Three centuries of Swiss economic sentiment*

Marc Burri†

November 19, 2024

Abstract

There is a lack of consistent and well-measured Swiss business cycle indicators over long historical episodes. This paper fills this gap by constructing a business cycle indicator on quarterly frequency spanning from 1820 to 2021. Using textual data such as historical company records, newspapers, and business association reports, I develop a business cycle indicator, drawing on sentiment and count-based measures related to key economic concepts. This approach involves extensive data collection, surpassing existing datasets in scope and historical coverage. The composite indicator demonstrates strong correlations with real economic activity, effectively capturing historical downturns and expansions. I also employ it to estimate recession probabilities, shedding light on Switzerland's business cycle history, and ultimately, establish the first business cycle dating for Switzerland in the 19th and early 20th centuries.

Keywords: 19th century, 20th century, business cycles, Switzerland, qualitative coincident indicators, textual analysis, sentiment analysis

*For helpful comments and discussions, I thank David Ardia, Elliott Ash, Jean-Marie Grether, Martin Huber, Malin Jensen, Daniel Kaufmann, Bruno Lanz, Jason Lennard, Jannis Stefanopulos, Christian Stohr, Rebecca Stuart, Leif Anders Thorsrud, Philipp Wegmüller, Mark W. Watson, and participants at the IRENE PhD Meeting in Neuchâtel, the PhD Macro Workshop in Basel, the SSES Annual Congress in Lucerne, the IAAE Conference in Thessaloniki, and the SECO Brown Bag Seminar. I am thankful to Regina Gloor (Tamedia), Patrick Halbeisen (SNB), Martin Lüpold (SWA), Théophile Naito (scriptorium.bcu-lausanne.ch), Jürg Rütimann (AWP), Florian Steffen (e-newspaperarchives.ch) and Mathias Wiesmann (ZKB) for providing data. Nicola Francescutto and Idy Abdoul N'Dao provided excellent research assistance. I gratefully acknowledge the support from a UniNE Doc.Mobility grant to visit the Norwegian Business School (BI).

†University of Neuchâtel, Institute of Economic Research, Rue A.-L. Breguet 2, CH-2000 Neuchâtel, marc.burri@unine.ch

1 Introduction

"The farther backward you can look, the farther forward you are likely to see."

— Winston Churchill

Churchill's insight resonates profoundly with the objective of this chapter, which takes on the challenge of uncovering Switzerland's business cycle history. A story not entirely told due to a lack of well-measured economic data for the 19th and early 20th centuries. In the first instance, consistent and well-measured gross domestic product (GDP) statistics, which are often used to describe business cycle fluctuations, are missing. There is only a tentative estimate of annual GDP since 1851 and no information before that (see Historische Statistik der Schweiz HSSO, 2012k; Stohr, 2016). Moreover, real GDP estimates suffer from error-prone deflators that bias econometric estimates and descriptive statistics (Kaufmann, 2020). The lack of well-measured historical economic data is not unique to Switzerland.

Because hard data is difficult to measure retrospectively, I use written narratives in this chapter to fill in the gaps where traditional economic data is missing or inaccurate. Using textual data, I develop a quarterly business cycle indicator for Switzerland spanning three centuries: the 19th, 20th, and 21st. Given the impossibility of asking people and businesses from the past about their financial and economic situation, I rely on alternative sources like business association reports, business reports, and newspapers. Following recent contributions, I construct qualitative indicators for several economic topics related to the business cycle (see e.g. Chapter **??**, Thorsrud (2020) Bybee et al. (2023)). I assemble and digitize a unique dataset of historical documents relevant to business cycle fluctuations to accomplish this. One of the challenges associated with historical text sources is that they are very noisy, primarily due to issues related to quality. Therefore, I develop methods beyond natural language processing (NLP) techniques to extract relevant information. Finally, I aggregate the topic-specific indicators into a composite business cycle indicator.

The indicator effectively tracks major economic downturns, including the recessions of the late 20th and early 21st centuries and crises like the two World Wars and the Great Depression. Moreover, the indicator also sheds light on lesser-known economic disturbances in the 19th century, such as the Sonderbund War and the Franco-Prussian War. The indicator's ability to reflect business cycle fluctuations in Switzerland, particularly in recent decades, validates its accuracy and underscores its utility for economic analysis.

Furthermore, the chapter provides a first business cycle chronology for Switzerland in the 19th and early 20th centuries, a period previously lacking systematic economic analysis. My research reveals that Swiss recessions have become less frequent, aligning with broader European economic trends (Broadberry & Lennard, 2023). However, contrary to the European trend, I do not find that the duration of Swiss recessions exhibits a significant decrease over the sample periods (Broadberry & Lennard, 2023). This finding challenges some prevailing narratives about the nature of business cycles.

The foundational work of Thorp (1926) and Burns and Mitchell (1946) serves as a basis for my approach of using textual data to create a business cycle indicator and chronology. Thorp (1926) used narrative accounts to describe the business cycle, highlighting the importance of qualitative data when hard data is missing. Burns and Mitchell (1946) further developed this field by introducing a more systematic and empirical approach to identifying business cycles, emphasizing the analysis of a wide range of economic indicators. This chapter bridges these methodologies by integrating advanced computational techniques to convert qualitative textual records into a quantifiable time series akin to Burns and Mitchell's (1946) methodology. An advantage of my approach is that it involves less judgment and more data-driven analysis, making it more objective and replicable without reading the documents. Moreover, it can be applied consistently across different periods.[1]

More recently, Shiller (2017, 2019) highlighted that stories and public discourse shape economic trends and, possibly, cause economic fluctuations. Narratives, often rooted in credible business and media sources, serve dual purposes: they reflect the economic conditions of their time and have the potential to influence future economic decisions and policies. Shiller's findings about the importance of stories highlight how written information can offer valuable economic insights, suggesting its broad applicability in various economic studies.

The chapter is related to a growing body of research that uses textual data to measure

[1]This contrasts the methodology used for determining NBER recession dates, which varies over time (see e.g. Romer, 1994; Romer & Romer, 2020).

economic activity and sentiment.[2] The study by Van Binsbergen et al. (2024), which develops a 170-year-long measure of economic sentiment for the US, is most closely aligned with this chapter in its historical approach. Similarly, Kabiri et al. (2023) developed a monthly sentiment index for the United States from 1920 to 1934. For more contemporary periods, Burri and Kaufmann (2020), Bybee et al. (2023), Larsen and Thorsrud (2019), Shapiro et al. (2022), and Thorsrud (2020) have utilized text mining and machine learning to analyze newspaper articles, creating leading indicators and business cycle measures. Complementing these efforts, studies by Ardia et al. (2019), Barbaglia et al. (2023), Burri (2023), Ellingsen et al. (2022), and Kalamara et al. (2022) demonstrate the effectiveness of textual data in predicting various economic variables.

I make several contributions to this body of research. First, while previous research primarily focused on contemporary data over short periods, this chapter applies similar methodologies to long historical episodes. I introduce a quarterly business cycle indicator that is the most extensive record of Swiss business cycle fluctuations. Second, I use a broad range of historical documents relevant to business cycle fluctuations, including business reports, association documents, and archival material, diverging from related studies that primarily rely on well-structured and often categorized newspaper articles. This unique and diverse textual dataset enriches the analysis and makes it more complex, thus requiring adequate methodology. Therefore, I present a method for extracting meaningful insights from noisy and heterogeneous historical text sources. Finally, I establish the first business cycle dating for Switzerland in the 19$^{\text{th}}$ and early 20$^{\text{th}}$ centuries.

The remainder of this chapter is organized as follows. In the next section, I describe the textual data in more detail. Section 3 explains the methodology to create the business cycle indicator. In section 4, I evaluate the indicators and discuss the business cycle chronology. Section 5 conducts a series of robustness checks. The last section concludes.

2 Data

This chapter aims to construct a Swiss business cycle indicator using textual data for the 19$^{\text{th}}$ and early 20$^{\text{th}}$ century, a period for which accurate hard data is necessarily scarce.

[2]Textual data are also used for a variety of other purposes in economics, such as predicting stock returns (Hanna et al., 2020; Hirshleifer et al., 2023; Tetlock, 2007), measuring economic policy uncertainty (Ardia et al., 2021; Baker et al., 2016; Larsen, 2021), identifying monetary policy shocks (Aruoba & Drechsel, 2024; Ter Ellen et al., 2022), and measuring inflation expectations (Angelico et al., 2022; Binder, 2016; Larsen et al., 2021). Ash and Hansen (2023) and Beach and Hanlon (2022) provide an extensive review.

Therefore, I collected and digitized many historical documents, including company records and business association reports that potentially comprise useful information on business cycle fluctuations.[3] They comprise written information on a company's individual performance or the economy's aggregate performance. I complement this information with digitized newspaper articles. Newspapers often write about the economy or related topics like the labor market. Therefore, they can provide valuable information on the business cycle. In total, I collected 106 sources in German and French language.[4] Table 1 provides an overview of the number of sources per type, language, and frequency.[5] Overall, there are more newspapers than reports available. French business reports are more numerous than German ones, and newspapers and business association reports are more numerous in German. Moreover, most sources are available at a higher frequency than quarterly.

Table 1 — Number of text sources per type, language, and frequency

	Newspapers	Business reports	Business association reports
German	44	8	7
French	28	16	3
Frequency	$> Q$	$< Q$	$< Q$

Notes: The table shows the number of sources per type, language, and frequency. $> Q$ means the sources are available at a higher frequency than quarterly, and $< Q$ means the sources are available at a lower frequency than quarterly.

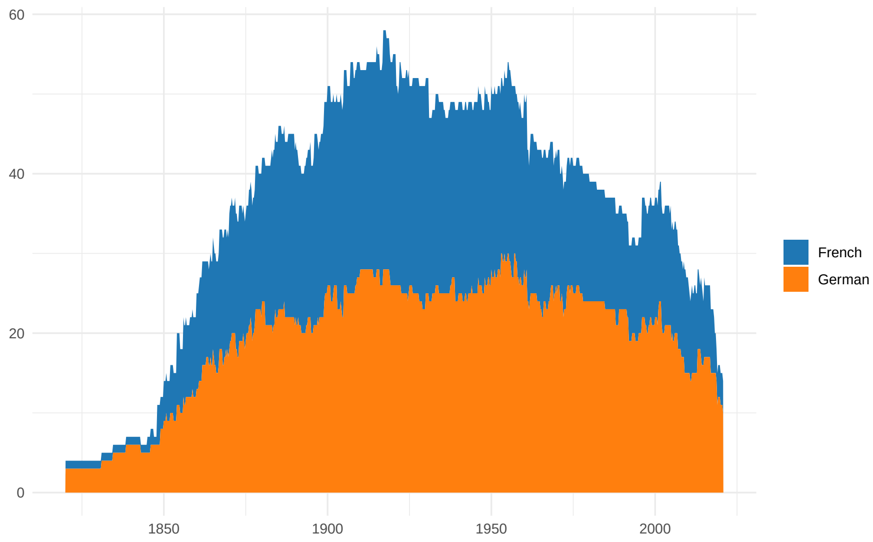
Figure 1 provides information on how many sources are available for each period. The number of sources increases over time, especially after 1850. It reaches its peak around 1925 and then decreases again. Due to the historical focus of this project, I did not include more recent sources in the analysis. The Figure shows that, on average, around 40 sources are available. Moreover, the data is balanced between French and German sources. The number of sources before 1850 is below 10. However, this should not be a problem as two important Swiss newspapers are available for this period. Namely, the *Neue Zürcher Zeitung* (NZZ) from the German part of Switzerland and *L'Express* from the French part of Switzerland.

[3] See table 1 in the Appendix for a comprehensive list of all data sources. Appendix B provides a detailed description of the data collection process.

[4] Even though it is one of the Swiss national languages, I did not include textual data in Italian in the analysis. The reason is that there are too few sources available.

[5] A source is the same publication available over a more extended period. This can be a newspaper (e.g., NZZ, daily 1820 - 2020) or a business report (e.g., Credit Suisse, annually 1895 - 2016)

Figure 1 — Number of sources



Notes: The graph shows the number of available sources per quarter for both languages, French and German.

3 Development of a historical business cycle indicator

In this section, I describe the methodology used to extract information from text, create topic-specific indicators, and ultimately create a composite indicator of the Swiss business cycle. The method of constructing the business cycle indicator combines tools from several fields. First, I use natural language processing (NLP) techniques to extract information from the textual data. Second, I use time series econometrics methods to correct for data anomalies. Finally, I use machine learning techniques to combine the different indicators into a composite business cycle indicator. The mapping from the individual text sources to a business cycle indicator involves several steps. To provide an overview, I illustrate these steps in a schematic diagram in Figure 1 in the Appendix.

3.1 Extracting economic signals from text

The texts contain much information, such as advertisements, tables, and figures, irrelevant to the business cycle indicator. Therefore, I use layout parsing techniques to identify and remove these elements (see section B in the Appendix). Moreover, as is standard in the literature, I remove punctuation marks, numbers, and special characters like Extensible Markup Language (XML) and Hypertext Markup Language (HTML) tags. I also remove stopwords, that is, words that are not informative, typically conjunctions such as “or” and “if” (see, e.g. Thorsrud, 2020). The stop words for both French and German are provided by Feinerer and Hornik (2019). Moreover, for each text, I count terms appearing in a German-French lexicon available from dict.cc, containing roughly 60’000 words. If a text is supposed to be in French, I only keep it if 20% of its terms are included in the French lexicon and more French than German words are identified. This filters texts with poor Oprical Character Recognition (OCR) quality or those not in German or French.

Only a fraction of the remaining texts, however, contain information about the business cycle. In the literature, it is common to use a topic modeling algorithm to classify texts into topics to further refine the selection (see, e.g. Thorsrud, 2020). However, with the amount of data available for this study, it would require a lot of computational power. Therefore, I use a simple keyword-based method to create indicators for hand-selected economic topics (see Chapter **??**). This method creates indicators by using two different approaches. The first is a count-based approach, where the indicator is given by simply counting terms related to specific topics. And second, a keyword-in-context (KWIC) approach in which topics are defined by keywords and sentiment is extracted from a few words surrounding the keywords (Luhn, 1960).[6] I apply these two approaches to create indicators for each topic and each source.

The keywords for eleven topics (three count-based and eight KWIC-based) are selected by reading through business association reports for all available periods.[7] This is important, especially for the 19$^{\text{th}}$ century when language might have been very different from today and different terms described the same topics. Reading the texts makes sure I capture potentially changing language over time. The French keywords are obtained by translating the German keywords using ChatGPT and reading French texts (OpenAI, 2023a). Using these keywords, I create a new text corpus in which each document consists of a keyword and its 15 preceding and 15 following words.

The count-based indicator measures the number of documents associated with a specific topic. I use a straightforward methodology to calculate sentiment scores that resembles business cycle indicators based on firm surveys. Using existing dictionaries, I classify all words into positive, neutral, and negative words. Then, I compute the sentiment

[6]Some papers also refer to this method as aspect-based sentiment analysis (see e.g Barbaglia et al., 2023).

[7]The complete list of topics with keywords chosen is available in Tables 3 and 4 the Appendix. Moreover, I provide information about the number of identified keywords over time in Figure 6 in the Appendix.

score for each document, that is, the share of positive minus the share of negative words. Finally, I create quarterly (for sources with quarterly or higher publication frequency) and yearly (for sources with lower than quarterly publication frequency) indicators for every topic and source. The count-based indicators are calculated as the sum of all identified keywords. The KWIC-based indicators are calculated as the average sentiment score over all documents.

To define positive and negative words, I follow Shapiro et al. (2022) and combine existing dictionaries that are proven to capture economic sentiment. The German lexicon combines the dictionaries developed by Remus et al. (2010) (see, e.g. Burri & Kaufmann, 2020) and a translation of the Loughran and McDonald (2011) lexicon (see, e.g. Ardia et al., 2019). The French lexicon consists of translations of the same two dictionaries and the dictionary developed by Abdaoui et al. (2017).

More formally, let $\mathcal{K}_j$ be the list of keywords for topic j, $\mathcal{P}$ the list of positive and $\mathcal{N}$ the list of negative words. Then $w_{j,i,t,d} = (w_{j,i,t,d,-15}, w_{j,i,t,d,-14}, \ldots, w_{j,i,t,d,0} \in \mathcal{K}_j, \ldots, w_{j,i,t,d,15})$ denotes the list of terms in document d at date t for topic j and source i. The count-based indicators are then calculated as

$$s_{j,i,t} = \sum_{d} \mathbb{1}(w_{j,i,t,d,0} \in \mathcal{K}_j) \tag{1}$$

The document-level sentiment score is given by

$$s_{j,i,t,d} = \frac{\sum_n \mathbb{1}(w_{j,i,t,d,n} \in \mathcal{P}) - \sum_n \mathbb{1}(w_{j,i,t,d,n} \in \mathcal{N})}{|w_{j,i,t,d}|}. \tag{2}$$

Figure 5 in the Appendix provides a more intuitive example of calculating the document-level sentiment score. Finally, KWIC-based indicators, $s_{j,i,t}$, for a given topic j and source i are calculated as a simple average of the sentiment scores

$$s_{j,i,t} = \frac{\sum_d s_{j,i,t,d}}{|s_{j,i,t,d}|}. \tag{3}$$

3.2 Creating topic-specific indicators

The source-level indicators suffer from several deficiencies. First, the quality of the OCR often changes over time. This is particularly relevant for the 19th and early 20th centuries, for which the archives often received the records from various sources of varying quality. This can also be interpreted as a change in the measurement error – the higher the quality of the OCR, the lower the measurement error. Second, the publication frequencies of the sources change over time – increased publication frequency means a higher information density and, therefore, lower measurement error. Third, due to technological advancements, the length of the articles and reports tends to increase over time. Finally, some indicators contain missing values, and most sources are only available for a limited period. These points potentially lead to indicators with trends and structural breaks in mean and variance. To address these issues, a six-step procedure is proposed to rectify anomalies in the data.[8] A comprehensive description of the procedure is provided in Section C of the Appendix.

1. If the frequency is annual, interpolate missing observations using Stineman's method (Stineman, 1980). This ensures that the indicator can be temporally disaggregated to quarterly frequency in step 5.

2. Detrend the indicator using Locally Estimated Scatterplot Smoothing (LOESS) with a bandwidth of 0.7 (Cleveland, 1979).

3. Detect structural breaks in mean and variance using a binary segmentation algorithm (Killick & Eckley, 2014). Split the indicator into segments at the detected breakpoints.

4. Normalize each segment, that is subtract mean and divide by standard deviation.[9]

5. If the frequency is annual, temporally disaggregate the indicator to quarterly frequency (Dagum & Cholette, 2006).

6. Remove outliers (observations more than three standard deviations away from the mean).

[8] Of course, other possibilities exist to interpolate or detrend the data. For example, Schorfheide and Song (2015) use mixed-frequency VAR models to account for varying frequencies of the underlying data. Moreover, Canova (1994, 1998) show that business cycle facts vary widely between detrending methods. Exploring these possibilities is beyond the scope of this chapter, and I leave these possibilities for future research.

[9] I show in Appendix D that with time-varying measurement error, this is preferable to not normalizing.

Finally, I aggregate the corrected source-level indicators into an overall indicator for each topic. To effectively summarize the information content of the data and eliminate idiosyncratic noise, I estimate a factor model in static form. Therefore, the source-level indicators follow

$$s_{j,i,t} = \lambda_{j,i} f_{j,t} + e_{j,i,t} \tag{4}$$

where $s_{j,i,t}$ is the corrected source-level indicator for topic j and source i, $\lambda_{j,i}$ is the factor loading for source i and topic j, $f_{j,t}$ is the common factor for topic j at time t, and $e_{j,i,t}$ is the idiosyncratic component. Finally, I use the first principal component as the topic-specific indicator, $S_{j,t}$ for topic j.

The advantage of using a factor model is that it allows for summarizing the information in a large data matrix with a small number of common factors. As in Chapter **??**, factors and loadings can be estimated through principal components, under the assumption that the idiosyncratic components are only weakly serially and cross-sectionally correlated (Bai & Ng, 2013; Stock & Watson, 2002).[10]

3.3 Estimation of composite indicator

The final step is to combine the topic-specific indicators into one composite indicator for the business cycle. For this, I fit the topic-specific indicators to a measure of the business cycle on a very recent sample where it is widely acknowledged that the business cycle is captured well. Then, I use the estimated coefficients to backcast the business cycle. As the measure for the business cycle, I use the output gap.[11] The topic-specific indicators, as well as four lags and four leads, are used as predictors.

Following Bybee et al. (2023), shrinkage methods are employed to estimate the coefficients. Specifically, I use elastic net regression. A primary strength of an elastic net model lies in its ability to handle collinearity effectively. Traditional regression

[10] I interpolate missing values using an EM algorithm (Stock & Watson, 2002), after standardizing the data to have zero mean and unit variance. For interpolating the data, I use one factor ($r = 1$) as there were converging issues with $r > 1$. Finally, I use the first principal component of the interpolated data set. Using least squares instead of the EM algorithm, I estimate the factor model with missing values as a robustness test. Even though the method is simpler than the EM method, it struggled with the sparsity of the data.

[11] To create a quarterly output gap, I use Hamilton (2018) filtered real GDP estimates from the State Secretariat for Economic Affairs (SECO). I also use estimates of the output gap provided by the SECO and the SNB for robustness. The results are very similar.

Figure 2 — Quarterly output gap split into training and test set

Notes: The red line represents quarterly output gap estimated using the method proposed by Hamilton (2018). The training set is used to estimate elastic net models with different combinations of hyperparameters. The gray shaded area is the test set used to evaluate the models and select the best combination of hyperparameters. The blue line is the fitted output gap using the best model.

methods can struggle to provide reliable coefficient estimates in such cases. The elastic net, however, combines L1 (Lasso) and L2 (Ridge) regularization techniques, creating a balance that addresses collinearity and encourages sparsity in the model. This means it can automatically select a subset of important predictors while shrinking the coefficients of less important ones toward zero. By doing so, the elastic net enhances the interpretability of the model and improves its predictive performance by reducing overfitting. Before estimating the model, I normalize all variables. The objective function of the elastic net is given by

$$\text{Minimize:} \quad \frac{1}{2T_{Train}} \sum_{t=1990Q1}^{T} (y_t - S'_{j,t}\beta)^2 + \lambda \sum_{k=1}^{p} \left(\alpha|\beta_k| + \frac{1-\alpha}{2}\beta_k^2 \right) \tag{5}$$

where T_{Train} is the number of observations in the training set, T is the date of the last observation in the data, p the number of predictors, y_t is the output gap, β is the vector of coefficients, $S_{j,t}$ is the vector of topic-specific indicators, λ is the penalty parameter, and α is the mixing parameter. The mixing parameter, α, controls the relative weight of the L1 and L2 penalties. When $\alpha = 0$, the penalty is an L2 penalty (ridge regression); when $\alpha = 1$, it is an L1 penalty (lasso regression). The elastic net penalty is a convex combination of the L1 and L2 penalties. Figure 2 illustrates how optimal values of λ and α are chosen. I perform a grid search over a range of λ and α values, estimate a model on a training set, and select the combination that minimizes the mean squared error (MSE) on a test set. The test set spans 1980 - 1989, and the training set 1990-2022. Using the best-performing model, the composite indicator is calculated as the fitted output gap from 1821 to 2021.

$$S_t = S'_{j,t}\hat{\beta}. \tag{6}$$

The blue line in Figure 2 shows the model fit on the training set and the prediction on the test set. The model performs reasonably well – the blue line lies close to the red line, which represents the actual output gap.

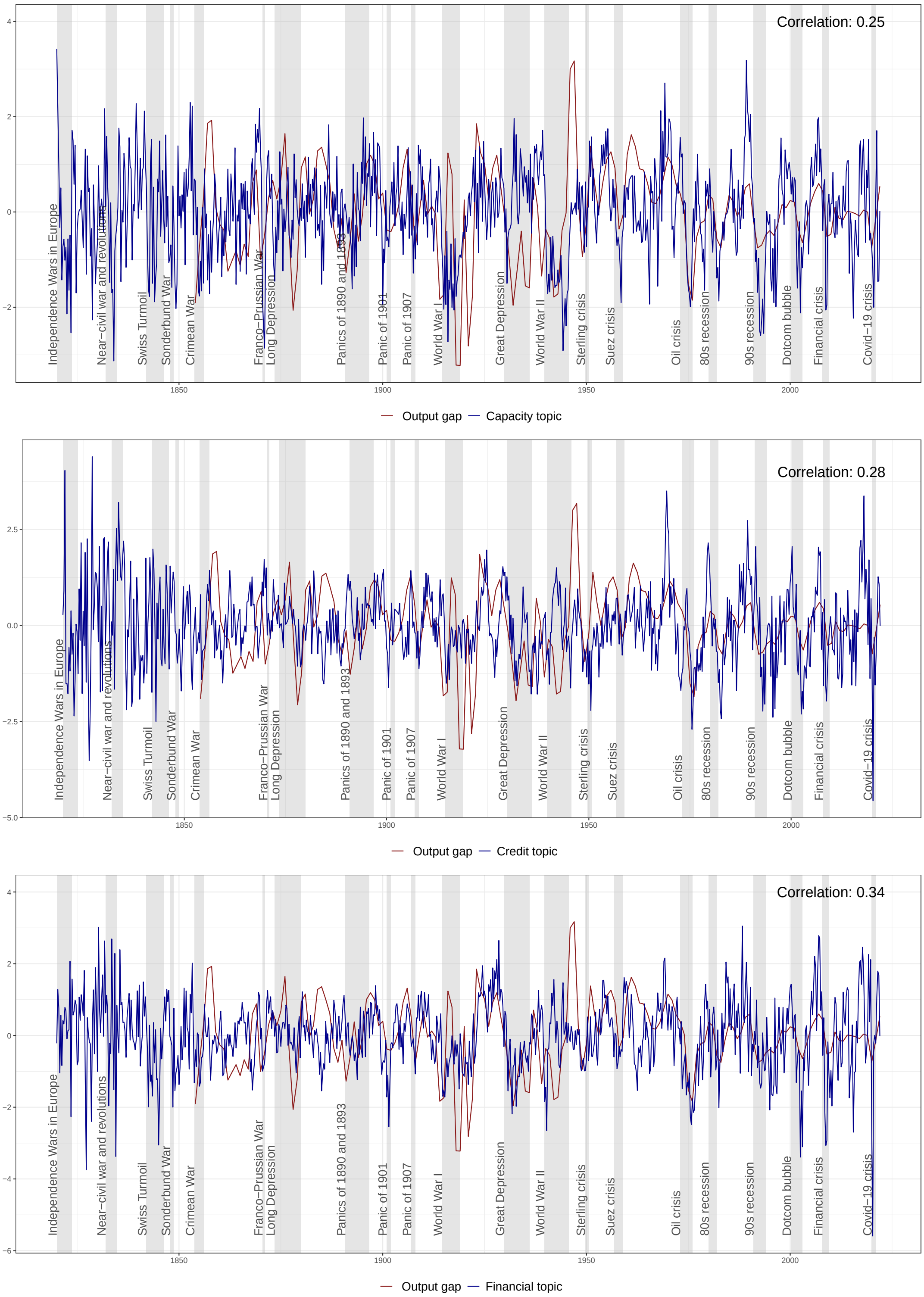
4 Evaluation of the indicators

In this section, I evaluate the business cycle indicator. Because reliable data is missing in the 19$^{\text{th}}$ and early 20$^{\text{th}}$ centuries, I rely on descriptive analysis and compare downswings in the indicator to narratives of economic crises and wars. First, I present the sentiment-based and count-based topic-specific indicators. Second, I show the composite indicator and discuss which topic-specific indicators contribute the most to the indicator. Then, I show the results from a correlation analysis. Finally, I present a business cycle chronology for Switzerland in the 19$^{\text{th}}$ and early 20$^{\text{th}}$ centuries.

