Asia-Pacific Journal of Accounting & Economics

ISSN: 1608-1625 (Print) 2164-2257 (Online) Journal homepage: www.tandfonline.com/journals/raae20

# Identifying the drivers of economic uncertainty perception in China: a news-based approach

Lena Marie Hufnagel & Ralf Metzler

**To cite this article:** Lena Marie Hufnagel & Ralf Metzler (25 Sep 2024): Identifying the drivers of economic uncertainty perception in China: a news-based approach, Asia-Pacific Journal of Accounting & Economics, DOI: 10.1080/16081625.2024.2405494

**To link to this article:** https://doi.org/10.1080/16081625.2024.2405494

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

View supplementary material

Published online: 25 Sep 2024.

Submit your article to this journal

Article views: 917

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=raae20

ASIA-PACIFIC JOURNAL OF ACCOUNTING & ECONOMICS
https://doi.org/10.1080/16081625.2024.2405494

Routledge
Taylor & Francis Group

OPEN ACCESS   Check for updates

# Identifying the drivers of economic uncertainty perception in China: a news-based approach

Lena Marie Hufnagel[a] and Ralf Metzler[b]

[a]Institute of Journalism, TU Dortmund University, Dortmund, Germany; [b]Mercator School of Management, University of Duisburg-Essen, Duisburg, Germany

**ABSTRACT**

Relying on news media as a reflection of public sentiment, we build the Uncertainty Perception Indicator (UPI) to measure perceived economic uncertainty in China. Applying Latent Dirichlet Allocation to $n = 5{,}600$ news articles, we decompose UPI into three components: Uncertainty relating to the real economy, financial markets, and politics. We find that on average, overall uncertainty perception is mainly driven by uncertainty about the real economy. This is especially true since 2018, when around 60 percent of the corpus was related to the real economy. We also find that stock price exposure is highest for real economic uncertainty.

**ARTICLE HISTORY**
Received 30 May 2023
Accepted 22 July 2024

**KEYWORDS**
Media effects; economic uncertainty; latent dirichlet allocation; asset pricing; cross-section of stock returns

## 1. Introduction

Economic uncertainty can have severe ramifications for various economic outcomes. This is exemplified by the current sudden rise in inflation across western countries, which forces central banks to increase interest rates substantially, making investments financed by debt more expensive for firms and reintroducing interest rate risk to the market. In uncertain times, demand decreases and firms have incentives to postpone or cancel investment and hiring decisions (Bernanke 1983), which in turn increases unemployment (Leduc and Zheng 2015). Besides these real-economic consequences, financial markets become more volatile in the face of high uncertainty. Many empirical studies even use the stock market volatility to proxy for economic uncertainty (Bekaert and Hoerova 2016; Kelly, Pastor, and Veronesi 2016).

The sources of economic uncertainty are diverse and can relate directly to macroeconomic factors in the real economy, like demand or consumption. Uncertainty can also stem from financial markets, if market participants do not confide in the financial system or there is negative investor sentiment. A third source for economic uncertainty is policy uncertainty. Since Baker, Bloom, and Davis (2016) introduced their news-based index, which attempts to measure Economic Policy Uncertainty (EPU) relying on uncertainty-related media coverage, many studies investigate its relationship with macroeconomic and firm-level outcomes (e.g. Chen, Jiang, and Tong 2017; J. Wu et al. 2018).

In this paper, we contribute to the empirical literature investigating economic uncertainty by applying a fine-grained measurement of economic uncertainty perception to the Chinese economy.

**CONTACT** Lena Marie Hufnagel lena.hufnagel@tu-dortmund.de Institute of Journalism, TU Dortmund University, Emil-Figge-Strasse 50, Dortmund 44227, Germany

Accepted by Junbo Wang upon the recommendation by Song Xuan Tam
Supplemental data for this article can be accessed online at https://doi.org/10.1080/16081625.2024.2405494

