor structure justify our decision to presume that a greater number and diversity actors in crisis coverage will automatically also indicate a greater number and diversity of actors that participate in crisis discourse. The chance of a recorded actor to participate in crisis discourse is consistently 50% or greater. Aggregating the entire coverage of an entire year and finding a greater diversity of actors can thereby serve as an indicator of greater diversity of actors that participate in crisis discourse. Any noise and imprecision will be dealt with by the aggregation procedure and the large number of articles per year.

The findings regarding the degree of crisis intensity allow the conclusion that our procedure of identifying crisis coverage using a set of crisis keywords was successful. It also makes clear, however, that not all crisis coverage as defined and operationalized by our keywords can be regarded full-blown crisis coverage in the style of the major collectively memorable crises like the COVID-19 crisis. This was one of the reasons for us to develop the distinction between *crisis labelling* and *crisis news waves*, where the articles that are part of a *crisis news wave* come closer to the full-blown crisis coverage typical for the major collectively memorable public crises.

Figure 7.7: Share of actors appearning in crisis coverage that participate in crisis discourse

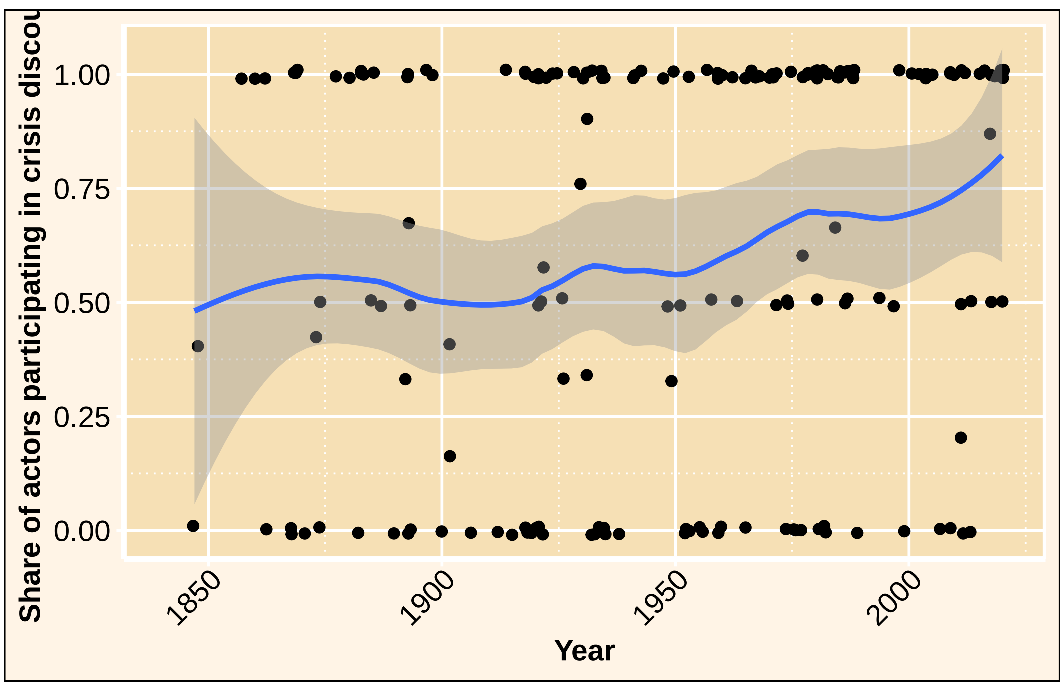


Table 7.1: Coding agreement in the manual coding

Construct	Agreement	$\kappa_{Brennan\&Prediger}$	Scale type
Crisis criteria	.54	.54 (!!)	Binary
Nature of the crisis	.98	.96	Binary
Current status	.94	.80	Ordinal (5)
Future status	.90	.63 (!!)	Ordinal (5)
Response by decision-makers	.94	.88	Binary
Response by third parties	.93	.86	Binary
Response by bystanders	.93	.87	Binary
Material symptoms	.95	.87	Ordinal (3)
Immaterial symptoms	.95	.88	Ordinal (3)
Systemic symptoms	.98	.94	Ordinal (3)
Penetration	.84	.28 (!!)	Ordinal(10)
Causation	.87	.74 (!)	Binary
Career Stage	.88	.76 (!)	Binary
Cond. development	.84	.69 (!)	Binary
Benchmarks	.77	.53 (!!)	Binary
Challenge: concern	.83	.67 (!)	Binary
Challenge: crisis	.89	.78 (!)	Binary
Challenge: labelling	.95	.90	Binary
Challenge: material	.89	.79 (!)	Binary
Challenge: systemic	.91	.82	Binary
Challenge: diagnosis	.76	.53 (!!)	Binary
Challenge: response by decision-makers	.74	.49 (!!)	Binary
Challenge: response by bystanders	.84	.69 (!)	Binary
Challenge: response by third parties	.72	.45 (!!)	Binary
Geography: Contitent Symptoms/ origins/ solutions	.96	.89	Binary
Geography: Country Symptoms/ origins/ solutions	.99	.97	Binary
Geography: Country appearance	.97	.94	Binary
Newswork	.97	.95	Binary
Crisis degree	.90	.72 (!)	Interval(10)
Actors			
Appearance	.97	.93	Binary
Speakers	.97	.95	Binary
Protagonists	.99	.97	Binary

Table 7.2: Actors mentioned and actors participating in crisis discourse

	All articles % (n=246)	Articles that mention actors % (n=168)
Mentions actors	68.3	100.0

	All articles % (n=246)	Articles that mention actors % (n=168)
Crisis treatment	44.7	65.5
Crisis diagnosis	33.3	48.8
Crisis labelling	19.1	28.0
Crisis discourse	53.3	78.0

Table 7.3: Diversity of actor categories in crisis crisis discourse

	Period			
	1785-1899	1900-49	1950-99	2000-20
Actor categories count				
Crisis labelling	7	10	18	19
Crisis diagnosis	17	15	22	25
Crisis treatment	17	27	29	20
Crisis discourse	26	31	35	30
1-Gini				
Crisis labelling	0.077	0.091	0.177	0.197
Crisis diagnosis	0.170	0.128	0.184	0.261
Crisis treatment	0.158	0.219	0.235	0.166
Crisis discourse	0.202	0.214	0.247	0.238

7.4 Study 3: Contrasting the vocabulary in the crisis and the noncrisis corpus

7.4.1 Methods and data

We conducted a keyness analysis for the vocabulary used in the crisis corpus (the data used in the analysis of our main study) and the noncrisis corpus (a stratified random sample of ca. 800 news articles from the *Times* archive per year). It contrasts the vocabulary of the two corpora and determines which words best discriminate between the two corpora. This method is fully data-driven, i.e. we do not predefine which words or concepts should be focused on or prioritized in the keyness analysis. We considered only terms that appeared at least 100 times across the two corpora.

7.4.2 Results

The top 100 words discriminating between the crisis and the noncrisis corpus were assigned to twelve categories (the same word could be assigned to several categories):

1. Political power ("minister", "state", "election", "authority")
2. Threat ("fall", "cholera", "mess", "attack", "serious", "bad", "tragedy")
3. Non-normalcy, uncertainty, and judgment ("situation", "likely", "believe", "may", "possible", "realise", "face", "question")
4. Temporal development, ("yesterday", "tomorrow", "rise", "moment", "today", "increase", "remain")
5. Finances and economy ("cut", "currency", "inflation","industry","investor")
6. People and unity ("community", "public", "people", "staff", "nation","union", "us", "unite", "majority")
7. Information ("fact", "realise", "report", "message", "warn")
8. Action/response ("plan", "effort", "measure", "action", "support")
9. Negotiations ("conference", "committee", "deal", "demand", "agreement")
10. Values ("peace", "oil", "profit", "community")
11. Systems ("pension", "system", "nhs")
12. Emotions ("fear", "hope", "cheer")

7.4.3 Discussion and Conclusion

These systematic differences in vocabulary between the two corpora make clear that the keyword search produced a very specific corpus that sharply contrasts with a random sample of news coverage in *The Times*. These differences, which emerged in a fully data-driven procedure, correspond excellently with the working definition of crisis laid out in our study.

The two main criteria of threat and non-normalcy occur as major categories. We view *temporal development* and *information* as related to the *non-normalcy* component (information and tracking development over time implies that there is uncertainty and the hope that things will normalize in the future). We view *political power*, *finances and economy*, *action/response*, *negotiations*, *values*, and *systems* components as linked to the *threat* component as it specifies what is feared and/or that unusually strong measures are considered (most likely, to combat the threat). Finally, the *positive* and *negative emotions* express the seriousness of the threat and the relief as normalization occurs.

Overall, we view this as a strong indication of the validity of our corpus building procedure.

Chapter 8

Appendix H: Details on Text Preprocessing

The preprocessing of the newspaper texts that were retrieved involved the following steps:

- *all versions* Read all text files and extract their ID numbers.
- *all versions* Read the metadata and match them with the text files using the ID number.
- *all versions* Remove doublettes where the headline *and* the first 100 characters of two articles overlap with 95% or more in their headline and their first 100 words (or more technically: which have a string distance of less than 5% of the maximum string distance).
- *intermediate step* Extract entities. Use named entity recognition (spacyr) on a 10000 text sample of the raw text data and extract all entities that appear twice or more often; these are added to the English dictionary to avoid deleting names of, e.g., organizations, persons, locations, states, works of arts, legal documents.
- *all versions* Remove special characters. Remove all characters that are not normal alphabetic letters, numbers, punctuation signs.
- *only versions with p* Remove punctuation.
- *only versions with s* Spellcheck: Substitute each word that is not part of the dictionary (with entities included) with the most similar word in the dictionary.
- *only versions without u* Lowercase: Make all characters lowercase.
- *only versions with l* Lemmatize: Substitute declinated or derived words with their word root.
- *all versions* Create representations as data frame (tx), corpus (cp), tokens (tk), quanteda document–feature matrix (qdfm), stm document–feature matrix (sdfm).
- *only versions with c* Remove stopwords for the English language
- *only versions with t1000* Remove words that occur in less that 1/1000 of texts.
- *only versions with t100* Remove words that occur in less that 1/100 of texts.

Table 8.1: Text representations with upper/lowercase, punctuation, spellcheck. Each is available as data.frame, corpus, tokens, quanteda dfm, stm dfm

Capitalization	Raw	Punctuation removed	Spellchecked	Punctuation removed, spellchecked
Original case	tms_X_u	tms_X_pu	tmx_X_su	tms_X_spu
Lowercase	tms_X	tms_X_p	tmx_X_s	tms_X_sp

Table 8.2: Text representations (all lowercase) with lemmatization, stopwords removal, and/or removal of infrequent words. Each is available as quanteda dfm, stm dfm

Version	Full	min $\frac{1}{1000}$ texts	min $\frac{1}{100}$ texts
No punctuation, ...			
...lemmatization	tms_X_pl	tms_X_pl_t1000	tms_X_pl_t100
...stopwords removed	tms_X_pc	tms_X_pc_t1000	tms_X_pc_t100
...lemmatized, stopwords removed	tms_X_plc	tms_X_plc_t1000	tms_X_plc_t100
...spellchecked, stopwords removed	tms_X_psc	tms_X_psc_t1000	tms_X_psc_t100
...spellchecked, lemmatized, stopwords removed	tms_X_plsc	tms_X_plsc_t1000	tms_X_plsc_t100

R packages used: `tidyverse` (data management), `stringr` (text changes), `hunspell` (spellchecking), `quanteda` (text processing), `stringdist` (assess how different words are from each other), `pbapply` (apply functions with progress bar), `spacyr` (named entity recognition, lemmatization), `stm` (structural topic modeling), `lemmar` (lemmatization).

Chapter 9

Appendix J: Checking for Doublettes

9.1 Purpose

The purpose of this appendix is to describe the procedure used for article doublette detection and elimination when building the corpus. Doublettes only occurred in the Factiva-based part of the text corpus, which are the articles published 2015–2020.

9.2 Procedure

- We checked for duplicates by calculating the string distance between all headlines relative to headline length (number of characters); per headline, we listed all headlines that matched the target article's headline perfectly (100%).
- We then compared the full texts of the target articles with the other articles that had a 95+%-similar headline. If the texts were more than 95% similar, we regarded it as a doublette.
- We then removed only the first doublette for each target article for which one or several doublettes were identified. This successfully removed all doublettes, as a repetition of the procedure after doublettes had been removed showed.
- Overall, this led to the removal of 1895 articles (out of the initial 28584 we had downloaded).

9.3 Code

```
library(stringdist)
library(quanteda)
library(stringr)
library(matrixStats)

# Load text data

load("c://users//stefange//onedrive - ntnu//2-data//crisis_collab//texts//tms_tx.RData")

# Filter pre-2015 data to speed up the procedure (no doublettes present in pre-2015 data)

hl <- subset(tms_tx,year>2014)$headline
tx <- (subset(tms_tx,year>2014)$text)

# Let progress bar only update the screen at every 1000 iterations
```

```
mseq <- seq(1,30000,1000)

# Compute length of headlines and texts

ddiag <- data.frame(hl=subset(tms_tx,year>2014)$headline)
ddiag$hl_length <- nchar(hl)
ddiag$tx_length <- nchar(tx)

# Create a matrix of all headlines

hld <- matrix(NA,ncol=length(ddiag$hl),nrow=length(ddiag$hl))
rownames(hld) <- ddiag$hl
colnames(hld) <- ddiag$hl

# Compute the string distance of all headlines, takes approximately 10min

for (i in 1:length(ddiag$hl)){
    hld[i,] <- stringdist(a=ddiag$hl,b=ddiag$hl[i])
    hld[i,i] <- NA
    if (i%in%mseq) {
        print(i)
        flush.console()
    }
}

# Find the minimal string distance for each headline, headline similarity, and check whether any headlines

# hllength <- t(matrix(rep(ddiag$hl_length,each=28584),ncol=28584,nrow=28584))

# hldmax <- rowMaxs(hld,na.rm=TRUE)
# hldmaxmat <- t(matrix(rep(hldmax,each=28584),ncol=28584,nrow=28584))

# hldrel <- ((hldmaxmat-hld)/hldmaxmat)


ddiag$min_strdist <- rowMins(hld,na.rm=TRUE)
ddiag$number_identical <- rowSums(hld==0,na.rm=TRUE)
ddiag$hl_similarity <- (((ddiag$hl_length-ddiag$min_strdist)/ddiag$hl_length))
ddiag$hl_similarity95 <- ddiag$hl_similarity>0.95

pos_dou <- matrix(NA,ncol=max(ddiag$number_identical),nrow=dim(ddiag)[1])
    rownames(pos_dou) <- ddiag$hl

sd_dou <- matrix(NA,ncol=max(ddiag$number_identical),nrow=dim(ddiag)[1])
    rownames(sd_dou) <- ddiag$hl

for (i in 1:dim(pos_dou)[1]){
    if (ddiag[i,"min_strdist"]==0){
        strdst <- stringdist(a=rownames(pos_dou)[i],b=ddiag$hl[-i])
        zeropos <- which(strdst==0)
        pos_dou[i,1:length(zeropos)] <- zeropos
    }
}
```

```
for (i in 1:dim(sd_dou)[1]){
    for (j in 1:dim(sd_dou)[2]){
        if (is.na(pos_dou[i,j])) {
            sd_dou[i,j] <- ddiag[i,"tx_length"]}
        if (!is.na(pos_dou[i,j])) {
            sd_dou[i,j] <- stringdist(tx[i],tx[pos_dou[i,j]])}
        }
    if (i%in%mseq) {
        print(i)
        flush.console()
    }
}

ddiag$min_textdist <- rowMins(sd_dou,na.rm=TRUE)
ddiag$min_textdist <- ifelse(ddiag$min_textdist>ddiag$tx_length | ddiag$min_textdist==9999,ddiag$tx_length
ddiag$hlsim <- (ddiag$hl_length-ddiag$min_strdist)/ddiag$hl_length
ddiag$txsim <- (ddiag$tx_length-ddiag$min_textdist)/ddiag$tx_length

ddiag$article_position <- 1:dim(ddiag)[1]
ddiag$doublette_position <- pos_dou[,1]

ddiag$kill <- ifelse(ddiag$article_position%in%unique(ddiag$doublette_position),1,0)


### Round 2

ddiag2 <- subset(ddiag,kill==0)

hld2 <- matrix(NA,ncol=length(ddiag2$hl),nrow=length(ddiag2$hl))
rownames(hld2) <- ddiag2$hl
colnames(hld2) <- ddiag2$hl

for (i in 1:length(ddiag2$hl)){
    hld2[i,] <- stringdist(a=ddiag2$hl,b=ddiag2$hl[i])
    hld2[i,i] <- NA
    if (i%in%mseq) {
        print(i)
        flush.console()
    }
}

ddiag2$min_strdist <- rowMins(hld2,na.rm=TRUE)
ddiag2$number_identical <- rowSums(hld2==0,na.rm=TRUE)

pos_dou2 <- matrix(NA,ncol=max(ddiag2$number_identical),nrow=dim(ddiag2)[1])
    rownames(pos_dou2) <- ddiag2$hl

sd_dou2 <- matrix(NA,ncol=max(ddiag2$number_identical),nrow=dim(ddiag2)[1])
    rownames(sd_dou2) <- ddiag2$hl

for (i in 1:dim(pos_dou2)[1]){
    if (ddiag2[i,"min_strdist"]==0){
        strdst <- stringdist(a=rownames(pos_dou2)[i],b=ddiag2$hl[-i])
        zeropos <- which(strdst==0)
```

```
        pos_dou2[i,1:length(zeropos)] <- zeropos
    }
}

for (i in 1:dim(sd_dou2)[1]){
    for (j in 1:dim(sd_dou2)[2]){
        if (is.na(pos_dou2[i,j])) {
            sd_dou2[i,j] <- ddiag2[i,"tx_length"]}
        if (!is.na(pos_dou2[i,j])) {
            sd_dou2[i,j] <- stringdist(tx[i],tx[pos_dou2[i,j]])}
        }
    if (i%in%mseq) {
        print(i)
        flush.console()
    }
}

ddiag2$min_textdist <- rowMins(sd_dou,na.rm=TRUE)
ddiag2$min_textdist <- ifelse(ddiag2$min_textdist>ddiag2$tx_length | ddiag2$min_textdist==9999,ddiag2$tx_le
ddiag2$hlsim <- (ddiag2$hl_length-ddiag2$min_strdist)/ddiag2$hl_length
ddiag2$txsim <- (ddiag2$tx_length-ddiag2$min_textdist)/ddiag2$tx_length

ddiag2$article_position <- 1:dim(ddiag2)[1]
ddiag2$doublette_position <- pos_dou[,1]

ddiag2$kill <- ifelse(ddiag2$article_position%in%unique(ddiag2$doublette_position),1,0)

table(ddiag$min_strdist==0,ddiag$min_textdist==0)

table(ddiag$hlsim==1,ddiag$txsim==1)

save(ddiag,file="doublette_diagnosis.RData")
```

Chapter 10

Appendix K: Structural Topic Modeling and Its Validation

10.1 Determining the number of topics (k) for the structural topic modeling

We defined a set of explorative structural topic models (Roberts, Stewart, and Tingley 2019) with "year" as covariate. [1]

Due to the long period and the wide topical spectrum we expect, we reckoned with a high number of topics. We started with 10 topics, stepwise increased them by 10, up all the way to 300 (10, 20, 30, ..., 300). Promising solutions between 200 and 250 were examined in more detail by including every other possible solution (200, 202, 204, ..., 250).