4.1 Characteristics of topic-specific indicators

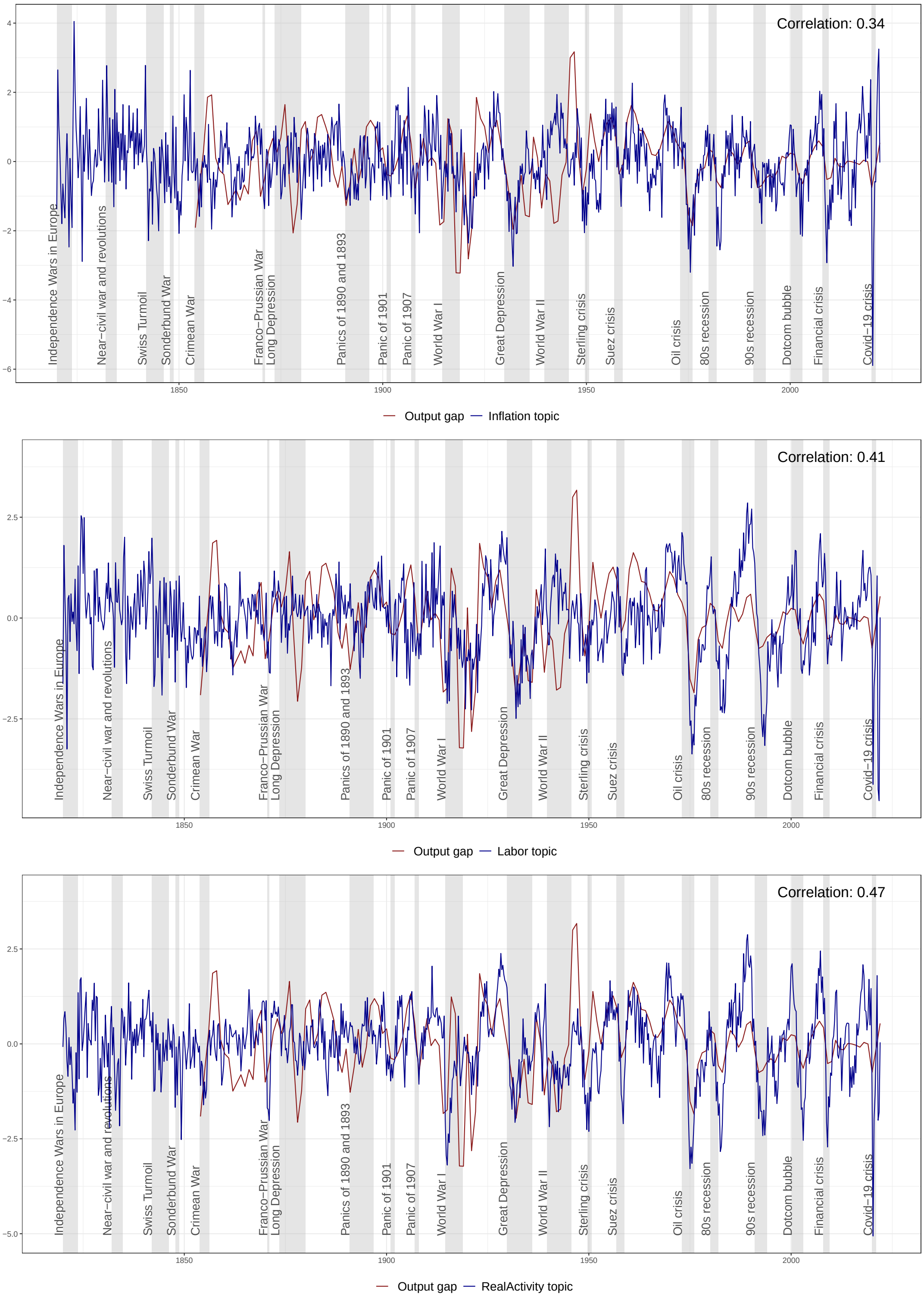
The sentiment-based topic-specific indicators are depicted in Figure 3. Gray-shaded areas highlight wars, crises, and recessions related to the Swiss or the global economy. Two observations stand out. First, the indicators are highly correlated with each other. This is unsurprising as the topics are related to the business cycle and, therefore, should move together. The correlation matrix in Table 7 in the Appendix confirms this. These correlations underline the importance of using a model that can handle collinearity, like the elastic net, to combine the indicators.

Figure 3 — Sentiment-based topic-specific indicators



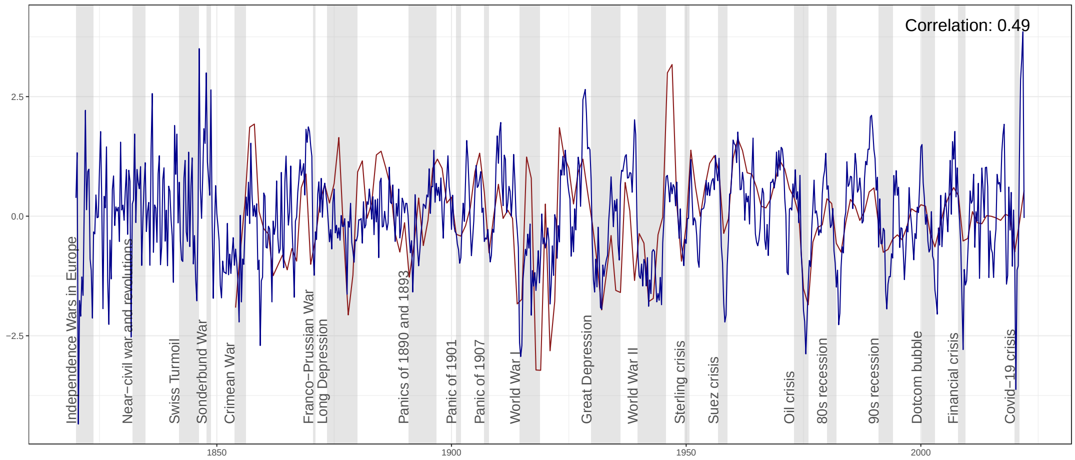
Notes: These graphs show sentiment-based topic-specific indicators and the output gap. Gray-shaded areas highlight wars, crises, and recessions related to the Swiss or the global economy.

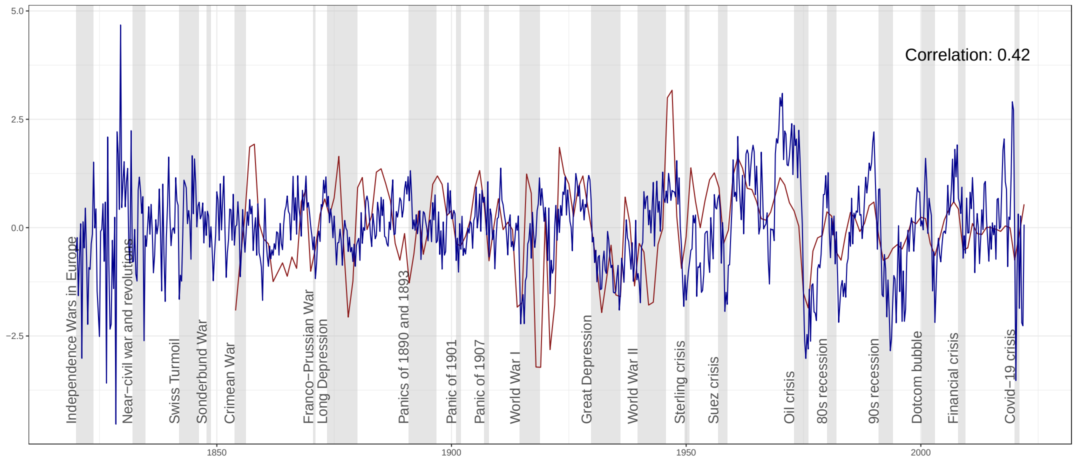
Figure 3 — Sentiment-based topic-specific indicators, continued from previous page



Notes: These graphs show sentiment-based topic-specific indicators and the output gap. Gray-shaded areas highlight wars, crises, and recessions related to the Swiss or the global economy.

Figure 3 — Sentiment-based topic-specific indicators, continued from previous page





Notes: These graphs show sentiment-based topic-specific indicators and the output gap. Gray-shaded areas highlight wars, crises, and recessions related to the Swiss or the global economy.

Second, the variance of the indicators seems to increase over time. This may seem counterintuitive given that the overall economy, including GDP growth, exhibited reduced volatility during the Great Moderation period. The quarterly Net Economic Sentiment for the US developed by Van Binsbergen et al. (2024) covering 1850 - 2020 shows a similar increase in volatility over time. However, they do not discuss this behavior in their paper. Nevertheless, there are possible explanations for it. First, it may be related to increased media sensationalism. This refers to the practice of presenting news stories or events in a way that exaggerates their importance or sensationalizes aspects of the story to attract attention and generate public interest. Media sensationalism might have intensified over the past century due to technological, economic, and social factors. Second, the data from earlier periods may contain higher measurement errors, leading to lower volatility by construction. In practical terms, when normalizing the identified segments of the source-level indicators, a higher noise-to-signal ratio in the early data can lead to a more pronounced signal reduction during normalization. This results in a dampened signal in the aggregated indicator, where some noise is effectively canceled. Appendix D provides a more detailed explanation of this phenomenon.

The count-based topic-specific indicators are depicted in Figure 4. They are scaled to have a positive correlation with the business cycle. Overall, the coincident correlations of these indicators are weaker than those of the sentiment-based indicators. Moreover, the increase in volatility over time is not present for the Boom and War indicators. For the Recession indicator, it is there. Although I also use other keywords like "crisis" for this indicator, this might be explained because the word "recession" has been less used (or did not exist in this context) in the earlier periods.[12] Gray-shaded areas in the graph for the war indicator highlight important wars. The indicator captures most of these combat actions well.

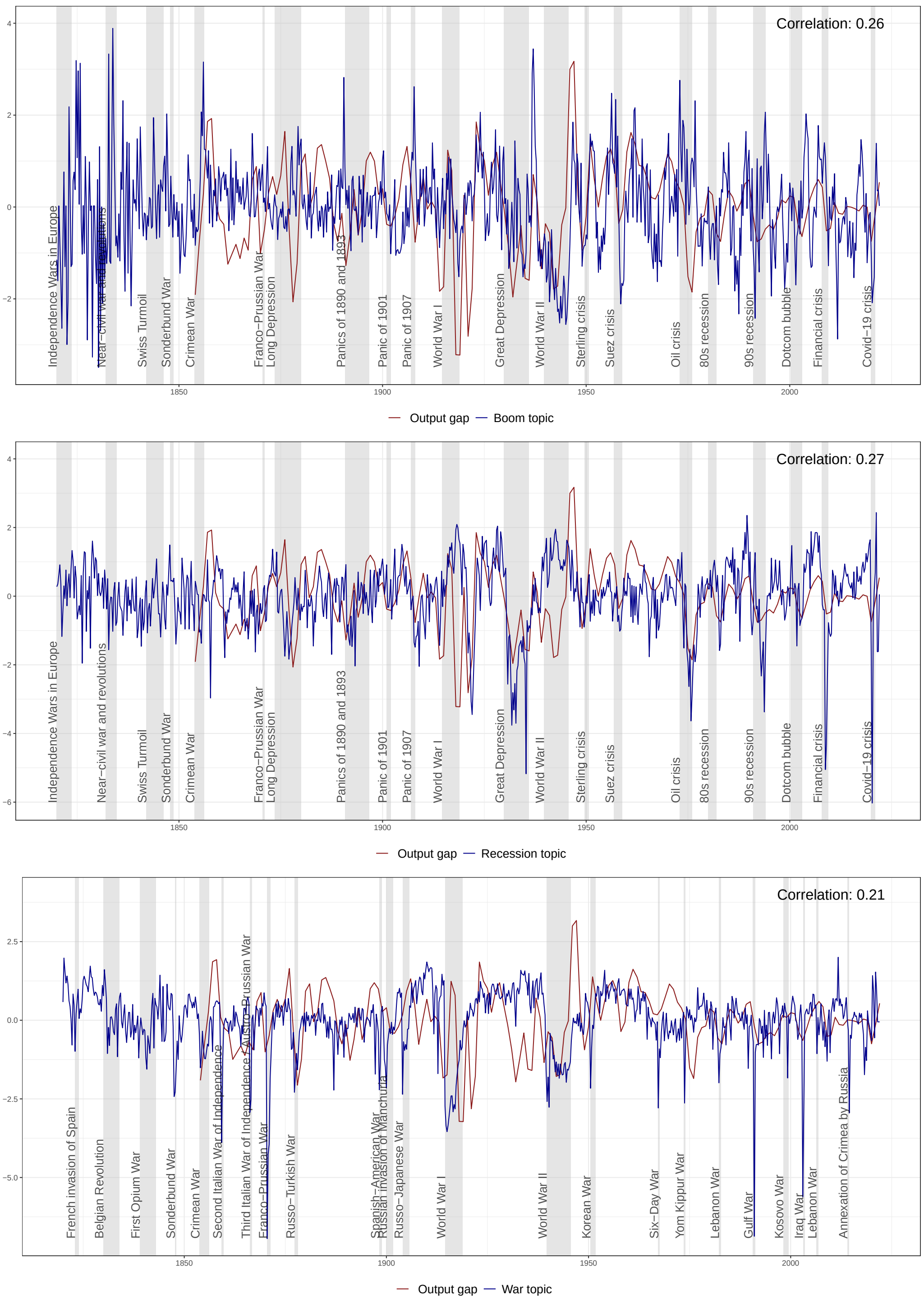
4.2 Characteristics of composite indicator

Figure 5 depicts the composite indicator with the black solid line. A few observations stand out. First, the indicator tracks economic downswings, given by the gray-shaded areas, very well.[13] It captures all the major recessions the Swiss economy faced in the last decades, like the Oil crisis from 1974, the recessions from the beginning of the 80s and 90s, the Great Financial Crisis (GFC) and the Covid–19 crisis. Moreover, the indicator responds to the two World Wars and the Great Depression. It is also able to capture crises in the 19$^{\text{th}}$ century like the Sonderbund War from 1847, the Franco-Prussian War from 1870, the Long Depression from 1873, and the Panic from 1890 (also known as

[12] According to the Oxford English Dictionary, the word "recession" was only increasingly used in the second half of the 20$^{\text{th}}$ century.

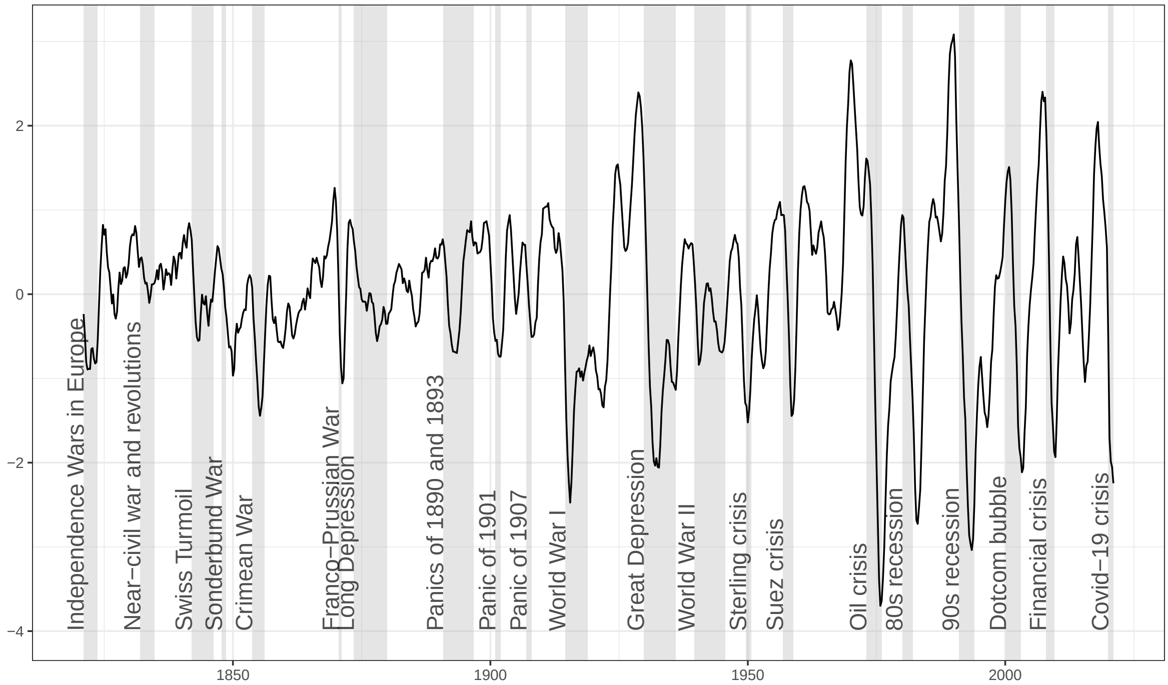
[13] For Switzerland, no official business cycle dating like from the NBER for the US exists. Therefore, these downswings are based on national and international narratives of crises and wars (see, e.g. Church & Head, 2013)

Figure 4 — Count-based topic-specific indicators



Notes: These graphs show count-based topic-specific indicators and the output gap. Indicators are scaled to have a positive correlation with the business cycle. Gray-shaded areas highlight wars, crises, and recessions related to the Swiss or the global economy.

Figure 5 — Historical business cycle indicator



Notes: The black line represents the composite historical business cycle indicator. Gray-shaded areas indicate global and Swiss crises, recessions, and wars.

the Baring crisis).[14] The indicator also captures Swiss-specific turmoil from 1842 as Switzerland experienced significant tensions between Liberals and Catholics. Second, as the underlying indicators, the composite indicator also gets more volatile over time. This is unsurprising as the indicator is a linear combination of the topic-specific indicators.

Which topic-specific indicators contribute the most to the composite indicator? To answer this question, Table 2 shows the ten most important predictors (highest coefficients in absolute values) from the elastic net regression. The most important predictors are the indicators for Wages, coincident and leading Labor, lagged Capacity utilization, lagged Trade, as well as Real activity and Inflation. Because Capacity

[14]The Sonderbund War was a civil conflict in Switzerland. It happened because seven Catholic cantons formed the Sonderbund in 1845 to protect their interests from a centralization of power. The war ended with the Sonderbund's defeat. This led to Switzerland becoming a federal state and marked the end of a period of political change in the country (Church & Head, 2013). The Franco-Prussian War was a conflict between the Second French Empire and the North German Confederation. As a neutral state, Switzerland was not directly involved in the conflict. However, due to its proximity, the war significantly impacted the Swiss economy (Church & Head, 2013). The Baring crisis was a financial crisis that occurred in 1890. It was caused by the near-collapse of Barings Bank, one of the most important financial institutions in the world at the time. The crisis had a significant impact on the global economy.

Table 2 — Ten most important predictors

Predictor	Coefficient
Wages	0.0212
Labor	0.0211
Capacity lag 3	0.0203
Capacity lag 2	0.0194
Trade lag 3	0.0193
Capacity lag 1	0.0186
Labor lead 1	0.0180
Real activity	0.0178
Inflation	0.0177
Capacity	0.0169

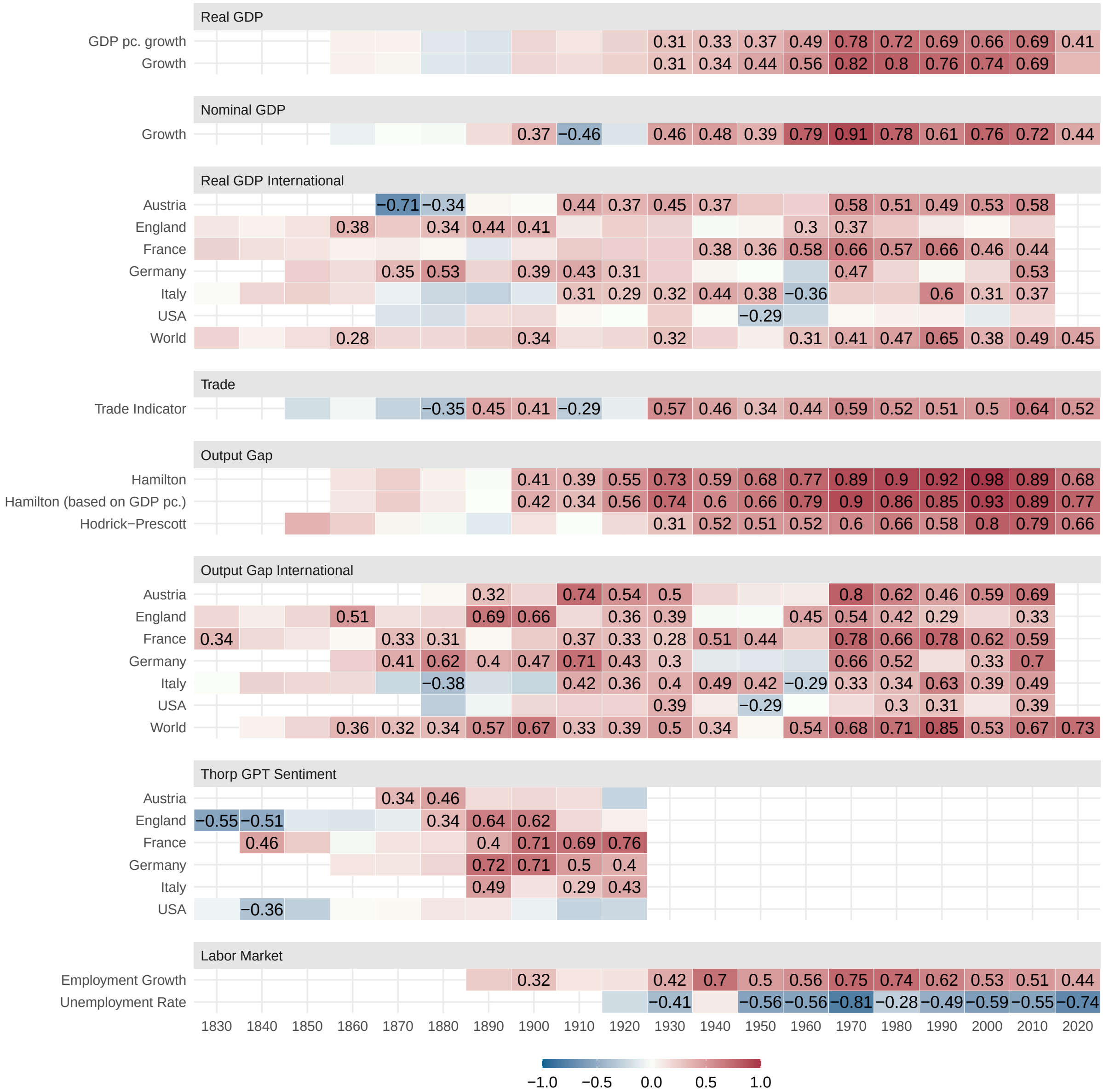
Notes: The table shows the ten most important predictors (highest coefficients in absolute values) from the elastic net regression.

utilization is directly related to the output gap, it is not surprising that it is one of the most important predictors. The count-based indicators seem less important.

Figure 6 shows the rolling correlations between the indicator and a selected set of variables with a fixed window size of 20 years. By construction, the indicator is highly correlated with the output gap, with an overall correlation of around 0.5. The correlation is highest during the late 20th and 21st centuries, with a correlation coefficient of around 0.9. The correlation is lower during the early 20th century but with around 0.5 still substantial. There is still some correlation for the 19th century, but it is not statistically different from zero. A similar picture emerges for labor market data, trade, and real and nominal GDP growth. The correlation is high during the late 20th and 21st centuries and lower for the 19th and early 20th centuries. The lower correlation in the early sample does not necessarily imply that the indicator is less accurate. It may be due to higher measurement errors in 19th-century data, which pushes the coefficients towards zero.

In the late 20th and early 21st centuries, the indicator shows high correlations between the output gap and real GDP growth of neighboring countries. This correlation is less pronounced in the USA and England yet remains substantial. In contrast, during the 19th century, the correlation diminishes for all countries but maintains statistical significance in France, Germany, and England. This pattern indicates a higher synchronization of the Swiss business cycle with international trends in the 20th century compared to the 19th century. These observations align with the findings of Broadberry and Lennard (2023), who report increasing synchronization of business cycles over time. Specifically, they find that 25% of the potential correlations between European countries' GDP growth were significantly positive from 1870 to 1950. This proportion rose to 66% from 1950 to 2000. However, these findings could also be driven by measurement errors.

Figure 6 — Rolling correlations with 20 year window



Notes: The graph shows rolling coincident correlations with a fixed window size of 20 years between the indicator on the left and the composite indicator. The period considered is given by the year on the x-axis plus and minus ten years. Only statistically significant (on a 10% level) correlations are labeled. Correlations with fewer than ten observations are not shown.

The broad picture is confirmed by the correlations with the sentiment indicators based on Thorp's (1926) description of the state of the economy. These indicators were generated using OpenAI's (2023b) GPT-3.5, as no sentiment indicators exist for the 19th century.[15] For all countries except the USA, there is a substantial correlation with the business cycle indicator. The strongest correlations are observed for Germany, France, and England.

Before 1850 there are a few significant positive correlations with variables from France. This makes sense since Switzerland was trading a lot with its larger neighbor sharing a common language. However, Switzerland was also highly dependent on southern Germany. However, there are no positive significant correlations with variables from other countries. This divergence could be explained by unique internal political and social conflicts following the Congress of Vienna in 1815. The period until the foundation of modern Switzerland in 1848 was characterized by pronounced divisions between liberal and conservative cantons, culminating in the Sonderbund War of 1847. These internal disputes, centered on governance, federal authority, and religious differences, fostered a climate of economic uncertainty and a focus on domestic concerns. Consequently, economic sentiment in Switzerland might have been insulated from broader European trends during this period (see, e.g. Church & Head, 2013).

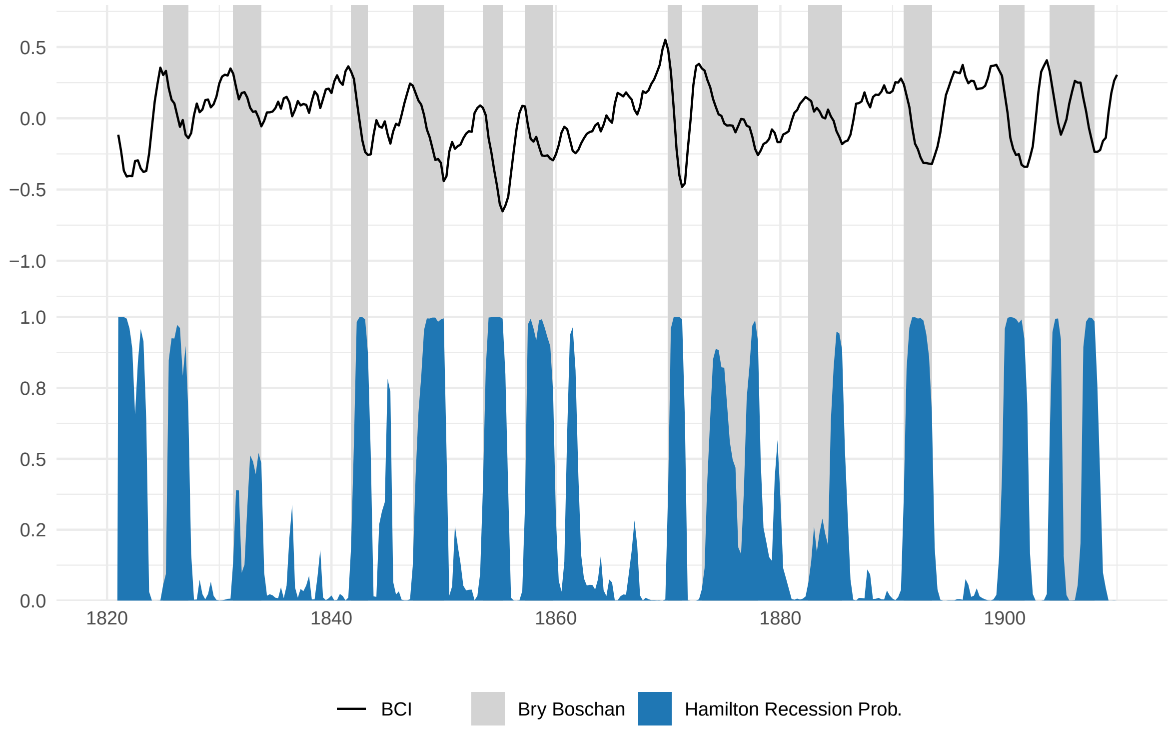
4.3 Business cycle dating

It is generally acknowledged that the modern business cycle features long expansions combined with short recessions (see, e.g. Romer & Romer, 2020). Due to data limitations, systematic data-driven analysis of business cycles could, until recently, be only conducted on very modern data (Broadberry & Lennard, 2023). However, with the progress in the quantification of economic activity, Broadberry and Lennard (2023) analyzed the nature of business cycles of nine European countries. Their main finding is that recessions got less frequent and shorter over time.