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

Specifically, we use the Uncertainty Perception Indicator (UPI) developed by Müller et al. (2018) to proxy for the perception of economic uncertainty using the share of newspaper articles that relate to economic uncertainty. We apply the *topic*[1] modeling approach of Latent Dirichlet Allocation (LDA) to a corpus of $n = 5{,}600$ uncertainty-related articles published in South China Morning Post (SCMP) between 2000 and 2020. We decompose overall perceived economic uncertainty into three different components (uncertainty related to the real economy, financial markets, and politics) to better understand which type of uncertainty drives overall perceived economic uncertainty at different points in time. Further, we study which type of perceived uncertainty has the largest effect on stock prices, and for which type investors require the largest risk premium. Looking at the entire observation period, we find that uncertainty about the real economy is the most pronounced type of perceived economic uncertainty in China, accounting for an overall share of 42.8 percent between 2000 and 2020 (see Table 1). The predominance of uncertainty stemming from real economic developments is particularly striking in the year 2018, when around 60 percent of uncertainty *topics* were related to the real economy, while only 15 to 25 percent related to financial markets or politics (see Figure 3). Accordingly, using Fama-Macbeth regressions, we find that stock prices are most exposed to changes in real economic uncertainty. Stocks that are in the decile portfolio most exposed to real economic uncertainty earn 12.5 percentage points lower monthly excess returns if real economic uncertainty rises by one standard deviation.

To date, most studies that use news media data to proxy for economic uncertainty use the Economic Policy Uncertainty Index (EPU) by Baker, Bloom, and Davis (2016) (e.g. Chen, Jiang, and Tong 2017; J. Wu et al. 2018). However, EPU has two key disadvantages: First, it is too narrowly defined to only measure uncertainty if it is related to some policy decision. Second, Baker, Bloom, and Davis (2016) partly rely on manual content analysis to construct EPU. In this study, we investigate a similar, but more advanced news-based index: The Uncertainty Perception Indicator (UPI) developed by Müller et al. (2018). Rather than economic uncertainty itself, UPI is designed to measure its *perception* using newspaper articles that mention economic uncertainty and which we understand as a reflection of public sentiment. Unlike EPU, UPI includes non-policy-related economic uncertainty in its measurement. Moreover, UPI construction employs *topic* modeling methodology, which does not only facilitate the processing of large amounts of text data, but can in our context also be used to investigate the sources of economic uncertainty perception. Using *topic*

**Table 1.** Overview of *topics* ($k = 18$), labels and contents grouped by *uncertainty factors*, shares are given in percent.

| Uncertainty Factor | Share | Content |
|---|---|---|
| **UPI real economy** | **42.8** | |
| subtopics: | | |
| T2 housing & real estate | 4.31 | housing, property market, investment |
| T8 business | 4.9 | risk management, recruitment, job market, entrepreneurship |
| T9 mainland economy | 10.88 | (macro)economic performance, manufacturing, consumption |
| T12 trade war | 6.57 | US sanctions, negotiations, Made in China 2025 |
| T15 energy & resources | 4.83 | energy production, demand & consumption, natural resources |
| T16 trade | 3.39 | import/export activities, trade, shipping |
| T17 travel | 3.91 | tourism & shopping, hotel industry, event industry |
| T18 technology | 4.01 | digitalisation, 5G, tech giants, semiconductors, robotics |
| **UPI financial markets** | **16.66** | |
| subtopics: | | |
| T3 banking & finance | 5.62 | monetary policy, banking system, (foreign) capital, loans |
| T5 stock market | 6.16 | trader sentiment, stock prices, volatility |
| T13 capital market | 4.88 | investor sentiment & investment activity, bonds & equities |
| **UPI politics** | **24.91** | |
| subtopics: | | |
| T4 hK politics | 6.24 | HK governance, democracy, protests, role of Beijing |
| T6 international politics | 6.44 | global crises, IR, climate change, summits (BRICS, G20) |
| T10 mainland politics | 5.78 | leadership, internal struggles, national security, stability |
| T11 regional security | 6.45 | nonproliferation, SCS, military & defense, regional initiatives |

Note: The shares don’t add up to 100 percent as four out of 18 *topics* have been excluded from further analysis.

modeling, we can distinguish three types of economic uncertainty perception related to either the real economy, financial markets, or politics. These three *uncertainty factors* enable us to analyze which kind of perceived uncertainty drove overall economic uncertainty perception at a particular point in time. As UPI has only been constructed for and applied to the German economy so far, we extend its use to the Chinese economy, which plays a key role in international trade and which is currently the second-largest economy in the world (International Monetary Fund 2021). We find that with an overall share of roughly 42 percent, uncertainty perception related to the real economy is on average the most pronounced type of perceived uncertainty, as compared to uncertainty perception related to financial markets (16.66 percent) and politics (24.91 percent).[2] This has especially been the case since 2018, with politics and the financial market becoming relatively less important sources of perceived uncertainty in China. Moreover, our analysis reveals that the US-China negotiations regarding tariffs and trade has been the main reason for this increase in UPI real economy.