The number of topics (k) was determined based on two main criteria: push the dispersion of residuals towards 1 (which would indicate a correct specification of the topic model) while maximizing exclusivity (meaning that high-probability words in one topic are do not high-probability words in any other topic(s); a word is typical only for one topic, not several); heldout likelihood and semantic coherence were considered as side-criteria. These criteria were prioritized because the goal of the analysis was to describe the texts as good as possible with topics that are clearly distinct and interpretable as two different topical domains. Only 190 or more topics are considered as the dispersion of residuals is too high in models with $k < 190$. Exclusivity suggests extracting between 250 or more topics results in optimal exclusivity. Heldout-likelihood suggest that extracting 250 or more topics is optimal. Finally, looking at semantic coherence shows that going beyond 250 topics leads to substantial deterioration of semantic coherence without producing any benefits in other indicators. Hence, the 250-topic solution figured as optimal when viewing these four criteria simultaneously. Since the topics are further inspected and grouped manually (and some meaningless topics were expected), the exact number of topics was less consequential; care was taken to not choose a too low number of topics that conceal meaningful differences Figure 10.1.

10.2 Overview of the 250 topics solution

[1] We did not include time as a continuous predictor to speed up the fitting procedure and to ensure that topic estimates are only based on the words published in the articles.

Table 10.1: STM topic solution and our intermediate and top-level topic categorizations

ID	Detailed topic	Intermediate category	Top-level category
X1	depth/world/international/develop/countr	GEO/GEOPOLITICS	GEO
X2	inner/turmoil/circle/throw/urban/core/up	POL/REBELLION	POL
X3	pretax/profit/dividend/earnings/sale/tur	ECO/BUSINESS	ECO
X4	diary/newspaper/news/press/sunday/journa	PUB/NEWS	PUB
X5	2.30/parliament/debate/house/opposition/	POL/PARLIAMENT	POL
X6	evans/michael/richard/scott/jackson/davi	(invalid)	NA
X7	8.00/9.00/10.00/7.00/11.00/6.00/5.00	(invalid)	NA
X8	cholera/death/die/case/poison/dying/fata	EPI/CHOLERA	EPI
X9	dow/street/close/wall/jones/point/friday	ECO/STOCKEXCHAN	ECO
X10	signor/italian/italy/socialist/rome/chri	GEO/ITALY	GEO
X11	flu/disease/epidemic/virus/infection/pan	EPI/INFLUENZA	EPI
X12	opus/theatre/musical/festival/music/come	LEI/FESTIVAL	LEI
X13	bp/oil/mexico/shell/spill/petroleum/offs	ECO/RAW/FOSSIL	ECO
X14	experiment/case/upon/may/therefore/fact/	SCI/EVIDENCE	SCI
X15	cent/per/1.3/compare/average/fall/percen	ECO/BUSINESS	ECO
X16	ferry/disaster/victim/inquiry/sink/berea	DIS/SHIP	DIS
X17	map/problem/threaten/solution/solve/boun	PUB/PROBLEM	PUB
X18	pupil/school/child/parent/boy/girl/prima	EDU/SCHOOL	EDU
X19	palestinian/israel/israeli/jewish/jerusa	GEO/ISRAEL	GEO
X20	fi/wvas/wvith/wvhich/thle/anid/thie	(invalid)	NA
X21	pasha/cabinet/resignation/ministry/minis	POL/CABINET	POL
X22	dam/site/mile/lake/metre/square/pipe	DIS/CONSTRUCTIO	DIS
X23	0/e/j/h/t/w/s	(invalid)	NA
X24	influenza/year/month/last/week/figure/ag	EPI/INFLUENZA	EPI
X25	gp/nhs/health/patient/doctor/healthcare/	HEA/HEALTHCARE	HEA
X26	smart/space/launch/tip/scrap/satellite/s	SCI/SPACE	SCI
X27	collapse/cliff/fall/hole/crack/brink/sav	DIS/COLLAPSE	DIS
X28	barclay/financial/regulator/bailout/regu	ECO/FINANCE	ECO
X29	tv/review/catastrophe/comprehensive/sham	LEI/TVSHOWS	LEI
X30	movie/s/1992/1990/1987/1991/1989	(invalid)	NA
X31	boeing/airline/aircraft/airport/passenge	TRA/AIRTRAVEL	TRA
X32	barrel/price/oil/producer/commodity/prod	ECO/RAW/FOSSIL	ECO
X33	pension/scheme/saving/retirement/fund/pe	WEL/PENSIONS	WEL
X34	steamer/vessel/passenger/steamship/cargo	TRA/OCEANTRAVEL	TRA
X35	cyprus/agreement/talk/negotiation/meet/y	GEO/NEGOTIATION	GEO
X36	puzzle/5/7/6/8/9/3	(invalid)	NA
X37	korea/north/korean/northern/south/lee/pe	GEO/KOREA	GEO
X38	tin/trade/exchange/stock/trader/market/t	ECO/RAW/METAL	ECO
X39	howard/c/b/q/w/h/clarke	(invalid)	NA
X40	sheep/farmer/farm/land/agricultural/agri	ECO/RAW/AGRICUL	ECO
X41	bus/road/transport/traffic/cross/lane/hi	TRA/ROADTRAVEL	TRA
X42	theresa/brexit/eu/referendum/uk/britain’	GEO/BREXIT	GEO
X43	colliery/explosion/mine/pit/gas/miner/ma	DIS/MINE	DIS
X44	tt/thc/bv/thev/ihe/ot/arc	(invalid)	NA
X45	equivalent/increase/rise/rate/level/aver	ECO/BUSINESS	ECO
X46	fire/blaze/flame/burn/fireman/smoke/brig	DIS/FIRE	DIS
X47	russian/ukraine/russia/moscow/kremlin/we	GEO/RUSSIA	GEO
X48	wine/fine/bottle/quality/glass/taste/fru	ECO/RAW/ALCOHOL	ECO
X49	mental/social/care/community/welfare/eld	WEL/CARE	WEL
X50	hong/asian/kong/asia/latin/brazil/global	GEO/ASIA&LATIN	GEO
X51	graduate/skill/recruitment/management/wo	EDU/CAREER&HR	EDU

Table 10.1: STM topic solution and our intermediate and top-level topic categorizations

ID	Detailed topic	Intermediate category	Top-level category
X52	police/constable/officer/arrest/policema	DOM/POLICE	DOM
X53	steel/industry/production/industrial/pla	ECO/INDUSTRY/ST	ECO
X54	cathedral/st/castle/rev/abbey/dean/james	OTH/HERITAGE	OTH
X55	cancer/dr/brain/blood/illness/symptom/do	HEA/CANCER&ILL	HEA
X56	migrant/eu/refugee/immigration/asylum/im	DOM/IMMIGRATION	DOM
X57	itv/t/advertise/martin/y/ion/ft	(invalid)	NA
X58	academy/award/oxford/cambridge/prize/ent	EDU/UNIVERSITY	EDU
X59	zealand/new/australian/australia/sydney/	GEO/AUSTRALIA&N	GEO
X60	beef/ban/subsidy/measure/impose/restrict	ECO/REGULATION	ECO
X61	isnt/didnt/thats/thing/youre/doesnt/like	(invalid)	NA
X62	election/nomination/vote/candidate/poll/	POL/ELECTION	POL
X63	covid/pandemic/coronavirus/lockdown/viru	EPI/COVID	EPI
X64	nurse/hospital/patient/doctor/medical/be	HEA/HEALTHCARE	HEA
X65	pakistan/muslim/afghanistan/islamic/mili	GEO/AFGHANISTAN	GEO
X66	audit/report/inquiry/commission/committe	DOM/INQUIRY	DOM
X67	carriage/railway/engine/station/passenge	TRA/RAILTRAVEL	TRA
X68	tho/cruiser/bo/havo/aro/wero/thero	(invalid)	NA
X69	lifeboat/boat/captain/crow/passenger/dec	DIS/SHIPWRECK	DIS
X70	herr/german/germany/berlin/bonn/germanos	GEO/GERMANY	GEO
X71	borrower/loan/debt/mortgage/lender/credi	ECO/FINANCE	ECO
X72	66/july/june/april/october/september/30	(invalid)	NA
X73	prison/prisoner/sentence/crime/jail/crim	DOM/CRIME	DOM
X74	operative/committee/conference/resolutio	POL/COMMITTEE	POL
X75	diamond/bid/suit/beer/trick/z/heart	(invalid)	NA
X76	francisco/san/pacific/el/boston/tale/bay	GEO/USA	GEO
X77	lawrence/john/james/george/william/thoma	(invalid)	NA
X78	pope/catholic/roman/priest/protestant/ca	REL/CHURCH	REL
X79	tanker/port/ton/cargo/dock/tonnage/freig	TRA/CARGO	TRA
X80	submarine/naval/navy/admiral/fleet/admir	MIL/NAVY	MIL
X81	yen/dollar/currency/sterling/tokyo/deval	ECO/CURRENCY	ECO
X82	arabia/saudi/iran/gulf/iraq/oil/ruler	GEO/ARABIA	GEO
X83	lehman/2008/2009/ofthe/2007/2010/uk	ECO/FINANCE	ECO
X84	illustrate/england/english/article/engli	GEO/ENGLAND	GEO
X85	europe/exclusive/member/membership/join/	GEO/EU	GEO
X86	stadium/club/football/liverpool/fan/shef	LEI/FOOTBALL	LEI
X87	nigeria/federal/governor/canada/commonwe	GEO/CANADA&NIGE	GEO
X88	defendant/court/law/legal/justice/lawyer	DOM/JUSTICE	DOM
X89	fever/epidemic/disease/sanitary/medical/	EPI/FEVER	EPI
X90	cbi/inflation/recession/rate/economic/un	ECO/NATIONAL	ECO
X91	ambulance/service/emergency/office/staff	DIS/EMERGENCY	DIS
X92	hungarian/constitution/chamber/constitut	POL/CONSTITUTIO	POL
X93	museum/art/exhibition/gallery/artist/col	LEI/ART&EXHIBIT	LEI
X94	rubber/dividend/company/profit/sharehold	ECO/BUSINESS	ECO
X95	toy/christmas/december/january/november/	LEI/CHRISTMAS&T	LEI
X96	ulster/ireland/irish/unionist/dublin/bel	GEO/IRELAND	GEO
X97	hedge/investment/investor/fund/asset/equ	ECO/FINANCE	ECO
X98	pig/animal/cattle/meat/slaughter/cow/bre	ECO/RAW/AGRICUL	ECO
X99	bridge/engineer/repair/section/work/conc	INF/CONSTRUCTIO	INF
X100	sir/obedient/letter/faithfully/editor/se	(invalid)	NA
X101	peking/china/chinese/japanese/japan/chin	GEO/CHINA	GEO
X102	mr/trump/secretary/yesterday/tell/speak/	POL/STATEMENT	POL

Table 10.1: STM topic solution and our intermediate and top-level topic categorizations

ID	Detailed topic	Intermediate category	Top-level category
X103	striker/union/worker/strike/labour/emplo	LAB/STRIKE	LAB
X104	coroner/inquest/jury/witness/inquiry/dec	DOM/CRIME	DOM
X105	nuclear/power/station/electricity/plant/	NRG/NUCLEAR	NRG
X106	hostage/bomb/attack/terrorist/security/t	DOM/TERRORISM	DOM
X107	sugar/colony/empire/imperial/colonial/do	GEO/COLONIES&EM	GEO
X108	hurricane/island/wind/storm/weather/wave	DIS/STORM	DIS
X109	shakespeare/play/tragedy/stage/audience/	LEI/THEATER	LEI
X110	ad/r/news/weather/regional/murder/host	PUB/NEWS	PUB
X111	billion/goldman/analyst/global/sachs/us/	ECO/FINANCE	ECO
X112	delhi/india/indian/congress/government/t	GEO/INDIA	GEO
X113	church/bishop/archbishop/clergy/christia	REL/CHURCH	REL
X114	vaccine/research/test/scientist/science/	SCI/VACCINE	SCI
X115	franc/french/france/paris/frenchman/fran	GEO/FRANCE	GEO
X116	channel/4/44/3/4s/desperate/miserable	WEL/POVERTY	WEL
X117	market/pharmaceutical/sector/uk/growth/c	ECO/INDUSTRY/PH	ECO
X118	british/britain/1956/britainy/briton/ove	GEO/BRITAIN	GEO
X119	pet/dog/horse/hunt/wild/cat/ride	LEI/PETS	LEI
X120	bulgaria/constantinople/prince/sultan/po	GEO/TURKEY&BULG	GEO
X121	nato/military/defence/missile/weapon/arm	GEO/ALLIANCE	GEO
X122	plastic/use/machine/pack/fit/type/tool	ENV/PLASTIC	ENV
X123	kate/v/x/1/campbell/mm/ii	(invalid)	NA
X124	heathrow/holiday/travel/operator/tour/tr	TRA/AIRTRAVEL	TRA
X125	milk/food/egg/cook/bread/vegetable/meal	ECO/RAW/AGRICUL	ECO
X126	iraqi/iraq/un/humanitarian/council/secre	GEO/IRAQ&HUMANI	GEO
X127	bin/drive/driver/lorry/licence/wheel/van	TRA/ROADTRAVEL	TRA
X128	6d/mess/co/subscription/2s/5s/2d	(invalid)	NA
X129	eec/european/brussels/summit/community/c	GEO/EU	GEO
X130	francis/austria/rustrian/vienna/emperor/	GEO/AUSTRIA	GEO
X131	lisbon/correspondent/today/dr/nov/dec/oc	GEO/PORTUGAL	GEO
X132	princess/prince/queen/royal/duke/palace/	POL/ROYAL	POL
X133	lloyde/shareholder/share/merger/takeover	ECO/BUSINESS	ECO
X134	11.30/news/10.30/7.30/6.00/6.30/9.30	(invalid)	NA
X135	imf/eurozone/economic/economy/euro/monet	ECO/NATIONAL	ECO
X136	tunnel/project/contract/construction/inf	INF/CONSTRUCTIO	INF
X137	wheat/crop/cotton/grain/harvest/corn/mil	ECO/RAW/AGRICUL	ECO
X138	background/east/middle/eastern/west/west	GEO/REGIONS	GEO
X139	people/homeless/home/live/many/thousand/	WEL/HOUSING	WEL
X140	hollywood/film/love/music/pop/story/song	LEI/MOVIES&POP	LEI
X141	textile/export/import/trade/surplus/paym	ECO/INDUSTRY/TE	ECO
X142	plague/native/province/peasant/populatio	WEL/PEASANTRY	WEL
X143	parry/relationship/marriage/couple/partn	FAM/MARRIAGE	FAM
X144	barry/smith/wilson/jenkins/harold/ian/br	(invalid)	NA
X145	jew/democracy/political/politic/politici	POL/SYSTEM	POL
X146	library/society/association/institute/in	PUB/ASSOCIATION	PUB
X147	cuba/unite/state/u.s/kingdom/american/at	GEO/USA&UK&CUBA	GEO
X148	soviet/warsaw/communist/moscow/poland/un	GEO/SOVIET	GEO
X149	landlord/property/house/tenant/estate/re	WEL/HOUSING	WEL
X150	olympic/game/black/white/event/host/spon	LEI/OLYMPICS	LEI
X151	coffee/tea/bag/bean/estate/acre/tap	ECO/RAW/AGRICUL	ECO
X152	earthquake/town/village/destroy/damage/b	DIS/EARTHQUAKE	DIS
X153	sexual/woman/ms/sex/female/abuse/male	DOM/CRIME	DOM

Table 10.1: STM topic solution and our intermediate and top-level topic categorizations

ID	Detailed topic	Intermediate category	Top-level category
X154	argentine/treaty/argentina/government/co	GEO/ARGENTINA	GEO
X155	company/insolvency/business/firm/corpora	ECO/BUSINESS	ECO
X156	tue/sit/sun/u/wed/15/noon	(invalid)	NA
X157	syrian/refugee/rebel/troop/guerrilla/civ	GEO/SYRIA	GEO
X158	thames/wale/west/yorkshire/welsh/midland	GEO/BRITAIN	GEO
X159	forgive/much/seem/can/even/may/perhaps	(invalid)	NA
X160	gold/currency/silver/exchange/payment/mo	ECO/CURRENCY	ECO
X161	bless/us/life/god/human/love/shall	REL/RELIGION	REL
X162	partis/party/liberal/coalition/conservat	POL/PARTYPOLITI	POL
X163	petersburg/telegram/despatch/inst/telegr	GEO/RUSSIA	GEO
X164	job/redundant/worker/employee/unemployme	LAB/UNEMPLOYMEN	LAB
X165	insurer/insurance/compensation/loss/clai	ECO/FINANCE	ECO
X166	battalion/army/troop/regiment/soldier/mi	MIL/ARMY	MIL
X167	donor/aid/fund/relief/charity/donation/a	WEL/CHARITY&AID	WEL
X168	fish/sea/coast/beach/ocean/bay/shore	ECO/RAW/FISHERY	ECO
X169	literary/book/write/read/writer/story/au	LEI/LITERATURE	LEI
X170	singapore/executive/chairman/director/bo	ECO/BUSINESS	ECO
X171	cricket/team/sport/player/race/match/win	LEI/CRICKET	LEI
X172	student/university/professor/college/aca	EDU/UNIVERSITY	EDU
X173	alcohol/drink/eat/sleep/fat/diet/weight	HEA/HEALTHYLIFE	HEA
X174	telecom/million/group/uk/fullyear/operat	INF/TELECOM	INF
X175	thriller/star/comedy/drama/sky/film/murd	LEI/MOVIES	LEI
X176	quote/day/hour/week/next/saturday/fortni	(invalid)	NA
X177	king/monarch/majesty/crown/throne/monarc	POL/ROYAL	POL
X178	bonus/pay/money/payment/cash/cost/fee	ECO/BUSINESS	ECO
X179	gladstone/cheer/hon/laughter/m.p/hear/ge	POL/PM	POL
X180	ash/tree/wood/garden/bird/forest/cloud	ENV/WOOD&TREES	ENV
X181	coach/train/rail/driver/passenger/signal	TRA/RAILTRAVEL	TRA
X182	shopper/retailer/retail/customer/shop/co	ECO/SERVICE/RET	ECO
X183	boris/brown/cameron/chancellor/osborne/g	POL/PM	POL
X184	max/47/phillips/lion/tiie/elliott/hie	(invalid)	NA
X185	city/bradford/centre/birmingham/manchest	GEO/BRITAIN	GEO
X186	bank/deutsch/banker/liquidity/deposit/en	ECO/FINANCE	ECO
X187	council/kensington/london/county/borough	GEO/BRITAIN	GEO
X188	petrol/supply/fuel/shortage/crisis/deliv	ECO/RAW/FOSSIL	ECO
X189	shipbuilding/government/economic/policy/	ECO/REGULATION	ECO
X190	0.5/recession/gdp/growth/output/slowdown	ECO/NATIONAL	ECO
X191	obama/president/washington/congress/repu	POL/USPRESIDENT	POL
X192	cinema/picture/box/screen/film/photograp	LEI/MOVIES	LEI
X193	belgian/minister/government/prime/crisis	POL/PM	POL
X194	aberdeen/scotland/scottish/glasgow/edinb	GEO/SCOTLAND	GEO
X195	mexican/upon/shall/law/congress/power/le	POL/PARLIAMENT	POL
X196	african/south/africa/cape/southern/racia	GEO/SOUTHAFRICA	GEO
X197	copper/ton/market/price/stock/mine/metal	ECO/RAW/METAL	ECO
X198	pub/hotel/restaurant/park/tourist/visito	LEI/TOURISM&RES	LEI
X199	memorial/sympathy/message/funeral/tribut	PUB/EVENTS	PUB
X200	software/computer/technology/phone/inter	SCI/IT	SCI
X201	sword/man/upon/character/friend/statesma	PUB/OBITUARY	PUB
X202	ship/cruise/vessel/sail/liner/voyage/sea	TRA/OCEANTRAVEL	TRA
X203	angeles/american/york/america/california	GEO/USA	GEO
X204	foe/enemy/victory/defeat/attack/ally/upo	MIL/ENEMY&BATTL	MIL