How do these findings compare to the Swiss business cycle? For the 19th and early 20th centuries, there is, to the best of my knowledge, no business cycle chronology available. The Economic Cycle Research Institute (ECRI) provides a chronology of business cycle turning points for Switzerland starting in 1956. The ECRI chronology is more judgemental, considering several indicators, including output, employment, income, and sales (Glocker & Wegmueller, 2020). Moreover, Siliverstovs (2011) uses an approach based on Markov-switching models, and Glocker and Wegmueller (2020) use a Markov-switching dynamic factor model (MS-DFM) to date business cycle turning

[15] Section B.2.2 in the Appendix provides a detailed description of the methodology to create the Thorp GPT Sentiment indicators.

Figure 7 — Indicator and recession classification 1820 - 1910



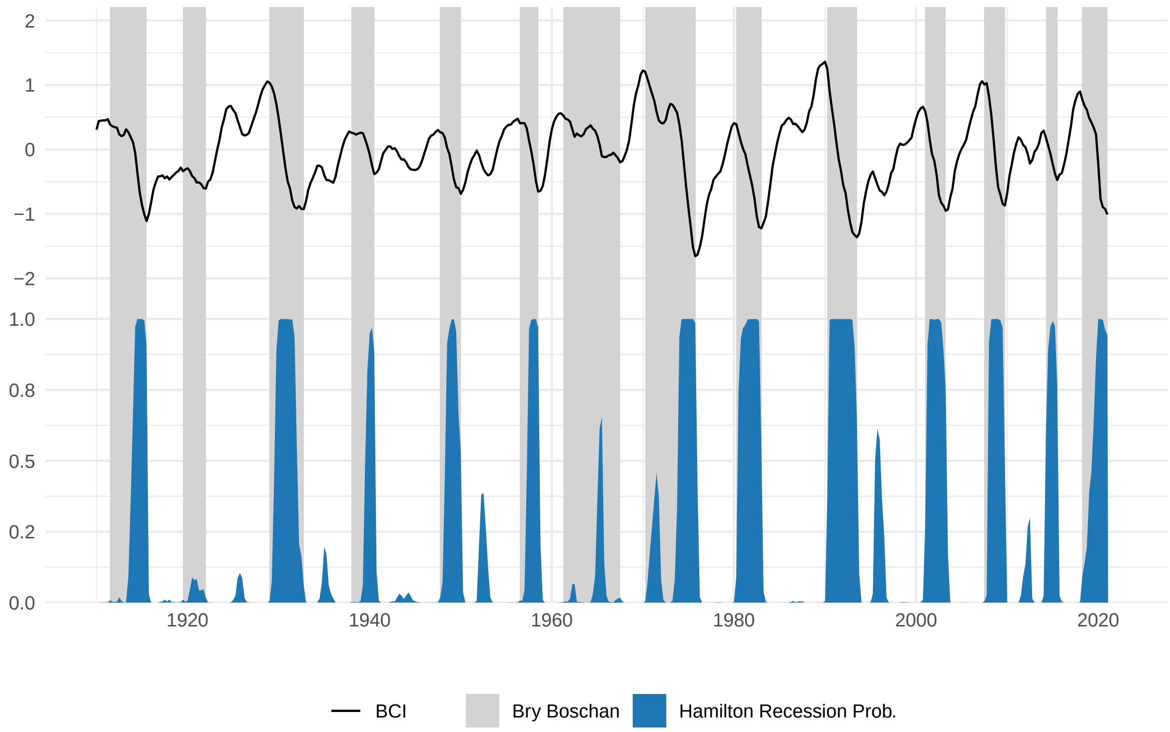
Notes: The graph shows the composite indicator given by the black solid line together with smoothed recession probabilities obtained by Hamilton's (1989) Markow-Switching autoregression model. The gray shaded areas indicate recessions obtained by an adapted Bry and Boschan (1971) algorithm.

points for Switzerland starting in 1980. Both identify turning points that are broadly consistent with those determined by the ECRI. Therefore, I mainly use the ECRI chronology for comparison.

To address the lack of a business cycle chronology for Switzerland during the 19$^{\text{th}}$ and early 20$^{\text{th}}$ centuries, this section utilizes the developed business cycle indicator to identify and date business cycle turning points. As Romer and Romer (2020) explain, recessions are not simply random categorizations of macroeconomic outcomes. Instead, they represent critical macroeconomic moments characterized by a rapid and significant deviation of economic activity from its normal state. Establishing a business cycle chronology for the 19$^{\text{th}}$ and early 20$^{\text{th}}$ centuries is thus crucial for understanding these economic shifts.

I use the Markov-Switching autoregression model proposed by Hamilton (1989) to estimate recession probabilities. This aligns with Romer and Romer (2020) who show that "recession periods emerge clearly from a Markov-switching model". The model is

Figure 8 — Indicator and recession classification 1910 - 2021



Notes: The graph shows the composite indicator given by the black solid line together with smoothed recession probabilities obtained by Hamilton's (1989) Markow-Switching autoregression model. The gray shaded areas indicate recessions obtained by an adapted Bry and Boschan (1971) algorithm.

given by

$$y_t = \mu_{s_t} + \phi y_{t-1} + \varepsilon_t \tag{7}$$

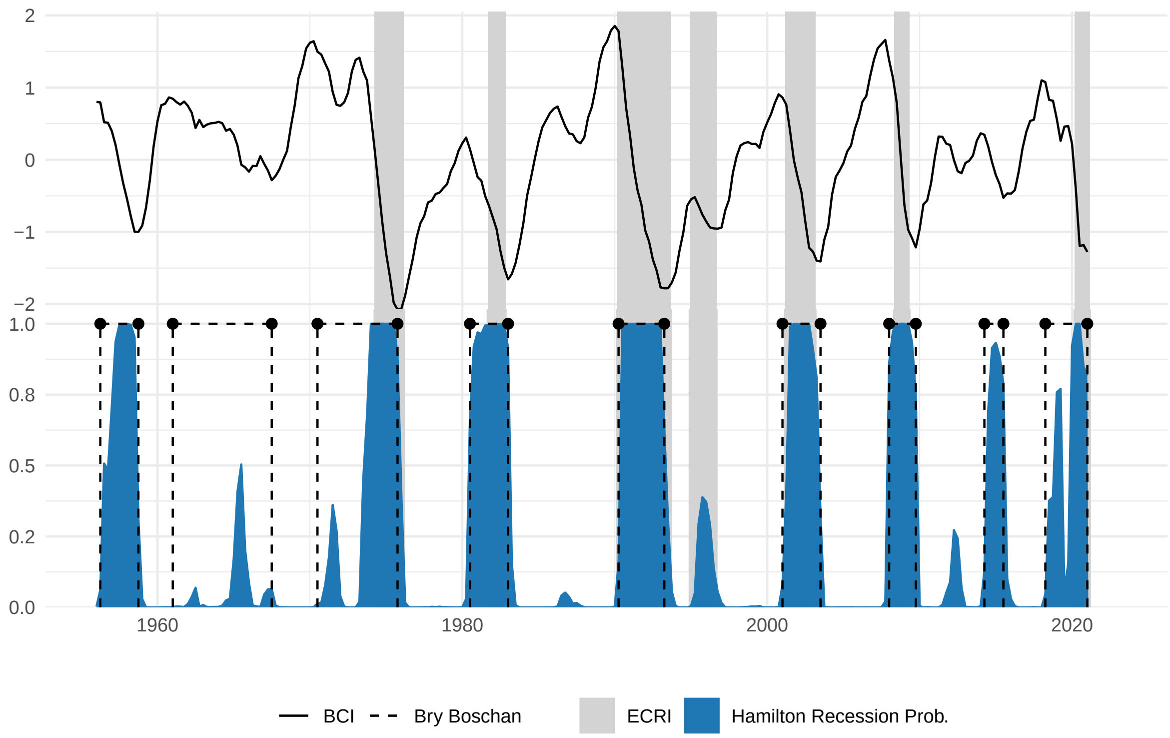
with $\varepsilon_t \sim N(0, \sigma^2)$ and where s_t is the realization of a two-state Markov chain with

$$\Pr(s_t = j|s_{t-1} = i, s_{t-2} = k, \ldots, y_{t-1}, y_{t-2}, \ldots) = \Pr(s_t = j|s_{t-1} = i) = p_{ij} \tag{8}$$

where $i, j = 0, 1$. Within this framework, one can label s_t = 0 and s_t = 1 as the expansion and recession states at time t. I estimate the model on two samples because the indicator is less volatile in the 19th century. First, from 1820 to 1910, and second, from 1910 to

2021.[16]

Figure 9 — Comparison with ECRI dating



Notes: The graph shows the composite indicator given by the black solid line together with smoothed recession probabilities obtained by Hamilton's (1989) Markow-Switching autoregression model. The black dashed lines indicate recessions obtained by an adapted Bry and Boschan (1971) algorithm. The gray-shaded areas are recession dates obtained from ECRI.

For robustness, I also use the nonparametric algorithm proposed by Harding and Pagan (2002) to date business cycle turning points. The algorithm is a quarterly adoption of Bry and Boschan's (1971) algorithm for monthly data. However, since the algorithm is designed for data in levels, I increased the number of quarters within which a local maximum or minimum has to occur to 10 quarters.[17]

An issue when utilizing detrended data arises because it tends to categorize so-called growth recessions as contractions. In the literature, these are not generally recognized as actual recessions (Broadberry & Lennard, 2023). To address this, I refine the identified recession periods, ensuring that their duration aligns with negative values in the

[16]Estimating the model on different samples is not uncommon. For example, Romer and Romer (2020) estimate the model with data from 1929 to 1947 and 1948 to 2019.

[17]This aligns with the general understanding that business cycles can vary but often fall within a 2 to 10-year range. See Harding and Pagan (2002) for more details on the algorithm.

business cycle indicator. The censored business cycle dating is very similar to the original one. All recessionary episodes identified using the Markov-switching model are listed in the Appendix in table 8.

Table 3 — Frequency, duration, and number of Swiss recessions

	Algorithm	Recession (Peak to trough) 1820-1910	1911-1950	1951-2022
Frequency	Hamilton (1989)	32.5	17.5	25.3
	Bry and Boschan (1971)	50.6	40.0	38.8
	BL2023 (Swiss GDP)	25.4	45.0	15.3
	ECRI			19.3
	OECD			39.0
	Technical recession			15.7
Duration	Hamilton (1989)	6.5	7.0	7.1
	Bry and Boschan (1971)	14.0	12.8	12.1
	BL2023 (Swiss GDP)	5.0	8.0	6.3
	ECRI			7.4
	OECD			22.5
	Technical recession			2.7
Number	Hamilton (1989)	18.0	4.0	10.0
	Bry and Boschan (1971)	13.0	5.0	9.0
	BL2023 (Swiss GDP)	12.0	9.0	7.0
	ECRI			7.0
	OECD			13.0
	Technical recession			10.0

Notes: Frequency is the share in percent of quarters in a given phase. Duration is the average number of quarters in a given phase. The ECRI dating starts in 1956, OECD dating starts in 1960, and quarterly GDP to calculate technical recessions is available from 1980. Broadberry and Lennard (2023) (BL2023 Swiss GDP) classify recessions as negative real annual GDP growth. Annual real Swiss GDP growth is available from 1852.

The results of the estimated recessionary episodes are depicted in Figures 7 and 8. I associate values of the regime probability above 0.5 with a recession.[18] The two regimes identified appear to correspond with phases of recessions and expansions, with the model's high recession probability periods showing considerable alignment with the recessions delineated by the Bry-Boschan algorithm. However, the Bry-Boschan procedure tends to overestimate recessions' duration.

[18] The recession threshold of 0.5 is arbitrary and chosen to be in line with Hamilton (1989). However, using different thresholds like, e.g., 0.8 proposed by Romer and Romer (2020) would not change the results significantly because the estimated probability is typically close to either 0 or 1.

Table 4 — Frequency, duration, and number of Swiss expansions

	Algorithm	Expansion (Trough to peak) 1820-1910	1911-1950	1951-2022
Frequency	Hamilton (1989)	67.5	82.5	74.7
	Bry and Boschan (1971)	49.4	60.0	61.2
	BL2023 (Swiss GDP)	74.6	55.0	84.7
	ECRI			80.7
	OECD			61.0
	Technical recession			84.3
Duration	Hamilton (1989)	13.8	34.5	21.3
	Bry and Boschan (1971)	14.1	20.4	19.4
	BL2023 (Swiss GDP)	13.5	10.5	34.3
	ECRI			29.9
	OECD			33.2
	Technical recession			13.5
Number	Hamilton (1989)	18.0	4.0	10.0
	Bry and Boschan (1971)	13.0	5.0	9.0
	BL2023 (Swiss GDP)	12.0	9.0	7.0
	ECRI			7.0
	OECD			13.0
	Technical recession			10.0

Notes: Frequency is the share in percent of quarters in a given phase. Duration is the average number of quarters in a given phase. The ECRI dating starts in 1956, OECD dating starts in 1960, and quarterly GDP to calculate technical recessions is available from 1980. Broadberry and Lennard (2023) (BL2023 Swiss GDP) classify recessions as negative real annual GDP growth. Annual real Swiss GDP growth is available from 1852.

Figure 9 compares my classification and the recession episodes identified by the ECRI. There is a significant overlap between the recessionary regime probabilities (illustrated in blue) and the ECRI's recession dating (indicated by the gray-shaded area). The Hamilton model tends to classify recessions more often. It additionally identifies with high probability recessions in 1958, known as the Eisenhower recession, and in 2015, coinciding with the SNB's termination of the minimum exchange rate. Additionally, the model suggests with less certainty another brief recessionary period in the 1960s.[19] The fact that we can attribute the additional recessions to well-known economic events suggests that the Markov-switching model identifies recessions more reliably than the

[19] I also compare my classification to the recession dates provided by the Organisation for Economic Co-operation and Development (OECD), recessionary episodes based on the technical definition for recessions (two consecutive quarters of negative GDP growth), and the definition by Broadberry and Lennard (2023) - classifying periods with negative annual GDP growth as recessions. The classifications show a high degree of overlap and are depicted in Figures 7, 8 and 9 in the Appendix.

existing ECRI classification.

How does the modern Swiss business cycle compare to the business cycle in the 19th century? Tables 3 and 4 show the frequency, duration, and number of identified recessions and expansions for periods 1820 to 1910, 1911 to 1950, and 1951 to 2022. Recessions in the 19th century were more frequent than in the 20th and 21st centuries. The share of quarters in recession decreases no matter which classification is used. However, the duration of recessions does not change significantly over the periods considered. Expansions, in contrast, get longer over time. The average duration of expansions increases from 13.8 quarters in the period up to 1910 to 21.3 quarters since 1951. Tables 5 and 6 in the Appendix compare these numbers to the European (Broadberry & Lennard, 2023) and the British business cycle (Broadberry et al., 2023). The results on the frequency and the expansions are in line with the findings on the European business cycle, albeit less pronounced. However, the duration of recessions is not. Broadberry and Lennard (2023) find that the duration of recessions in Europe decreases over time. On the other hand, Broadberry et al. (2023) confirm that the duration of recessions in the UK stays constant over time. Therefore, the widespread belief that recessions got shorter is not fully supported by the data.

5 Robustness

In this section, I perform a series of robustness checks. In particular, I show that the keyword-based algorithm delivers similar results as a rating based on OpenAI's (2023b) LLM GPT-3.5. Moreover, I analyze the sensitivity of the indicator concerning different aggregation techniques and different subsets of the data.

5.1 Keyword-based algorithm

Given the emergence of LLMs, the question of why these models are not employed for extracting information from the text corpus is naturally raised. There are two main reasons. First, the cost of utilizing these models is a significant factor. For instance, OpenAI's (2023b) GPT-3.5 model incurs a charge of 0.001 USD per 1000 tokens[20]. With a conservative estimate of the average text length being only 20 tokens, processing 100 million texts would approximate a total expense of around 2 million USD. Second, it would take significant time to process all the texts. Nevertheless, it would be interesting to compare the keyword-based algorithm with LLMs for future research.

[20]This corresponds to approximately 750 words.

Due to the high costs associated with LLMs, their use for analyzing my text corpus is not feasible. Nonetheless, to ascertain the efficacy of the keyword-based algorithm, I also apply the algorithm to Thorp's (1926) work on business cycles (i.e., his descriptions of the state of the economy). I then compare the indicators created this way to those derived with GPT-3.5 as described in section B.2.2 in the Appendix. To be precise, GPT-3.5-based indicators use the texts in English. For the keyword-based algorithm, the English texts are first translated to German and French.[21]

As shown in Figure 10, the correlations between the GPT-3.5-based and keyword-based indicators range between 0.5 and 0.8 for the countries under consideration. This suggests that the keyword-based algorithm delivers similar results as LLMs. Whether one or the other method is superior is beyond the scope of this study and should be addressed in future research. However, the results suggest that the keyword-based algorithm is a viable alternative to LLMs for constructing business cycle indicators.

5.2 Aggregation techniques

There are countless possibilities for constructing a business cycle indicator from textual data. I compare the baseline indicator to several alternatives to check that it is not sensitive to the aggregation technique.

Other aggregation methods to composite indicator. Instead of using an elastic net model to aggregate the topic-specific indicators into the composite indicator, I use a

- Weighted average based on keyword importance in each topic:

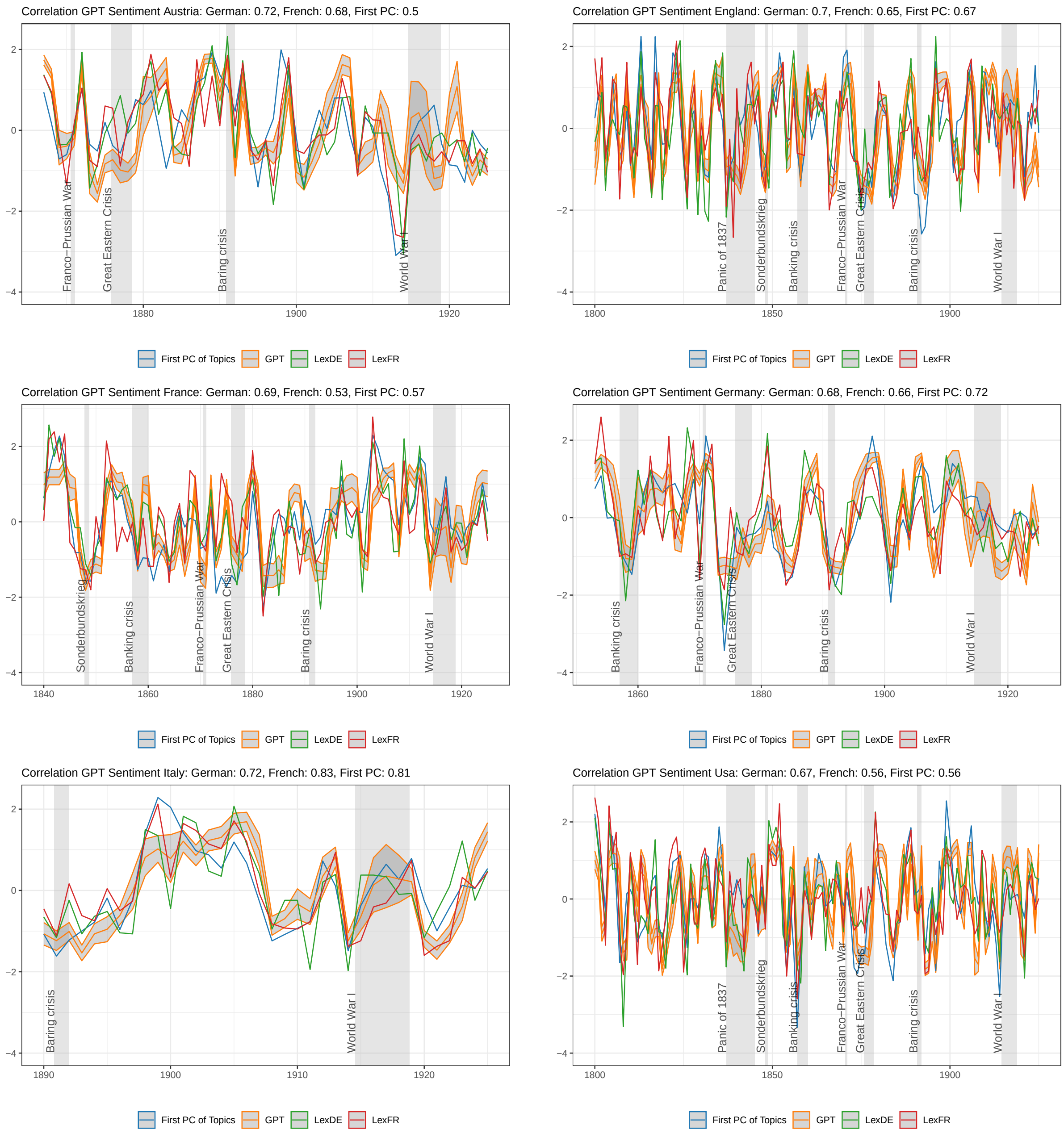
$$S_t = \sum_{j=1}^{P} \omega_{j,t} s_{j,t} \tag{9}$$

  where S_t denotes the composite indicator in quarter t, P is the total number of topics. $\omega_{j,t}$ and $s_{j,t}$ represent the weight derived by the number of identified keywords and the topic-specific indicator, respectively, for topic j in quarter t.

- Markov-Switching Dynamic Factor Model (MS-DFM) brought forward by Kim (1994), Diebold and Rudebusch (1996), Kim and Yoo (1995) and Chauvet (1998).[22] In these models, an unobservable regime-switching variable governs the common

[21] To translate the texts, I use ChatGPT (OpenAI, 2023a).
[22] I thank Philipp Wegmüller for providing his programs to estimate MS-DFMs.

Figure 10 — Comparison of lexicons and methods



Notes: These graphs show several text-based sentiment indicators based on Thorp's (1926) description. In orange, the normalized sentiment indicator based on OpenAI's (2023b) GPT-3.5 model and one standard deviation confidence bands. In green, the sentiment indicator based on a German translation of Thorp (1926) and the German lexicon. In red, the sentiment indicator based on a French translation of Thorp (1926) and the French lexicon. In blue, the first principal component from indicators based on both translations and the keyword-based algorithm outlined in section 3. Gray-shaded areas represent crises.

component.[23] The model comprises one factor (based on screeplot in Figure 10) and five lags (based on Bayesian Information Criterion (BIC)).[24] Using an MS-DFM allows to check the sensitivity of the indicator and the sensitivity of the business cycle dating.

Other aggregation methods to topic-specific indicators. Instead of using a static factor model to aggregate the sources to a topic-specific indicator, I

- use a weighted average based on keyword importance in each source:

$$S_{j,t} = \sum_{i=1}^{N_t} \omega_{j,i,t} s_{j,i,t} \tag{10}$$

  where $S_{j,t}$ denotes the topic-specific indicators. N_t is the total number of sources available in quarter t. $\omega_{j,i,t}$ and $s_{j,i,t}$ represent the weight derived by the number of identified keywords and the source-level indicator, respectively, for a specific topic j and source i in quarter t.

- use a fixed effects regression model à la Shapiro et al. (2022). I estimate the quarter fixed effects and use these as the indicator ($S_{j,t} = \hat{f}_{j,t}$) from the following regression

$$s_{j,i,t} = f_{j,t} + f_{j,i} + f_{j,l} + f_{j,f} + f_{j,i} \times f_{j,p} + \varepsilon_{j,i,t} \tag{11}$$

  where $s_{j,i,t}$ is the sentiment score in quarter t for a specific topic j and source i. $f_{j,t}$ is a sample quarter fixed effect, $f_{j,i}$ a source fixed effect, $f_{j,l}$ a language fixed effect, $f_{j,f}$ a frequency fixed effect, $f_{j,i} \times f_{j,p}$ a source$\times$type fixed effect. The type is either a report or a news article. Allowing for all these other fixed effects besides the quarter fixed effects ensures that the index is independent of changes over time in the sample's composition across newspapers and reports versus regular articles. Moreover, it controls for differences in the sentiment scores across sources, languages, and frequencies. This might be important because the sentiment scores

[23]Camacho et al. (2015) show that performing the estimation in one step is superior to estimating a Markov-switching process on the factor in a sequential step.
[24]See section E in the Appendix for more details.

differ considerably across sources, types, frequencies, and languages.

- treat the identified segments as separate indicators in the static factor model. The reason for this robustness test is that in the baseline specification, the source-level sentiments are assumed to follow

$$s_{j,i,t} = (\lambda_{j,i}/\sigma_{j,i,t})f_{j,t} + (e_{j,i,t}/\sigma_{j,i,t}). \tag{12}$$

  The factor loadings $\lambda_{j,i}$ have been scaled by $1/\sigma_{j,i,t}$. If $\sigma_{j,i,t}$ is large in the early part of the sample, then principle components (which assumes a time-invariant value of $\lambda_{j,i}$) potentially will under-estimate the variance of $f_{j,t}$ in the early part of the sample.

Other adjustment methods. Instead of adjusting the source-level indicators according to the procedure described in section 3.2, I

- do not scale the sentiment-based indicators. I still identify structural breaks but only demean each segment. I do not change the calculation of the count-based indicators. See section C in the Appendix for more details.
- do not subtract a trend.

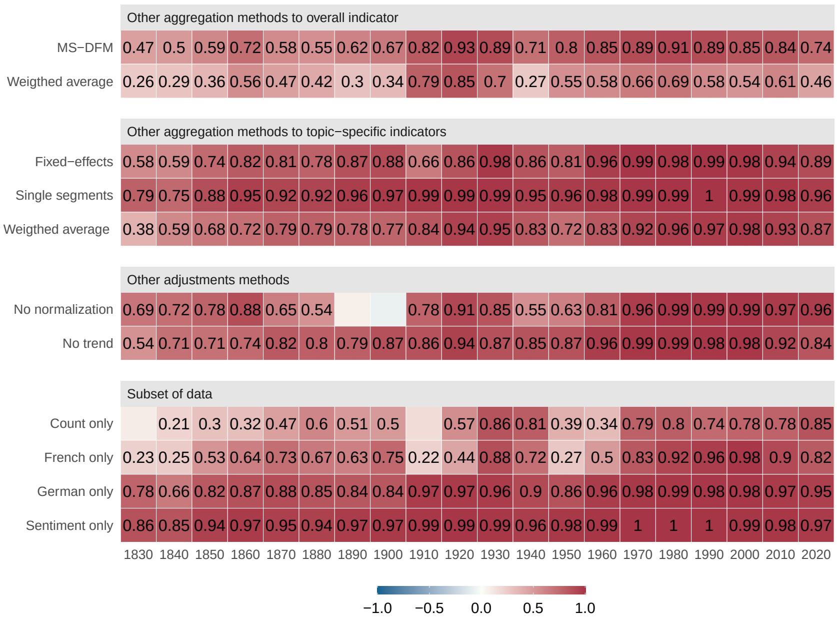
Subsets of data. Instead of using all available data, I calculate the indicator

- only using German texts.
- only using French texts.
- only using sentiment-based indicators.
- only using count-based indicators.