Previous studies find that uncertainty shocks have a negative aggregate effect on stock prices and severely increase stock price crash risk (Bali, Brown, and Tang 2017; Du et al. 2023). However, these studies are unable to distinguish between different types of uncertainty. We, however, can measure the economic uncertainty aversion of stock market investors, to investigate which type of perceived uncertainty leads to the most pronounced stock market reaction. Applying two-stage Fama-Macbeth regressions which are standard in the asset pricing literature, we find that stock prices are most exposed to perceived uncertainty related to the real economy. Stocks that are in the decile portfolio most exposed to real economic uncertainty earn 12.5 percentage points lower monthly excess returns if real economic uncertainty rises by one standard deviation. However, when it comes to pricing economic uncertainty perception, investors require the largest risk premium for holding stocks that are negatively exposed to perceived financial market uncertainty, meaning that they prefer to hold stocks that are positively correlated with the uncertainty sentiment of investors in financial markets.

## 2. Literature review

### 2.1. Measuring economic uncertainty

In economics and finance, risk is regularly defined as quantifiable uncertainty about some future economic outcome, like the riskiness of a stock being reflected by the volatility of past returns. A precise risk assessment is therefore a prediction exercise and requires accurate assumptions about the probability distribution for future economic states. However, many economic decision processes are not only subject to quantifiable risks, but to further uncertainty about the true probability distribution for the different outcomes. This second type of uncertainty is unforecastable by definition. Some authors, like Jurado, Ludvigson, and Ng (2015) and Rossi and Sekhposyan (2015) argue that the definition of economic uncertainty should revolve around the question if the economy has become less predictable, thus focusing on the idea that uncertainty is unforecastable.

Ideally, we would like to calculate the total economic uncertainty by aggregating the uncertainty involved in every individual decision situation for each agent in the economy. As this is not feasible, the empirical economic and financial literature relies on three main approaches to quantify economic uncertainty. First, some studies proxy economic uncertainty by the volatility of the stock market, either by the realized volatility of stock returns, or by the implied volatility via option prices (Bekaert and Hoerova 2016; Kelly, Pastor, and Veronesi 2016). In the context of implied volatility, the CBOE Volatility Index (VIX) is a measure that receives much attention among practitioners and has been used to proxy economic uncertainty in the empirical literature as well. Some authors point out that the VIX does not only represent economic uncertainty, but also contains information about risk aversion (Bali and Zhou 2016; Bekaert and Hoerova 2016). Therefore, the degree of risk aversion contained in the VIX must be estimated separately and subtracted from the VIX to obtain an estimate for the residual, which represents the true economic uncertainty. This measure is often called the variance premium (as introduced by Carr and Wu 2009).

Second, some studies proxy economic uncertainty by the cross-sectional dispersion of either firm-level variables (e.g. profits) or the dispersion of analyst forecasts (Barron, Harris Stanford, and Yu 2010; Clement, Frankel, and Miller 2003). An obvious shortcoming of this economic uncertainty approximation is that these dispersions may be partly due to factors unrelated to uncertainty. For example, analyst disagreement may increase or decrease over time even if the true degree of uncertainty in the economy remains constant.

As the above-mentioned economic uncertainty proxies may be subject to different conceptual and estimation errors, a third strand of the literature uses economic uncertainty indexes that are based on the frequency of newspaper coverage. These indexes can be constructed even for countries in which alternative economic uncertainty measures are hard to obtain. The underlying assumption is that the uncertainty sentiment in an economy is reflected by the number of articles that mention economic uncertainty. Therefore, these news-based indexes measure uncertainty *perception* rather than uncertainty as such, which still is a useful indicator for actual levels of uncertainty in an economy because the impression of living in more uncertain times can already be a reason for economic decision-makers to be more cautious. Phrased differently, if a market participant expects a particular market outcome with high certainty, but they receive the information of high uncertainty of other market participants, they may themselves become less certain, e.g. because they fear firm investment may decrease. The most prominent news-based uncertainty index is the economic policy uncertainty index of Baker, Bloom, and Davis (2016), which attempts to measure the degree of policy uncertainty in an economy from frequency counts of newspaper articles that contain certain related keywords. Specifically, EPU for the United States is measured by searching ten of the leading US newspapers and counting the number of articles that contain the following terms:

'uncertain' OR 'uncertainty'

AND 'economy' OR 'economic'

AND 'Congress' OR 'deficit' OR 'Federal Reserve' OR 'legislation' OR 'regulation' OR 'White House'

EPU is standardized to account for the variation in total number of articles per month. Baker, Bloom, and Davis (2016) confirm that the peaks in EPU coincide with major events like wars, presidential elections, terrorist attacks or banking crises. In addition, EPU shocks are associated with higher stock price volatility and can be harmful to overall macroeconomic performance.