Table 10.1: STM topic solution and our intermediate and top-level topic categorizations

ID	Detailed topic	Intermediate category	Top-level category
X205	flood/water/river/rain/mud/drown/pump	DIS/FLOOD	DIS
X206	knife/man/murder/shoot/young/suicide/you	DOM/CRIME	DOM
X207	misery/class/man/labour/work/prosperity/	WEL/POVERTY	WEL
X208	build/victorian/tower/building/design/ar	INF/CONSTRUCTIO	INF
X209	madrid/spain/spanish/portugal/portugues/	GEO/IBERIA	GEO
X210	ice/snow/mountain/expedition/cold/temper	ENV/CLIMATE	ENV
X211	strife/policy/political/country/situatio	POL/POLICYMAKIN	POL
X212	teacher/education/teach/educational/grad	EDU/SCHOOL	EDU
X213	jordan/arab/lebanon/syria/iraq/guerrilla	GEO/MIDDLEEAST	GEO
X214	athens/greek/greece/turkish/turkey/turk/	GEO/GREECE&TURK	GEO
X215	ford/car/motor/vehicle/engine/manufactur	ECO/INDUSTRY/CA	ECO
X216	coal/miner/wage/federation/employer/work	LAB/WAGES	LAB
X217	iranian/foreign/diplomatic/regime/sancti	GEO/IRAN	GEO
X218	stun/go/back/see/come/minute/away	(invalid)	NA
X219	canal/egypt/egyptian/suez/cairo/colonel/	GEO/EGYPT&SUEZ	GEO
X220	tobacco/tax/budget/deficit/expenditure/r	ECO/RAW/TOBACCO	ECO
X221	rescuer/kill/injure/dead/survivor/rescue	DIS/EMERGENCY	DIS
X222	universal/daily/register/compulsory/virt	PUB/NEWS	PUB
X223	pollution/local/authority/plan/area/envi	ENV/POLLUTION	ENV
X224	drug/medicine/dose/chemical/treatment/to	HEA/DRUG&MEDICI	HEA
X225	danish/dutch/sweden/belgium/denmark/swed	GEO/SCANDINAVIA	GEO
X226	flower/dress/wear/hair/blue/colour/cloth	LEI/FASHION	LEI
X227	thatcher/mp/tory/labour/blair/mrs/conser	POL/PM	POL
X228	famine/million/population/country/poor/p	WEL/FAMINE&POVE	WEL
X229	rio/de/la/le/des/brazilian/du	GEO/SOUTHAMERIC	GEO
X230	protester/president/leader/protest/oppos	POL/PROTEST	POL
X231	heap/oblock/catastrophe/halfpast/aud/mor	DIS/CATASTROPHE	DIS
X232	virgin/sky/bt/franchise/guy/bang/110	(invalid)	NA
X233	ftse/market/investor/stock/share/index/p	ECO/STOCKEXCHAN	ECO
X234	peer/lord/bill/amendment/legislation/cla	POL/PARLIAMENT	POL
X235	indonesia/region/area/regional/southeast	GEO/SOUTHEASTAS	GEO
X236	rating/credit/bond/yield/crunch/federal/	ECO/NATIONAL	ECO
X237	rig/safety/accident/incident/inspection/	DIS/OILRIG	DIS
X238	page/collins/col/continue/warn/article/1	(invalid)	NA
X239	swiss/league/geneva/switzerland/internat	GEO/SWITZERLAND	GEO
X240	baby/child/family/mother/age/father/pare	FAM/CHILDREN	FAM
X241	carbon/energy/gas/climate/environmental/	ENV/CLIMATE	ENV
X242	ann/mrs/wife/husband/miss/lady/friend	PUB/OBITUARY	PUB
X243	bbc/television/radio/broadcast/interview	PUB/BBC	PUB
X244	fraud/allegation/investigation/accuse/sc	POL/INVESTIGATI	POL
X245	auction/sell/sale/buy/buyer/price/purcha	ECO/BUSINESS	ECO
X246	pilot/air/fly/helicopter/flight/land/mac	TRA/AIRTRAVEL	TRA
X247	omit/question/reply/statement/matter/ask	POL/STATEMENT	POL
X248	vietnam/war/peace/nation/churchill/armam	GEO/VIETNAM	GEO
X249	revenge/paper/document/print/plot/white/	PUB/PUBLICATION	PUB
X250	gamble/much/take/one/make/now/time	OTH/RISK	OTH

10.3 Comparison of 120 and 250 topics solutions

Table 19.2: K=120 topic solution

Topic	word1	word2	word3	word4	word5	word6	word7
Topic 1	rea	ing	tion	ment	com	con	gov
Topic 2	diari	page	newspap	paper	letter	articl	sunday
Topic 3	ani	door	safeti	danger	must	risk	defect
Topic 4	map	cross	red	border	flag	frontier	tape
Topic 5	cent	per	gilt	rate	averag	higher	percentag
Topic 6	striker	court	legal	justic	case	judg	claim
Topic 7	flower	dont	say	didnt	get	peopl	think
Topic 8	market	withstand	price	mortgag	rate	buyer	buy
Topic 9	illustr	fact	differ	interest	natur	articl	method
Topic 10	pretax	profit	â£m	dividend	pre-tax	ã,â£m	turnov
Topic 11	cholera	death	epidem	sanitari	case	mortal	diseas
Topic 12	ã,â£	fund	compani	director	board	truste	chairman
Topic 13	opec	oil	energi	price	fuel	barrel	gas
Topic 14	nervous	electr	breakdown	power	telephon	station	cabl
Topic 15	cup	race	team	sport	player	win	won
Topic 16	movi	film	min	screen	cinema	camera	video
Topic 17	equit	report	commiss	inquiri	committe	investig	regul
Topic 18	flu	drug	virus	bird	test	human	scientist
Topic 19	tow	oclock	yard	morn	hour	minut	half-past
Topic 20	signor	parti	socialist	elect	democrat	coalit	vote
Topic 21	bank	depositor	loan	debt	credit	financi	banker
Topic 22	dow	share	market	index	stock	investor	ftse
Topic 23	dedic	surviv	georg	memori	lloyd	honour	award
Topic 24	pasha	cabinet	minist	resign	chamber	ministri	-day
Topic 25	teacher	council	local	author	counti	borough	grant
Topic 26	oecd	economi	inflat	econom	growth	gdp	rate
Topic 27	influenza	year	week	month	last	number	januari
Topic 28	tho	thb	upon	havo	aud	men	england
Topic 29	lord	ite	bill	hous	cheer	amend	parliament
Topic 30	percent	sep	dec	jan	feb	swiss	sweden
Topic 31	ferri	disast	said	survivor	rescu	victim	yesterday
Topic 32	golf	hotel	holiday	travel	tourist	garden	tour
Topic 33	â€¢	ofth	onlin	â–º	internet	websit	â€œthe
Topic 34	ounc	gold	currenc	bours	exchang	silver	discount
Topic 35	tuc	union	wage	worker	strike	labour	employ
Topic 36	cruiser	troop	telegram	regiment	despatch	battalion	colonel
Topic 37	palestinian	israel	isra	arab	egypt	lebanon	syria
Topic 38	million	tempus	pre-tax	full-year	group	profit	â,¬
Topic 39	wool	export	trade	price	import	textil	manufactur
Topic 40	olymp	today	weekend	tomorrow	talk	said	tonight
Topic 41	fri	sat	sun	oxford	wed	straight	mat
Topic 42	titan	messag	sympathi	disast	express	terribl	telegram
Topic 43	dam	water	flood	river	rain	lake	tree
Topic 44	theatr	king	queen	princ	royal	art	princess
Topic 45	ini	thc	wvas	tho	wvith	thle	thie
Topic 46	kong	china	chines	hong	asia	asian	singapor
Topic 47	nhs	servic	sector	govern	public	social	health
Topic 48	coach	train	railway	rail	passeng	carriag	driver
Topic 49	cyprus	greek	greec	turkish	turkey	athen	island

Topic	word1	word2	word3	word4	word5	word6	word7
Topic 50	davi	said	inquiri	repli	ask	wit	question
Topic 51	rover	compani	manufactur	busi	redund	supplier	job
Topic 52	shaft	coal	miner	collieri	pit	mine	men
Topic 53	â€”	bite	global	downturn	will	brown	decad
Topic 54	heap	bridg	wall	roof	collaps	floor	stone
Topic 55	background	east	middl	suez	canal	eastern	west
Topic 56	earthquak	villag	town	kill	homeless	injur	peopl
Topic 57	wine	retail	shop	store	food	christma	supermarket
Topic 58	pig	farmer	agricultur	farm	cattl	meat	milk
Topic 59	graduat	job	employ	recruit	busi	manag	work
Topic 60	fever	district	board	upon	tho	town	aud
Topic 61	ira	labour	tori	ireland	ulster	blair	mps
Topic 62	lifeboat	boat	vessel	steamer	crew	captain	ship
Topic 63	femal	women	young	men	age	woman	girl
Topic 64	pension	tax	budget	incom	deficit	pay	revenu
Topic 65	boe	airlin	aircraft	air	airport	flight	pilot
Topic 66	tin	stock	exchang	metal	copper	commod	contract
Topic 67	iceland	euro	eurozon	imf	â‚¬	debt	greec
Topic 68	itv	bbc	news	televis	radio	advertis	broadcast
Topic 69	divorc	law	famili	marriag	lawyer	abus	legal
Topic 70	constabl	polic	offic	arrest	riot	chief	crowd
Topic 71	sudan	rebel	guerrilla	refuge	muslim	violenc	troop
Topic 72	submarin	naval	navi	defenc	fleet	admiralti	ship
Topic 73	commend	american	unit	state	america	washington	world
Topic 74	tonnag	ton	ship	shipbuild	tanker	port	cargo
Topic 75	korea	south	africa	african	north	korean	coloni
Topic 76	cathedr	sir	lord	ladi	mayor	john	hall
Topic 77	hedg	invest	investor	fund	equiti	asset	bond
Topic 78	spill	nuclear	gas	plant	fish	sea	chemic
Topic 79	cancer	research	professor	studi	scienc	scientist	technolog
Topic 80	beef	european	eec	europ	brussel	britain	communiti
Topic 81	ukrain	soviet	russian	russia	moscow	poland	communist
Topic 82	francisco	insur	premium	underwrit	california	san	lloyd
Topic 83	herr	german	germani	berlin	bonn	von	hitler
Topic 84	malta	meet	committe	confer	deleg	propos	resolut
Topic 85	famin	aid	relief	food	chariti	donat	fund
Topic 86	thame	london	scotland	wale	scottish	midland	glasgow
Topic 87	blaze	fire	explos	flame	burn	smoke	firemen
Topic 88	pupil	school	children	educ	student	parent	colleg
Topic 89	steel	industri	product	output	plant	invest	manufactur
Topic 90	stadium	club	footbal	liverpool	manchest	fan	sheffield
Topic 91	librari	build	hous	properti	project	rent	site
Topic 92	nil	dividend	compani	debentur	railway	sharehold	stock
Topic 93	pope	church	cathol	bishop	archbishop	roman	religi
Topic 94	comedi	star	film	drama	sky	music	sport
Topic 95	co-op	societi	member	organis	associ	group	institut
Topic 96	lehman	billion	sharehold	goldman	â‚¬	barclay	merger
Topic 97	derek	said	chairman	director	board	sir	yesterday
Topic 98	coffe	crop	wheat	sugar	grain	harvest	price
Topic 99	rhodesia	govern	polici	british	minist	britain	agreement
Topic 100	jail	prison	alleg	sentenc	trial	crimin	prosecut
Topic 101	raymond	white	black	game	colour	card	racial
Topic 102	hungarian	constitut	parti	liber	govern	polici	dissolut

Topic	word1	word2	word3	word4	word5	word6	word7
Topic 103	cotton	econom	countri	trade	tariff	expenditur	reduct
Topic 104	nurs	hospit	patient	doctor	medic	health	bed
Topic 105	inquest	mrs	coron	wife	daughter	juri	death
Topic 106	vaccin	diseas	epidem	infect	virus	outbreak	health
Topic 107	pakistan	india	indian	delhi	provinc	bombay	calcutta
Topic 108	jobless	recess	growth	unemploy	sector	output	quarter
Topic 109	motorist	car	motor	vehicl	road	driver	transport
Topic 110	diamond	sale	bid	sell	sold	buy	takeov
Topic 111	yen	japan	dollar	japanes	currenc	sterl	devalu
Topic 112	zealand	new	york	australia	canada	feder	canadian
Topic 113	bueno	presid	senat	congress	spain	spanish	argentina
Topic 114	what	can	seem	even	will	much	thing
Topic 115	iranian	iraq	iran	obama	reagan	nato	bush
Topic 116	balkan	bulgarian	russia	russian	constantinopl	princ	austrian
Topic 117	rob	polit	democraci	peopl	social	revolut	power
Topic 118	gaull	franc	french	pari	itali	italian	leagu
Topic 119	peke	armi	troop	militari	war	enemi	soldier
Topic 120	cuba	cuban	kennedi	missil	crisi	presid	weapon

10.4 Validation of STM solution against manual coding

For validating the STM solution against manual coding, we conducted a manual content analysis of a stratified random sample (5 time periods defined the strata: 1785–1849, 1850–1899, 1900–1949, 1950–1999, 2000–2020) of 242 articles from the crisis corpus.

The 250 topics solution and the 98 topic categories of the manual coding procedure were aggregated to 20 broader topic areas of which three topics appeared sufficiently frequently in the 242 articles: *Economic topics* (91 articles, 38%), *Disaster and accident topics* (65 articles, 27%), and *Health topics* (including epidemics) (59 articles, 24%); all other topic areas together account for only 27 articles (11%), which makes any agreement analysis unfeasible for these topics.

The data structure of manual coding is quite different from that of structural topic modeling. In structural topic model, each article has a score on all 250 topics that express the composition of topics in the article (or their probability) such that they account for the vocabulary used as best as possible. If we sum the scores of all topics that we manually classified as *economic topics*, it that sum expresses the share of economics topics in the article. In manual coding, coders had to pick one primary topic that relates to the central crisis depicted in the article, leading to a binary topic-to-article assignment.

To make both data structures compatible, we defined a threshold at 10% summed topic probability to treat a topic as "present" in the article, according to the STM. This way, we obtain a binary score in both the STM topic classification (which, however, is not exclusive; theoretically, up to 9 topics could have a topic probability >10%) and in the manual topic classification (where coders were restricted to choosing a single primary topic).

The results show substantial to good classification performance of the simplified STM, if we treat human coding results as the gold standard. Economic topics appeared 91 times, according to human coders, and 115 times, according to the STM result (and 77 agreements). Disaster topics appeared 65 times according to human coders, and 57 times according to the STM result (and 46 agreements). Epidemics topics appeared 59 times according to human coders, and 68 times according to the STM result (and 51 agreements).

Figure 10.1: Diagnostics for determining the number of topics (k) to extract for a structural topic model

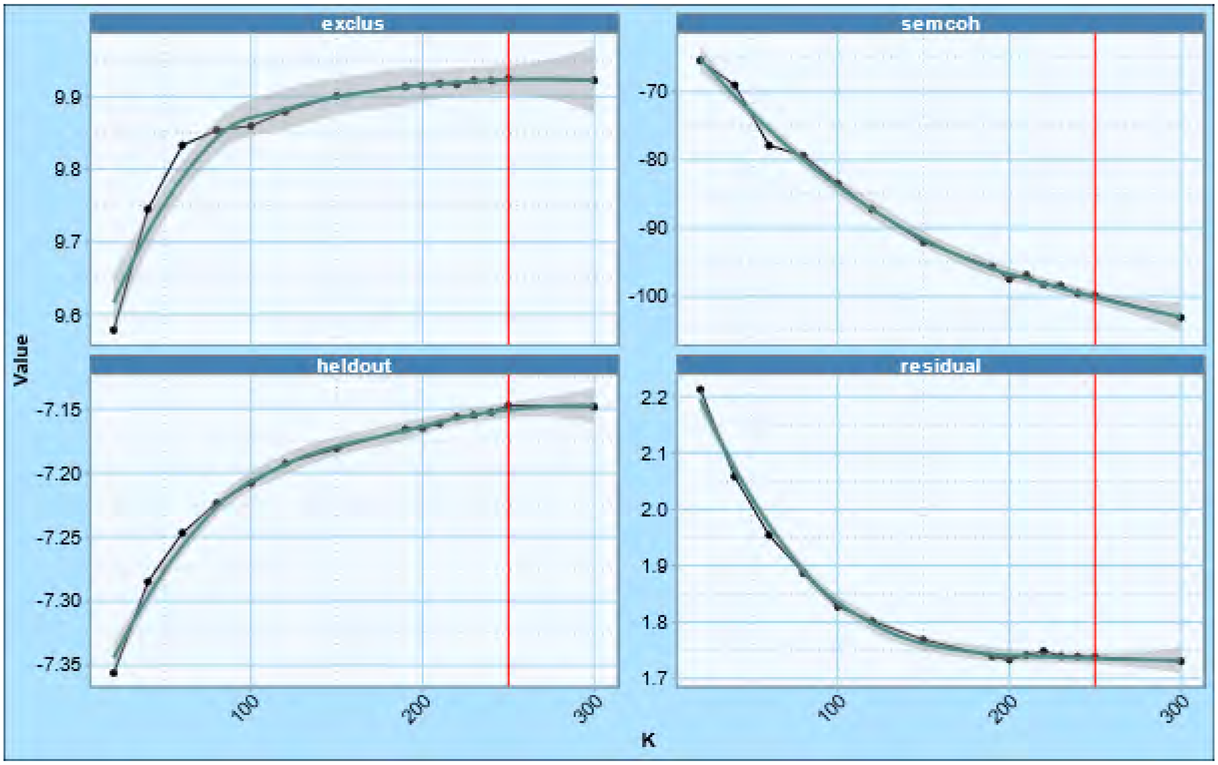


Table 10.3: Raw results of STM and manual human coding

Topic	Human coder result	STM result Present	Not Present
Economy	Present	77 (TP)	14 (FN)
	Not Present	48 (FP)	87 (TN)
Disaster	Present	46 (TP)	19 (FN)
	Not Present	11 (FP)	150 (TN)
Epidemics	Present	51 (TP)	8 (FN)
	Not Present	17 (FP)	150 (TN)

Table 10.4: Manual classification of topics compared to topic modeling solution (aggregated to main topic categories)

Topic	Occurrances	Sensitivity / Recall	Specificity	Precision	F1	Correlation	$\alpha_{\text{Krippendorff}}$
Economy	91	.846	.644	.616	.713	.380	.451
Disaster	65	.708	.932	.807	.754	.559	.664
Epidemics	59	.864	.898	.750	.803	.735	.727
Others	27	—	—	—	—	—	—

Chapter 11

Appendix L: Details on automatic crisis news wave (CNW) detection and its validation

We developed and applied a procedure for automatically and inductively discovering CNWs and automatically label them with keywords.