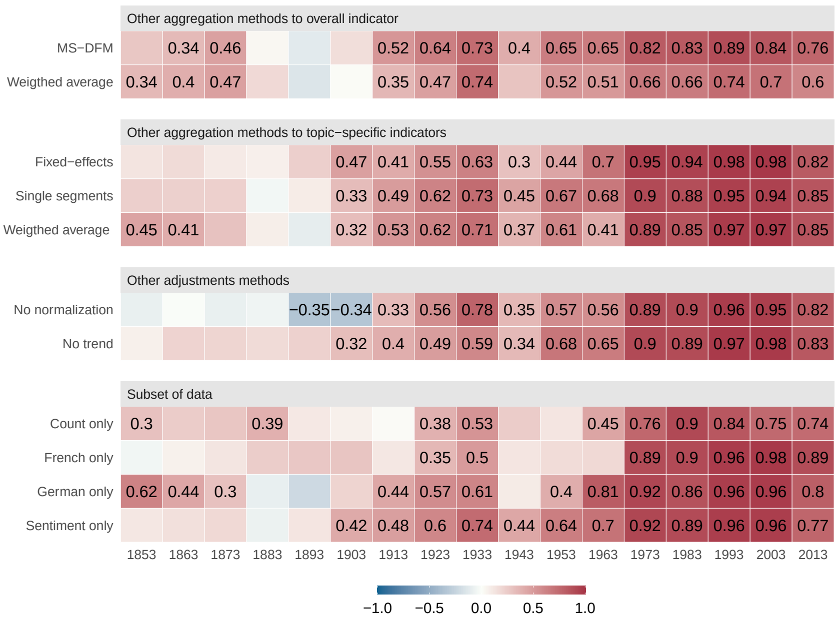
The alternative indicators are very similar, as shown in Figures 11 and 12 in the Appendix. However, there is some more dispersion before World War I. Nevertheless, all alternative indicators capture major downturns like the Franco-Prussian war from 1870 or the Baring crisis from 1890. There is a substantial correlation between the alternative indicators and the baseline indicator. The correlation between the alternative indicators and the output gap is similar to that between the baseline indicator and the output gap. Finally, the recession probabilities obtained with the MS-DFM in Figure 11

Figure 11 — Correlations of alternative indicators

a) Correlation with baseline composite indicator



b) Correlation with output gap



Notes: The graph shows rolling correlations with a fixed window size of 20 years between the indicator on the left and the composite indicator (in panel a) and the output gap (in panel b). The period considered is given by the year on the x-axis plus minus ten years. Only statistically significant (on a 10% level) correlations are labeled.

in the Appendix show substantial overlap with the dating obtained using Hamilton's (1989) model. These analyses demonstrate that the indicator is not overly sensitive to specific aggregation techniques or data limitations.

6 Concluding remarks

It is well known that measurement problems in hard economic data distort business cycle facts in the 19th century. For Switzerland, existing GDP series have "the characteristics of a rough estimate at best" (Historische Statistik der Schweiz HSSO, 2012k) and are available only at an annual frequency.

To overcome this problem, this chapter proposes a novel approach to measuring business cycle fluctuations over long historical episodes using textual data. I collect a large body of texts relevant to business cycle fluctuations. Using textual analysis, I create a quarterly business cycle indicator for Switzerland from 1821 to 2021. Based on this indicator, a business cycle chronology for Switzerland is established.

The business cycle indicator successfully captures major economic downturns, including the two World Wars, the Great Depression, and various financial crises. It strongly correlates with existing hard economic data in Switzerland, particularly in the second half of the 20th century and the early 21st century. This correlation, although somewhat weaker, extends to earlier periods. These lower correlations do not necessarily imply that the indicator does not accurately represent economic fluctuations in the 19th and early 20th centuries. Instead, they may reflect higher measurement errors in the data from earlier periods, which push the correlation toward zero. Therefore, a low correlation is what we would expect under the assumption that hard economic data in the 19th century is measured with higher measurement error. Comparing the indicator with potentially uncollected hard data to verify its accuracy would be an interesting avenue for future research.

The business cycle dating is consistent with the one by a dating committee for the 20th century. For the 19th century, the dating is consistent with narratives about wars and crises. Moreover, the results show that Swiss recessions have become less frequent, aligning with trends observed in other European nations. In contrast to other European countries, the duration of Swiss recessions does not show a significant reduction.

Because hard data and qualitative data are subject to measurement error, it is crucial to combine different data sources to obtain a reliable measure of the business cycle.

Therefore, a promising line for future research would be to combine error-prone hard and qualitative data to obtain a more accurate measure of the business cycle.

References

Abdaoui, A., Azé, J., Bringay, S., & Poncelet, P. (2017). FEEL: A french expanded emotion lexicon. *Language Resources and Evaluation*, *51*(3), 833–855. https://doi.org/10.1007/s10579-016-9364-5

Angelico, C., Marcucci, J., Miccoli, M., & Quarta, F. (2022). Can we measure inflation expectations using twitter? *Journal of Econometrics*, *228*(2), 259–277. https://doi.org/10.1016/j.jeconom.2021.12.008

Ardia, D., Bluteau, K., & Boudt, K. (2019). Questioning the news about economic growth: Sparse forecasting using thousands of news-based sentiment values. *International Journal of Forecasting*, *35*(4), 1370–1386. https://doi.org/10.1016/j.ijforecast.2018.10.010

Ardia, D., Bluteau, K., & Kassem, A. (2021). A century of economic policy uncertainty through the french–canadian lens. *Economics Letters*, *205*, 109938. https://doi.org/10.1016/j.econlet.2021.109938

Aruoba, S. B., & Drechsel, T. (2024). Identifying monetary policy shocks: A natural language approach. *NBER Working Paper No. 32417*. https://doi.org/10.3386/w32417

Ash, E., & Hansen, S. (2023). Text algorithms in economics. *Annual Review of Economics*, *15*(1), 659–688. https://doi.org/10.1146/annurev-economics-082222-074352

Bai, J., & Ng, S. (2013). Principal components estimation and identification of static factors. *Journal of Econometrics*, *176*(1), 18–29. https://doi.org/10.1016/j.jeconom.2013.03.007

Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, *131*(4), 1593–1636. https://doi.org/10.1093/qje/qjw024

Barbaglia, L., Consoli, S., & Manzan, S. (2023). Forecasting with economic news. *Journal of Business & Economic Statistics*, *41*(3), 708–719. https://doi.org/10.1080/07350015.2022.2060988

Beach, B., & Hanlon, W. W. (2022). Historical newspaper data: A researcher's guide and toolkit. *NBER Working Paper No. 30135*.

Binder, C. C. (2016). Estimation of historical inflation expectations. *Explorations in Economic History*, *61*, 1–31. https://doi.org/10.1016/j.eeh.2016.01.002

Bolt, J., & van Zanden, J. L. (2020). Maddison style estimates of the evolution of the world economy. a new 2020 update. *Maddison-Project Working Paper*.

Broadberry, S., Chadha, J. S., Lennard, J., & Thomas, R. (2023). Dating business cycles in the united kingdom, 1700–2010. *The Economic History Review*, *76*(4), 1141–1162. https://doi.org/10.1111/ehr.13238

Broadberry, S., & Lennard, J. (2023). European business cycles and economic growth, 1300-2000. *Economic History Working Papers 120364*. https://ideas.repec.org/p/ehl/wpaper/120364.html

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., . . . Amodei, D. (2020, July 22). Language models are few-shot learners. http://arxiv.org/abs/2005.14165

Bry, G., & Boschan, C. (1971). *Cyclical analysis of time series: Selected procedures and computer programs*. National Bureau of Economic Research; Columbia University Press.

Burns, A. F., & Mitchell, W. C. (1946). *Measuring business cycles*. National Bureau of Economic Research, Inc.

Burri, M. (2023). Do daily lead texts help nowcasting GDP growth? *IRENE Working Papers 23-02*.

Burri, M., & Kaufmann, D. (2020). A daily fever curve for the swiss economy. *Swiss Journal of Economics and Statistics*, *156*(1), 6. https://doi.org/10.1186/s41937-020-00051-z

Bybee, L., Kelly, B. T., Manela, A., & Xiu, D. (2023). Business news and business cycles. *Journal of Finance, Forthcoming*. http://dx.doi.org/10.2139/ssrn.3446225

Camacho, M., Perez-Quiros, G., & Poncela, P. (2015). Extracting nonlinear signals from several economic indicators. *Journal of Applied Econometrics*, *30*(7), 1073–1089. https://doi.org/10.1002/jae.2416

Canova, F. (1994). Detrending and turning points. *European Economic Review*, *38*(3), 614–623. https://doi.org/10.1016/0014-2921(94)90097-3

Canova, F. (1998). Detrending and business cycle facts. *Journal of Monetary Economics*, *41*(3), 475–512. https://doi.org/10.1016/S0304-3932(98)00006-3

Chauvet, M. (1998). An econometric characterization of business cycle dynamics with factor structure and regime switching. *International Economic Review*, *39*(4), 969. https://doi.org/10.2307/2527348

Church, C. H., & Head, R. C. (2013, May 23). *A concise history of switzerland* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139013765

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, *74*(368), 829–836. https://doi.org/10.1080/01621459.1979.10481038

Dagum, E. B., & Cholette, P. A. (2006). *Benchmarking, temporal distribution, and reconciliation methods for time series* (Vol. Lecture notes in statistics). Springer.

Denton, F. T. (1971). Adjustment of monthly or quarterly series to annual totals: An approach based on quadratic minimization. *Journal of the American Statistical Association*, *66*(333), 99–102. https://doi.org/10.1080/01621459.1971.10482227

Diebold, F. X., & Rudebusch, G. D. (1996). Measuring business cycles: A modern perspective. *The Review of Economics and Statistics*, *78*(1), pp. 67–77.

Edwards, A. W. F., & Cavalli-Sforza, L. L. (1965). A method for cluster analysis. *Biometrics*, *21*(2), 362. https://doi.org/10.2307/2528096

Ellingsen, J., Larsen, V. H., & Thorsrud, L. A. (2022). News media versus FRED-MD for macroeconomic forecasting. *Journal of Applied Econometrics*, *37*(1), 63–81. https://doi.org/10.1002/jae.2859

Feinerer, I., & Hornik, K. (2019). *Tm: Text mining package*. https://CRAN.R-project.org/package=tm

Glocker, C., & Wegmueller, P. (2020). Business cycle dating and forecasting with real-time swiss GDP data. *Empirical Economics*, *58*(1), 73–105. https://doi.org/10.1007/s00181-019-01666-9

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, *57*(2), 357. https://doi.org/10.2307/1912559

Hamilton, J. D. (2018). Why you should never use the hodrick-prescott filter. *The Review of Economics and Statistics*, *100*(5), 831–843. https://doi.org/10.1162/rest_a_00706

Hanna, A. J., Turner, J. D., & Walker, C. B. (2020). News media and investor sentiment during bull and bear markets. *The European Journal of Finance*, *26*(14), 1377–1395. https://doi.org/10.1080/1351847X.2020.1743734

Harding, D., & Pagan, A. (2002). Dissecting the cycle: A methodological investigation. *Journal of Monetary Economics*, *49*(2), 365–381. https://doi.org/10.1016/S0304-3932(01)00108-8

Hirshleifer, D., Mai, D., & Pukthuanthong, K. (2023). War discourse and the cross section of expected stock returns. *NBER Working Paper No. 31348*. https://doi.org/10.3386/w31348

Historische Statistik der Schweiz HSSO. (2012a). Ausfuhrmengen nach warenarten und ausfuhrmengenindizes 1851–1913. Retrieved September 19, 2023, from https://hsso.ch/de/2012/l/2

Historische Statistik der Schweiz HSSO. (2012b). Bilanz des aussenhandels und zollerträgnisse nach warenarten 1886–1992. Retrieved September 19, 2023, from https://hsso.ch/de/2012/l/3

Historische Statistik der Schweiz HSSO. (2012c). Bruttoinlandprodukt nach verwendungsarten in preisen von 1929 und nominal, 1890-1948. Retrieved September 19, 2023, from https://hsso.ch/de/2012/q/16a

Historische Statistik der Schweiz HSSO. (2012d). Bruttoinlandprodukt nach verwendungsarten zu preisen von 1990 und nominal, 1948-2005. Retrieved September 19, 2023, from https://hsso.ch/de/2012/q/16b

Historische Statistik der Schweiz HSSO. (2012e). Einfuhrmengen nach warenarten und einfuhrmengenindizes 1851–1913. Retrieved September 19, 2023, from https://hsso.ch/de/2012/l/1a

Historische Statistik der Schweiz HSSO. (2012f). Erwerbstätige, wochen-, jahresarbeitszeit 1890-2005. Retrieved September 19, 2023, from https://hsso.ch/de/2012/f/29a

Historische Statistik der Schweiz HSSO. (2012g). Monatliche und vierteljährliche ausfuhrmengen januar 1924 bis april 1967 und monatlicher und vierteljährlicher ausfuhrmengenindex von august 1974 bis oktober 1987. Retrieved September 19, 2023, from https://hsso.ch/de/2012/l/5a

Historische Statistik der Schweiz HSSO. (2012h). Monatliche und vierteljährliche ausfuhrwerte 1924–1992. Retrieved September 19, 2023, from https://hsso.ch/de/2012/l/8

Historische Statistik der Schweiz HSSO. (2012i). Monatliche und vierteljährliche einfuhrmengen januar 1924 bis dezember 1987. Retrieved September 19, 2023, from https://hsso.ch/de/2012/l/4

Historische Statistik der Schweiz HSSO. (2012j). Monatliche und vierteljährliche einfuhrwerte 1924–1992. Retrieved September 19, 2023, from https://hsso.ch/de/2012/l/7

Historische Statistik der Schweiz HSSO. (2012k). Nominales und reales BIP und bruttowertschöpfung nach branchen 1851-1913. Retrieved September 19, 2023, from https://hsso.ch/2012/q/1a

Historische Statistik der Schweiz HSSO. (2012l). Stellensuchende und arbeitslosenquote nach geschlecht im jahresmittel 1913-1995. Retrieved September 19, 2023, from https://hsso.ch/de/2012/f/18a

Horvath, L. (1993). The maximum likelihood method for testing changes in the parameters of normal observations. *The Annals of Statistics*, *21*(2). https://doi.org/10.1214/aos/1176349143

Kabiri, A., James, H., Landon-Lane, J., Tuckett, D., & Nyman, R. (2023). The role of sentiment in the US economy: 1920 to 1934. *The Economic History Review*, 3–30. https://doi.org/10.1111/ehr.13160

Kalamara, E., Turrell, A., Redl, C., Kapetanios, G., & Kapadia, S. (2022). Making text count: Economic forecasting using newspaper text. *Journal of Applied Econometrics*, *37*(5), 896–919. https://doi.org/10.1002/jae.2907

Kaufmann, D. (2020). Is deflation costly after all? the perils of erroneous historical classifications. *Journal of Applied Econometrics*, *35*(5), 614–628. https://doi.org/10.1002/jae.2762

Killick, R., & Eckley, I. A. (2014). **changepoint** : An *R* package for changepoint analysis. *Journal of Statistical Software*, *58*(3). https://doi.org/10.18637/jss.v058.i03

Kim, C.-J. (1994). Dynamic linear models with markov-switching. *Journal of Econometrics*, *60*(1), 1–22. https://doi.org/10.1016/0304-4076(94)90036-1

Kim, M.-J., & Yoo, J.-S. (1995). New index of coincident indicators: A multivariate markov switching factor model approach. *Journal of Monetary Economics*, *36*(3), 607–630. https://doi.org/10.1016/0304-3932(95)01229-X

Larsen, V. H., & Thorsrud, L. A. (2019). The value of news for economic developments. *Journal of Econometrics*, *210*(1), 203–218. https://doi.org/10.1016/j.jeconom.2018.11.013

Larsen, V. H., Thorsrud, L. A., & Zhulanova, J. (2021). News-driven inflation expectations and information rigidities. *Journal of Monetary Economics*, *117*, 507–520. https://doi.org/10.1016/j.jmoneco.2020.03.004

Larsen, V. H. (2021). Components of uncertainty. *International Economic Review*, *62*(2), 769–788. https://doi.org/10.1111/iere.12499

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of Finance*, *66*(1), 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x

Luhn, H. P. (1960). Key word-in-context index for technical literature. *American Documentation*, *11*(4), 288–295. https://doi.org/10.1002/asi.5090110403

Mariano, R. S., & Murasawa, Y. (2010). A coincident index, common factors, and monthly real GDP. *Oxford Bulletin of Economics and Statistics*, *72*(1), 27–46. https://doi.org/10.1111/j.1468-0084.2009.00567.x

OpenAI. (2023a). ChatGPT. Retrieved October 1, 2023, from https://www.chatgpt.com/

OpenAI. (2023b). GPT-3.5 API. Retrieved October 1, 2023, from https://www.platform.openai.com/

Picard, F., Robin, S., Lavielle, M., Vaisse, C., & Daudin, J.-J. (2005). A statistical approach for array CGH data analysis. *BMC Bioinformatics*, *6*(1), 27. https://doi.org/10.1186/1471-2105-6-27

Remus, R., Quasthoff, U., & Heyer, G. (2010). SentiWS - a publicly available german-language resource for sentiment analysis. *Proceedings of the seventh international conference on language resources and evaluation (LREC'10)*.

Romer, C. D. (1994). Remeasuring business cycles. *The Journal of Economic History*, *54*(3), 573–609.

Romer, C. D., & Romer, D. H. (2020). NBER recession dates: Strengths, weaknesses, and a modern upgrade. *mimeo*.

Sax, C., & Steiner, P. (2013). Temporal disaggregation of time series. *The R Journal*, *5*(2), 80. https://doi.org/10.32614/RJ-2013-028

Schorfheide, F., & Song, D. (2015). Real-time forecasting with a mixed-frequency VAR. *Journal of Business & Economic Statistics*, *33*(3), 366–380. https://doi.org/10.1080/07350015.2014.954707

Scott, A. J., & Knott, M. (1974). A cluster analysis method for grouping means in the analysis of variance. *Biometrics*, *30*(3), 507. https://doi.org/10.2307/2529204

Sen, A., & Srivastava, M. S. (1975). On tests for detecting change in mean. *The Annals of Statistics*, *3*(1). https://doi.org/10.1214/aos/1176343001

Shapiro, A. H., Sudhof, M., & Wilson, D. J. (2022). Measuring news sentiment. *Journal of Econometrics*, *228*(2), 221–243. https://doi.org/https://doi.org/10.1016/j.jeconom.2020.07.053

Shiller, R. J. (2017). Narrative economics. *American Economic Review*, *107*(4), 967–1004. https://doi.org/10.1257/aer.107.4.967

Shiller, R. J. (2019). *Narrative economics: How stories go viral & drive major economic events*. Princeton University Press.

Siliverstovs, B. (2011). Dating business cycles in a historical perspective: Evidence for switzerland. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1886347

Stineman, R. W. (1980). A consistently well behaved method of interpolation. *Creative Computing*, *6*(7), 54–57.

Stock, J. H., & Watson, M. W. (1989). New indexes of coincident and leading economic indicators. *NBER Macroeconomics Annual*, *4*, 351–394. https://doi.org/10.1086/654119

Stock, J. H., & Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, *20*(2), 147–162. https://doi.org/10.1198/073500102317351921

Stock, J. H., & Watson, M. W. (2016). Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics. *Handbook of macroeconomics* (pp. 415–525). Elsevier. https://doi.org/10.1016/bs.hesmac.2016.04.002

Stohr, C. (2016). Trading gains: New estimates of swiss GDP,1851 to 2008. *Economic History Working Papers No. 67032*.

Stohr, C. (2017). Das schweizer bruttoinlandprodukt: Methoden, daten und internationale vergleiche. *mimeo*. https://doi.org/10.13140/RG.2.2.29030.11848

Ter Ellen, S., Larsen, V. H., & Thorsrud, L. A. (2022). Narrative monetary policy surprises and the media. *Journal of Money, Credit and Banking*, *54*(5), 1525–1549. https://doi.org/10.1111/jmcb.12868

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, *62*(3), 1139–1168. https://doi.org/10.1111/j.1540-6261.2007.01232.x

Thorp, W. L. (1926). *Business annals*. National Bureau of Economic Research, Inc.

Thorsrud, L. A. (2020). Words are the new numbers: A newsy coincident index of the business cycle. *Journal of Business & Economic Statistics*, *38*(2), 393–409. https://doi.org/10.1080/07350015.2018.1506344

Van Binsbergen, J., Bryzgalova, S., Mukhopadhyay, M., & Sharma, V. (2024). (almost) 200 years of news-based economic sentiment. *NBER Working Paper No. 32026*. https://doi.org/10.3386/w32026

A Supplementary material

Table 1 — Textual data sources

Publication	Name	Source	Availability	Lang.	Freq.	# texts
AGB	Zeitschrift Schweizer Arbeitgeber	SWAB	1907 - 2014	de	M	124.94
ARC	ArcInfo	e-newspaper archives.ch	2018 - 2021	fr	D	143.05
AWP	AWP Financial News	AWP	2001 - 2020	de	D	1860.85
BAN	Briger Anzeiger	e-newspaper archives.ch	1899 - 1933	de	BW	204.22
BAZ	Basler Zeitung	Tamedia	2000 - 2021	de	D	323.94
BEOL	Berner Oberländer	Tamedia	2018 - 2021	de	D	76.22
BILA	Bilan	Tamedia	1996 - 2021	fr	D	18.08
BNN	Bündner Nachrichten	e-newspaper archives.ch	1885 - 1892	de	D	50.42
BR_DE_GB	Annual Report Swiss Confederation	Swiss Confederation	1849 - 2018	de	Y	306.57
BR_de_SR	Annual Report Government Account	Swiss Confederation	1849 - 2021	de	Y	3193.07
BR_FR_GB	Annual Report Swiss Confederation	Swiss Confederation	1848 - 2018	fr	Y	3170.38
BU	Der Bund	Tamedia	1995 - 2021	de	D	800.47
BUR	Bote vom Untersee und Rhein	e-newspaper archives.ch	1900 - 2018	de	BW	620.46
BZ	Berner Zeitung	Tamedia	1996 - 2021	de	D	1051.75
CMV	Aktiv : CMV/FCOM : Gemeinsames Gewerkschaftsmagazin	e-newspaper archives.ch	1930 - 1998	de	BM	40.83
CS_BUL	Credit Suisse Bulletin	Swiss National Library	1895 - 2016	de	Y	3121.80
DBB	Der Bund	e-newspaper archives.ch	1850 - 1994	de	D	2865.65
Démocrate	Le Démocrate	Scriptorium	1855 - 1999	fr	D	4141.31
DMR	Der Murtenbieter	e-newspaper archives.ch	1854 - 2010	de	BW	651.18
EDP	Engadiner Post	e-newspaper archives.ch	1893 - 1930	de	TW	92.23

Continued on next page

Table 1 – continued from previous page

Publication	Name	Source	Availability	Lang.	Freq.	# texts
EIZE	Eidgenössische Zeitung	e-newspaper archives.ch	1838 - 1864	de	D	73.19
EXR	FAN - L'express : feuille d'avis de Neuchâtel	e-newspaper archives.ch	1738 - 2018	fr	D	8410.07
FAAV	Feuille d'avis du district d'Avenches	Scriptorium	1882 - 2010	fr	D	5750.28
FCS	Feuille commerciale de Sierre et du district	e-newspaper archives.ch	1918 - 1930	fr	BW	50.74
FDV	Journal et feuille d'avis du Valais	e-newspaper archives.ch	1903 - 1968	fr	TW	1311.36
FUW	Finanz und Wirtschaft	Tamedia	2000 - 2021	de	D	93.47
FZG	Freiburger Nachrichten	e-newspaper archives.ch	1864 - 2006	de	D	2638.32
GAV	Gazette du Valais	e-newspaper archives.ch	1855 - 1922	fr	TW	258.07
GBL	Geschäftsblatt für den obern Teil des Kantons Bern	e-newspaper archives.ch	1876 - 1938	de	BW	339.40
GDB	Gazette de Berne	e-newspaper archives.ch	1692 - 1797	fr	BW	69.08
GTR	Grütlianer	e-newspaper archives.ch	1852 - 1925	de	BW	229.89
HAZ	Handels-Zeitung	SWAB	1861 - 1894	de	D	36.38
HEU	24 Heures	Tamedia	1996 - 2021	fr	D	535.01
IMP	L'impartial	e-newspaper archives.ch	1881 - 2018	fr	D	7102.04
IND	L'indicateur = Der Anzeiger	e-newspaper archives.ch	1914 - 1923	fr	W	48.32
ISB	Intelligenzblatt für die Stadt Bern	e-newspaper archives.ch	1834 - 1922	de	D	579.98
JC	Journal du district de Cossonay	Scriptorium	1899 - 2007	fr	D	3353.52
JM	Journal de Morges	Scriptorium	1894 - 2017	fr	D	9032.77
KK_CA	Reports Kommission für Konjunkturbeobachtung	SNB	1953 - 1983	de	Y	7.81
KK_ES	Econmic Situation report Kommision für Konjunkturbeobachtung	SNB	1932 - 2001	de	Y	47.15

Continued on next page

Table 1 – continued from previous page

Publication	Name	Source	Availability	Lang.	Freq.	# texts
KK_MIT	Mitteilungen Kommission für Konjunkturbeobachtung	SNB	1933 - 2007	de	Y	145.69
LAT	Langenthaler Tagblatt	Tamedia	2019 - 2021	de	D	26.99
LB	Der Landbote	Tamedia	1998 - 2021	de	D	446.08
LBP	Le bien public	e-newspaper archives.ch	1879 - 1888	fr	TW	70.05
LCE	Le confédéré	e-newspaper archives.ch	1861 - 2009	fr	W/BW	701.46
LCG	Le confédéré de Fribourg	e-newspaper archives.ch	1848 - 1907	fr	BW	169.57
LCR	Le chroniqueur	e-newspaper archives.ch	1854 - 1881	fr	TW	96.29
LEJ	Le Jura	e-newspaper archives.ch	1852 - 1970	fr	BW	182.35
LES	L'essor	e-newspaper archives.ch	1906 - 2015	fr	M	51.82
LFM	Le Franc-Montagnard	e-newspaper archives.ch	1898 - 2020	fr	BW/TW	533.68
LGE	La Gruyère	e-newspaper archives.ch	1882 - 1930	fr	TW	158.79
LLE	La liberté	e-newspaper archives.ch	1871 - 2012	fr	D	6975.28
LSL	La Suisse libérale	e-newspaper archives.ch	1881 - 1982	fr	W	304.98
MIG	Construire : hebdomadaire du capital à but social	e-newspaper archives.ch	1944 - 2004	fr	W	399.06
MIM	Wir Brückenbauer : Wochenblatt des sozialen Kapitals	e-newspaper archives.ch	1943 - 2004	de	W	479.30
NVB	Nidwaldner Volksblatt	e-newspaper archives.ch	1866 - 1991	de	BW	236.52
NVE	Le nouvelliste	e-newspaper archives.ch	1904 - 2016	fr	D	7505.93
NZG	Neue Zuger Zeitung	e-newspaper archives.ch	1846 - 1891	de	BW	25.10
NZN	Neue Zürcher Nachrichten	e-newspaper archives.ch	1895 - 1991	de	D	1412.94
NZZ	Neue Zürcher Zeitung	e-newspaper archives.ch	1780 - 2020	de	D	3698.76