While Baker, Bloom, and Davis (2016)'s original index is a suitable alternative to measure economic policy uncertainty, it also suffers from several limitations. As the authors concede, their approach raises potential concerns about newspaper reliability, accuracy, bias, and consistency (Baker, Bloom, and Davis 2016, 1595) and Jurado, Ludvigson, and Ng (2015) point out that EPU is not an objective measure of economic uncertainty. In addition, as Müller and Hornig (2020) fail to invoke the respective established concepts from communication science such as framing, even though their index relies on media content as a proxy for economic uncertainty. Another limitation of EPU is that it is partly based on manual content analysis, which is not only time-consuming and labour-intensive (Gavaldon 2017), but also prone to human error. In contrast, computer-assisted approaches can detect hidden thematic structures in texts (Maier et al. 2018) and are applicable to large collections of documents. In this spirit, Müller et al. (2018) have constructed a news-based measure of perceived economic uncertainty based on the computer-assisted *topic* modeling technique Latent Dirichlet Allocation (LDA): The uncertainty perception indicator. UPI has two major advantages over EPU: First, it applies a broader query. While EPU only considers policy-related articles, UPI includes all articles that contain the following terms:

'uncertain' OR 'uncertainty'

AND 'economy' OR 'economic'

EPU is therefore a subset of UPI. Including the perception of non-policy-related economic uncertainty is important because not all economic uncertainty is related to policy decisions. For example, the occurrence of a natural catastrophe may induce uncertainty in an economy, even without any related policy decisions taking place. The second advantage of UPI is the application of *topic* modeling, which does not only facilitate the processing of large amounts of text data, but which can also be used to investigate the sources of economic uncertainty perception more thoroughly. Using LDA, Müller and Hornig (2020) identify thematically associated word clusters, so-called *topics*, which they sort into three distinctive categories, the so-called *uncertainty factors*: *UPI Politics*, *UPI Real Economy*, and *UPI Financial Markets*. This allows the authors to quantify the degree to which each category drives overall levels of economic uncertainty perception.

While the EPU index has been constructed for and applied to various countries over the past years, the concept of UPI is relatively new. The Dortmund Center for Data-based Media Analysis (DoCMA) is currently working on constructing UPI for several countries. But so far, published research only includes working papers on UPI Germany (e.g. Müller and Hornig 2020; Müller, Rieger, and Hornig 2021). We contribute to this body of research by conducting a UPI analysis for China. Specifically, we add to the empirical literature by identifying the sources of economic uncertainty perception in China, thus answering the following two research questions:

(1) Which *uncertainty factor* drives the changes in the overall level of perceived economic uncertainty: *UPI politics*, *UPI financial markets* or *UPI real economy*?
(2) Which individual *topics* are driving the *uncertainty factors* and thereby the overall level of perceived economic uncertainty?

### 2.2. Market reactions to economic uncertainty

Uncertainty aversion is not only a trait regularly assumed for agents in economic models, but has been abundantly documented empirically. The famous Equity Premium Puzzle coined by Mehra and Prescott (1985) describes the inability of economic models including the Consumption-based Capital Asset Pricing Model (CCAPM) to explain the empirically observable high difference between equity returns and safe government bond returns. This difference has historically averaged over six percent, implying extreme degrees of uncertainty aversion (Mankiw and Stephen 1991). This aversion for economic uncertainty is also exemplified by the constant presence of economic narratives in the media and society in general. On the one hand, economic narratives offer an explanation for a society, period or events, and on the other hand, they often involve advice on how to behave (Shiller 2020). Shiller identifies major recurring economic narratives, which can be classified by the type of uncertainty they address: First, there are narratives relating to the real economy like the narrative of artificial intelligence replacing almost all jobs. Second, there are financial market narratives like the narrative of stock market bubbles. Third, there are political narratives like the narratives of the wage-price spiral and evil labor unions. In the following chapters we will decompose perceived levels of uncertainty measured by the frequency of uncertainty-related media coverage into these three sub-types.

To date, a broad variety of studies have examined the macroeconomic and financial market implications of economic uncertainty shocks. We focus on studies which investigate the market reactions to EPU, as this is the measure closest related to UPI. Brogaard and Detzel (2015) analyse the impact of EPU on asset prices in the United States. They first confirm the countercyclical nature of EPU by means of standardized regressions of EPU on economic state variables such as term spread and dividend-yield. The authors also find that high levels of EPU will lead to an increase in forecasted three-month abnormal excess returns. More importantly, EPU shocks command a negative risk premium in the cross-section of stock returns suggesting that investors will demand lower returns to hold assets that hedge against increases in EPU. Bali, Brown, and Tang (2017) also examine the relationship between economic uncertainty and the cross-section of stock returns.[3] They estimate Fama-Macbeth regressions to identify the exposure of any given stock traded on the NYSE, Amex and Nasdaq between 1972 and 2014. Stocks that are positively exposed to increases in economic uncertainty for a given month earn significantly lower returns in the following month, as uncertainty-averse investors are willing to accept lower returns for these types of stocks. This confirms the results of Brogaard and Detzel (2015).