11.1 Preprocessing

After estimating the structure of the 250 topics (which words go with which topic) and the share of each topic in the texts (which topics are found in which texts to what extent), we compiled day-by-day news coverage time series of topics (not news stories). For example, if exactly two articles a1, a2 were published on day d about topic t, and the share of the topic in the articles was 75% or 0.75 for a1 and 40% or 0.4 for a2, it would mean that topic t on day d gets a salience score of 0.75+0.4=1.15. This can be interpreted as the equivalent of 1.15 full-length articles devoted to the topic.

Across all 250 topics, the sum of salience scores equals the total number of articles analyzed on that day (i.e., all articles with CL published in the «Times» on that day). Dividing the salience score by the total number of articles on that day results in a score of the topic's share of the newspaper's CL on that day.

11.2 Finding potential CNWs

To find phases of extraordinary amounts of crisis coverage (i.e., CNWs), we computed different baseline values, which were moving averages of news coverage about a topic: a 90-day («quarterly»), a 180-day («semi-annually»), a 365-day («annual»), a 730-day («bi-annual»), a 1461-day («quinquennial»), and a 3652-day («decennial») moving average.

The observed coverage was also smoothed by computing a 30-day («monthly») moving average. Coverage of a topic at any point in time (in the 30-day moving averages) was compared with all the six baselines (quarterly, semi-annual, annual, biannual, quinquennial, decennial). If current coverage (30-day MA) was above one of these moving averages the detector would record a potential crisis news wave (pCNW). It would record when the 30-day MA moved above one of the baselines for the first time (start of the pCNW) and when it fell below all six of the baselines again for the first time (end of the pCNW). The «volume» of the pCNW (how many articles above the baseline were observed) , the «duration» of the pCNW (temporal distance between end and start), the average «intensity» (volume divided by duration), the «peak intensity» (the maximum distance between the 30-day MA and the baselines, averaged across the five baselines) and the «variability» of coverage during the pCNW was recorded for each pCNW. For computing volume, intensity, maximum intensity, we added the above-baseline and the below-baseline coverage of the issue. This makes the simplifying assumption that if a crisis strikes in one issue, all coverage about that issue during the crisis news wave is considered crisis-related. For instance, in a major financial crisis, all financial issues coverage with crisis labelling is considered as belonging to that specific financial crisis even though some articles may also deal with other financial crises.

11.3 Filtering CNWs

But this still leaves us with many minor instances of above-baseline coverage. The total number of pCNWs totaled 102,504. We defined a public perception threshold (Neuman, 1990) at which a change in volume of coverage about a topic is regarded salient enough that a substantial share of the public is likely to register and experience it as an unusual intensity. This threshold was defined relative to the volume of news coverage in a newspaper: if only a few articles are published per day, the threshold for registering a change may be very low; if newspapers include hundreds of articles per day, a stronger increase will be necessary to draw attention. We defined the threshold at 1.825% of total news coverage on a day, or 1 article for each 20'000 articles published per year; if this number was below 0.5 articles, the threshold was fixed at 0.5 articles. This means that for 1785 (8.67 articles per day), the threshold would be at 0.158 above-baseline articles but was set to 0.5 articles due to the absolute minimum; it rose to 5.427 in 2020. These thresholds reduced the number of pCNWs from 102,504 to 1'022, which we treated as CNWs.

11.4 Automatic labelling of CNWs

After finding the start and end dates (plus the topics) of the CNWs, the vocabulary of topic-related articles (the share of the topic needed to be 0.5% or more) published during the CNW with the vocabulary in articles about the topic published in 1000 days before the CNW (pre-during comparison) and in articles about the topic published in 1000 days after the CNW (during-post comparison).

11.5 Manual validation of CNWs

Based on the extracted keywords, one most-representative text and the headlines of the ten most-representative texts, a human coder labelled a random selection of 300 CNWs with a unique crisis event label, such as *Oil crisis 1973/74* or *Titanic Disaster 1912*.

274 out of 300 crisis news waves (91%) could be clearly assigned to a single key event or crisis, in 17 crisis news waves (6%) there were several key events that the surge in coverage jointly or partially traced back to, and 9 crisis news waves (3%) could not be connected to any recognizable key event. This attests to a very low false positive rate. Depending on the strictness of assessment, it could be estimated between 3% and 9%. The agreement between two coders who coded 50 randomly selected crisis news waves for the intensity of threat and non-normalcy (based on the headlines and significant terms) was $PA = .92$, $\kappa_{\text{Brennan \& Prediger}} = .68$ and αKrippendorff $= .53$. The relatively low αKrippendorff reflects the high prevalence of the highest crisis intensity category, which is often scored as chance agreement according to the αKrippendorff calculates chance agreement. Coders also rated their confidence that the label they picked accurately captures the underlying critical condition or event. This rating of confidence produced agreement of $PA = .96$, $\kappa_{\text{Brennan \& Prediger}} = .83$ and αKrippendorff $= .64$. Again, the relatively low αKrippendorff reflects the high prevalence of high certainty ratings. The mean confidence rating (coder 1) was 3.38 (on a 0–4 scale) and the mean crisis intensity (coder 1) was 3.12 (on a 0–4 scale).

Table 11.1: Coders' crisis intensity scores

Crisis intensity scores					
	Coder 2				
Coder 1	0	1	2	3	4
0	1	1	1	1	0
1	0	1	0	1	2
2	0	0	1	1	3
3	0	0	0	3	3
4	0	0	2	4	25

Table 11.2: Coders' condition–wave link confidence scores

Critical condition–wave link confidence scores					
	Coder 2				
Coder 1	0	1	2	3	4
0	1	1	1	1	0
1	0	1	0	1	2
2	0	0	1	1	3
3	0	0	0	3	3
4	0	0	2	4	25

50 randomly chosen crisis news waves were labelled by another coder to establish intercoder agreement in the assignment of crisis news wave labels. Comparing the crisis news wave labels assigned by the two coders, 48 out of 50 labels referred to the same underlying event or condition (or agreed that the crisis news wave did not treat a recognizable event or condition), amounting to coder agreement of 96% (since chance agreement is virtually impossible in this task, we do not correct for chance agreement).

The question how many significant public crises we overlook was approached by selecting historical events that most likely led to a crisis news wave, and check if this event is also discovered by our procedure. We defined four different tasks and report the hit rate for this task (a) general historical events across domains, (b) epidemics, (c) mining accidents, (d) train accidents. The three domain-specific tasks were formulated to check for performance on less obvious cases and to allow drawing from more or less complete lists of events of a similar kind.

The results show that the procedure reliably identifies CNWs around major historical critical events, and CNWs around major epidemics, mining accidents, and train accidents when train accidents were a novel phenomenon. The low hit rate on smaller mining accidents and later train accidents most likely reflect that these events did not produce news waves large enough to be detected by the procedure, in line with the interpretation that these news waves reflect a public feeling of a deep crisis in most cases. We conclude that the procedure has an acceptable beta error. Overall, we can expect a low number of false positives and false negatives.

Table 11.3: Testing the validity of crisis news wave detection

Task	Targets	Hits	Hit rate	Explanation
Major historical events across topics	20	19	95%	Procedure robustly identified crisis news waves of clearly historical dimensions. The failure to find the 1978/79 crisis results from the fact that the Times was not printed due to the strikes.
All epidemics in the UK with an estimated number of deaths >5000	13	10	77%	Consistent identification performance
Mining accidents with more than 100 deaths	28	13	46%	Insufficient newsworthiness to trigger crisis news waves consistently
Mining accidents with more than 200 deaths	8	6	75%	Greater consistency as rarer, more newsworthy events are selected
Train accidents with more than 10 deaths	78	11	14%	Insufficient newsworthiness to trigger crisis news waves consistently when the type of event was established
Train accidents with more than 10 deaths before 1871	10	8	80%	More consistent triggering of crisis news waves as this type of event was new

11.6 Crisis news wave examples

300 annotated crisis news waves (those in brackets could not assigned to specific events):

- 1832 Cholera Epidemic (Paris/London)
- 1832 Cholera in Paris
- 1842 Spanish Ministerial Crisis
- 1845 French Agricultural Crisis
- 1845 Ministerial Crisis in England
- 1845 Railroad Crisis
- 1846 Affairs of Spain
- 1849 Cholera Epidemic
- 1849 Debate on Theories of Cholera
- 1851 Ministerial Crisis in England (Lord John Russell)
- 1852 RMS Amazon Shipwreck
- 1853 Cholera Epidemic
- 1853 Debate on Theories of Cholera
- 1853 Railway Accidents (various minor)
- 1854 Cholera Epidemic
- 1854 Cholera in North England
- 1854 Cholera outbreak report
- 1854 Railroad disasters report
- 1857 Fiscal/Banking Crisis Scotland
- 1861 London and Brighton Railway Disaster
- 1864 Shipwreck at Fifeness/Glasgow Harbor/SS Iowa
- 1866 Debate on Theories of Cholera
- 1868 Abergele rail disaster/report
- 1868 Abergele Railway Disaster
- 1870 Italian Revolution
- 1871 Explosives: Gun-cotton
- 1872 Constitutional Crisis in Germany and France

- 1872 Constitutional Crisis in Prussia/Ministerial Crisis in Hungary
- 1873 RMS Atlantic Shipwreck
- 1873 Ville du Havre Shipwreck
- 1873 Wigan Railway Accident
- 1873 Wigan Railway Accident/series of accidents
- 1874 Shipton Railway Accident
- 1874 Shipton Railway Accident/series of accidents
- 1874 Shipton Train Crash
- 1876 (various shipwrecks)
- 1876 Debate on Theories of Epidemics
- 1876 Great Queensland Shipwreck
- 1876 Ministerial Crisis in France
- 1877 (various shipwrecks)
- 1877 Avelanche/Forest Ship Collision
- 1877 Government Crisis in France
- 1877 Various Shipping Disasters
- 1878 Apedale/Unity Brook/Kersley Colliery Disasters
- 1878 Princess Alice Shipwreck
- 1878 Yellow Fever/series of accidents
- 1879 Tay-Bridge Disaster
- 1880 Eastern Crisis
- 1880 Kibworth Railway Accident
- 1880 The Eastern Crisis
- 1881 Ministerial Crisis in France
- 1881 RMS Teuton Shipwreck
- 1881 Teuton Shipwreck
- 1882 Crisis in Egypt
- 1882 Egypt Crisis (Navy)/Anglo-Egyptian War
- 1883 Cimbria Shipwreck
- 1883 Crisis in France/Spain
- 1883 SS Daphne Disaster/series of accidents/Cholera outbreaks
- 1884 Daniel Steinmann shipwreck
- 1884 Daniel Steinmann/State of Florida Shipwrecks
- 1885 Cholera in Spain
- 1885 Cholera in Spain/Eastern Crisis
- 1885 Eastern Crisis/Bulgarian Crisis
- 1885 Ministerial Crisis (Gladstone)
- 1885 Ministerial Crisis/colliery explosions/Cholera outbreaks
- 1886 Cholera in Italy
- 1886 Greek Crisis/Eastern Crisis
- 1886 Ministerial Crisis (Salisbury)
- 1886/87 Life Boat Disasters
- 1887 Ministerial Crisis in France
- 1887 Ministerial Crisis in France
- 1887 Scarlet Fever Epidemic
- 1887 Scarlet Fever Pandemic (London), strain on hospitals
- 1887 Series of Fires, explosions, epidemics
- 1889 Japanese Cabinet Crisis
- 1889 Penicuick Pit Disaster/Antwerp/Mossfield
- 1889-* 1890 Mauricewook/Morfa/Pontypool Colliery Disasters
- 1890 (invalid/Shipping)
- 1890 Asiatic Flu/Russian Flu
- 1890 HMS Serpent/Ertugrul Shipwrecks
- 1890 Influenza Epidemic

- 1890 Ministerial Crisis in Portugal
- 1890 Shipping/Financial Crisis
- 1891 Cabinet Crisis in Norway and Italy
- 1891 Utopia shipwreck
- 1892 Bokhara shipwreck/SS Romania shipwreck
- 1892 Cabinet Crisis in Italy, France, and Greece
- 1892 Cholera in Europe
- 1892 Cholera in Paris
- 1893 Cholera
- 1893 Crisis in Serbia/Austria
- 1893 Thornhill Colliery Explosion/Coal Trade Crisis
- 1893 Victoria Shipwreck/Cholera/Coal Trade Crisis
- 1894 (invalid)
- 1894 Ministerial Crisis in Belgium
- 1894 Ministerial14 Crisis in Hungary, Italy, and Bulgaria
- 1894/95 Resignation of the French President
- 1895 Venezuelan Crisis
- 1897 Eastern Crisis (Greco-Turkish War)
- 1897 Ministerial Crisis in Italy/Crisis in Austria
- 1897 Typhoid Epidemic (Maidstone)
- 1898 Crisis in France/Ministerial Crisis in Italy
- 1899 Crisis in Hungary
- 1899 Crisis in Hungary
- 1899 Crisis in the Church/Bullfinch disaster/Transvaal Crisis
- 1899 French Cabinet Crisis
- 1899 Transvaal Crisis
- 1900 Boxer Rebellion
- 1900 Boxer Rebellion (Navy)
- 1900 Boxer Rebellion (shipping of troops)
- 1901 Cabinet Crisis in Spain, Romania, and Italy
- 1901 Crisis in the British Industry (strikes)
- 1901 The Russie Shipwreck
- 1903 Bulgarian Cabinet Crisis/Balkan Crisis
- 1903 Crisis in the Far East
- 1903 Crisis in the Far East (pre Russo-Japanese War)
- 1903 Hungarian Crisis
- 1903 Hungarian Crisis/Crisis in Greece
- 1903 Hungarian Crisis/Epidemic at Cambridge
- 1904 Hungarian Crisis
- 1904 Russo-Japanese War
- 1905 Swedish-Norwegian Conflict
- 1907 Banking Crisis (Panic of * 1907)
- 1907 Education crisis
- 1907 Shewsbury Rail Disaster
- 1908 Crisis in Eastern Europe
- 1908 Wigan Colliery Disaster
- 1909 Greek Crisis/Hungarian Crisis
- 1910 Hawes Junction Rail Crash
- 1910 Houlton/Pretoria Pit Disaster (Bolton)
- 1910 Rat Plague in East Anglia/* 1911 Whitehaven Colliery Disaster
- 1910 Whitehaven Pit Disaster
- 1912 RMS Titanic Shipwreck
- 1913 Aisgill Rail Accident/* 1913 Mine Disasters
- 1913 Colchester Railway Accident

- 1913 Senghenydd Colliery Disaster/Sealing Shipwreck/Empress of Ireland Shipwreck
- 1913–1914 Irish Crisis
- 1914 Pre-War Crises/Ireland
- 1914 Resignation of Egyptian, Italian, and Japanese Cabinets
- 1914 Sealing Disaster (SS Newfoundland)/Empress of Ireland Disaster/War Preparations
- 1914 Senghenydd Colliery Disaster
- 1916 Greek Crisis/Collision in the Irish Channel
- 1918 Spanish Flu
- 1923 Cabinet Crises in Germany, Poland, Netherlands, Portugal/Bengal Crisis
- 1925 Belgian Cabinet Crisis
- 1930 R101 Airship Disaster
- 1931 British Honduras Disaster (hurricane)/* 1931 Currency Crisis
- 1931 Financial Crisis
- 1931 Financial Crisis
- 1931 German Crisis
- 1931: UK General Election (MacDonald, National)
- 1938 Nazi Threat/War Preparations
- 1938 Sudeten Crisis (negotiations)
- 1938 Sudeten Crisis/Munich Agreement
- 1944 Post-Nazi Government Crises (Italy, Belgium, Norway, Poland, Greece)
- 1947 Cholera in Punjab and Egypt
- 1947 Coal Crisis (negotiations)
- 1947 Coal/Energy/Industry Crisis
- 1947 Currency/Financial Crisis
- 1947 Energy and Industry Recovery
- 1947 Financial Crisis/Marshall Plan
- 1947 Industry/Steel/Coal Crisis
- 1947 Industry/Steel/Coal Crisis (2)
- 1947 Postwar Reverberations
- 1950 Creswell Colliery Disaster
- 1956 Suez Crisis
- 1957 Asian Flu
- 1957 Recession
- 1958 Taiwan Straits Crisis
- 1961 Berlin Crisis
- 1962 Cuba Crisis
- 1964 Cyprus Crisis
- 1967 Foot and Mouth Disease
- 1968 Money Talks (revaluation of German Mark)
- 1971 Currency Crisis/Bretton Woods Crisis
- 1971 Currency Revaluations
- 1971 Monetary Crisis
- 1973 Currency crisis
- 1973 Oil Crisis
- 1973 Oil Crisis
- 1973 Oil Crisis
- 1973 Oil Crisis
- 1973 Oil Crisis
- 1973 Oil Crisis (Japan)
- 1973 Oil Crisis: Inflation
- 1973 Oil Crisis: Regulations/Restrictions/Rationing
- 1973 Oil Crisis: Steel Industry
- 1974 Banking Crisis
- 1974 Turkish Invasion of Cyprus

- 1975 Recession
- 1976 Recession
- 1978 (invalid)
- 1978 Economic Problems
- 1979 (invalid)
- 1979 Afghanistan War/Iran Hostage Crisis
- 1979 Afghanistan-Soviet War
- 1979 Afghanistan-Soviet War/* 1979 Iranian Hostage Crisis
- 1979 Iran Hostage Crisis
- 1979 Iranian Hostage Crisis
- 1979 Irianian Hostage Crisis/* 1979 Afghanistan War
- 1979 Mount Erebus Disaster
- 1979 Recession
- 1979 Societ-Afghan War/Pakistan Crisis
- 1979 Steel Strikes
- 1980 Afghanistan War/Iran Hostage Crisis
- 1980 Cyprus Negotiations
- 1980 Recession
- 1980 Recession
- 1980 Recession
- 1980 Steel Quotas Negotiations
- 1980 Strikes/Protests in Poland
- 1980/81 Unemployment Surge/Recession
- 1981 Recession
- 1982 (invalid)
- 1982 Argentina Crisis
- 1982 Falkland War
- 1982 Falkland War: Peruvian Peace Initiative
- 1982 Flu & Pertussis Epidemics/Falkland War
- 1982 Recession
- 1982 Falkland War: Tory Wins at Local Elections
- 1983 Oil Output
- 1985 Boeing Disaster (British Airtours 28M
- 1985 Tin Crisis
- 1986 Israel Cabinet Crisis/Chernobyl
- 1987 Ferry Disaster (MS Herald of Free Enterprise)
- 1987 NHS Crisis (hospital beds)
- 1988 King’s Cross Underground Fire Report
- 1988 Lockerbie Air Disaster (PanAm 103 Crash)
- 1990-1992 Recession
- 1991 Recession
- 1991 Recession
- 1991 Theatre Season
- 1992 Extreme Weather
- 1992 Recession
- 1992 Theatre Season
- 1992 (Personal Crises)
- 1993 Theatre Season
- 1995 Theatre Season
- 1997 Asian Financial Crisis
- 1997 Hong Kong Crisis
- 1998 (invalid)
- 1998 Asian Crisis
- 1998 Flu Epidemic