Continued on next page

Table 1 – continued from previous page

Publication	Name	Source	Availability	Lang.	Freq.	# texts
OIZ	Die Gewerkschaft : das Magazin der Gewerkschaft Textil, Chemie, Papier	e-newspaper archives.ch	1901 - 1992	de	BM	51.08
OTB	Oberländer Tagblatt	e-newspaper archives.ch	1909 - 1961	de	D	371.66
SGZ	St. Galler Zeitung	e-newspaper archives.ch	1831 - 1881	de	D	173.81
SLB	Seeländer Bote	e-newspaper archives.ch	1850 - 1904	de	TW	75.30
SMZ	SMUV-Zeitung / Schweizerischer Metall- und Uhrenarbeitnehmer-Verband	e-newspaper archives.ch	1902 - 2001	de	M	44.69
SNB_GB	Annual Report SNB	SNB	1908 - 2018	de	Y	736.58
SNB_GMB	Money Market Report SNB	SNB	1945 - 1985	de	Y	58.09
SNB_QB	Quarterly Report SNB	SNB	1967 - 2019	de	Q	46.70
SWA_BER_Vorort	Annual Report Vorort	SWAB	1878 - 1976	de	Y	2063.75
SWA_GB_BCG	Annual Report Banque du Commerce Geneve	SWAB	1845 - 1907	fr	Y	85.54
SWA_GB_BCN	Annual Report Banque cantonale Neuchâteloise	SWAB	1883 - 1989	fr	Y	166.83
SWA_GB_BCV	Annual Report Banque cantonale vaudoise	SWAB	1854 - 1949	fr	Y	25.48
SWA_GB_BCVL	Annual Report Banque cantonale du valais	SWAB	1856 - 1871	fr	Y	11.27
SWA_GB_BDG	Annual Report Banque de Genève	SWAB	1858 - 1930	fr	Y	36.24
SWA_GB_BGG	Annual Report Banque Glane/Gruyere	SWAB	1888 - 1960	fr	Y	4.45
SWA_GB_BHCG	Annual Report Banque hypothécaire du canton de Genève	SWAB	1848 - 1950	fr	Y	108.56
SWA_GB_BPG	Annual Report Banque populaire de la Gruyère	SWAB	1864 - 1968	fr	Y	58.07

Continued on next page

Table 1 – continued from previous page

Publication	Name	Source	Availability	Lang.	Freq.	# texts
SWA_GB_CAIB	Annual Report Crédit agricole et industriel de la Broye	SWAB	1867 - 1965	fr	Y	103.68
SWA_GB_CCIG	Annual Report Chambre de Commerce de Genève	SWAB	1865 - 1930	fr	Y	80.12
SWA_GB_CEB	Annual Report Caisse d'epargne de Bassecourt	SWAB	1883 - 1988	fr	Y	59.58
SWA_GB_CECL	Annual Report Caisse d'epargne et de crédit Lausanne	SWAB	1878 - 1956	fr	Y	35.91
SWA_GB_CEPL	Annual Report Caisse d'Epargne et de Prévoyance de Lausanne	SWAB	1905 - 1960	fr	Y	7.80
SWA_GB_CF	Annual Report Credit foncier vaudois	SWAB	1901 - 1956	fr	Y	41.84
SWA_GB_CHCF	Annual Report Caisse Hypo Fribourg	SWAB	1854 - 1909	fr	Y	4.58
SWA_GB_CHCV	Annual Report Caisse hypothécaire Cantonale vaudoise	SWAB	1860 - 1900	fr	Y	34.26
SWA_GB_CME	Annual Report Caisse Mutuel pour l'Epargne	SWAB	1874 - 1920	fr	Y	17.42
SWA_GB_CVCI	Annual Report Chambre vaudoise du commerce et de l'industrie	SWAB	1904 - 1929	fr	Y	2.90
SWA_GB_Helveti	Annual Report Helvetia	SWAB	1859 - 1957	de	Y	4.18
SWA_GB_SLKB	Annual Report Spar- und Leihkasse Bern	SWAB	1858 - 1960	de	Y	46.59
SWA_GB_SwissR	Annual Report Swiss RE	SWAB	1864 - 1960	de	Y	33.26
SWA_GB_ZVAO	Annual Report Arbeitgeberverband	SWAB	1910 - 1960	de	Y	94.63
TA	Tages-Anzeiger	Tamedia	1996 - 2021	de	D	994.33
TAA	Täglicher Anzeiger für Thun und das Berner Oberland	e-newspaper archives.ch	1877 - 1907	de	D	68.06

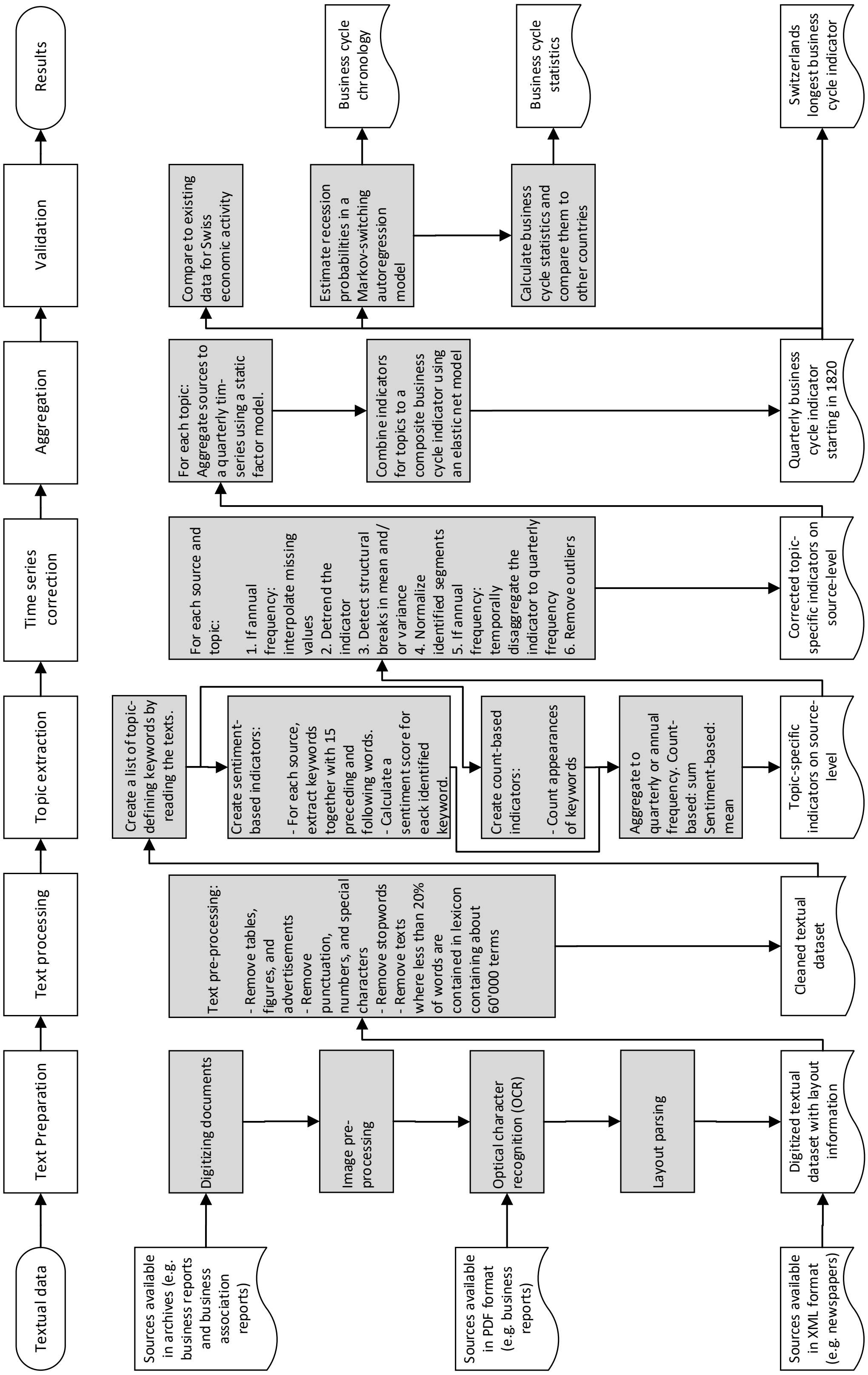
Continued on next page

Table 1 – continued from previous page

Publication	Name	Source	Availability	Lang.	Freq.	# texts
TCP	FTCP : organe de la Fédération du personnel du textile, de la chimie et du papier	e-newspaper archives.ch	1951 - 1979	fr	BM	23.40
TDG	Tribune de Genève	Tamedia	1996 - 2021	fr	D	592.38
THT	Thuner Tagblatt	Tamedia	2019 - 2021	de	D	28.61
TSB	Tagblatt der Stadt Biel	e-newspaper archives.ch	1865 - 1900	de	D	115.44
VHT	VHTL-Zeitung / Gewerkschaft Verkauf, Handel, Transport, Lebensmittel	e-newspaper archives.ch	1904 - 2004	de	W/BM	21.98
WAB	Walliser Bote	e-newspaper archives.ch	1861 - 2008	de	D	2494.40
ZFZ	Zürcherische Freitagszeitung	e-newspaper archives.ch	1705 - 1914	de	W	53.95
ZGN	Zuger Nachrichten	e-newspaper archives.ch	1886 - 1900	de	TW	18.80
ZHUL	Zürcher Unterländer	Tamedia	2011 - 2021	de	D	179.52
ZKB_GB	Annual Reports Zürcher Kantonalbank	ZKB	1870 - 2016	de	Y	733.79
ZSZ	Zürichsee-Zeitung	Tamedia	2001 - 2021	de	D	549.34
ZVB	Zuger Volksblatt	e-newspaper archives.ch	1861 - 1900	de	BW	32.11
ZWB	Zürcherisches Wochenblatt	e-newspaper archives.ch	1801 - 1842	de	BW	33.43

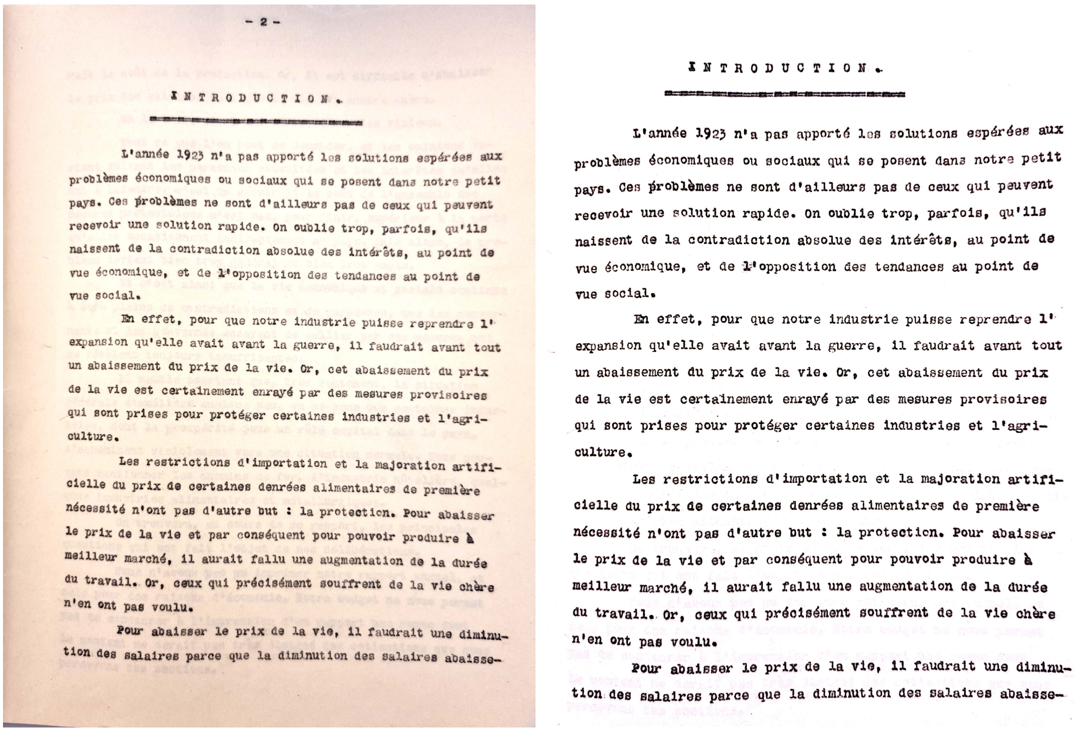
Notes: Publications with source SWAB were digitized by ourselves (except AGB and HAZ).
Frequencies: D: daily, TW: three a week, BW: twice a week, W: weekly, BM: twice a month, M: monthly, Q: quarterly, Y: yearly.
Number of texts is in thousands.

Figure 1 — Schematic diagram of the study's workflow



Notes: The diagram shows the schematic workflow of the study. Wavy boxes denote different states of the textual data. Gray-shaded boxes represent tasks that process the data.

Figure 2 — Scanned page, a pre-processed image, and the OCR text: Chambre vaudoise du commerce et de l'industrie, 1923.



\- 2 -

INTRODUCTION.

L'année 1923 n'a pas apporté les solutions espérées aux problèmes économiques ou sociaux qui se posent dans notre petit pays. Ces problèmes ne sont d'ailleurs pas de ceux qui peuvent recevoir une solution rapide. On oublie trop, parfois, qu'ils naissent de la contradiction absolue des intérêts, au point de vue économique, et de l'opposition des tendances au point de vue social.

En effet, pour que notre industrie puisse reprendre l'expansion qu'elle avait avant la guerre, il faudrait avant tout un abaissement du prix de la vie. Or, cet abaissement du prix de la vie est certainement enrayé par des mesures provisoires qui sont prises pour protéger certaines industries et l'agriculture.

Les restrictions d'importation et la majoration artificielle du prix de certaines denrées alimentaires de première nécessité n'ont pas d'autre but : la protection. Pour abaisser le prix de la vie et par conséquent pour pouvoir produire à meilleur marché, il aurait fallu une augmentation de la durée du travail. Or, ceux qui précisément souffrent de la vie chère n'en ont pas voulu.

Pour abaisser le prix de la vie, il faudrait une diminution des salaires parce que la diminution des salaires abaisse-

Notes: Source: Schweizerisches Wirtschaftsarchiv Basel.

Result from OCR of figure 2
INTRODUCTION*.

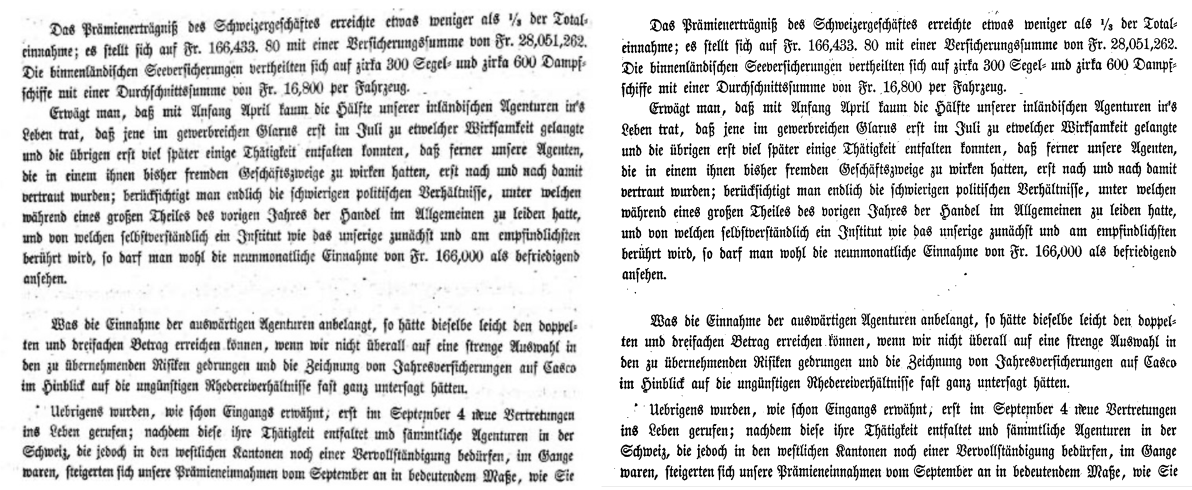
L'année 1923 n'a pan apporté les solutions espérées aux problèmes économiques ou sociaux qui so posent dans notre petit pays* Ces problèmes ne sont d'ailleurs pas de ceux qui peuvent recevoir une solution rapide, on ou'olie trop, parfois, qu'Ils naissent de la contradiction absolue des intérêts, au point de vue économique, et de *3?'opposition des tendances au point de vue social»

En effet, pour que notre industrie puisse reprendre l'expansion qu'elle avait avant la guerre, il faudrait avant tout un abaissement du prix de la vie. Or, cet abaissement du prix de la vie est certainement enrayé par des mesures provisoires qui sont prises pour protéger certaines industries et l'agriculture.

Les restrictions d'importation et la majoration artificielle du prix de certaines denrées alimentaires de première nécessité n'ont pas d'autre but : la protection. Pour abaisser le prix de la vie et par conséquent pour pouvoir produire à meilleur marché, il aurait fallu une augmentation de la durée du travail. Or, ceux qui précisément souffrent de la vie chère n'en ont pas voulu.

Pour abaisser le prix de la vie, il faudrait une diminution des salaires parce que la diminution des salaires abaisse-

Figure 3 — Scanned page, a pre-processed image, and the OCR text: Erster Geschäftsbericht, Helvetia, 1860.



Das Prämienerträgniß des Schweizergeschäftes erreichte etwas weniger als 1/3 der Totaleinnahme; es stellt sich auf Fr. 166,433. 80 mit einer Versicherungssumme von Fr. 28,051,262. Die binnenländischen Seeversicherungen vertheilten sich auf zirka 300 Segel- und zirka 600 Dampfschiffe mit einer Durchschnittssumme von Fr. 16,800 per Fahrzeug.

Erwägt man, daß mit Anfang April kaum die Hälfte unserer inländischen Agenturen in's Leben trat, daß jene im gewerbreichen Glarus erst im Juli zu etwelcher Wirksamkeit gelangte und die übrigen erst viel später einige Thätigkeit entfalten konnten, daß ferner unsere Agenten, die in einem ihnen bisher fremden Geschäftszweige zu wirken hatten, erst nach und nach damit vertraut wurden; berücksichtigt man endlich die schwierigen politischen Verhältnisse, unter welchen während eines großen Theiles des vorigen Jahres der Handel im Allgemeinen zu leiden hatte, und von welchen selbstverständlich ein Institut wie das unserige zunächst und am empfindlichsten berührt wird, so darf man wohl die neunmonatliche Einnahme von Fr. 166,000 als befriedigend ansehen.

Was die Einnahme der auswärtigen Agenturen anbelangt, so hätte dieselbe leicht den doppelten und dreifachen Betrag erreichen können, wenn wir nicht überall auf eine strenge Auswahl in den zu übernehmenden Risiken gedrungen und die Zeichnung von Jahresversicherungen auf Casco im Hinblick auf die ungünstigen Rhedereiverhältnisse fast ganz untersagt hätten.

Uebrigens wurden, wie schon Eingangs erwähnt, erst im September 4 neue Vertretungen ins Leben gerufen; nachdem diese ihre Thätigkeit entfaltet und sämmtliche Agenturen in der Schweiz, die jedoch in den westlichen Kantonen noch einer Vervollständigung bedürfen, im Gange waren, steigerten sich unsere Prämieneinnahmen vom September an in bedeutendem Maße, wie Sie

Notes: Source: Schweizerisches Wirtschaftsarchiv Basel.

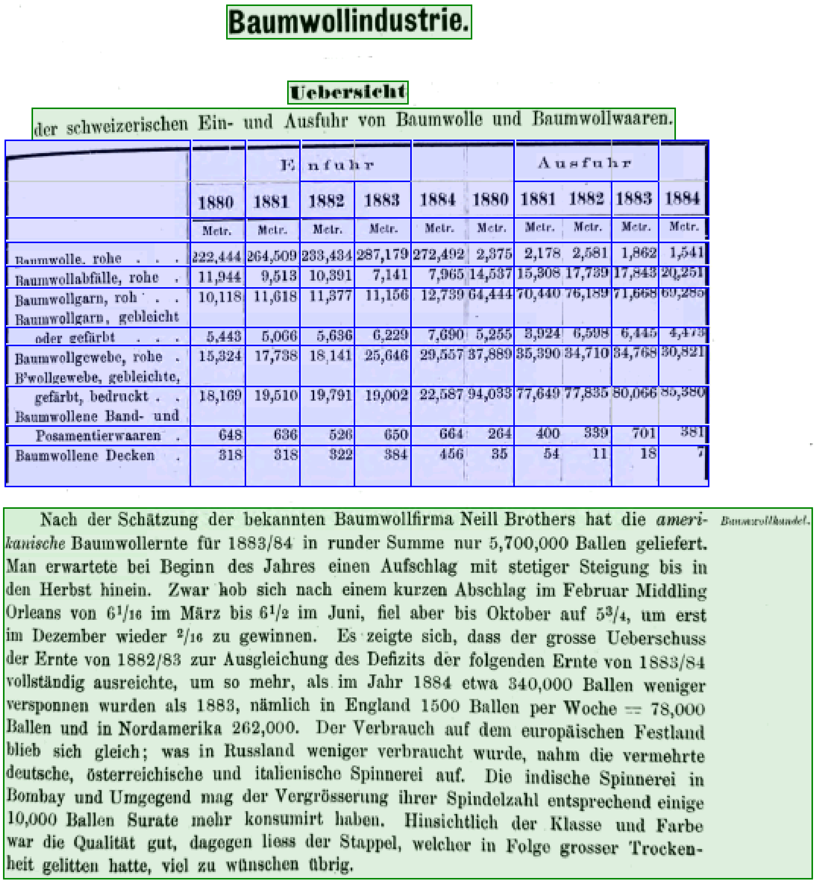
Result from OCR of figure 3

Das Prämienerträgniß des Schweizergeschäftes erreichte etwas weniger als !/; der Totaleinnahme; es stellt sich auf Fr. 166,433. 80 mit einer Versicherungssumme von Fr. 28,051,262. Die binnenländis hen Seeversicherungen vertheilten sich auf zirka 300 Segel- und zirka 600 Dampfschiffe mit einer Durchschnittssumme von Fr. 16,800 per Fahrzeug.

Erwägt man, daß mit Anfang April kaum die Hälfte unserer inländischen Agenturen in's Leben trat, daß jene im gewerbreichen Glarus erst im Juli zu etwelcher Wirksamkeit gelangte und die übrigen erst viel später einige Thätigkeit entfalten konnten, daß ferner unsere Agenten, die in einem ihnen bisher fremden Geschäftszweige zu wirken hatten, erst nach und nach damit vertraut wurden; berücsichtigt man endlich die schwierigen politischen Verhältnisse, unter welchen während eines großen Theiles des vorigen Jahres der Handel im Allgemeinen zu leiden hatte, und von welchen selbstverständlich ein Jnstitut wie das unserige zunächst und am empfindlichsten berührt wird, so darf man wohl die neunmonatliche Einnahme von Fr. 166,000 als befriedigend anjehen.

Was die Einnahme der auswärtigen Agenturen anbelangt, so hätte dieselbe leicht den doppelten und dreifachen Betrag erreichen können, wenn wir nicht überall auf eine strenge Auswahl in den zu übernehmenden Risiken gedrungen und die Zeichnung von Jahresversicherungen auf Casco im Hinbli> auf die ungünstigen Rhedereiverhältnisse fast ganz untersagt hätten. * "Vebrigens wurden, wie shon Eingangs erwähnt; erst im September 4 neue Vertretungen ins Leben gerufen; nachdem diese ihre Thätigkeit entfaltet und sämmtliche Agenturen in der Schweiz, die jedoch in den westlichen Kantonen no< einer Vervollständigung bedürfen, im Gange waren, steigerten sich unsere Prämieneinnahmen vom September an in bedeutendem Maße, wie Sie

Figure 4 — Parsing the layout: Handels- und Industrieverein Vorort, 1923



Baumwollindustrie.

Uebersicht

der schweizerischen Ein- und Ausfuhr von Baumwolle und Baumwollwaaren.

	Einfuhr					Ausfuhr				
	1880	1881	1882	1883	1884	1880	1881	1882	1883	1884
	Metr.	Metr.	Metr.	Metr.	Metr.	Metr.	Metr.	Metr.	Metr.	Metr.
Baumwolle, rohe . . .	222,444	264,509	233,434	287,179	272,492	2,375	2,178	2,581	1,862	1,541
Baumwollabfälle, rohe .	11,944	9,513	10,391	7,141	7,965	14,537	15,308	17,739	17,843	20,251
Baumwollgarn, roh . .	10,118	11,618	11,377	11,156	12,739	64,444	70,440	76,189	71,668	69,285
Baumwollgarn, gebleicht oder gefärbt . . .	5,443	5,066	5,636	6,229	7,690	5,255	3,924	6,598	6,445	4,473
Baumwollgewebe, rohe .	15,324	17,738	18,141	25,646	29,557	37,889	35,390	34,710	34,768	30,821
B'wollgewebe, gebleichte, gefärbt, bedruckt . .	18,169	19,510	19,791	19,002	22,587	94,033	77,649	77,835	80,066	85,380
Baumwollene Band- und Posamentierwaaren .	648	636	526	650	664	264	400	339	701	381
Baumwollene Decken .	318	318	322	384	456	35	54	11	18	7

Nach der Schätzung der bekannten Baumwollfirma Neill Brothers hat die *amerikanische* Baumwollernte für 1883/84 in runder Summe nur 5,700,000 Ballen geliefert. Man erwartete bei Beginn des Jahres einen Aufschlag mit stetiger Steigung bis in den Herbst hinein. Zwar hob sich nach einem kurzen Abschlag im Februar Middling Orleans von $6^1/_{16}$ im März bis $6^1/_2$ im Juni, fiel aber bis Oktober auf $5^3/_4$, um erst im Dezember wieder $^2/_{16}$ zu gewinnen. Es zeigte sich, dass der grosse Ueberschuss der Ernte von 1882/83 zur Ausgleichung des Defizits der folgenden Ernte von 1883/84 vollständig ausreichte, um so mehr, als im Jahr 1884 etwa 340,000 Ballen weniger versponnen wurden als 1883, nämlich in England 1500 Ballen per Woche = 78,000 Ballen und in Nordamerika 262,000. Der Verbrauch auf dem europäischen Festland blieb sich gleich; was in Russland weniger verbraucht wurde, nahm die vermehrte deutsche, österreichische und italienische Spinnerei auf. Die indische Spinnerei in Bombay und Umgegend mag der Vergrösserung ihrer Spindelzahl entsprechend einige 10,000 Ballen Surate mehr konsumirt haben. Hinsichtlich der Klasse und Farbe war die Qualität gut, dagegen liess der Stappel, welcher in Folge grosser Trockenheit gelitten hatte, viel zu wünschen übrig. *Baumwollhandel.*

Notes: This figure illustrates that it is possible to parse the layout (i.e., titles, tables and paragraphs) of a given text. Source: Schweizerisches Wirtschaftsarchiv Basel.