- 2000 (invalid)
- 2001 Foot and Mouth Disease
- 2004 (invalid)
- 2005 Bird Flu
- 2007 Mortgage Crisis (Northern Rock)
- 2008 Financial Crisis
- 2008 Financial Crisis
- 2008 Financial Crisis
- 2008 Financial Crisis
- 2008 Financial Crisis
- 2008 Financial Crisis (Pensions)
- 2008 Financial Crisis: Bailouts
- 2008 Financial Crisis: Job market for graduates
- 2008 Lehman Brothers
- 2008 Lehman Brothers (Oil Prices)
- 2008 Financial Crisis (Personal Crises)
- 2009 (invalid)
- 2009 Swine Flu/Mexican Flu
- 2010 Deepwater Horizon Disaster
- 2011 European Debt Crisis
- 2011 European Debt Crisis
- 2011 European Debt Crisis
- 2011 European Debt Crisis (Personal crises)
- 2014 NHS crisis (A&E capacity)
- 2014/2015 Migrantion Crisis
- 2014–2016 Oil Price Fall
- 2015 Migration Crisis
- 2015 Migration Crisis
- 2015 Migration Crisis (Greece)
- 2015 Migration Crisis (stress on NHS)
- 2015 NHS Crisis (A&E capacity)
- 2015 Migration Crisis (Personal crises)
- 2016 BHS Collapse (retailer)/NHS Pensions
- 2016 Brexit Referendum
- 2017 NHS Crisis (staff shortages)
- 2017/18 Flu Season and NHS Crisis
- 2018 BHS Report (Business auditors)
- 2018–2019 Brexit Negotiations
- 2019 Brexit Consequences
- 2020 COVID Crisis
- 2020 COVID Crisis (Pensions Crisis)
- 2020 COVID Crisis/strain on NHS
- 2020 Covid Pandemic
- 2020 Covid Pandemic (A&E capacity)
- 2020 COVID-19 Crisis
- 2020 COVID-19 Crisis
- 2020 COVID-19 Crisis: Auditing shortages
- 2020 COVID-19 Crisis: Hospitals and Long-term effects
- 2020 COVID-19 Crisis: Job market requirements
- 2020 COVID-19 Crisis: Psychosocial consequences
- 2020 COVID-19 Crisis: Regulations/Restrictions/Rationing
- 2020 COVID-19 Crisis: Strain on hospitals
- 2020 COVID-19 Pandemic
- 2020 Recession

- 2020 Stock market crisis
- 2020 COVID-19 Crisis
- 2020 COVID-19 Crisis (Personal crises)

Chapter 12

Appendix M: Details on the Named Entity Recognition procedure

12.1 Purpose

This appendix details the procedure (1) for finding unique organizations mentioned in the crisis corpus using named entity recognition and (2) for checking that organizations mentioned in crisis coverage engage in crisis discourse frequently such that we can presume that a larger number of unique organizations that appear in crisis coverage will coincide with a larger number of unique organizations that engage in crisis discourse, when aggregated across an entire year.

12.2 Procedure

For named entity recognition with the `spacyr` package in R (Benoit and Matsuo 2020), we used a text representation with minimal manipulations because features like punctuation and uppercase letters, which create unwanted noise in several other procedures, are important cues for identifying proper names of organizations, persons, countries, etc. We used `tms_tx_u` representation.

We separated the text corpus into chunks organized by publication date: Before 1800, 1800–1899, 1900–1924, 1925–1949, 1950–1969, 1970–1979, 1980–1989, 1990–1999, 2000–2009, 2010–2014, 2015–2017, 2018–2020. For each of these corpora, the spacy parser annotated each word with its grammatical function(s), and the kind of entity that the algorithm assumes that it represents. The entities that are distinguished are:

1. Numeric entities: cardinal numbers (`CARDINAL`), ordinal numbers (`ORDINAL`), quantities (`QUANTITY`), money (`MONEY`), percentages (`PERCENT`), dates (`DATE`), time (`TIME`).
2. Other entities: Persons (`PER`), nationalities/religious or political groups (`NORP`), facilities and buildings (`FAC`), organizations (`ORG`), geopolitical entities (`GPE`), non-GPE locations (`LOC`), products (`PRODUCT`), events, (`EVENT`), works of art (`WORK_OF_ART`), law (`LAW`), language (`LANGUAGE`).

In this analysis, we concentrate on the ORG entities only that represent spacy's best guess that the word represents an organization.

First, we create a year-by-year list of all unique organizations that are mentioned two or more times in that year. This threshold of two mentions removes most erroneous guesses from the list. The length of that list provides the number of unique organizations mentioned in crisis coverage during the respective year. We presume that many of these mentions represent participation in the crisis frame contest.

Second, we divide the number of unique organizations mentioned in crisis coverage by the total number of articles published in *The Times* in that year. This way, the indicator captures the density of mentions of organizations in a crisis context. Given that the likelihood that the newspaper chooses to cover the efforts of a crisis frame sponsor is constant, this translates into a greater activity of organizations as crisis frame sponsors or competitors in a crisis framing contest. We do not divide by the number of CL articles to prevent any artificially exaggerated statistical linkage between the indicator of crisis frame sponsor activity (independent variable) and CL salience, CNW salience, and CNW count (dependent variables).

Third, we multiply the indicator with 1000, making it the *Number of unique crisis frame sponsors mentioned per 1000 published articles*.

12.3 Validation

12.3.1 Crisis and noncrisis corpus

We conducted a small manual coder agreement test that contrasts a stratified random selection of 25 articles from our crisis corpus with 25 articles that are a stratified random sample from the full *The Times* archive. The coders were blind as to whether an article came from the crisis corpus or the random corpus. They were then asked to classify

- Which actors/speakers/protagonists appear in coverage, and do they contribute to crisis discourse?

Not all actors that are being mentioned in an article in the crisis corpus definitively contribute to crisis discourse—there is a lot of uncertainty in many cases; approximately half the actors that appear are classified as contributing to crisis discourse with “high” or “very high” certainty. However, almost all actors appearing in the random corpus clearly do not contribute to a crisis discourse of some sort (see Appendix G).

12.3.2 Crisis corpus

We conducted a manual coding of 246 randomly selected articles from the crisis corpus as training and validation data for building a classifier.

The manifold categories that were coded mostly had good (above .80) or satisfactory (.67-.80) inter-coder agreement, with few categories falling below .67. We mark all categories where results may be less than satisfactory with double exclamation marks.

Table 12.1: Coding agreement in the manual coding

Construct	Agreement	Kappa_BP	Scale type
Crisis criteria	.54	.54 (!!)	Scale type
Nature of the crisis	.98	.96	Scale type
Current status	Agreement	.80	Scale type
Future status	Agreement	.63 (!!)	Scale type
Response by decision-makers	.94	.88	Scale type
Response by third parties	.93	.86	Scale type
Response by bystanders	.93	.87	Scale type
Material symptoms	.95	.87	Scale type
Immaterial symptoms	.95	.88	Scale type
Systemic symptoms	.98	.94	Scale type
Penetration	.84	.28 (!!)	Scale type
Causation	.87	.74 (!)	Scale type
Career Stage	.88	.76 (!)	Scale type
Cond. development	.84	.69 (!)	Scale type
Benchmarks	.77	.53 (!!)	Scale type
Challenge: concern	.83	.67 (!)	Scale type

Construct	Agreement	Kappa_BP	Scale type
Challenge: crisis	Agreement	.78 (!)	Scale type
Challenge: labelling	Agreement	.90	Scale type
Challenge: material	Agreement	.79 (!)	Scale type
Challenge: systemic	Agreement	.82	Scale type
Challenge: diagnosis	Agreement	.53 (!!)	Scale type
Challenge: response by decision-makers	Agreement	.49 (!!)	Scale type
Challenge: response by bystanders	Agreement	.69 (!)	Scale type
Challenge: response by third parties	Agreement	.45 (!!)	Scale type
Geography: Contitent Symptoms/ origins/ solutions	.98	.93	Scale type
Geography: Country Appearance	.93	.82	Scale type
Geography: Country	.98	.92	Scale type
Newswork	.98	.92	Scale type
Crisis degree	.98	.92	Scale type
Actors			
Appearance	.96	.93	Scale type
Speakers	.80	.60 (!!)	Scale type
Protagonists	.86	.73 (!)	Scale type

12.3.3 Can actors in crisis labelling coverage be presumed to engage in crisis labelling?

Most actors that are mentioned in crisis labelling coverage also participate in crisis discourse in some way. 68% of articles include some actors, and out of these 168 articles, 78% some actor engages in crisis discourse. Out of all 350 actors recorded, 210 (60%) engaged in crisis discourse. This corresponds to the findings from the corpus contrast study where roughly half the actors mentioned in the crisis corpus engaged in crisis discourse; in the random corpus, almost no actors engaged in crisis discourse. We observed a substantial upward trend in crisis discourse participation of actors that were mentioned. 1785-1899, we 56% of actors recorded in the crisis corpus participated in crisis discourse; 1900-1949, the share was at 50%; then, it rose to 65% (1950-1999), and stood at 76% 2000-2020 Figure 12.1 (see also Appendix G for methodological details).

At the same time, more different actor categories appear in news coverage during the period of study, with 9% of the available actor categories appearing in the 1785-1899 sample; 11% of actor categories appeared in the 1900-1949 period; 23% of actor categories appeared in the 1950–1999 period; and 24% of actor categories appeared in the 2000-2020 period.

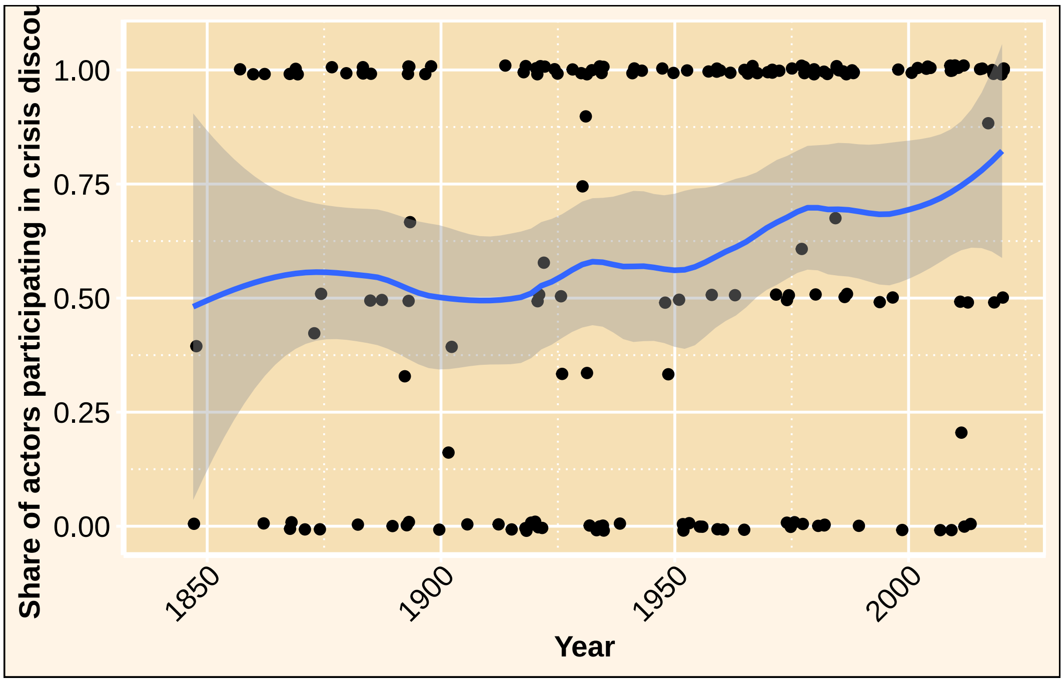
This justifies our decision to identify actor mentions in crisis coverage and presume that a greater number of different actors will participate in crisis discourse over time—we observe a greater diversity of actors, and the chance of a recorded actor to participate in crisis discourse is 50% or greater. Aggregating the entire coverage of an entire year and finding a greater diversity of actors can thereby serve as an indicator of greater diversity of actors that participate in crisis discourse.

Table 12.2: Actors mentioned and actors participating in crisis discourse

	All articles % (n=246)	Articles that mention actors % (n=168)
Mentions actors	68.3	100.0

	All articles % (n=246)	Articles that mention actors % (n=168)
Crisis treatment	44.7	65.5
Crisis diagnosis	33.3	48.8
Crisis labelling	19.1	28.0
Crisis discourse	53.3	78.0

Figure 12.1: Share of actors appearing in crisis coverage that participate in crisis discourse



Chapter 13

Appendix N: External data sources—Media Penetration

13.1 Purpose

This appendix complies the sources for the data on media penetration and the calculations used to arrive at a global indicator of media penetration that is used in the study.

13.2 Newspaper penetration (NP)

After 1950, the Audit Bureau of Circulation (ABC) provides definitive figures on newspaper circulation; before that, we rely on a mixture of ABC data, data collected by advertisers, and data collected for collecting stamp tax. The years covered are 1838, 1852, 1863, 1910, 1921, 1930, 1939, 1947, 1956, 1961, 1966, 1976, 1980, 1987, 1992, 1997, 2000-2020. For the years not covered, we used linear interpolation. The data from these sources was compiled and we collected it via the Wikipedia page "List_of_newspapers_in_the_United_Kingdom_by_circulation".

We checked the original sources for a small subset of the numbers.

13.3 TV penetration (TVP)

The share of UK households with TV sets (1956-2019) was documented by "Closer.ac.uk", a longitudinal data hosting service hosted by the University College London's Social Research Institute; it is a collaboration involving UK Data Service and The British Library.

13.4 Internet penetration (IP)

We used the World Bank's data repository to retrieve the International Telecommunication Union's data on Internet usership in the UK, given as percentage of the population. We used these proportions (maximum value: 2020: 0.95) as an indicator of Internet penetration.

13.5 Aggregation to media penetration (MP)

According to our model, the vital part of media penetration is what share of the population can be expected to receive regular updates on political matters. A conservative estimate of that is the maximum of newspaper penetration, TV penetration, or Internet penetration in the respective year. This estimate is conservative because it presumes that there is full overlap between the higher-penetration sources and all the lower-penetration sources such that the full reach equals "only" the reach of the highest-penetration source.

$$MP_{year} = \max\left(NP_{year}, TVP_{year}, IP_{year}\right)$$

Chapter 14

Appendix O: Calculation of public spending diversity

14.1 Purpose

This appendix provides descriptives of the UK budget structure and its development, plus equations and example calculations of the diversity measure used in the study.

14.2 Data and Example Calculation

UK budget and spending data (classified into 18 budget categories) are available from 1790–2020, covering almost the entire period of study (1785–2020). As examples for illustrating the calculation of the diversity of public spending, we display the data for 1800, 1850, 1900, 1950, 2000, and 2020.

The four-step calculation of the diversity score is given below, working backwards. "i" is the index for the c=18 budget categories. Budget categories need to be sorted in ascending order before the calculation.

$$Diversity_{year} = 1 - Gini_{year}^{unbiased}$$

$$Gini_{year}^{unbiased} = Gini_{year} \cdot \frac{1}{1 - \sum_i^c w^2} = Gini_{year} \cdot 1.059$$

$$Gini_{year} = \frac{2}{\frac{1}{18}} \cdot \sum_{i=1}^{c} \frac{1}{18} \cdot \left(x_i - \frac{1}{18}\right) \cdot \hat{f}_i = 36 \cdot \sum_{i=1}^{c} \frac{\hat{f}_i}{18} \cdot \left(x_i - \frac{1}{18}\right)$$

$$\hat{f} = \left(\frac{-8.5}{18}, \frac{-7.5}{18}, \frac{-6.5}{18}, \frac{-5.5}{18}, \frac{-4.5}{18}, \frac{-3.5}{18}, \frac{-2.5}{18}, \frac{-1.5}{18}, \frac{-0.5}{18}, \frac{0.5}{18}, \frac{1.5}{18}, \frac{2.5}{18}, \frac{3.5}{18}, \frac{4.5}{18}, \frac{5.5}{18}, \frac{6.5}{18}, \frac{7.5}{18}, \frac{8.5}{18}\right)$$

The coefficients are calculated using `DescTools::Gini` in `R`.

Table 14.1: Distribution of spending across budget categories (rounded), and calculated Gini coefficients and public spending diversity measures

Year	1800	1850	1900	1950	2000	2020
Data	—	—	—	—	—	—
Economy	0	0	7	13	2	1
Agriculture	0	0	0	0	0	1
Resources	0	0	0	0	1	0
Pensions	0	0	0	11	24	24
Health	0	0	3	10	18	24
Education	0	0	0	0	5	7
Military	97	67	57	37	20	13
Family	0	0	0	0	0	0
Unemployment	0	2	1	2	1	1
Transport	0	0	0	0	1	1
Housing	0	0	1	11	1	1
Recreation	0	0	1	0	2	1
Environment	3	14	5	2	4	3
Government	0	10	16	11	10	11
Welfare	0	0	0	0	12	2
Security	0	0	10	13	10	16
Economic Aid	0	0	0	0	0	0
Result	—	—	—	—	—	—
Gini	.997	.950	.839	.708	.664	.698
Diversity (1-Gini)	.003	.050	.161	.292	.336	.302

Chapter 15

Appendix P: Correlation matrix of predictors and criteria

15.1 Purpose

This appendix provides the correlation matrix for all variables included in the analyses in the main study.

15.2 Correlation matrix

Table 15.1 display partial correlations above the diagonal and zero-order correlation coefficients below the diagonal.

Table 15.1: Correlation matrix of predictors and criteria

	1 Yr	2 PSI	3 PSD	4 MPen	5 MAut	6 CFSA	7 CLS	8 CNWS	9 nCNW
1 Year	—	.416	.647	.0940	.354	-.101	.314	.568	.247
2 Public spending intensity	.782	—	-.691	.746	-.198	.446	-.163	-.583	-.153
3 Public spending diversity	.907	.592	—	.583	-.240	.450	-.186	-.634	-.060
4 Media Penetration	.928	.868	.863	—	.155	-.358	.133	.448	-.053
5 Media Autonomy	.620	.452	.516	.566	—	.730	-.998	-.885	.069
6 Crisis frame sponsor activity	.784	.484	.691	.659	.667	—	-.764	-.330	.137
7 CL salience	.843	.591	.746	.739	.706	.934	—	-.860	.076
8 CNW salience	.260	.115	.239	.200	.309	.431	.478	—	-.004
9 CNW count	.258	.037	.240	.150	.232	.484	.507	.909	—

Chapter 16

Appendix Q: Predicting CL salience, CNW salience, and CNW count

16.1 Purpose

This appendix provides the full tables of the regression models from which the regression models presented in the study were picked. In addition, it presents model comparison tests that allow to check which models lead to improvement in model fit.

16.2 Regression tables and model comparisons

16.2.1 Crisis labelling salience

16.2.1.1 Model series

Table 16.1 presents the full series of models that predict crisis labelling salience

Table 16.1: CL Salience - Complete Model Series

	M1.0	M1.1	M1.2	M1.3	M1.4	M1.5	M1.6	M1.7
(Intercept)	1.377***	-0.089	0.260***	0.297***	0.130**	-0.025	-0.113*	-0.152*
	(0.066)	(0.071)	(0.068)	(0.073)	(0.039)	(0.097)	(0.054)	(0.060)
Pervasiveness of media logic								
Media Penetration			1.203***			-0.155	-0.301	-0.325
			(0.111)			(0.359)	(0.199)	(0.199)
Media Autonomy			8.526***			8.335***	2.345***	2.239***
			(0.923)			(0.871)	(0.550)	(0.553)
Public spending								
Public Spending Intensity				1.632***		1.216+	1.344***	1.059**
				(0.371)		(0.654)	(0.363)	(0.407)
Public Spending Diversity				5.133***		4.095***	1.563***	1.019+
				(0.437)		(0.756)	(0.434)	(0.560)
Crisis frame sponsors								

Table 16.1: CL Salience - Complete Model Series

	M1.0	M1.1	M1.2	M1.3	M1.4	M1.5	M1.6	M1.7
Crisis Frame Sponsor Activity					0.404***		0.311***	0.298***
					(0.010)		(0.014)	(0.016)
Years since 1785		0.012***						0.002
		(0.001)						(0.001)
Num.Obs.	236	236	236	236	236	236	236	236
R2	0.000	0.710	0.667	0.591	0.872	0.709	0.911	0.912
R2 Adj.	0.000	0.709	0.664	0.587	0.871	0.704	0.909	0.910
AIC	677.9	387.6	422.1	470.9	195.6	394.5	117.2	116.8
BIC	684.8	398.0	436.0	484.8	206.0	415.3	141.4	144.5
Log.Lik.	-336.931	-190.809	-207.058	-231.452	-94.807	-191.249	-51.585	-50.389
F		573.257	233.704	168.300	1587.154	140.736	470.373	394.639
RMSE	1.01	0.54	0.58	0.65	0.36	0.54	0.30	0.30

Note: ^^ + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

16.2.1.2 Model comparison

Model comparisons for crisis labelling salience using Wald F-tests:

Table 16.2: Model comparisons—predicting CL salience

Models compared	Change to previous model	df (change)	p value
Null model			
M1.0 to M1.1	Adds time	2(1)	<.001 ***
M1.0 to M1.2	Adds media logic	3(2)	<.001 ***
M1.0 to M1.3	Adds public spending	3(2)	<.001 ***
M1.0 to M1.4	Adds crisis frame sponsor activity	2(1)	<.001 ***
Media logic only model (ML)			
M1.2 to M1.5	Adds gov. spending	5(2)	<.001 ***
Public spending only model (PS)			
M1.3 to M1.5	Adds media logic	5(2)	<.001 ***
ML and PS model			
M1.5 to M1.6	Adds crisis frame sponsor activity (CFSA)	6(1)	<.001 ***
ML, GP, and CFSA model			
M1.6 to M1.7	Adds time (T)	7(1)	.128 ^{ns}

16.2.2 Crisis news wave salience

16.2.2.1 Model series

Table 16.3 presents the full series of models that predict crisis news wave salience.

Table 16.3: CNW Salience - Complete Model Series

	M2.0	M2.1	M2.2	M2.3	M2.4	M2.5	M2.6	M2.7
(Intercept)	0.159***	-0.012	-0.018	0.049	-0.059	-0.032	-0.054	-0.020
	(0.025)	(0.048)	(0.043)	(0.042)	(0.037)	(0.064)	(0.061)	(0.068)
Pervasiveness of media logic								
Media Penetration			0.034			-0.034	-0.071	-0.050
			(0.069)			(0.238)	(0.226)	(0.227)
Media Autonomy			2.201***			2.110***	0.581	0.673
			(0.578)			(0.578)	(0.624)	(0.629)
Public spending								
Public Spending Intensity				-0.111		-0.225	-0.192	0.054
				(0.213)		(0.434)	(0.412)	(0.463)
Public Spending Diversity				0.840***		0.567	-0.080	0.390
				(0.252)		(0.501)	(0.492)	(0.637)
Crisis frame sponsors								
Crisis Frame Sponsor Activity					0.071***		0.079***	0.091***
					(0.010)		(0.015)	(0.018)
Years since 1785		0.001***						-0.002
		(0.000)						(0.001)
Num.Obs.	236	236	236	236	236	236	236	236
R2	0.000	0.068	0.096	0.058	0.186	0.111	0.202	0.207
R2 Adj.	0.000	0.064	0.089	0.050	0.182	0.095	0.185	0.186
AIC	220.5	206.0	200.6	210.4	174.0	200.8	177.2	177.8
BIC	227.5	216.4	214.4	224.3	184.4	221.6	201.5	205.5
Log.Lik.	-108.268	-99.993	-96.295	-101.201	-83.996	-94.411	-81.608	-80.917
F		16.998	12.441	7.190	53.441	7.196	11.661	9.956
RMSE	0.38	0.37	0.36	0.37	0.35	0.36	0.34	0.34

Note: ^^ + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

16.2.2.2 Model comparison

Model comparisons for crisis news wave salience using Wald F-tests:

Table 16.4: Model comparisons—predicting CNW salience

Models compared	Change to previous model	df (change)	p value
Null model			
M2.0 to M2.1	Adds time	2(1)	<.001 ***
M2.0 to M2.2	Adds media logic	3(2)	<.001 ***
M2.0 to M2.3	Adds public spending	3(2)	<.001 ***
M2.0 to M2.4	Adds crisis frame sponsor activity	2(1)	<.001 ***
Media logic only model (ML)			

Models compared	Change to previous model	df (change)	p value
M2.2 to M2.5	Adds public spending	5(2)	.158 ^{ns}
Public spending only model (PS)			
M2.3 to M2.5	Adds media logic	5(2)	.001 **
ML and PS model			
M2.5 to M2.6	Adds crisis frame sponsor activity (CFSA)	6(1)	<.001 ***
ML, PS, and CFSA model			
M2.6 to M2.7	Adds time (T)	7(1)	.248 ^{ns}

16.2.3 Crisis news wave count

16.2.3.1 Model series

Table 16.5 presents the full series of models that predict crisis news wave count.

Table 16.5: CNW Count – Complete Model Series

	M3.0	M3.1	M3.2	M3.3	M3.4	M3.5	M3.6	M3.7
(Intercept)	1.184***	0.669***	0.924***	0.976***	0.479***	0.821***	0.716***	0.476***
	(0.064)	(0.122)	(0.114)	(0.105)	(0.090)	(0.163)	(0.132)	(0.142)
Pervasiveness of media logic								
Media Penetration			0.091			-0.402	-0.578	-0.726
			(0.184)			(0.606)	(0.489)	(0.476)
Media Autonomy			3.013+			2.404	-4.784***	-5.434***
			(1.537)			(1.470)	(1.350)	(1.320)
Public spending								
Public Spending Intensity				-1.917***		-1.455	-1.301	-3.043**
				(0.530)		(1.103)	(0.891)	(0.973)
Public Spending Diversity				3.516***		3.887**	0.848	-2.474+
				(0.625)		(1.274)	(1.064)	(1.338)
Crisis frame sponsors								
Crisis Frame Sponsor Activity					0.228***		0.373***	0.294***
					(0.023)		(0.033)	(0.038)
Years since 1785		0.004***						0.012***
		(0.001)						(0.003)
Num.Obs.	236	236	236	236	236	236	236	236
R2	0.000	0.092	0.032	0.120	0.292	0.130	0.436	0.471
R2 Adj.	0.000	0.088	0.023	0.112	0.289	0.115	0.424	0.457
AIC	666.1	645.3	662.5	640.0	586.5	641.1	540.9	527.7
BIC	673.0	655.7	676.3	653.8	596.9	661.9	565.2	555.4
Log.Lik.	-331.047	-319.641	-327.245	-315.979	-290.252	-314.560	-263.474	-255.867
F		23.748	3.814	15.867	96.642	8.659	35.556	34.007
RMSE	0.98	0.94	0.97	0.92	0.83	0.92	0.74	0.72

Note: ^^ + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

16.2.3.2 Model comparisons

Model comparison for crisis news wave count using Wald F-tests

Table 16.6: Model comparisons—predicting CNW count

Models compared	Change to previous model	df (change)	p value
Null model			
M3.0 to M3.1	Adds time	2(1)	<.001 ***
M3.0 to M3.2	Adds media logic	3(2)	.023 *
M3.0 to M3.3	Adds public spending	3(2)	<.001 ***
M3.0 to M3.4	Adds crisis frame sponsor activity	2(1)	<.001 ***
Media logic only model (ML)			
M3.2 to M3.5	Adds public spending	5(2)	<.001 ***
Public spending only model (PS)			
M3.3 to M3.5	Adds media logic	5(2)	.249 ^{ns}
ML and PS model			
M3.5 to M3.6	Adds crisis frame sponsor activity (CFSA)	6(1)	<.001 ***
ML, PS, and CFSA model			
M3.6 to M3.7	Adds time (T)	7(1)	<.001 ***

16.2.4 Crisis frame sponsor activity

16.2.4.1 Model series

Table 16.7 presents the full series of models that predict crisis frame sponsor activity.

Table 16.7: Crisis frame sponsor activity – Complete Model Series

	M4.0	M4.1	M4.2	M4.3	M4.4
(Intercept)	3.085***	0.667***	0.884***	0.282	-0.649**
	(0.152)	(0.181)	(0.190)	(0.259)	(0.241)
Pervasiveness of media logic					
Media Penetration		2.305***		0.470	-0.186
		(0.293)		(0.961)	(0.820)
Media Autonomy		20.088***		19.245***	11.522***
		(2.450)		(2.332)	(2.142)
Public spending					
Public Spending Intensity			1.893*	-0.411	-6.473***
			(0.959)	(1.751)	(1.619)
Public Spending Diversity			12.080***	8.135***	-5.938**
			(1.132)	(2.023)	(2.269)
Years since 1785					0.043***
					(0.004)
Num.Obs.	236	236	236	236	236
R2	0.000	0.561	0.486	0.609	0.719

Table 16.7: Crisis frame sponsor activity – Complete Model Series

	M4.0	M4.1	M4.2	M4.3	M4.4
R2 Adj.	0.000	0.557	0.481	0.602	0.713
AIC	1072.7	882.6	919.8	859.2	783.1
BIC	1079.6	896.4	933.7	879.9	807.4
Log.Lik.	-534.360	-437.277	-455.924	-423.577	-384.574
F		148.738	109.967	89.917	117.696
RMSE	2.33	1.54	1.67	1.46	1.23

Note: ^^ + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

16.2.4.2 Model comparisons

Model comparisons for crisis frame sponsor activity using Wald F-tests

Table 16.8: Model comparisons—predicting crisis frame sponsor activity

Models compared	Change to previous model	df (change)	p value
Null model			
M4.0 to M4.1	Adds media logic	3(2)	<.001 ***
M4.0 to M4.2	Adds gov. spending	3(2)	<.001 ***
Media logic only model (ML)			
M4.1 to M4.3	Adds gov. spending	5(2)	<.001 ***
Gov. spending only model (GS)			
M4.2 to M4.3	Adds media logic	5(2)	<.001 ***
ML and GS model			
M4.3 to M4.4	Adds time (T)	6(1)	<.001 ***

Chapter 17

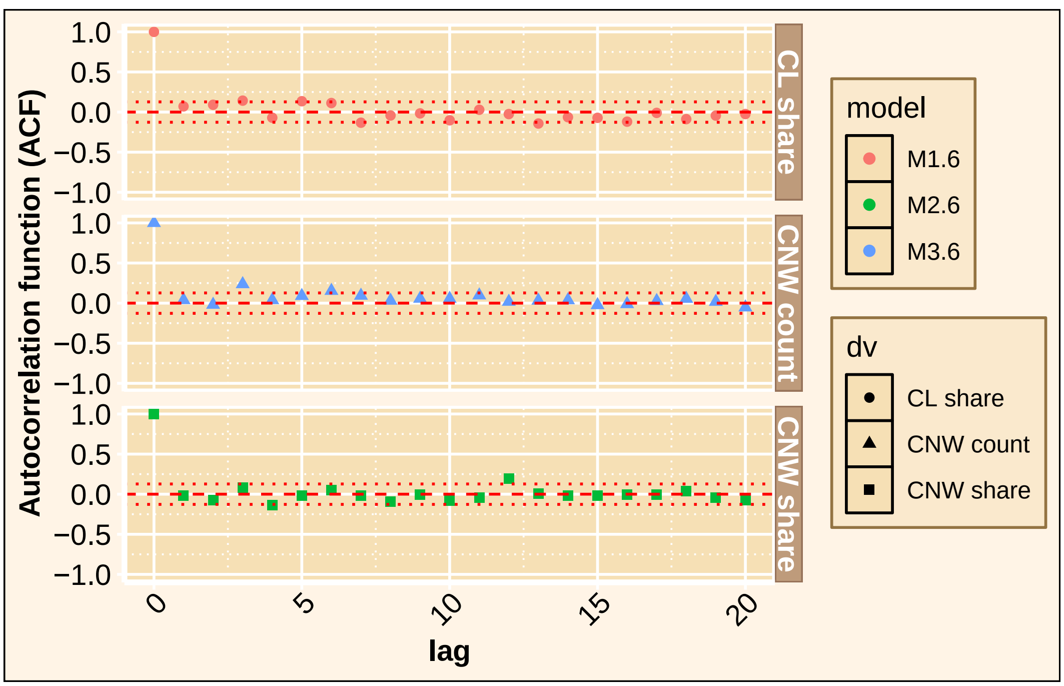
Appendix R: Tests for Autocorrelation and Stationarity

17.1 Non-stationarity and Autocorrelation tests

We tested the models for predicting CL salience, CNW salience, and CNW count for non-stationarity and autocorrelation.

There are no indications of autocorrelation Figure 17.1 or partial autocorrelation Figure 17.2.

Figure 17.1: Autocorrelation functions (ACF) for M1.6, M2.6, and M3.6



These visual results are confirmed by Box-Ljung tests for autocorrelation Table 17.1.

Figure 17.2: Partial Autocorrelation functions (PACF) for M1.6, M2.6, and M3.6

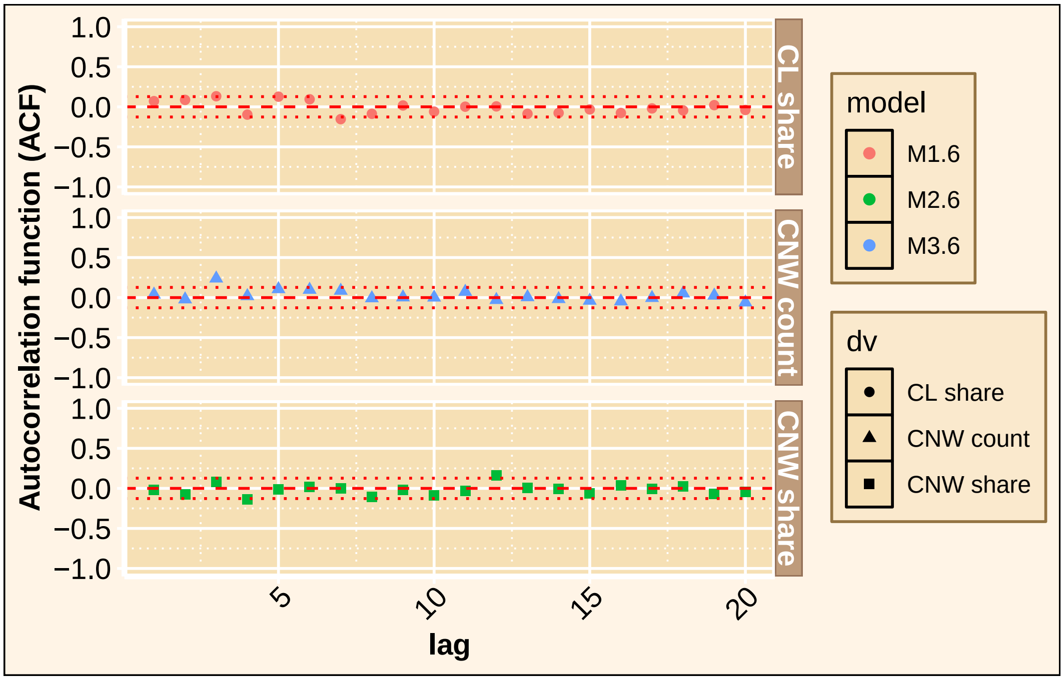


Table 17.1: Results of Ljung-Box tests of autocorrelation

Box_Ljung	Alternative_hypothesis	P_value
86.844	Autocorrelation present	0.823
96.944	Autocorrelation present	0.568
110.798	Autocorrelation present	0.216

We also tested for stationarity of the time series using KPSS tests (for level and trend stationarity, respectively) and Augmented Dickey–Fuller (ADF) tests, which all signal that the data can be treated as stationary.

Table 17.2: Results of Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests of stationarity

KPSS	p	Alternative_hypothesis
0.033	>0.10	Trend is not stationary
0.036	>0.10	Trend is not stationary
0.118	>0.10	Trend is not stationary
0.035	>0.10	Level is not stationary
0.040	>0.10	Level is not stationary
0.375	0.088	Level is not stationary

Table 17.3: Results of Augmented Dickey–Fuller (ADF) tests of stationarity

Dickey_Fuller	p	Alternative_hypothesis
-4.875	<0.01	Time series is stationary
-5.595	<0.01	Time series is stationary

Table 17.3: Results of Augmented Dickey–Fuller (ADF) tests of stationarity

Dickey_Fuller	p	Alternative_hypothesis
-4.405	<0.01	Time series is stationary

17.2 Automatic ARIMA model identification

Auto-detection of appropriate ARIMA(n,p,q) modeling for these models suggested to include moving averages of the first order (ARIMA(0,0,1)) for models M1.6 and M2.6 and ARIMA(0,1,2) for M3.6.