Table 2 — Validation data sources

Publication	Name	Source	Availability	Freq.
HSSODEF	Real GDP (deflator-adjusted)	Historische Statistik der Schweiz HSSO (2012k)	1851 - 1890	Y
HSSOCPI	Real GDP (CPI-adjusted)	Historische Statistik der Schweiz HSSO (2012k)	1851 - 1890	Y
HSSO	Nominal GDP	Historische Statistik der Schweiz HSSO (2012k)	1851 - 1890	Y
HSSO16A	Real and nominal GDP	Historische Statistik der Schweiz HSSO (2012c)	1890 - 1948	Y
HSSO16B	Real and nominal GDP	Historische Statistik der Schweiz HSSO (2012d)	1948 - 2005	Y
MADDISON	Real GDP	Bolt and van Zanden (2020)	1820 - 2018	Y
FSOGDP	Real and nominal GDP	Federal Statistical Office (FSO)	1948 - 2021	Y
SECOGDP	Real and nominal GDP	State Secretariat for Economic Affairs (SECO)	1980 - 2022	Q
CHAINR	Chained real GDP	Maddison, HSSO, FSO, SECO	1851 - 2021	Y
CHAINN	Chained nominal GDP	HSSO, FSO, SECO	1851 - 2021	Y
RGDPWORLD	Real GDP world		1830 - 2022	Y
RGDPCapita	Real GDP per capita	HSSO	1851 - 2021	Y
OutputGapHP	HP-filtered CHAINR	see CHAINR	1851 - 2021	Y
OutputGapHam	Hamilton-filtered CHAINR	see CHAINR	1851 - 2021	Y
GDPSTOHR	Real and nominal GDP	Stohr (2016)	1851 - 2008	Y
RGDPINT	Real GDP international (Austria, England, France, Germany, Italy, USA)	Bolt and van Zanden (2020)	1851 - 2020	Y
IMPVOL	Import volumes	Historische Statistik der Schweiz HSSO (2012e)	1851 - 1913	Y
EXPVOLBAI	Export volumes index (Bairoch)	Historische Statistik der Schweiz HSSO (2012a)	1851 - 1913	Y
EXPVOLBER	Export volumes index (Bernegger)	Historische Statistik der Schweiz HSSO (2012a)	1851 - 1913	Y
IMPVAL	Import values	Historische Statistik der Schweiz HSSO (2012b)	1886 - 1992	Y
EXPVAL	Export values	Historische Statistik der Schweiz HSSO (2012b)	1886 - 1992	Y

Continued on next page

Table 2 – continued from previous page

Publication	Name	Source	Availability	Freq.
IMPVOLM	Import volumes	Historische Statistik der Schweiz HSSO (2012i)	1924 - 1987	M
EXPVOLM	Export volumes	Historische Statistik der Schweiz HSSO (2012g)	1924 - 1967	M
IMPVALM	Import values	Historische Statistik der Schweiz HSSO (2012j)	1924 - 1992	M
EXPVALM	Export values	Historische Statistik der Schweiz HSSO (2012h)	1924 - 1992	M
IMPVALEZV	Import values	Federal Office for Customs and Border Security (FOCBS)	1988 - 2022	M
EXPVALEZV	Export values	Federal Office for Customs and Border Security (FOCBS)	1988 - 2022	M
TradeIndicator	First principal component of trade related indexes	HSSO, FOCBS	1851 - 2022	Y
THORPGPT	Created sentiment indicators (Austria, England, France, Germany, Italy, USA)	Thorp (1926), GPT-3.5	1800 - 1925	Y
HSSOUNEMP	Unemployment rate	Historische Statistik der Schweiz HSSO (2012l)	1913 - 1995	Y
HSSOEMP	Employment growth	Historische Statistik der Schweiz HSSO (2012f)	1890 - 2005	Y
FSOUNEMP	Unemployment rate	FSO	1970 - 2021	M
FSOEMP	Employment growth	FSO	1998 - 2021	M
ECRIREC	Recession indicator	Economic Cycle Research Institute (ECRI)	1956 - 2022	M
OECDREC	Recession indicator	Organisation for Economic Co-operation and Development (OECD), Fred: CHEREC	1960 - 2022	M

Notes: This table shows sources, available periods and frequency of the validation data. Most series are further spliced to a long time series spanning as many periods as possible.

Table 3 — Keywords defining economic concepts in German

Topic	Keywords (based on readings)	Method
Real activity	wirtschaft ware absatz nachfrage geschäft konsum waare fabrikant erlös umsatz markt industrie branche käufer unternehmer ernte ergebniss konjunktur kundschaft verkauf produktion dienstleistung verarbeitung gewerbe ertrag einnahmen ausgaben fabrikation bestellung versorgung materialbeschaffung einkäufer verlust konkurrenten fabrizieren fabrikat besteller werth neugründung materialien betriebsmittel materialeinkäufe jahresresultat geschäftsperiode werkstätten erfolg bestellungen eigenkosten produkte rohmaterialien einbusse fabric fabrik herstellung geschäftsgang wirtschaftsleben wirtschaftlich nachfrage geschäftslage marge angebot erträgnis rendite produzent vertrieb volkswirtschaft konjunkturrückgang verkaufsziffern kauflust kaufunlust geschäftsjahr kleinbetrieb bautätigkeit verbraucher konkurrenz erzeugnis konsum profit fremdenverkehr dienste dienstleistungen investition versorgungsmöglichkeit versorgungslage versorgungsschwierig bruttoinlandprodukt bruttosozialprodukt realwachstum wertschöpfung	KWIC
Trade	eingangszölle eingangszoll konkurrenzverhältnisse konkurrenzverhältnis einfuhr ausfuhr export import sendungen aussendung importeure exporteure handelsstatistik absatzfeld wettbewerb absatzgebiet zwischenhandel handel importhaus zollverhältnisse zollverhältnis handelsbilanz waarenverkehr warenverkehr waarenausfuhr warenausfuhr waareneinfuhr wareneinfuhr importhandel handelsverkehr zoll zolleinnahmen zölle weltbedarf fracht exportziffern gesammtexport gesamtexport taxen verkehrserleichterung bezugsquelle ausland generaltarif tarif einfuhrverbote zufuhren zufuhr grosshändler seefracht wasserweg welthandel weltverkehr güterstrom güteraustausch güterumschlag umschlagverkehr wagenverkehr	KWIC
Capacity	lager kornspeicher speicher ueberproduktion überproduktion vorräthe vorräte liefertermine lieferfrist vorrat vorrath aufträge lieferfristen lagerware lager depots bestellungen lieferungen wagenmangel	KWIC
Labor	arbeit erwerb beruf erwerbende arbeiter aufsichtspersonal arbeiterin angestellte arbeitskräfte beschäftigung arbeitszeit arbeitgeber arbeiterschaft ueberzeit überzeit arbeitstag arbeitsleistung ausbildung lehrlinge ueberzeitarbeit überzeitarbeit streik arbeitseinstellung arbeitsfeld thätigkeit tätigkeit arbeiterinnen personal	KWIC

Continued on next page

Table 3 – continued from previous page

Topic	Keywords (based on readings)	Method
Inflation	preis getreidepreise theuer teuer preisfall kostenpreise wechselkurs silberpreis preisaufschlag silberkurs preisbasis kurs preissturz baumwollpreise preisnotierung preisnotirung waarenpreise warenpreise kursschwankungen kostenpreis preissteigerung schleuderpreise abschlag materialpreise maschinenpreise unterbietung preisvorteile preisvortheile verkaufspreise rohpreise preiserhöhung entwerthung entwertung goldkurs geldvertheuerung geldverteuerung geldwerth geldwert vertheuerung verteuerung wechselkurse pari preise teuerung kaufkraft aufschlag inflation inflatorisch	KWIC
Wages	entgelt lohnerhöhung arbeitslöhne löhne arbeitslohn kaufkraft lohnverhältnisse lohnsätze lohnansätze stundenlöhne stundenlohn akkordlöhne akkordlohn akkord tagesverdienst löhnungen löhnung lohnverhältnis verdienstverhältnisse verdienstverältnis einkommen lohnniveau lohnstopp lohnpolitik	KWIC
Credit	kreditverhältnisse banknoten münzen einleger guthaben einzahlung rückzahlung prämien renten kapitalien verzinsung amortisation wechsel geldmarkt zinsen zins rendite diskontsätze disconto discontsätze diskontosätze diskontosatz diskontsatz geldkraft zinsfuss wechselgeschäft wechselkverkehr zinssätze zinssatz diskontopolitik geldinstitute lombardvorschüsse notenemission vorschüsse kontokorrent kreditwirtschaft kredit schuldbriefe kreditwesen leihmarkt emissionsbanken diskonto geldbedarf diskonti diskontoverkehr geldsuchenden emissionen geldverteuerung liquidität geldknappheit geldstand lombardsatz geldleihpreis geldmärkte kapitalmarkt anleihen hypothek depositen darlehen pfandbrief geldpolitik geldmenge	KWIC
Financial	kapitalvermehrung emissionskurs tageskurs agio kurse dividenden emittirt emittiert obligationen rentabilität konversion börsengeschäft kapital werthpapiere wertpapiere titelverkäufe papiere kurssteigerung portefeuille tratten ueberspekulation überspekulation entwerthung entwertung finanzcrisis finanzkrisis börse emission gründung aktie actie aktien kurs effekten märkte wechselkurs devisen valoren dividende wertschriften	KWIC
Recession	crisis krise rezession recession finanzcrisis finanzkrisis spekulationskrise krach krisis zahlungsfähigkeit fallimente bankerottir liegenschaftenkrisis konkurse schaden konjunkturrückgang weltbrand notstand valutasturz depression valutaschwierigkeit wirtschaftskatastrophe schäden liquidation liquidier hemmnisse zusammenbruch notlage katastrophe baisse	Count
Boom	hausse aufschwung hochkonjunktur prosperität erholungsperiode boom	Count

Continued on next page

Table 3 – continued from previous page

Topic	Keywords (based on readings)	Method
War	krieg konflikt putsch	Count

Notes: The table contains keywords used to identify economic concepts in German. The most right column indicates the method used to identify the keywords. KWIC stands for keyword-in-context, which means that a sentiment score is calculated from the words surrounding the keyword. Count indicates that the indicator is created by counting the appearances of these keywords. Terms embedded in \b are used to avoid counting words that contain the keyword, but are not the keyword itself. For example, \bertrag\b (in english: revenue) is used to avoid counting the word "ertragen" (endure) as a keyword. For the other words, wildcards are used, which means that all words containing e.g. "import" are counted as a keyword.

Table 4 — Keywords defining economic concepts in French

Topic	Keywords (based on readings)	Method
Real activity	économie marchandise vente \bdemande\b affaire consommation marchandis fabricant revenu horlogerie marchand industrie branche acheteur entrepreneur récolte marcheé conjoncture clientèle production \bservice\b \bachat\b rendement recettes dépenses fabrication commande approvisionnement perte concurrents fabriquer produit commanditaire matériaux achats commandes \bcoûts\b produits fabriq usine fabric économique demande récoltes marge \boffre\b \bmarché\b producteur distribution economie ventes economique \bacheter\b fiscal entreprise consommateur concurrence profit tourisme revenus services investissement produite pib \bcroissance\b	KWIC
Trade	importation exportation export import importateurs exportateurs commerc commerce douani marchandises douane fret taxes tarif approvisionnements approvisionnement grossistes marchandise transbordement	KWIC
Capacity	entrepôt silo stockage surproduction réserves livraison réserve commandes marchandise dépôts livraisons \bstock\b	KWIC
Labor	travail gain métier travailleurs ouvrier personnel ouvrière employée main-d'œuvre emploi employeur formation apprentis grève activité ouvrières	KWIC
Inflation	prix taux cours rabais dévalorisation monétaire renchérissement parité majoration inflation inflationniste	KWIC
Wages	rémunération salaire salaires salariaux accord rémunérations revenu	KWIC
Credit	crédit banque monnaie déposants solde versement remboursement primes rentes capitaux intérêt amortissement change monétaire intérêts rendement taux escompte lombard émission \bavances\b créance argent émissions liquidité capital emprunts hypothèque dépôts prêts	KWIC
Financial	capital spéculation \bémission\b cours agio dividendes émis obligations rentabilité conversion bourse titres papiers portefeuille trattes sur-spéculation dévalorisation financière financi émission fondation action actions \bcours\b effets \bmarchés\b devises valeurs dividende	KWIC
Recession	crise récession krach solvabilité faillites faillite dommages conflagration dépression catastrophe liquidation liquider effondrement baisse	Count
Boom	hausse expansion prospérité rétablissement boom	Count

War	guerre conflit putsch	Count

Notes: The table contains keywords used to identify economic concepts in French. The most right column indicates the method used to identify the keywords. KWIC stands for keyword-in-context, which means that a sentiment score is calculated from the words surrounding the keyword. Count indicates that the indicator is created by counting the appearances of these keywords. Terms embedded in \b are used to avoid counting words that contain the keyword, but are not the keyword itself. For example, \bdemande\b (in english: demand) is used to avoid counting the word "demander" (asking) as a keyword. For the other words, wildcards are used, which means that all words containing e.g. "import" are counted as a keyword.

Figure 5 — Document-level sentiment score

Lead text of FUW from March 6, 2020

Before cleaning
Das Coronavirus trifft die Schweizer Wirtschaft ab dem Frühjahr mit voller Kraft. Volkswirte stimmen auf ein schwaches zweites Quartal ein. Konsum, Tourismus und Exportindustrie leiden bereits. Der Bundesrat sollte die Kurzarbeit ausweiten.

After cleaning
coronavirus trifft schweizer wirtschaft frühjahr voller kraft volkswirte stimmen schwaches zweites quartal konsum tourismus exportindustrie leiden bereits bundesrat kurzarbeit ausweiten

---

In English

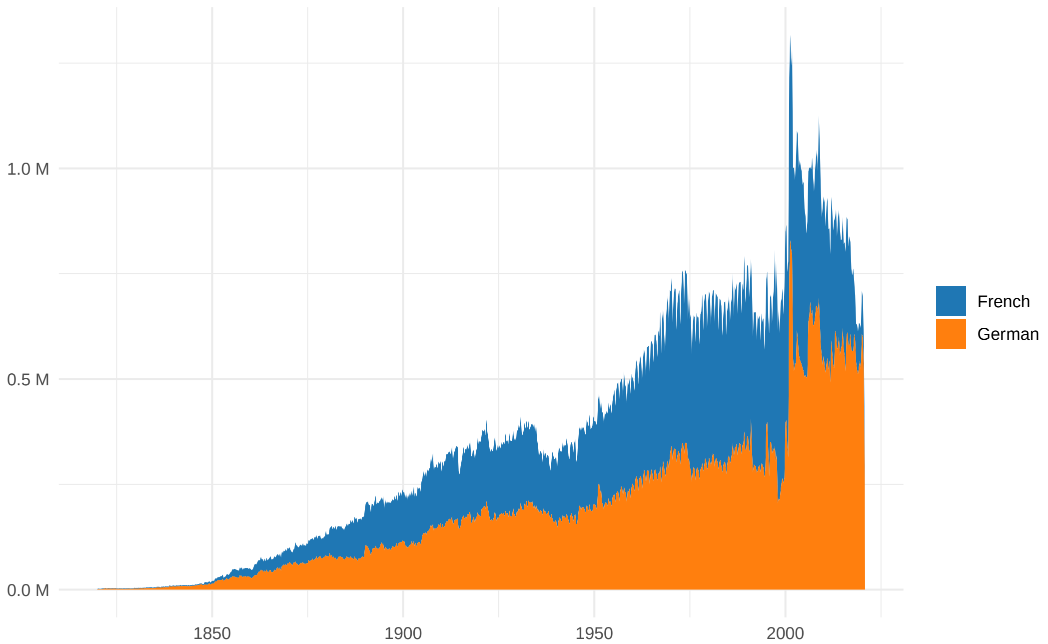
Before cleaning
The corona virus is hitting the Swiss economy with full force from the spring. Economists are predicting a weak second quarter. Consumption, tourism and the export industry are already suffering. The Federal Council should extend short-time work.

After cleaning
corona virus hitting swiss economy spring full force economists predicting weak second quarter consumption tourism export industry already suffering federal council extend short time work

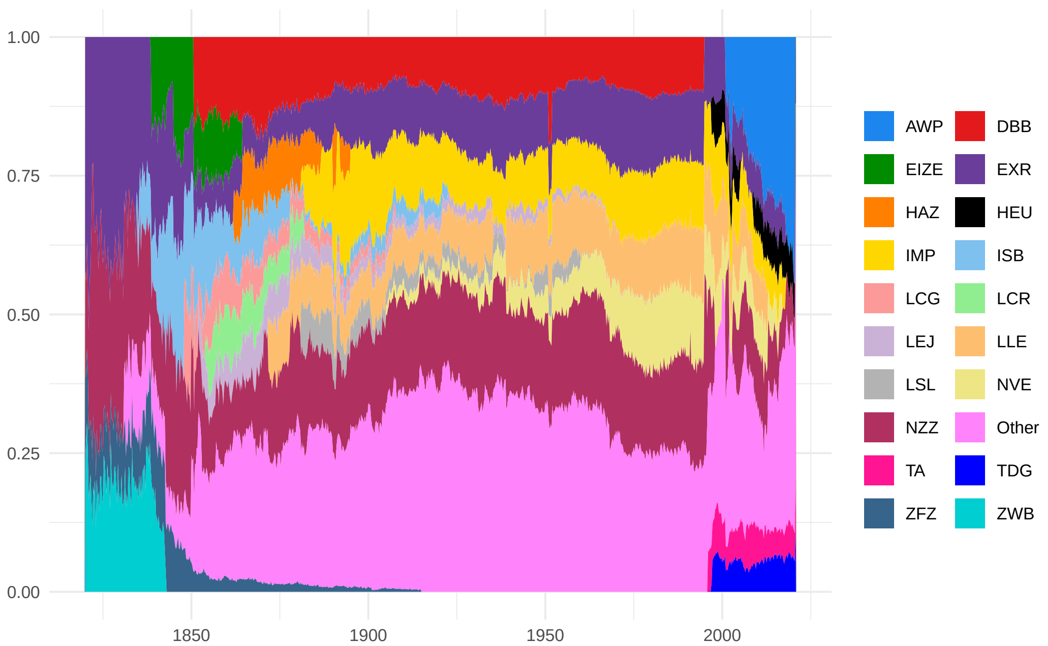
Notes: Example of how document-level sentiment scores for two topics are calculated based on a article teaser from FUW. For the general economy topic that is defined by the keyword (in blue) *wirtschaft*, the number of negative words (in red) is subtracted from the number of positive words (in green) within the ten preceding and following words from the keyword, and this result is divided by the total number of words. In this case, the sentiment score is $S_{t,d,\text{economy}} = (1-1)/14 = 0$. Note that there are only 14 words in the denominator because the keyword is close to the beginning of the text. The same method is applied to calculate the sentiment score for other topics, such as the industry topic, which in this example is given by $S_{t,d,\text{industry}} = (1-2)/19 = -0.05$.

Figure 6 — Number of identified keywords and important sources

a) Number of identified keywords



b) Most important sources



Notes: The graph shows the number of identified keywords for each language in Panel a). Panel b) shows the most important (by most identified keywords) sources and their share.

Table 5 — Frequency, duration, and number of Swiss compared to European business cycles

	Algorithm	Recession (Peak to trough) 1800-1870	1870-1950	1950-2000	Expansion (Trough to peak) 1800-1870	1870-1950	1950-2000
Frequency	Hamilton (1989)	32.3	25.3	21.5	67.7	74.7	78.5
	Bry and Boschan (1971)	49.3	45.9	36.5	50.7	54.1	63.5
	BL2023 (Swiss GDP)	35.0	32.5	14.0	65.0	67.5	86.0
	ECRI			19.2			80.8
	BL2023 (European GDP)	38.0	28.4	5.9	62.0	71.6	94.1
Duration	Hamilton (1989)	7.6	6.5	7.2	17.0	19.6	23.0
	Bry and Boschan (1971)	11.4	15.1	14.4	14.6	16.9	21.8
	BL2023 (Swiss GDP)	5.6	6.5	7.0	9.6	13.2	36.0
	ECRI			8.7			32.1
	BL2023 (European GDP)	6.4	7.2	4.0	9.6	17.2	54.8
Number	Hamilton (1989)	8.0	13.0	6.0	9.0	12.0	6.0
	Bry and Boschan (1971)	7.0	11.0	5.0	8.0	10.0	6.0
	BL2023 (Swiss GDP)	5.0	16.0	4.0	5.0	16.0	4.0
	ECRI			4.0			4.0

Notes: Frequency is the share of quarters in a given phase in percent. Duration is the average number of quarters in a given phase. Dating based on Hamilton (1989) and Bry and Boschan (1971) starts in 1820, real Swiss GDP growth is available from 1852, and the ECRI dating starts in 1956. Broadberry and Lennard (2023) (BL2023 European GDP) classifiy recessions as negative GDP per capita growth.

Table 6 — Frequency, duration, and number of Swiss compared to UK business cycles

	Algorithm	Recession (Peak to trough) 1817-1908	1909-1947	1948-2009	Expansion (Trough to peak) 1817-1908	1909-1947	1948-2009
Frequency	Hamilton (1989)	33.4	13.5	27.4	66.6	86.5	72.6
	Bry and Boschan (1971)	51.6	35.9	39.9	48.4	64.1	60.1
	BL2023 (Swiss GDP)	25.9	43.6	17.7	74.1	56.4	82.3
	ECRI			21.9			78.1
	B2023 (UK GDP)	22.8	25.6	11.3	77.2	74.4	88.7
Duration	Hamilton (1989)	6.9	7.0	7.6	14.6	35.7	23.2
	Bry and Boschan (1971)	14.0	13.5	12.6	14.1	20.4	21.1
	BL2023 (Swiss GDP)	5.0	8.5	6.3	13.7	10.7	33.3
	ECRI			7.9			27.7
	B2023 (UK GDP)	5.2	10.0	7.2	17.6	29.2	55.2
Number	Hamilton (1989)	17.0	3.0	9.0	17.0	3.0	9.0
	Bry and Boschan (1971)	13.0	4.0	8.0	13.0	5.0	7.0
	BL2023 (Swiss GDP)	12.0	8.0	7.0	12.0	8.0	7.0
	ECRI			6.0			6.0
	B2023 (UK GDP)	16.0	4.0	4.0	16.0	4.0	4.0

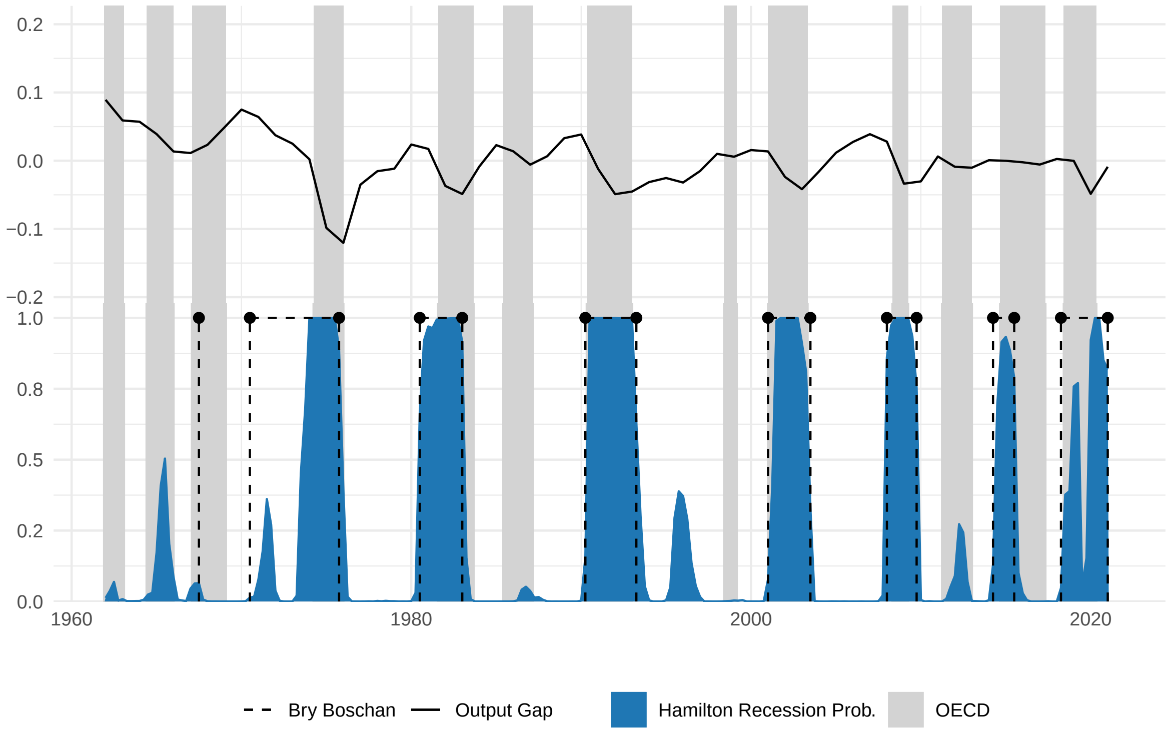
Notes: Frequency is the share of quarters in a given phase in percent. Duration is the average number of quarters in a given phase. Dating based on Hamilton (1989) and Bry and Boschan (1971) starts in 1820, real Swiss GDP growth is available from 1852, and the ECRI dating starts in 1956. Broadberry et al. (2023) (B2023 UK GDP) use discretion to classify phases in economic activity. BL2023 classifies recessions as negative GDP per capita growth.

Table 7 — Coincident correlations of topic-specific indicators

	Boom	Capacity	Credit	Financial	Inflation	Labor	RealActivity	Recession	Trade	Wages
Boom										
Capacity	0.14***									
Credit	0.08**	0.30***								
Financial	0.02	0.37***	0.46***							
Inflation	0.11***	0.31***	0.43***	0.55***						
Labor	0.09**	0.36***	0.43***	0.43***	0.47***					
RealActivity	0.23***	0.60***	0.50***	0.56***	0.58***	0.67***				
Recession	-0.11***	0.06*	0.20***	0.35***	0.33***	0.40***	0.35***			
Trade	0.20***	0.42***	0.37***	0.46***	0.44***	0.43***	0.62***	0.23***		
Wages	0.11***	0.26***	0.34***	0.41***	0.35***	0.53***	0.51***	0.31***	0.34***	
War	0.06	0.30***	0.09**	0.18***	0.05	0.02	0.29***	-0.13***	0.26***	0.03

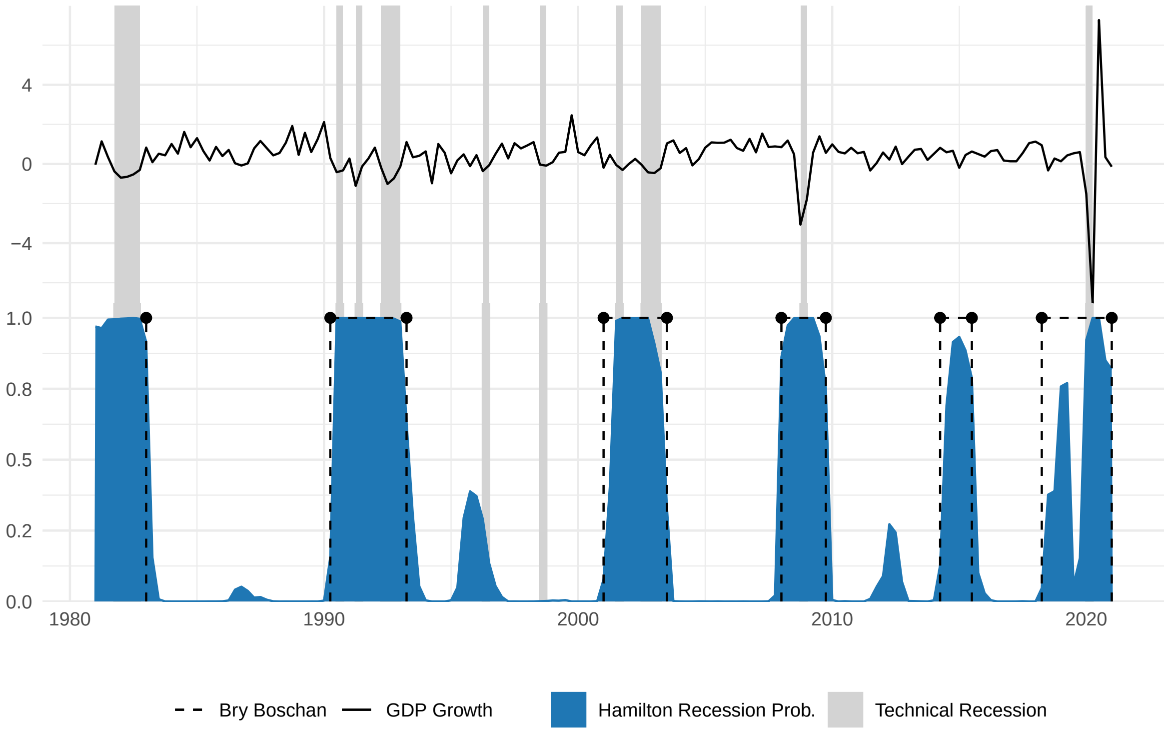
Notes: The table shows the coincident correlations between the topic-specific indicators. All indicators are scaled to have a positive correlation with the output gap. Significance levels are as follows: ***: 1%, **: 5%, *: 10%.

Figure 7 — Comparison with OECD dating



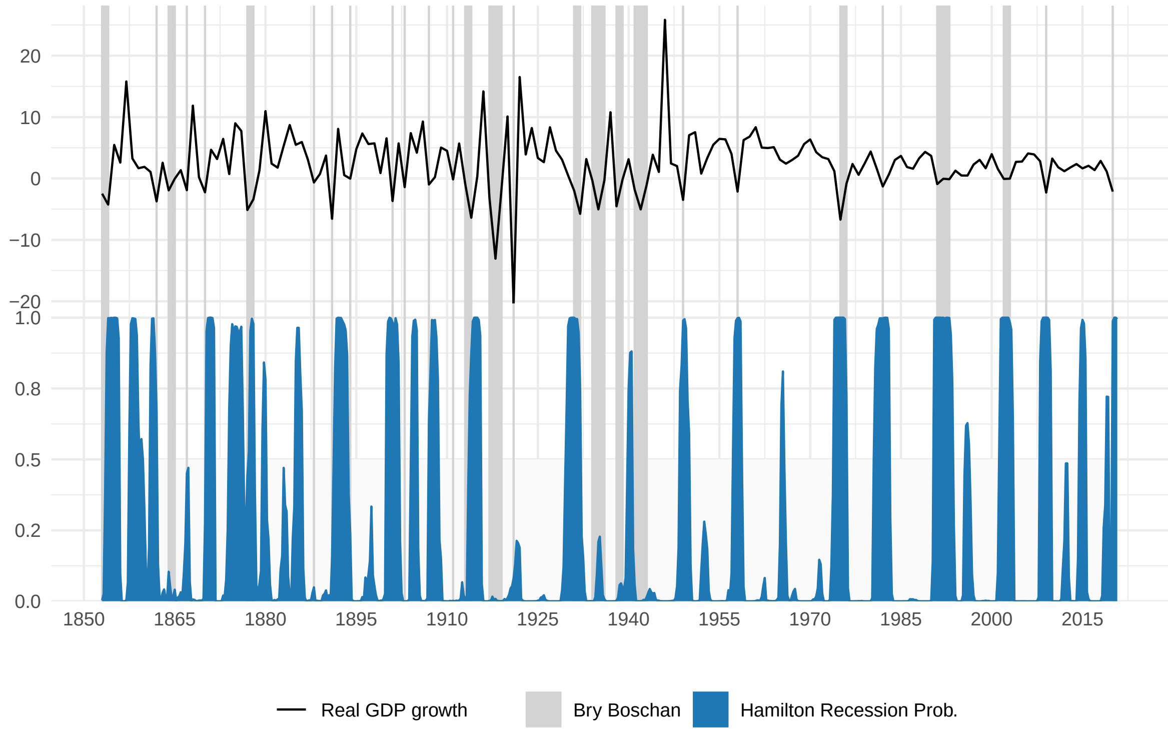
Notes: This graph shows the output gap given by the black solid line together with smoothed recession probabilities obtained by Hamilton's (1989) Markow-Switching autoregression model. The black dashed lines indicate recessions obtained by an adapted Bry and Boschan (1971) algorithm. The gray-shaded areas are recession dates obtained from the OECD.

Figure 8 — Comparison with technical recession



Notes: This graph shows quarterly real GDP growth given by the black solid line together with smoothed recession probabilities obtained by Hamilton's (1989) Markow-Switching autoregression model. The black dashed lines indicate recessions obtained by an adapted Bry and Boschan (1971) algorithm. The gray shaded areas indicate technical recessions.

Figure 9 — Comparison with BL2023 algorithm based on annual GDP growth



Notes: This graph shows annual real GDP growth given by the black solid line and smoothed recession probabilities obtained by Hamilton's (1989) Markow-Switching autoregression model. The gray shaded areas indicate contractions obtained with the algorithm proposed by Broadberry and Lennard (2023).

Table 8 — Quarterly turning points in Switzerland, 1821 - 2021

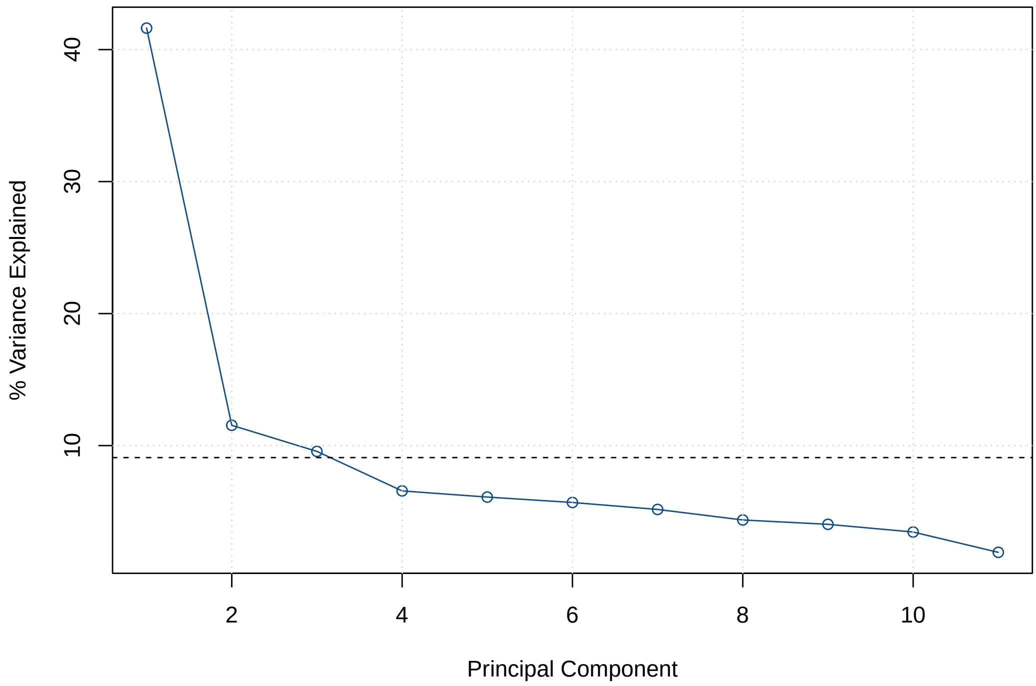
Censored	
Peak	Trough
1821 Q1	1823 Q3
1825 Q3	1827 Q2
1842 Q1	1843 Q3
1845 Q1	1845 Q2
1847 Q4	1850 Q2
1853 Q4	1855 Q3
1857 Q3	1859 Q4
1861 Q1	1861 Q4
1870 Q2	1871 Q3
1873 Q4	1875 Q3
1877 Q1	1878 Q1
1879 Q4	1879 Q4
1884 Q3	1885 Q4
1891 Q2	1893 Q3
1900 Q1	1902 Q1
1904 Q1	1905 Q1
1907 Q1	1908 Q2
1914 Q1	1915 Q3
1929 Q4	1932 Q1
1939 Q4	1940 Q3
1948 Q3	1950 Q1
1957 Q3	1958 Q3
1965 Q2	1965 Q3
1974 Q1	1975 Q4
1980 Q3	1983 Q1
1990 Q3	1993 Q3
1995 Q3	1996 Q1
2001 Q2	2003 Q2
2008 Q1	2009 Q3
2014 Q2	2015 Q3
2019 Q3	2021 Q1

Notes: Quarterly truning points of Swiss business cycles obtained from Hamilton's (1989) Markov-switching autoregression model with censoring (i.e. identified recessions where the underlying indicator is positive over the whole time span are removed).

Not censored	
Peak	Trough
1821 Q1	1823 Q3
1825 Q1	1826 Q3
1842 Q2	1843 Q2
1848 Q1	1850 Q2
1853 Q4	1855 Q4
1857 Q3	1859 Q1
1859 Q3	1859 Q4
1861 Q1	1862 Q1
1870 Q2	1871 Q3
1874 Q1	1876 Q2
1877 Q2	1878 Q1
1879 Q3	1880 Q1
1885 Q1	1886 Q1
1891 Q3	1893 Q3
1900 Q1	1902 Q1
1904 Q2	1905 Q1
1907 Q1	1908 Q3
1913 Q4	1915 Q3
1929 Q4	1932 Q1
1940 Q1	1940 Q3
1948 Q3	1950 Q1
1957 Q2	1958 Q3
1965 Q2	1965 Q3
1974 Q1	1976 Q1
1980 Q3	1983 Q1
1990 Q3	1993 Q3
1995 Q4	1996 Q2
2001 Q2	2003 Q3
2008 Q1	2009 Q4
2014 Q3	2015 Q3
2019 Q1	2019 Q2
2020 Q1	2021 Q1

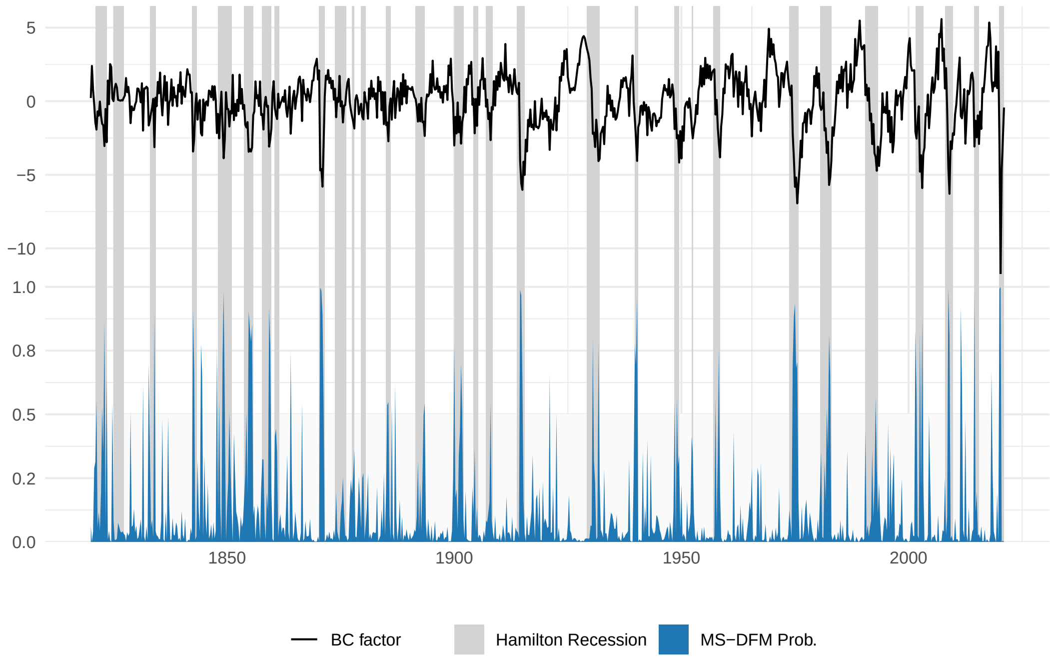
Notes: Quarterly truning points of Swiss business cycles obtained from Hamilton's (1989) Markov-switching autoregression model.

Figure 10 — Scree plot for MS-DFM used for robustness



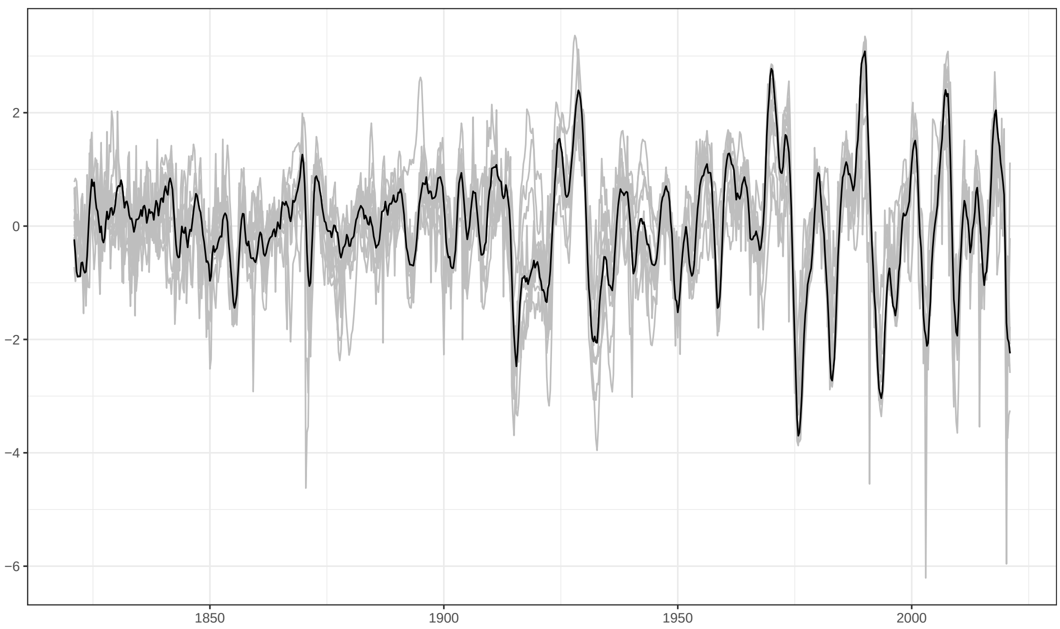
Notes: This graph shows the marginal contribution of each additional factor to R^2.

Figure 11 — Recession probabilities derived from MS-DFM



Notes: This graph shows the business cycle factor (black solid line) together with smoothed recession probabilities (blue) obtained with the MS-DFM. The gray shaded areas indicate recessions obtained with Hamilton's (1989) Markow-Switching autoregression model with the baseline indicator.

Figure 12 — Comparison to indicators based on other aggregation techniques, adjustment methods, and subsets of data.



Notes: This graph shows the baseline indicator (black solid line) together with several indicators based on other aggregation techniques, adjustment methods, and subsets of data (gray lines). See section 5 for more details.

B Data collection

This Appendix first discusses how I processed the scanned images, put them in machine-readable format, and parsed their layout. Moreover, I present the data to validate the new business cycle indicator.

B.1 Textual data

The indicators for measuring business cycle fluctuations are based on textual analysis of historical documents.[25] Therefore, I collected and digitized a large body of historical documents that potentially comprise useful information on business cycle fluctuations. Annual reports of companies and reports of national and regional business associations stem mainly from the Wirtschaftsarchiv Basel, whereas newspapers primarily stem from e-newspaperarchives.ch and scriptorium.bcu-lausanne.ch. Most of the newspapers are already in machine-readable format. Therefore, I only had to apply optical character recognition (OCR) to the documents that I scanned myself and those where the OCR is of poor quality.[26] In total, I collected 105 sources[27] in German and French language.

B.1.1 Scanning documents

I only selected documents to scan that likely comprise information about the business cycle. These either concern a textual description of a company's situation (predominantly annual reports of companies) or a description of the state of the economy in different sectors (primarily reports of business associations). Figures 2-3 show examples of scanned pages from the annual report of the "Chambre vaudoise du commerce et de l'industrie" (1923) and the insurance company Helvetia (1860). From a substantive point of view, the paragraphs comprise a discussion of the weak industry sector and an assessment of the business situation conditional on the difficult political situation of the insurance company. Thus, firms and business associations regularly judge their economic situation and the overall situation of the economy in their writing.

[25] See table 1 in the Appendix for a comprehensive list of all data sources. Figure 1 provides information on how many sources were available for each period.

[26] Already digitized data are in PDF or METS/ALTO format. METS/ALTO files are digital files used to store information about documents, such as books or articles, in a structured way. METS describes the document's structure, like its chapters or pages. At the same time, ALTO contains the actual text of the document along with details about its layout, formatting, and coordinates of each element on the page. Especially for larger documents with many pages, these files can be resource-intensive to process with statistical software. As a result, loading METS/ALTO files can take a long time. I developed R routines to read METS/ALTO files in parallel, which might be helpful to researchers. Find more information here: https://marcburri.github.io/posts/2023/09/11/mets-alto-r/.

[27] A source is a publication over a more extended period. This can be a newspaper (e.g., NZZ, 1820 - 2020) or a business report (e.g., Credit Suisse, 1895 - 2016)

B.1.2 Image pre-processing

After scanning these reports, the files must be prepared for text recognition. I, therefore, set up an image pre-processing procedure that facilitates the quality of the scans and makes it easier to convert the documents into machine-readable format. This step includes cropping the images, turning them into black and white (eventually erasing shimmering text from the back side), removing the curvature of the text (stemming from bent book pages), removing speckles, and sharpening the contrast of the documents.[28] Figures 2-3 show that the quality of the pre-processed image is improved, which in turn improves the following OCR.

B.1.3 Optical text recognition (OCR)

Converting the images into machine-readable text format is a crucial step. Therefore, I use the Abbyy FineReader software based on machine-learning techniques, widely used for larger-scale digitizing projects. Figures 2 - 3 show the recognized text with that software. From a technical point of view, they demonstrate that it is possible to recognize text in high quality from different fonts and languages. We can readily convert regular font and gothic type into machine-readable text. However, the examples are not perfect. In particular, the software has trouble recognizing the punctuation marks correctly. But this is not a severe problem because they will be deleted in a further step anyway. Overall, these examples demonstrate, however, that it is possible to accurately convert scanned documents' images into text.

B.1.4 Layout parsing

We have different types of publications. Therefore, the length of texts differs. Annual reports might potentially be several hundred pages long. On the other hand, newspaper articles seldom exceed one page. Moreover, some advertisements and tables do not contain valuable information. Therefore, I also use the Abbyy software to parse the layout of publications, which has not been done before. Figure 4 shows an example of a parsed layout. Titles, paragraphs, and tables can be easily identified. This has several advantages. First, long texts from annual reports can be split into shorter paragraphs. Many of these do not contain any useful information. Second, tables and figures can be identified and filtered out.

[28] I mostly used the CamScanner application for mobile devices to conduct these steps. In a few cases, I also used ScanTailor.

B.2 Validation data

Validating the indicator is difficult, especially for the 19th century when data on real economic activity is scarce and inaccurately measured. Nevertheless, I collect as much real activity data as possible covering the 19th century to calculate correlations and validate the indicator's accuracy. In this section, I briefly describe the collected data for validation and explain why it should be correlated with the business cycle indicator. An extensive overview and data sources are listed in Table 2 in the Appendix.

B.2.1 Data on real activity

The indicator developed in this chapter is a measure of the business cycle. Therefore, it should be correlated with real economic activity. Real GDP is one of the most widely used measures of real economic activity. However, GDP growth rates only show how much the economy has grown from one period to the next, not whether that growth is above or below the economy's potential. Therefore, I mainly use the output gap as a validation measure. The output gap is the difference between actual and potential GDP. It measures the extent to which the economy operates above or below its potential. The output gap is a useful indicator of the health of the economy and the degree of inflationary pressures.

The output gap is calculated as suggested by Hamilton (2018). That is, I estimate an OLS regression of GDP in logs, y_{t+h}, on a constant and the $p = 2$ most recent values of y as of date t,

$$y_{t+h} = \beta_0 + \beta_1 y_t + \beta_2 y_{t-1} + v_{t+h} \tag{13}$$

where the residuals

$$\hat{v}_{t+h} = y_{t+h} - \hat{\beta}_0 - \hat{\beta}_1 y_t - \hat{\beta}_2 y_{t-1} \tag{14}$$

offer a reasonable way to construct the cyclical component, the output gap. As suggested by Hamilton (2018), $h = 2$ for annual data.

For Switzerland, annual GDP data in real and nominal terms for the 19th century is presented in Historische Statistik der Schweiz HSSO (2012k). The authors state that some sector estimates of value added have "the characteristics of a rough estimate at best". For instance, in relation to estimates of the wholesale and retail value-added, they "did not have sufficient data to reliably estimate the value added of this important branch". Based on this dataset, Stohr (2016) estimates an improved GDP series by refining the estimation of certain branches.[29] This estimate is a crucial reference point for the business cycle indicator. As the reference series, I use Stohr's (2016) single-deflated estimate from 1851-1890 and the double-deflated estimate from 1890 to 1947.[30] For the period from 1948 to 2022, I use the official GDP figures from the Swiss Federal Statistical Office (FSO).

Because Swiss GDP is not accurately measured for the 19th century and business cycles tend to be international, I also compare the business cycle indicator to estimates for real GDP growth and the output gap of various countries. In particular, to Switzerland's neighboring countries, the USA and England. The data is taken from the Maddison database (Bolt & van Zanden, 2020). Finally, I also use world GDP, Swiss GDP per capita, trade data, and Swiss labor market data as validation measures.

B.2.2 GPT-3.5 meets Business Annals: A new take on 19th century sentiments

The validation indicators discussed above are hard data, likely measured with error in the early periods. My business cycle indicator, in contrast, is a qualitative indicator based on sentiments from textual analysis. There is, however, no existing qualitative indicator for the 19th century that I can use for validation. In his book, Thorp (1926) provides a detailed narrative account of the business cycle in selected countries for every single year up to 1925. He describes the business cycle in terms of the state of the economy, the financial situation, and the state of the labor market. I, therefore, use his account to create a qualitative sentiment indicator for Switzerland's neighboring countries, England and the USA, using large language models (LLM) as validation measures.[31] In particular, I use OpenAI's (2023b) GPT-3.5 model and feed it repeatedly with ten years of Thorp's (1926) descriptions and ask it to rate the state of the economy in a given year from -5 to 5. The mean overall ratings serve as the sentiment indicator. Because the model does not always produce the same rating, the standard deviation over all estimates can be interpreted as confidence bands. The following section provides more details on creating these sentiment indicators.

Qualitative business cycle indicators based on consumer or business surveys became

[29]This series is available in the Maddison project database (Bolt & van Zanden, 2020).
[30]See Stohr (2016) and Stohr (2017) for more information about single- and double-deflated GDP
[31]Unfortunately Thorp (1926) does not provide a narrative account for Switzerland.

popular in the mid-20$^{\text{th}}$ century. Before, however, there were no systematic attempts to measure the business cycle using surveys. To overcome this, I propose to use Thorp's (1926) description of the state of the economy together with OpenAI's (2023b) GPT-3.5 language model to construct economic sentiment indicators for several countries in the 19$^{\text{th}}$ and early 20$^{\text{th}}$ centuries. In this section, I show step-by-step how this can be done.

Thorp's (1926) book Business Annals is a seminal work in the field of financial history and economic analysis. In this book, Thorp (1926) meticulously compiles and analyzes historical data related to business cycles, financial situations, and economic fluctuations. He aims to record business conditions and their impact on financial markets comprehensively. He provides a detailed narrative account of the business cycle in selected countries for every year from around 1830 to 1925. Thorp's work laid the foundation for systematically studying economic business cycles (Burns & Mitchell, 1946). Therefore, I use his account to create a qualitative sentiment indicator for Switzerland's neighboring countries, England and the USA.

GPT (Generative Pre-trained Transformer) models are a type of artificial intelligence model used for natural language understanding and generation tasks. GPT models work by pre-training on a large corpus of text to learn the patterns, grammar, and semantics of language. During pre-training, they learn to predict the next word in a sentence, which helps them capture contextual information. They generate text by probabilistically predicting the most likely next word based on the input context and the knowledge they have learned during pre-training (Brown et al., 2020).

> **Prompt**:
> Given the short economic descriptions from the year {start_year} to {end_year}, rate the state of the {country} economy from the text below on a scale of -5 (worst) to 5 (best). Take into account the surrounding years for relative judgment. Return a JSON with the year and your rating only. Text: {texts}

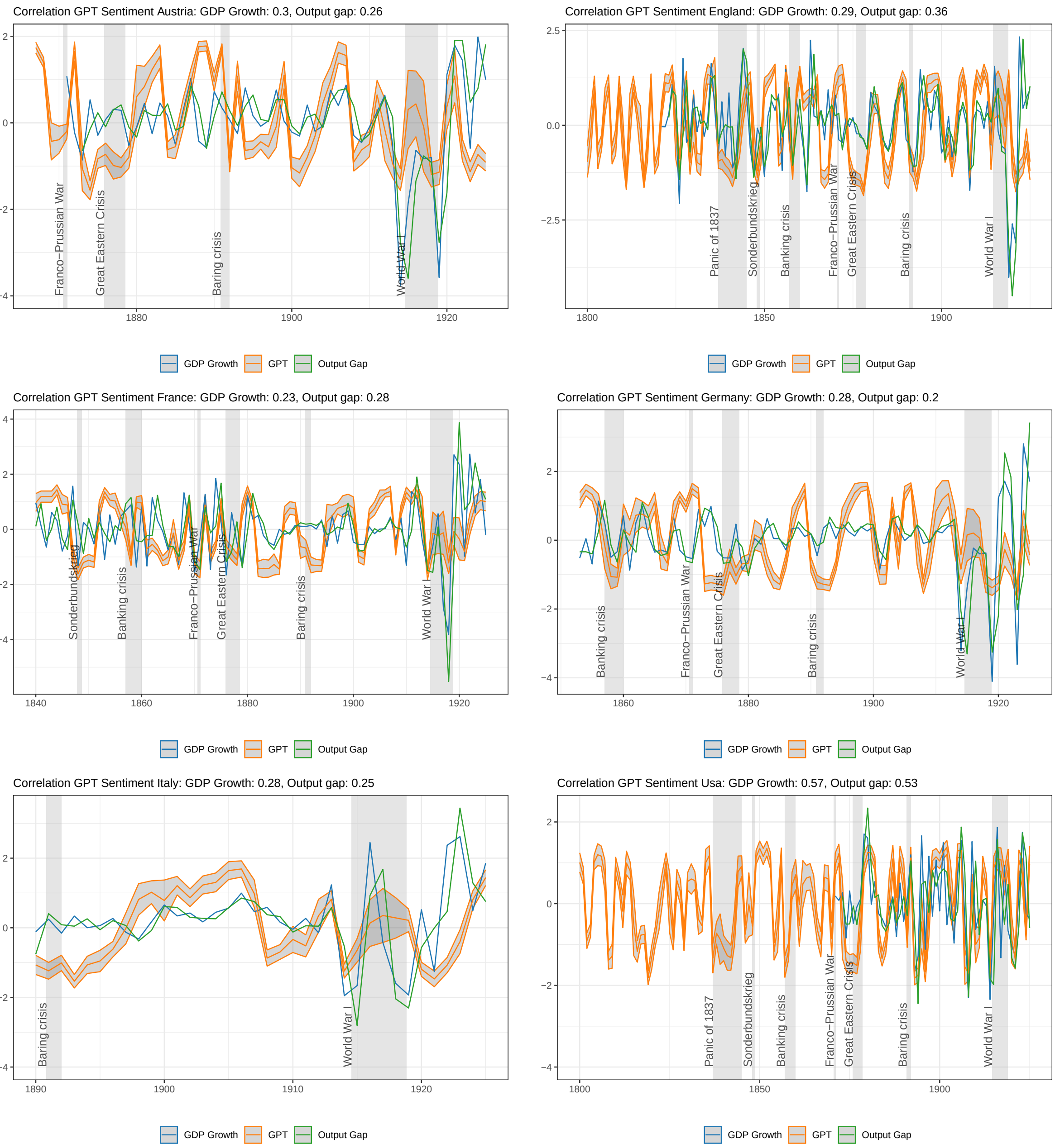
To create the sentiment indicators, I ask the model to rate the state of the economy on a scale of -5 (worst) to 5 (best) for every year until 1925. For relative judgment, the surrounding years should be taken into account. The exact prompt is depicted above. However, GPT-3.5 has a limited context window of 4'096 tokens. This means the model can roughly consider the last 3'000 words when generating the next word. Therefore, using the model in one batch is impossible because the text of Thorp's (1926) book is too long. Instead, I split the text into chunks of ten years and move forward by five years.

Hence, I get two ratings for every year. The average of the two ratings is the final rating for the year.

Providing the model with the same prompt multiple times does not necessarily produce the same output. This behavior is governed by the parameter 'temperature', which controls the randomness of the output. With a value of zero, the model becomes deterministic. I use the default value of one, which leads to a more diverse output. To get a more robust rating, I prompt the model 20 times and take the average of the 20 ratings. Figure 13 shows the normalized sentiment indicator (in orange) and one standard deviation confidence bands. Using the same degree of randomness every time allows for the interpretation of confidence bands as a measure of uncertainty. Most often, the model is quite confident about the state of the economy. However, during World War I, confidence bands become wider.