While this suggests that it might be possible to include an MA component, the ACF and PACF plots do not point in that direction. Also, the results from the ARIMA(0,0,1) models and the corresponding simple regression models do not lead to different results of our hypothesis tests (Table 17.4, Table 17.5, Table 17.6). We therefore continue with the simple regression models as they are more flexible when it comes to exploring mediation effects (H3) and comparing coefficients (H4)

Table 17.4: Results of simple regression versus ARIMA(0,0,1) model of CL salience

	Estimate	SE	P	ARIMAEstimate	ArimaSE	ArimaP
(Intercept)	-0.1129	0.0538	0.037	-0.0011	0.0006	0.062
Media Penetration	-0.3008	0.1993	0.133	-0.0027	0.0022	0.209
Media Autonomy	2.3454	0.5499	0.000	0.0262	0.0063	0.000
Public Spending Intensity	1.3439	0.3629	0.000	0.0128	0.0039	0.001
Public Spending Diversity	1.5630	0.4335	0.000	0.0164	0.0047	0.001
Crisis Frame Sponsor Activity	0.3112	0.0136	0.000	0.0030	0.0002	0.000

Table 17.5: Results of simple regression versus ARIMA(0,0,1) model of CNW salience

	Estimate	SE	P	ARIMAEstimate	ArimaSE	ArimaP
(Intercept)	-0.0541	0.0611	0.377	-0.0005	0.0006	0.393
Media Penetration	-0.0712	0.2264	0.753	-0.0006	0.0022	0.785
Media Autonomy	0.5806	0.6245	0.354	0.0051	0.0061	0.403
Public Spending Intensity	-0.1923	0.4121	0.641	-0.0021	0.0040	0.604
Public Spending Diversity	-0.0799	0.4923	0.871	-0.0008	0.0047	0.865
Crisis Frame Sponsor Activity	0.0795	0.0155	0.000	0.0008	0.0002	0.000

Table 17.6: Results of simple regression versus ARIMA(0,1,2) model of CNW count

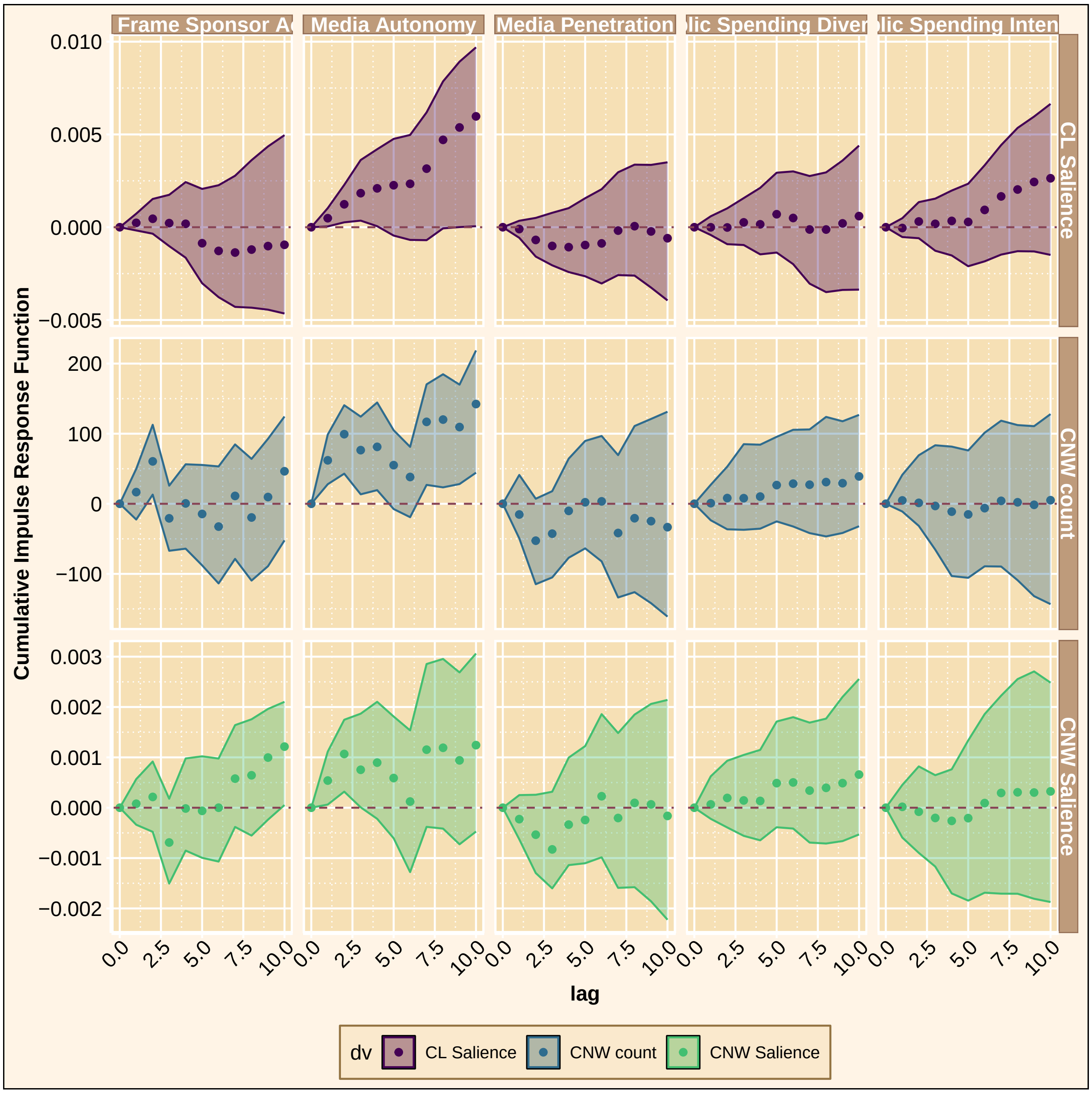
	Estimate	SE	P	ARIMAEstimate	ArimaSE	ArimaP
(Intercept)	0.7159	0.1321	0.000	0.7363	0.1320	0.000
Media Penetration	-0.5779	0.4892	0.239	-0.5269	0.4910	0.284
Media Autonomy	-4.7842	1.3496	0.000	-5.0210	1.3669	0.000
Public Spending Intensity	-1.3011	0.8907	0.145	-1.3704	0.8947	0.127
Public Spending Diversity	0.8484	1.0639	0.426	0.8348	1.0636	0.433
Crisis Frame Sponsor Activity	0.3735	0.0335	0.000	0.3706	0.0339	0.000

17.3 Vector autoregression models, Granger causality, and cumulative impulse-response functions

Table 17.7: Granger causality tests based on VAR models (all time lags <11 years)

Dependent variables	Independent variables	P values IV to DV	DV to IV	Simultaneous
	Media Autonomy	<.001***	<.001***	.959
	Media Penetr.	.050*	.402	.203
CL Salience	Publ. Spend. Int.	.186	.542	.575
	Publ. Spend. Div.	.011*	.066#	.065#
	Cr. Fr. Sp. Act.	.101	<.001***	<.001***
	Media Autonomy	<.001***	.565	<.001***
	Media Penetr.	.004**	.048*	.770
CNW Salience	Publ. Spend. Int.	.978	.569	.711
	Publ. Spend. Div.	.478	.503	.571
	Cr. Fr. Sp. Act.	<.001***	.052#	<.001***
	Media Autonomy	<.001***	.595	.604
	Media Penetr.	.011*	<.001***	.601
CNW Count	Publ. Spend. Int.	.996	.684	.903
	Publ. Spend. Div.	.844	.977	.647
	Cr. Fr. Sp. Act.	<.001***	<.001***	<.001***

Figure 17.3: Impulse–response functions based on Vector Autoregression Models



Chapter 18

Appendix S: Sensitivity analysis of mediation models

18.1 Purpose

This appendix documents the sensitivity analyses for the mediation models used to test hypotheses 3.1 and 3.2 in the main study.

18.2 Procedure

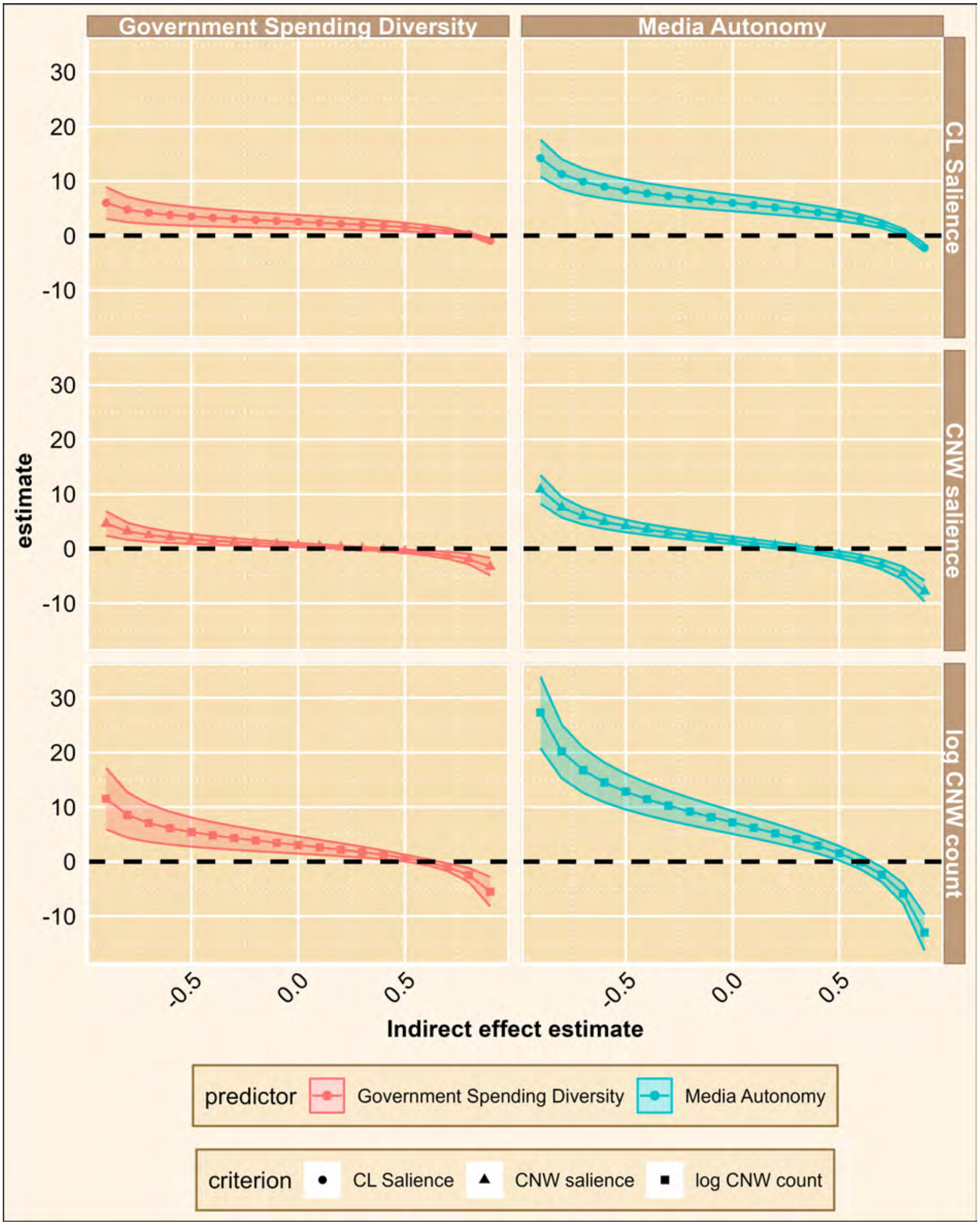
The test was conducted using the `mediation` package for `R` (Imai, Keele, and Yamamoto 2010). The sensitivity analysis tests hypothetical scenarios of how closely the mediator or the outcome is correlated with a hypothesized confounding variable. This is represented by the sensitivity parameter ρ on the x axis of the sensitivity analysis plots Figure 18.1.

18.3 Results and Conclusions

The indirect effects estimate (depicted on the y axis, along with confidence bands) are highly robust and signal a substantial indirect effect unless one presumes an extremely strong unobserved confounding variable affects both the mediator and the outcome. The likelihood of a spurious correlation is very low.

Note. Sensitivity analysis of mediation models.

Figure 18.1: Sensitivity analysis of mediation models



Chapter 19

Appendix T: Comparing the strength of effects on CL salience vs CNW salience

19.1 Purpose

This appendix provides details on the computation of standardization procedures used to ask the question whether effect sizes for CL salience are greater than effect sizes for CNW salience. It also includes the results in tabular form (which are presented as a figure in the main study).

19.2 Procedure

The effect of time on CL salience and CNW salience can appear different in size depending on which procedure one uses for standardizing the raw regression coefficients and make them comparable. Since CL salience has a greater average value ($M = 1.38\%$) than CNW salience ($M = 0.16\%$) and a greater variation ($SD = 1.01$ vs $SD = 0.38$), the raw regression coefficients may be exaggerating times' impact on CL salience relative to CNW salience; CNW salience is generally lower and varies less. This effect is obviously removed when applying z-standardization on CNW salience and CL salience. By substracting the average from the observed values and dividing the result by the standard deviation, means and standard deviations are the same for both variables. However, this may be an overcorrection that exaggerates times' impact on CNW salience relative to CL salience. As a middle ground, we standardized CL salience and CNW salience according to the Min-Max such that the lowest and highest observed values are defined as 0 or 1, respectively

We triangulate this question by using three different standardization techniques to compare regression coefficients across two different dependent variables: unstandardized coefficients (B_U, may exaggerate effects on CL salience relative to CNW salience), z-standardized coefficients (β_z, may exaggerate effects on CNW salience relative to CL salience), Min-Max-standardized (β_{MM}, intermediate solution).

19.3 Results

The results indicate that all effects on CL salience are stronger than those on CNW salience if unstandardized or Min-Max standardized coefficients are used. This also applies to z-standardized coefficients, with the exception of

Table 19.1: Comparing the strength of effects on CL salience and CNW salience. The results consistently show that effects on CL are substantially greater than on CNW (ratio>1)

Driver		Regression coefficient standardization			Coefficient of determination	
Predictor	DV	$B_{Unstd.}$	$\beta_{MinMax-std.}$	$\beta_{z-std.}$	η^2	$\eta^2 \div R^2$
Time	CL	1.248	0.236	1.234	.710	.779
		[1.145; 1.350]	[0.216; 0.255]	[1.133; 1.336]		
	CNW	0.146	0.044	0.381	.068	.337
		[0.076; 0.216]	[0.023; 0.065]	[0.199; 0.563]		
	Ratio	8.548^*	5.364^*	3.239^*	10.44	2.312
Frame spon.	CL	31.12	5.877	30.79	.694	.762
		[28.44; 33.81]	[5.370; 6.385]	[28.13; 33.44]		
	CNW	7.949	2.398	20.72	.103	.510
		[4.898; 11.00]	[1.478; 3.318]	[12.77; 28.67]		
	Ratio	3.915 *	2.451 *	1.486	6.738	1.494
Penetration	CL	120.3	22.72	119.0	.337	.370
		[98.53; 142.1]	[18.61; 26.83]	[97.47; 140.6]		
	CNW	3.411	1.029	8.891	.001	.005
		[-10.21; 17.03]	[-3.081; 5.139]	[-26.62; 44.40]		
	Ratio	35.27 *	22.08 *	13.39 *	337	74
Autonomy	CL	852.6	161.0	843.3	.268	.294
		[670.6; 1034.5]	[126.6; 195.4]	[663.4; 1023.3]		
	CNW	220.1	66.41	573.8	.059	.292
		[106.3; 333.9]	[32.08; 100.7]	[277.2; 870.4]		
	Ratio	3.873 *	2.424 *	1.470	4.542	1.01
Spending	CL	163.2	30.81	161.4	.077	.085
		[90.18; 236.2]	[17.03; 44.60]	[89.20; 233.6]		
	CNW	-11.12	-3.354	-28.98	.001	.005
		[-53.15; 30.92]	[-16.04; 9.328]	[-138.5; 80.6]		
	Ratio	-14.68 *	-9.187 *	-5.570 *	77	17
Diversity	CL	513.3	96.94	507.8	.372	.408
		[427.2; 599.5]	[80.68; 113.2]	[422.6; 593.0]		
	CNW	83.96	25.33	218.9	.046	.228
		[34.36; 133.56]	[10.37; 40.29]	[89.56; 348.1]		
	Ratio	6.114 *	3.827 *	2.320 *	8.087	1.789

Chapter 20

Appendix U: Predictive Performance when Ignoring time and ARIMA structure

20.1 Purpose and Method

This appendix provides illustrations of the predictive performance of the predictive models with various training and heldout periods Table 20.1. Each model series includes one, two, or all three of the following components: Time as a predictor (including ARIMA-modeling), text-external predictors (public spending and media logic indicators), and/or text-internal predictors (crisis-related organization density of coverage) Table 20.2.

Table 20.1: Overview over training and heldout period constellations

Constellation	Fitting/training period	Heldout/prediction period
A	1785–1899	1900–2020
B	1785–1949	1950–2020
C	1785–1974	1975–2020
D	1785–1989	1990–2020
E	1785–1999	2000–2020
F	1901–2020	1785–1900

Table 20.2: Model overview

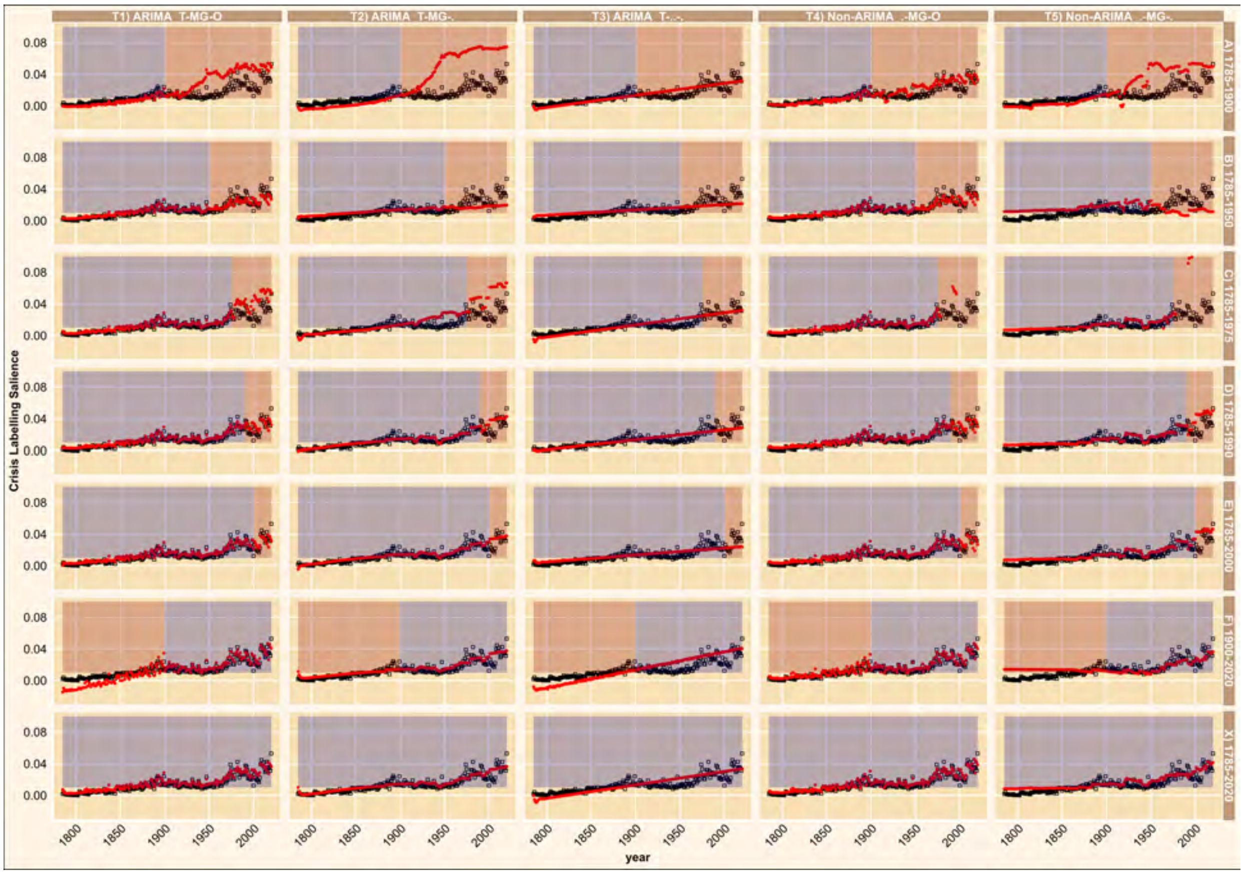
Model	ARIMA	Time	Mediatization and public Spending	Frame sponsors
T1[T.MG.O]	Yes	Yes	Yes	Yes
T2[T.MG./]	Yes	Yes	Yes	No
T3[T.//./]	Yes	Yes	No	No
T4[/.MG.O]	No	No	Yes	Yes
T5[/.MG./]	No	No	Yes	No

20.2 Predictive performance in different time-based, driver-based, and mixed models

Table 20.3: Model forecast performance in heldout period

Training period	Std. root-mean-square error [0–100] (rank) A 1785–1900	B 1785–1950	C 1785–1975	D 1785–1990	E 1785–2000	F 1900–2020	X 1785–2020
T1[T.MG.O]	4.2 (3)	1.4 (2)	3.7 (2)	1.7 (1)	1.7 (1)	1.9 (4)	0.8 (2)
T2[T.MG./]	8.6 (5)	2.7 (4)	5.6 (3)	2.5 (4)	2.3 (3)	0.7 (2)	1.0 (3)
T3[T.//./]	1.6 (1)	2.3 (3)	1.8 (1)	2.1 (3)	2.8 (4)	1.7 (3)	1.4 (5)
T4[/.MG.O]	1.7 (2)	1.3 (1)	36.0 (4)	1.8 (2)	2.0 (2)	0.7 (1)	0.8 (1)
T5[/.MG./]	5.3 (4)	3.8 (5)	100.0 (5)	3.5 (5)	3.5 (5)	1.9 (5)	1.3 (4)

Figure 20.1: Prediction of crisis labelling salience based time series models



Note. Blue periods are used for training the models. Red periods are held out to test the prediction performance. T1) ARIMA model with time (T), media logic (M), public spending (G), and organization density (O) as predictors. T2) ARIMA model with TMG components. T3) ARIMA model with T component only. T4) Non-ARIMA model with M, G, and O components. T5) Non-ARIMA model with M and G component only.

Adding time to the models M1.5, M2.5 and M3.5 (resulting in M1.6, M2.6, M3.6) only results in significant model improvement in M3.6 (logged CNW count) but not M1.6 (CL salience) and M2.6 (CNW salience). This suggests that our predictors can explain the over-time development to a point that we no longer need to know the time point at which an observation occurred to predict the level of crisis salience in news coverage.

We further investigate this idea—can we simply ignore the time if we know the current value of the driving factors?—by looking at five sets of predictive models of CL salience (we focus on this dependent variable for this exploration) that are trained on different time periods:

- T1: ARIMA model with all predictors (media logic, public spending, organizations, time)
- T2: Restricted ARIMA model (only media logic, public spending, time)
- T3: Pure ARIMA model (only time)
- T4: Non-ARIMA model with all predictors (media logic, public spending, organizations)
- T5: Restricted non-ARIMA model (only media logic, public spending).

The different time periods for training the models are: (A) 1785-1899 (early/shortest training), (B) 1785-1949 (early/short training), (C) 1785-1974 (early/moderate training), (D) 1785-1989 (early/long training), (E) 1785-1999 (early/longest training), and (F) 1901-2020 (late training). The corresponding heldout periods in which the trained models' predictive capacities are tested are 1900-2020, 1950-2020, 1975-2020, 1990-2020, 2000-2020, and 1785-1900 Table 20.1.

This analysis allows to check if models deteriorate if they are trained only to a subset of the time period (suggesting that the mechanism they model is not consistent over time) or that omitted some predictors leads to specific patterns of deterioration Table 20.3, Figure 20.1.

What we see is that T3 only predicts a linear trend with some noise but fails to capture the systematic ups and downs in CL salience—this also makes predictions fairly robust given the overall continuity of the time series.

T5 requires long training periods for proper predictive performance (good performance: D, E; bad performance: A, B, C, F) while T4 performs very well even with relatively short training periods (good performance: A, B, D, E, F), and is only confused if the period 1951-1975 is included in the training period as here an additional mechanism—the impact of media autonomy—emerges but the model requires more data points with high media autonomy to properly represent its impact.

T1, the most complete model, is very robust (good performance: B, D, E) but shows some weaknesses if training periods are relatively short (A, F) or stop during the transition period 1951-1975 (C).

T2, is much less robust, with good predictions in D, E, and F but problematic predictions in A, B, C. Overall, this suggests that high-performance predictions do not require knowledge of the time point (as in T1-T3) but that knowing about pervasiveness of media logic, structure of public spending, and organization density in coverage (T1 and T4) is necessary to arrive at high-performance predictions.

The models that (a) do not use ARIMA-modelling (T4 and T5) and (b) are trained in a period ends in the mid-1970s (1785-1975) show an interesting and revealing lapse: The estimates for the 1975-2020 period suddenly become extremely inaccurate in these models: they predict way too high levels of CL salience. We presume that in the 1950-1975 period a change in the mechanisms behind CL salience took place and the model did not yet have enough data and enough variation to properly model these developments. T1-T3 do not react as sensitively as the ARIMA part of the equation limits sudden jumps in predicted values and pushes towards continuity of the time series.

Chapter 21

Appendix V: Check for potential circularity of crisis frame sponsor measurement

21.1 Purpose

This appendix serves to discuss a potential circularity problem in the construction of the crisis frame sponsor activity indicator used in the study and the safety measures to minimize this problem.

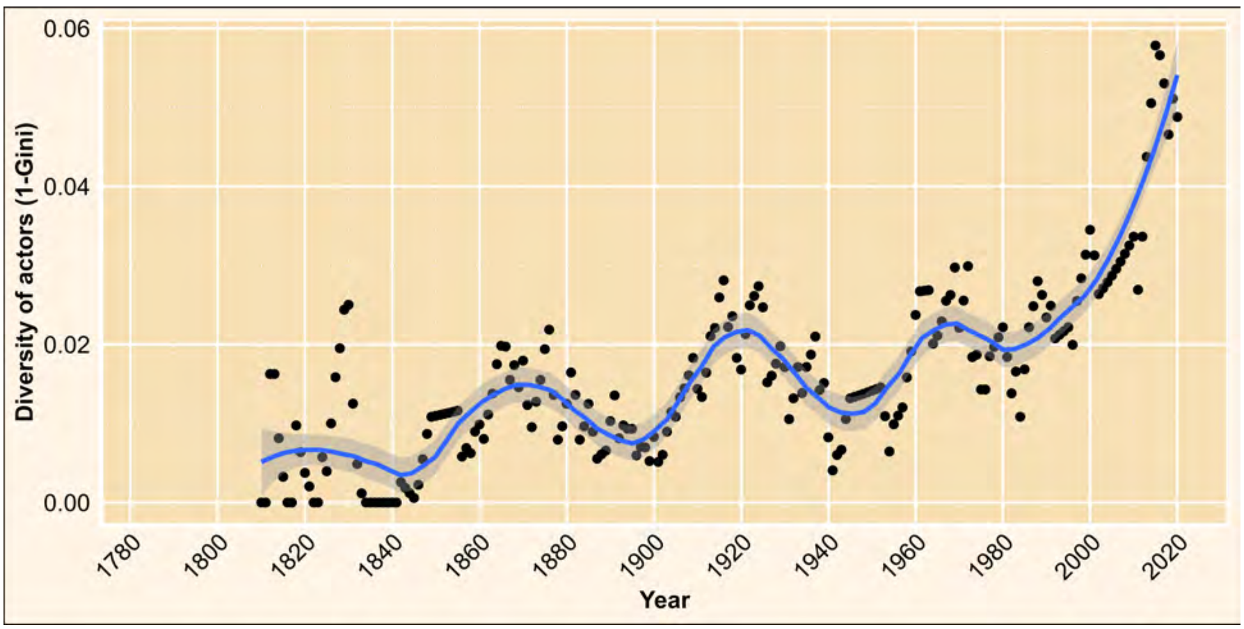
21.2 Problem statement

There is some circularity or co-dependency concerning the dependent variables CL salience and CNW salience on the one hand and the independent variable crisis frame sponsor density on the other hand. Both sets of variables depend on the number of articles with crisis labelling (otherwise, there can be no CL, no CNW, and no crisis-related mention of an organization). This means that, e.g., CL salience would produce a higher likelihood for more unique organizations to be mentioned in a crisis context, and a positive correlation might be spurious.

21.3 Problem minimization

We try to minimize that risk by checking which organizations were mentioned and which societal areas they belong to. This manual classification of a random selection of 100 organizations per 50-year period (1800-1849, 1850-1899,1900-1949,1950-1999,2000-2020) shows the thematic diversification of the organizations, with diversity scores increasing from period to period (0.07 – 0.13 – 0.18 – 0.28). Furthermore, the literature on the development of PR (Miller and Dinan 2000, 2008; L'Etang 2004) and lobbying (McGrath 2018; Jordan et al. 2012) in the UK match the timeline our data sketches, though the newspaper data provides much greater temporal resolution. This validates our measure of crisis frame sponsor activity and minimizes the risk of spuriousness or relationships.

Figure 21.1: Diversity of organizations mentioned in crisis coverage 1785–2020



Chapter 22

References

Baumgartner, Frank R., Bryan D. Jones, and John D. Wilkerson. 2002. "Studying Policy Dynamics." In *Policy Dynamics*, edited by Frank R. Baumgartner and Bryan D. Jones, 29–46. Chicago, IL: University of Chicago Press.

Bennett, W. Lance, Regina G. Lawrence, and Steven Livingston. 2007. *When the Press Fails: Political Power and the News Media from Iraq to Katrina*. Chicago, IL: University of Chicago Press.

Benoit, Kenneth, and Akitaka Matsuo. 2020. "Spacyr: Wrapper to the spaCy NLP Library."

Biggs, D, R. Biggs, V. Dakos, R. J. Scholes, and M. Schoon. 2011. "Are We Entering an Era of Concatenated Global Crises?" *Ecology and Society* 16 (2): art. 27.

Brecke, Peter. 2021. "Conflict Catalogue."

Butler, David, and Gareth Butler. 1994. *British Political Facts, 1900 - 1994*. 7. ed. Basingstoke, Hampshire: Macmillan.

"Cabinets and Administrations." 2005. In *The Routledge Companion to Britain in the Nineteenth Century, 1815-1914*. London: Routledge.

Carpentier, Vincent. 2018. "Expansion and Differentiation in Higher Education: The Historical Trajectories of the UK, the USA and France." Working {{Paper}} 33. Centre for Global Higher Education.

Centre for research on the Epidemiology of Disasters. 2021. "EM-DAT: The International Disaster Database." https://www.emdat.be/.

Chantrill, Christopher. 2023. "UK Public Spending."

"Chronology of Major Cabinet Changes." 2005. In *The Routledge Companion to Britain in the Nineteenth Century, 1815-1914*. London: Routledge.

Clark, Gregory. 2005. "The Condition of the Working Class in England, 1209–2004." *Journal of Political Economy* 113 (6): 1307–40. https://doi.org/10.1086/498123.

Cook, Chris, and Brendan Keith. 1984. *British Historical Facts 1830-1900*. Palgrave Macmillan UK.

Cook, Chris, and John Stevenson. 1980. *British Historical Facts, 1760-1830*. London: Macmillan.

Davenport, Romola. 2021. "Patterns of Death, 1800–2020: Global Rates and Causes." In *The Routledge History of Death Since 1800*, edited by Peter N. Strearns, 21–44. Routledge.

Duch, Raymond M., and Michaell A. Taylor. 1993. "Postmaterialism and the Economic Condition." *American Journal of Political Science* 37 (3): 747. https://doi.org/10.2307/2111573.

Estes, Carroll L. 1983. "Social Security: The Social Construction of a Crisis." *The Milbank Memorial Fund Quarterly. Health and Society* 61 (3): 445. https://doi.org/dc9hx8.

Goulbourne, Harry. 1991. *Ethnicity and Nationalism in Post-Imperial Britain*. Comparative Ethnic and Race Relations. Cambridge ; New York: Cambridge University Press.

Habermas, Jürgen. 1973. "What Does a Crisis Mean Today?" *Social Research* 40: 643–67. https://www.jstor.org/stable/40970159.

Haferkamp, Hans, and Neil J. Smelser, eds. 1992. *Social Change and Modernity*. Berkeley: Univ. of California Press.

Hay, Colin. 2016. "Good in a Crisis: The Ontological Institutionalism of Social Constructivism." *New Political Economy* 21 (6): 520–35. https://doi.org/ghmh4t.

Hidalgo, C. A., and R. Hausmann. 2009. “The Building Blocks of Economic Complexity.” *Proceedings of the National Academy of Sciences* 106 (26): 10570–75. https://doi.org/10.1073/pnas.0900943106.

Hilgartner, Stephen, and Charles L. Bosk. 1988. “The Rise and Fall of Social Problems: A Public Arenas Model.” *American Journal of Sociology* 94: 53–78. https://doi.org/br5nc2.

Hudson, Pat. 2014. *The Industrial Revolution*. Bloomsbury.

Imai, Kosuke, Luke Keele, and Teppei Yamamoto. 2010. “Identification, Inference and Sensitivity Analysis for Causal Mediation Effects.” *Statistical Science* 25 (1): 51–71. https://doi.org/10.1214/10-STS321.

Inglehart, Ronald, and Scott C. Flanagan. 1987. “Value Change in Industrial Societies.” *American Political Science Review* 81 (4): 1289–319. https://doi.org/10.2307/1962590.

Isik, Filiz. 2010. “An Entropy-Based Approach for Measuring Complexity in Supply Chains.” *International Journal of Production Research* 48 (12): 3681–96. https://doi.org/10.1080/00207540902810593.

James, Patrick. 1987. “Externalization of Conflict: Testing a Crisis-Based Model.” *Canadian Journal of Political Science* 20 (3): 573–98. https://doi.org/dd9cvq.

Jordan, Grant, Frank R. Baumgartner, John D. McCarthy, Shaun Bevan, and Jamie Greenan. 2012. “Tracking Interest Group Populations in the US and the UK.” In *The Scale of Interest Organization in Democratic Politics*, edited by Darren Halpin and Grant Jordan, 141–60. London: Palgrave Macmillan UK. https://doi.org/10.1057/9780230359239_7.

Kaukiainen, Y. 2001. “Shrinking the World: Improvements in the Speed of Information Transmission, c. 1820-1870.” *European Review of Economic History* 5 (1): 1–28. https://doi.org/10.1017/S1361491601000016.

L’Etang, Jacquie. 2004. *Public Relations in Britain: A History of Professional Practice in the Twentieth Century*. 1st ed. Routledge. https://doi.org/10.4324/9781410610812.

Lichtenstein, Sarah, Paul Slovic, Baruch Fischhoff, Mark Layman, and Barbara Combs. 1978. “Judged Frequency of Lethal Events.” *Journal of Experimental Psychology: Human Learning and Memory* 4 (6): 551–78.

Majima, Shinobu, and Mike Savage. 2007. “Have There Been Culture Shifts in Britain?: A Critical Encounter with Ronald Inglehart.” *Cultural Sociology* 1 (3): 293–315. https://doi.org/10.1177/1749975507082050.

McBeth, Mark K., Randy S. Clemons, Maria A. Husmann, Elizabeth Kusko, and Alethea Gaarden. 2013. “The Social Construction of a Crisis: Policy Narratives and Contemporary U.S. Obesity Policy: The Social Construction of a Crisis.” *Risk, Hazards & Crisis in Public Policy* 4 (3): 135–63. https://doi.org/ghmh4z.

McGrath, Conor. 2018. “British Lobbying in Newspaper and Parliamentary Discourse, 1800-1950: British Lobbying in Newspaper and Parliamentary Discourse, 1800-1950.” *Parliamentary History* 37 (2): 226–49. https://doi.org/10.1111/1750-0206.12363.

Miller, David, and William Dinan. 2000. “The Rise of the PR Industry in Britain, 1979-98.” *European Journal of Communication* 15 (1): 5–35. https://doi.org/10.1177/0267323100015001001.

———. 2008. *A Century of Spin: How Public Relations Became the Cutting Edge of Corporate Power*. London ; Ann Arbor, MI: Pluto Press.

Mitchell, Brian R. 2011. *British Historical Statistics*. Cambridge: Cambridge University Press.

O’Brien, Patrick Karl. 1996. “Path Dependency, or Why Britain Became an Industrialized and Urbanized Economy Long Before France.” *The Economic History Review* 49 (2): 213. https://doi.org/10.2307/2597914.

Office for National Statistics. 2015. “How Has Life Expectancy Changed over Time?”

Oneal, John R., and Anna Lillian Bryan. 1995. “The Rally ’Round the Flag Effect in U.S. Foreign Policy Crises, 1950-1985.” *Political Behavior* 17 (4): 379–401. https://doi.org/d6f646.

Ossewaarde, Marinus. 2018. “‘Crises of Modernity’ Discourses and the Rise of Financial Technologies in a Contested Mechanized World.” *Philosophy & Technology* 31 (1): 59–76. https://doi.org/10.1007/s13347-017-0255-5.

Roberts, Margaret E., Brandon M. Stewart, and Dustin Tingley. 2019. “Stm: An R Package for Structural Topic Models.” *Journal of Statistical Software* 91 (2). https://doi.org/10.18637/jss.v091.i02.

Ryland, Thomas, and Samuel H. Williamson. 2021. “What Was the U.K. GDP Then?” *MeasuringWorth*.

Schneider, Saundra K. 2007-06-05/2007-06-07. *Flirting with Disaster: Public Management in Crisis Situations: 10th Annual Emergency Management Education Conference*. Emmitsburg, MD.

Slovic, Paul, Baruch Fischhoff, and Sarah Lichtenstein. 1986. “The Psychometric Study of Risk Perception.” In *Risk Evaluation and Management*, edited by Vincent T. Covello, Joshua Menkes, and Jeryl Mumpower, 3–24. Boston, MA: Springer US. https://doi.org/10.1007/978-1-4613-2103-3_1.

Smelser, Neil J. 1967. "Sociological History: The Industrial Revolution and the British Working-Class Family." *Journal of Social History* 1 (1): 17–35.

Soroka, Stuart N., and Christopher Wlezien. 2010. *Degrees of Democracy: Politics, Public Opinion, and Policy*. Cambridge, MA: Univerity Press.

Strömbäck, Jesper. 2008. "Four Phases of Mediatization: An Analysis of the Mediatization of Politics." *International Journal of Press/Politics* 13 (3): 228–46. https://doi.org/chvcmh.

Thomas, R., and N. Dimsdale. 2017. "A Millennium of UK Data: Bank of England OBRA Dataset."

Tuchman, Gaye. 1978. "The News Net." *Social Research* 45: 253–76.

Tversky, Amos, and Daniel Kahneman. 1981. "The Framing of Decisions and the Psychology of Choice." *Science* 211 (4481): 453–58. https://doi.org/fj3z3r.

V-Dem Institute. 2022. "Democracy Report 2022: Autocratization Changing Nature?" Gothenburg: V-DEM Institute.

Watanabe, Kohei. 2021. "Latent Semantic Scaling: A Semisupervised Text Analysis Technique for New Domains and Languages." *Communication Methods and Measures* 15 (2): 81–102. https://doi.org/10.1080/19312458.2020.1832976.