Moreover, I compare the GPT sentiment indicators with real GDP growth (blue line) and the output gap (green line) calculated from the series provided by Bolt and van Zanden (2020). The correlation between the sentiment indicator and real GDP growth ranges from 0.23 to 0.57 for the countries under consideration. With the output gap, the correlation ranges from 0.2 to 0.53. The correlation with neighboring countries is somewhat lower than with USA or England. This might be because these two countries have a lot of good quality data available for the 19th century. The sentiment indicator is less erratic than real GDP growth and the output gap and, therefore, serves as an excellent alternative measure to validate the historical business cycle indicator.

Figure 13 — Thorp GPT sentiment compared to GDP growth and output gap



Notes: These graphs show the normalized sentiment indicator (in orange) based on Thorp's (1926) texts and OpenAI's (2023b) GPT-3.5 model together with one standard deviation confidence bands. The blue line is real GDP growth, and the green line is the output gap, calculated from the series provided by Bolt and van Zanden (2020). Gray shaded areas represent crises.

C Source-level indicator anomaly correction

In this section, I explain the six steps of the procedure to correct for anomalies in the source-level indicators in more detail. Before proceeding with the six steps, texts where less than 20% of the words are identified in a German-French lexicon are removed. The remaining texts are then aggregated to either quarterly or annual frequency. Texts from sources published more frequently than once per quarter are aggregated quarterly, while texts from sources with a publication frequency between quarterly and annually are aggregated annually. Count-based indicators are aggregated by summing the number of identified keywords, and sentiment-based indicators are aggregated by averaging the sentiment scores. I use examples of the indicator based on the *Credit Suisse Bulletin* (CS_BUL) published annually and the *Neue Zürcher Zeitung* (NZZ) published daily. For illustration, I show the sentiment-based real activity and the count-based recession topics. The same procedure is applied to all other source-level indicators. Figures 14 (unadjusted indicators) and 15 (adjusted indicators) show that the correlation with a measure for the business cycle increases substantially after applying the procedure. The following are the six steps:

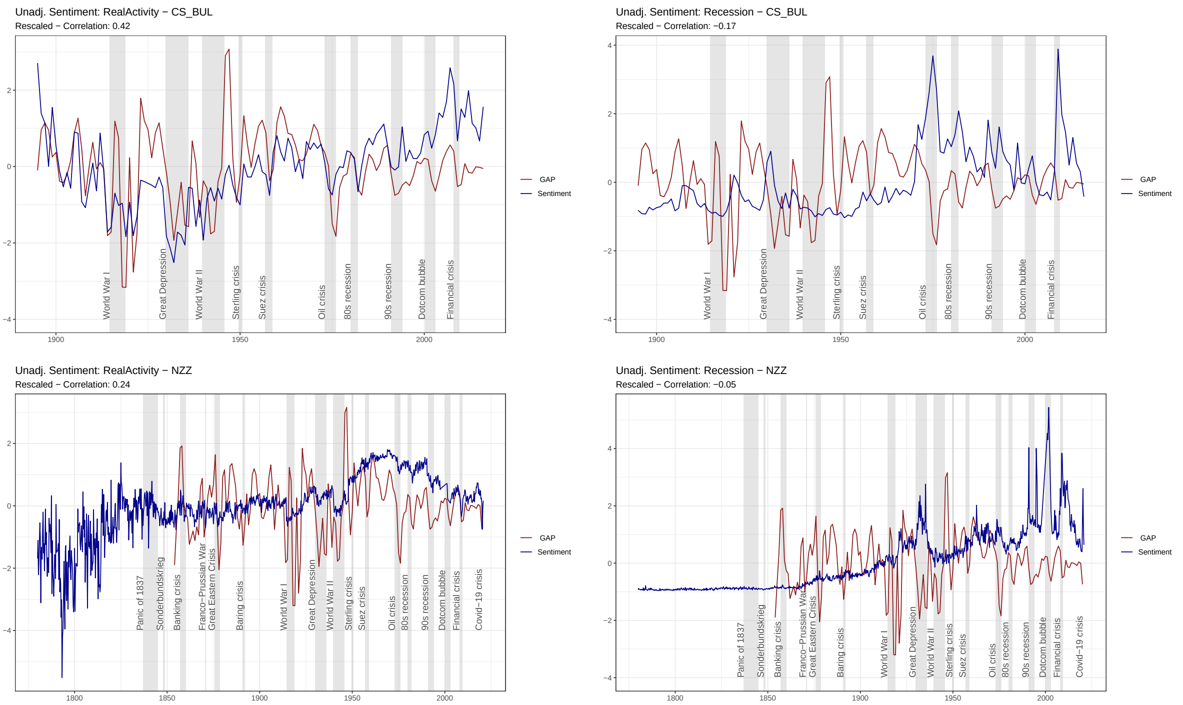
1. If the frequency is annual, interpolate missing observations using Stinemann interpolation (Stineman, 1980). This step is needed to ensure it is a regular time series before temporally disaggregating annual data to quarterly frequency.

2. Detrend the indicator using Locally Estimated Scatterplot Smoothing (LOESS) with bandwidth (or span) of 0.7 (Cleveland, 1979). LOESS (Locally Estimated Scatterplot Smoothing) is a statistical technique to create a smooth line through a scatterplot. This method selects a subset of data points and fits a local model, such as a linear or nonlinear function, to these points. This procedure is iteratively applied to each subset of the data, with the fitted models adapting to the specific characteristics of each section. The bandwidth, typically between 0 and 1, represents the proportion of the total data points in each local fit. A smaller bandwidth produces a more flexible, wiggly line, while a larger one produces a smoother line. A bandwidth of 0.7 is generally considered an effective compromise, balancing smoothness and adherence to the data points.

3. Detect structural breaks in mean and/or variance using a binary segmentation algorithm.[32] This algorithm originates from the work of Edwards and

[32] I use the implementation of Killick and Eckley (2014).

Figure 14 — Unadjusted source-level indicators



Notes: These graphs show the unadjusted source-level indicator (in blue) together with the output gap (in red). Gray shaded areas represent crises. Credit Suisse Bulletin at the top, NZZ at the bottom. Left the real activity topic, right the recession topic.

Cavalli-Sforza (1965), Scott and Knott (1974) and Sen and Srivastava (1975). The multiple parameter changepoint problem has been discussed by Horvath (1993) or Picard et al. (2005). The process initiates with a single changepoint test applied across the entire dataset to detect significant mean and/or variance changes. Upon identifying a changepoint, the data is split into two at the changepoint location. This test is then recursively conducted on each segment, continually splitting them at newly identified change points. The procedure repeats until no further significant changepoints are detected, adhering to criteria such as a predefined maximum number of changepoints or a minimum length for the segments. The outcome of this method is a division of the dataset into segments. The objective function the algorithm minimizes is given by

$$\sum_{i=1}^{m+1} \left[\mathcal{C} \left(y_{(\tau_{i-1}+1):\tau_i} \right) \right] + \beta m \tag{15}$$

where m is the number of breakpoints, $\mathcal{C}$ is a cost function (here, the negative log-likelihood), and βm is a linear penalty function. τ_i is the position of the i-th changepoint in the data. I assume the segments are normally distributed. Moreover, I allow a maximum of one breakpoint every ten years, with an overall cap of five breakpoints. Additionally, each segment is constrained to a minimum length of five years. This approach is based on the premise that the OCR quality of texts changes infrequently. Incorporating such prior information into the algorithm enhances its suitability for this particular application.

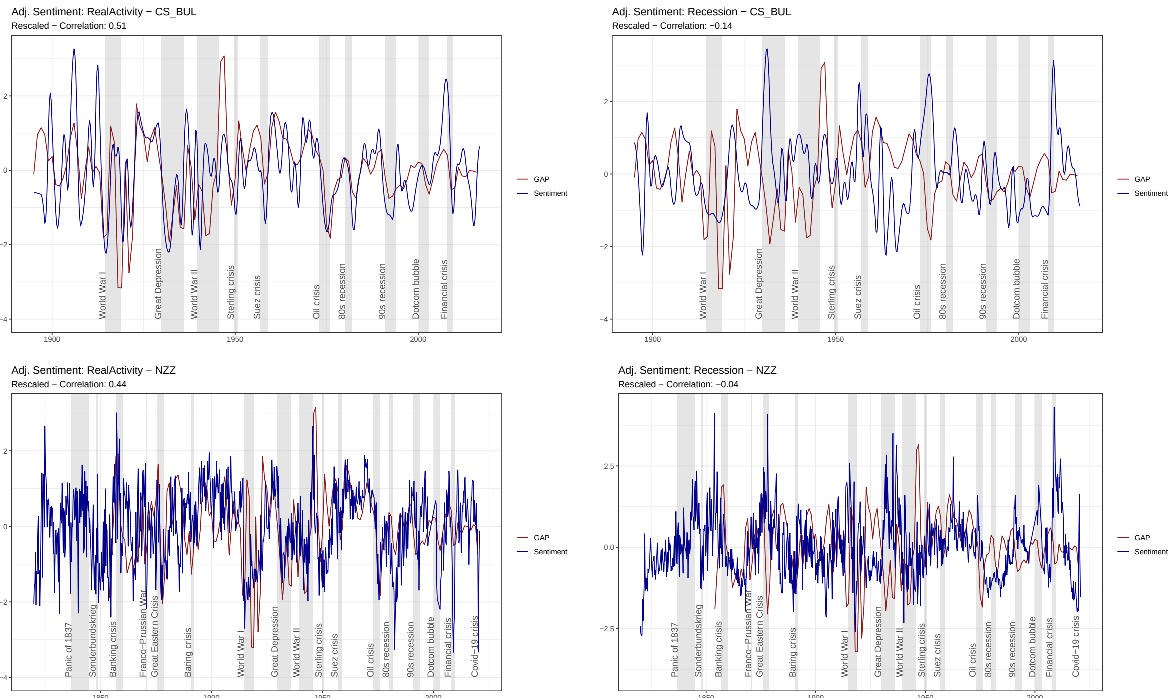
4. Normalize each identified segment.[33] The normalization is done by subtracting the mean and dividing by the standard deviation of the segment.

5. If the frequency is annual, temporally disaggregate the indicator to quarterly frequency using the method developed by Dagum and Cholette (2006), which is based on work by Denton (1971). I employ a straightforward method that does not require an indicator series. This method executes basic interpolation, adhering to the temporal additivity constraint as outlined by Sax and Steiner (2013). The temporal additivity constraint applied in this study mandates that the sum of the series at a lower frequency for count-based indicators must equal the corresponding series at a higher frequency. For sentiment-based indicators, it is the mean.

6. Remove outliers (observations more than three standard deviations away from the mean).

[33] I show in Appendix D that with time-varying measurement error, this is preferable to not normalizing.

Figure 15 — Adjusted source-level indicators



Notes: These graphs show the adjusted source-level indicator (in blue) together with the output gap (in red). Gray shaded areas represent crises. Credit Suisse Bulletin at the top, NZZ at the bottom. Left the real activity topic, right the recession topic.

D Normalization of time series with breaks

In this section, I show that normalizing data around identified breaks within the context of the dataset at hand is beneficial, particularly when dealing with a small number of available sources at the beginning of the sample.

The availability of data sources varies over time, with earlier periods having fewer digitized sources than later ones. Moreover, the measurement error associated with these source-level indicators is not static - it changes over time and tends to be larger in earlier periods. In most cases, the changes in the measurement error are abrupt and lead to structural breaks in the series. For instance, due to a sudden change in the optical character recognition (OCR) quality or because the frequency of the publication changes - increased publication frequency means a higher information density and, therefore, lower measurement error. This makes it essential to normalize the source-level indicators around these structural breaks.

The following section explains the challenges of normalization in the presence of time-varying measurement errors through a simplified illustration.

D.1 Simplified illustration

Suppose that a stochastic process determines the true but unobservable sentiment. For simplicity, I consider this process to be independent and identically distributed (iid), although it could also follow a different pattern, such as an AR(1):

$$s_t \sim iid(\mu, \sigma_s^2). \tag{16}$$

The sentiment observed from source i at time $t = 1, \ldots, T$, denoted $\hat{s}_{t,i}$, comprises the true sentiment plus some measurement error, which is also assumed to be iid:

$$\hat{s}_{t,i} = s_t + v_{t,i} \quad v_{t,i} \sim iid(0, \sigma_t^2). \tag{17}$$

I assume a constant magnitude for measurement error across all sources for simplicity. However, let us consider that from a certain time point, T_b, onwards, the magnitude of measurement error diminishes due to improvements in OCR quality or changes in publication frequency, for instance:

$$\sigma_t^2 = \begin{cases} \sigma_h^2 & \text{if } t \leq T_b \\ \sigma_l^2 & \text{if } t > T_b. \end{cases} \tag{18}$$

Here, σ_h^2 represents the higher measurement error variance before time T_b, and σ_l^2 is the lower variance after that. Overlooking this breakpoint and simply averaging out the indicators from all sources would lead to a higher variance in the high measurement error regime:

$$Var\left(\frac{1}{n}\sum_{i=1}^{n} \hat{s}_{t,i|t\leq T_b}\right) = \sigma_s^2 + \frac{1}{n}\sigma_h^2 > Var\left(\frac{1}{m}\sum_{i=1}^{m} \hat{s}_{t,i|t>T_b}\right) = \sigma_s^2 + \frac{1}{m}\sigma_l^2 \tag{19}$$

In this expression, n signifies the number of sources before T_b, and m is the number post T_b, typically with $m > n$. If both m and n are large enough, the difference in variance would be negligible. However, in the context of this project, the number of sources is limited, especially for the early sample, and the difference in variance is substantial. It follows that the level of the aggregated indicator is not comparable pre and post-break.

If this is the case, a nice indicator property would be to have the same variance across time for an application of business cycle dating utilizing a Markov-switching model. One potential strategy is to normalize the indicators from the pre-break and post-break periods separately before combining them.

$$\begin{gathered} Var\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\hat{s}_{t,i|t\leq T_b}-\mu}{\sqrt{\sigma_s^2+\sigma_h^2}}\right) < Var\left(\frac{1}{m}\sum_{i=1}^{m}\frac{\hat{s}_{t,i|t>T_b}-\mu}{\sqrt{\sigma_s^2+\sigma_l^2}}\right) \\ \frac{n^2\sigma_s^2+n\sigma_h^2}{n^2(\sigma_s^2+\sigma_h^2)} < \frac{m^2\sigma_s^2+m\sigma_l^2}{m^2(\sigma_s^2+\sigma_l^2)} \\ \frac{\sigma_s^2}{\sigma_s^2+\sigma_h^2}+\frac{\sigma_h^2}{n(\sigma_s^2+\sigma_h^2)} < \frac{\sigma_s^2}{\sigma_s^2+\sigma_l^2}+\frac{\sigma_l^2}{m(\sigma_s^2+\sigma_l^2)}. \end{gathered} \tag{20}$$

However, depending on the specific parameters, the combined indicator's resulting pre- and post-break variance could differ substantially. A significant disparity between m and n, coupled with a large $\frac{\sigma_h^2}{\sigma_l^2}$ ratio, results in a higher post-breakpoint variance. The intuition behind this result is that a large pre-break noise-to-signal ratio causes a more substantial downscaling of the signal during normalization, leading to a muted signal in the averaged indicator where some noise is canceled out.

D.2 Correlation with true sentiment

The ultimate goal is to end up with an indicator correlated as much as possible with the true sentiment over the entire sample. In the setup of the simple illustration above, the true sentiment is known. Therefore, I can calculate the correlation between the true sentiment and the aggregated indicator for different normalization strategies. If no normalization is applied before aggregation, the correlation is given by

$$\begin{aligned} \rho_{\text{NN}} &= Corr\left(s_t, \frac{T_b}{T}\frac{1}{n}\sum_{i=1}^{n}\hat{s}_{t,i|t\leq T_b}+\frac{T-T_b}{T}\frac{1}{m}\sum_{i=1}^{m}\hat{s}_{t,i|t>T_b}\right) \\ &= \frac{\sigma_s^2}{\sigma_s\sqrt{\sigma_s^2+\frac{T_b}{T}\frac{1}{n}\sigma_h^2+\frac{T-T_b}{T}\frac{1}{m}\sigma_l^2}} \end{aligned} \tag{21}$$

where $\frac{T_b}{T}$ and $\frac{T-T_b}{T}$ are the weights of the pre- and post-break period, respectively.

If normalization is applied to both segments individually before aggregation, the correlation is given by

$$\begin{aligned}\rho_{\text{SN}} &= Corr\left(s_t, \frac{T_b}{T}\frac{1}{n}\sum_{i=1}^{n}\frac{\hat{s}_{t,i|t\leq T_b}}{\sqrt{\sigma_s^2+\sigma_h^2}} + \frac{T-T_b}{T}\frac{1}{m}\sum_{i=1}^{m}\frac{\hat{s}_{t,i|t>T_b}}{\sqrt{\sigma_s^2+\sigma_l^2}}\right) \\ &= \frac{\sigma_s^2/\sqrt{\sigma_s^2+\frac{T_b}{T}\frac{1}{n}\sigma_h^2+\frac{T-T_b}{T}\frac{1}{m}\sigma_l^2}}{\sigma_s\sqrt{\frac{T_b}{T}\left(\frac{\sigma_s^2}{\sigma_s^2+\sigma_h^2}+\frac{\sigma_h^2}{n(\sigma_s^2+\sigma_h^2)}\right)+\frac{T-T_b}{T}\left(\frac{\sigma_s^2}{\sigma_s^2+\sigma_l^2}+\frac{\sigma_l^2}{m(\sigma_s^2+\sigma_l^2)}\right)}}.\end{aligned} \tag{22}$$

The question is whether normalization pays off, that is, whether $\rho_{\text{SN}} > \rho_{\text{NN}}$. This condition can be simplified to

$$\begin{aligned}\rho_{\text{SN}} &> \rho_{\text{NN}} \\ 1 &> \frac{T_b}{T}\left(\frac{\sigma_s^2}{\sigma_s^2+\sigma_h^2}+\frac{\sigma_h^2}{n(\sigma_s^2+\sigma_h^2)}\right)+\frac{T-T_b}{T}\left(\frac{\sigma_s^2}{\sigma_s^2+\sigma_l^2}+\frac{\sigma_l^2}{m(\sigma_s^2+\sigma_l^2)}\right).\end{aligned} \tag{23}$$

The right-hand side of this condition is a weighted average. Therefore, the weights sum up to one. Normalizing the segments is beneficial if the terms in brackets are lower than one.

$$\begin{aligned}1 &> \frac{\sigma_s^2}{\sigma_s^2+\sigma_h^2}+\frac{\sigma_h^2}{n(\sigma_s^2+\sigma_h^2)} \\ n(\sigma_s^2+\sigma_h^2)^2 &> (n\sigma_s^2+\sigma_h^2)(\sigma_s^2+\sigma_h^2) \\ \sigma_s^2+\sigma_h^2 &> \sigma_s^2+\frac{1}{n}\sigma_h^2.\end{aligned} \tag{24}$$

Hence, if $n > 1$, the left bracket is also lower than one. The same holds for the second bracket if $m > 1$. At least five sources are available in the early sample of the dataset used for this chapter. These findings suggest that normalizing the segments before aggregation is beneficial.

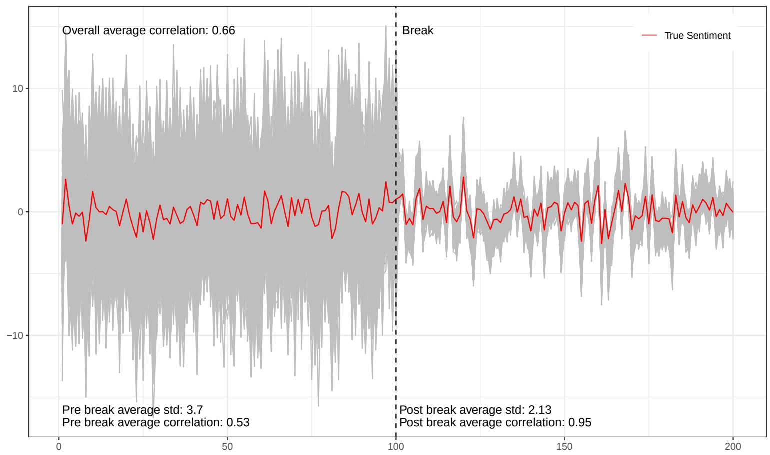
D.3 Simulation

To substantiate these findings, I carried out a simulation exercise. In this exercise, I compare the two aggregation methods discussed above. To mirror what is observed in the actual dataset, I employed the following parameter values:

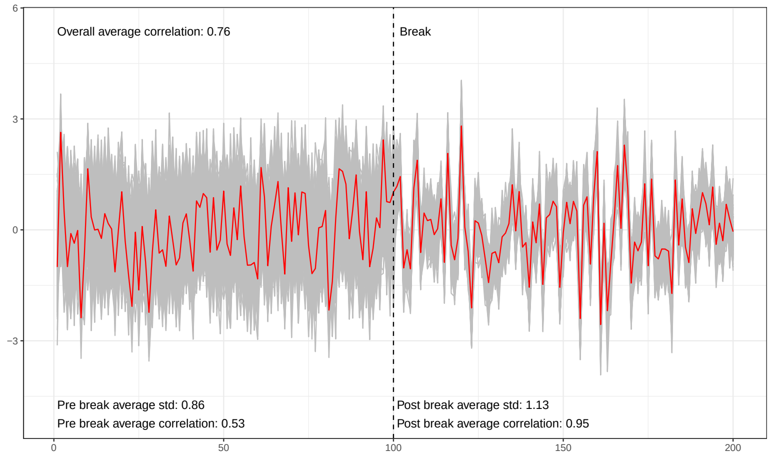
$$\sigma_s = 1, \quad \sigma_h = 10, \quad \sigma_l = 2, \quad n = 15,$$
$$m = 95, \quad t = 1, \ldots, 200, \quad T_b = 100.$$

Figure 16 — Simulation exercise for different normalization strategies

(a) Without normalization



(b) With normalization



I ran 1000 iterations to compute the average indicator for a set of 110 time series, conducting the procedure once without any normalization and once with normalization.

The outcomes are graphically represented in Figure 16: Panel (a) illustrates the scenario without normalization, where a gray line represents each of the 1000 simulations, and the true underlying series is traced in red. As anticipated, on average, the variance of the composite indicator is notably higher in the pre-break phase compared to the post-break phase.

Panel (b) of Figure 16 displays the results when the normalization is employed. The simulations reveal a compelling aspect: while the average correlation with the true series is essentially the same for both pre-break and post-break periods regardless of normalization, it is considerably higher across the entire timeline when the data is normalized. This enhancement in correlation serves as an extra argument to normalize the data.

E Markov-Switching Dynamic Factor Model

In this section, I describe the MS-DFM used for robustness.[34] MS-DFMs are pioneered by Kim (1994), Diebold and Rudebusch (1996), Kim and Yoo (1995) and Chauvet (1998). The model here closely follows Chauvet (1998). Estimating DFMs operates on the principle that an observed time series vector X_t can be split into two separate and orthogonal elements. First, the common components, often referred to as latent factors (f_t), encapsulate the joint movements among the observed variables in X_t. Second, the idiosyncratic component (e_t). These idiosyncratic components emerge from measurement errors and unique characteristics inherent in the data. The MS-DFM reads as follows:

$$X_t = \Lambda f_t + e_t \tag{25}$$

$$f_t = \mu_{s_t} + \sum_{p=1}^{P} A_p f_{t-p} + \eta_t \quad \eta_t \sim N(0, I) \tag{26}$$

$$e_t = \sum_{q=1}^{Q} C_q e_{t-q} + \varepsilon_t \quad \varepsilon_t \sim N(0, \Sigma) \tag{27}$$

where P is the lag number of the factor and Q is the lag order of the idiosyncratic component. s_t is the realization of a two-state Markov chain with

[34] See e.g. Mariano and Murasawa (2010) and Stock and Watson (1989, 2016) for prominent examples and further information on DFMs.

$$\Pr(s_t = j | s_{t-1} = i, s_{t-2} = k, \ldots, y_{t-1}, y_{t-2}, \ldots) = \Pr(s_t = j | s_{t-1} = i) = p_{ij} \tag{28}$$

where $i, j = 0, 1$. Within this framework, one can label s_t = 0 and s_t = 1 as the expansion and recession states at time t.

The specific MS-DFM in this application has one unobserved factor (See also scree plot 10), which is assumed to follow an AR(5) process (i.e. $P = 5$, based on Bayesian Information Criterion). The innovations are assumed to be independent (so that Σ is a diagonal matrix), and the error term associated with each equation is assumed to follow an independent AR(2) process (i.e. $Q = 2$). Therefore, the specification considered here is:

$$x_{i,t} = \lambda_i f_t + e_{i,t} \tag{29}$$

$$f_t = \mu_{s_t} + a_1 f_{t-1} + a_2 f_{t-2} + a_3 f_{t-3} + a_4 f_{t-4} + a_5 f_{t-5} + \eta_t \quad \eta_t \sim N(0, \sigma_f^2) \tag{30}$$

$$e_{i,t} = c_{i,1} e_{i,t-1} + c_{i,2} e_{i,t-2} + \varepsilon_{i,t} \quad \varepsilon_{i,t} \sim N(0, \sigma_i^2) \tag{31}$$

where i denotes one of the $n = 11$ indicators ($i = 1, \ldots, n$). For identification reasons, the variance of η_t, σ_f^2, is given by unity. For estimation, the model is cast into state-space representation and estimated using the Kalman filter.

$$y_t = H\xi_t + w_t \qquad w_t \sim NID(0, R) \tag{32}$$

$$\xi_t = \mu(s_t) + F\xi_{t-1} + v_t \qquad v_t \sim NID(0, Q). \tag{33}$$

I use the following definitions:

$$
\begin{aligned}
y_t &= [x_{1,t}, \dots, x_{n,t}]' && (34)\\
w_t &= 0_{(n\times 1)} && (35)\\
R &= 0_{(n\times n)} && (36)\\
\xi_t &= [f_t, \dots, f_{t-P}, u_{1,t}, \dots, u_{n,t}, u_{1,t-1}, \dots, u_{n,t-1}]' && (37)\\
H &= \begin{pmatrix} \lambda_1 & 0_{(1\times P-1)} & 1 & 0 & \dots & 0 & 0_{(1\times n)} \\ \lambda_2 & 0_{(1\times P-1)} & 0 & 1 & \dots & 0 & 0_{(1\times n)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \lambda_n & 0_{(1\times P-1)} & 0 & 0 & \dots & 1 & 0_{(1\times n)} \end{pmatrix} && (38)\\
v_t &= [\eta_t, 0_{(1\times P-1)}, \varepsilon_{i,t}, \dots, \varepsilon_{n,t}, 0_{(1\times n)},]' && (39)\\
diag(Q) &= [\sigma_f^2, 0_{(1\times P-1)}, \sigma_1^2, \dots, \sigma_n^2, 0_{(1\times n)}]' && (40)\\
F &= \begin{pmatrix} a_1 & \dots & a_P & 0_{(1\times n)} & 0_{(1\times n)} \\ I_{P-1} & \dots & 0_{(P-1\times 1)} & 0_{(P-1\times n)} & 0_{(P-1\times n)} \\ 0_{(n\times P)} & \dots & 0_{(n\times P)} & diag(c_{i,1}) & diag(c_{i,2}) \\ 0_{(n\times P)} & \dots & 0_{(n\times P)} & I_n & 0_{(n\times n)} \end{pmatrix} && (41)\\
\mu(s_t) &= [\mu_{s_t}, 0_{(1\times P-1)}, 0_{(1\times 2n)}]'. && (42)
\end{aligned}
$$

The estimation of the dynamic factor model with regime switching is carried out by maximizing its likelihood function. For estimating this model, I utilize the techniques developed by Kim (1994) and Chauvet (1998).