In the case of China, Du et al. (2023) regress several measures of stock price crash risk on EPU. They find a highly significant relationship, meaning that uncertainty shocks lead to a significant increase in the probability of stock price crashes. Chen, Jiang, and Tong (2017) further investigate the implications of economic policy uncertainty for the Chinese stock market. They observe a positive correlation between EPU and money growth, inflation, earnings-to-price ratio and stock market volatility. The authors conclude from their out-of-sample regressions that EPU possesses real-time predictive power (Chen, Jiang, and Tong 2017, 1268), which supports the hypothesis that the analysis of day-to-day news is useful for economic forecasting as suggested by Shiller (2020). Focusing on the influence of EPU on Chinese companies' overseas investment, J. Wu et al. (2018) find that firms tend to cut down overseas investment when economic policy uncertainty is high. This effect is stronger for state-owned enterprises (SOE) than for non-SOEs. Mirza and Ahsan (2020) show that EPU affects corporate strategic positioning and corporate risk, as elevated uncertainty levels increase both business and market risk of Chinese firms. The authors find that managers dynamically adjust their business strategies during uncertain times, depending on their level of risk aversion.

So far, there are no studies that investigate which type of perceived economic uncertainty market participants react to the strongest. We add to the literature by investigating the stock market reactions of investors to different types of perceived economic uncertainty. Specifically, we ask: For which *uncertainty factor* do investors require the largest risk premium: *UPI politics*, *UPI financial markets* or *UPI real economy*?

## 3. Data and methodology

### 3.1. Data and sample selection

The goal of this study is twofold: In a first step, we construct the news-based uncertainty perception indicator for China using a corpus of newspaper articles published in the print version of the South China Morning Post between 1 January 2000 and 31 December 2020. All articles were retrieved from LexisNexis database. In order to capture media coverage related to economic uncertainty, the original corpus was filtered employing a two-dimensional query suggested by Müller et al. (2018). Given that SCMP is based in Hong Kong, but our goal is to investigate uncertainty perception in mainland China, we modify accordingly and add the condition 'China' OR 'Chinese' to the search term to rule out any articles that do not deal with economic uncertainty in China[4]:

'China' OR 'Chinese'

AND 'economy' OR 'economic'

AND 'uncertain' OR 'uncertainty' OR 'uncertainties'

The filtering process yields a total of $n = 5{,}600$ articles that are relevant in terms of our research goal. For the purpose of *topic* analysis, we draw a representative sample of 900 articles from the corpus. We provide the methodological details in the following sections.

In a second step, we examine the economic uncertainty aversion of Chinese stock market investors using standard asset pricing methodology (see section 5.1 for details). In our two-step regression model, economic uncertainty perception is proxied by the UPI index developed in section 4.1 of this study. We download monthly stock returns of stocks contained in the most important Chinese stock index Shanghai SE Composite (SSE) ranging from January 2000 to December 2020 from Refinitiv Eikon/Datastream. We compute monthly excess returns for each stock using the risk-free rate, which is proxied by the US one-month treasury bill rate. We download four of the Fama-Fremch five factors (Fama and French 2015), namely size (SMB), book-to-market (HML), profitability (RMW) and investment (CMA) as well as the momentum factor

(MOM) from Kenneth French's data library (French 2022). For the market factor (MKT), we compute market excess returns directly from the SSE monthly closing values.

### 3.2. Methodology

In this study, we measure economic uncertainty perception in China and assess the influence of uncertainty perception on the Chinese stock market, by measuring the degree of economic uncertainty aversion of investors. The methodology is divided into two subsequent parts that build upon each other. First, we use computer-assisted *topic* modeling in the form of Latent Dirichlet Allocation to construct the uncertainty perception indicator. UPI measures economic uncertainty perception based on uncertainty-related news coverage and may function as a kind of early warning system to detect potential uncertainty shocks. Its construction and functionality are explained in section 4.1, while our findings on how uncertainty perception in China has changed over time are summarized in section 4.2. Second, we assess stock market reactions to uncertainty perception. To do so, we conduct a two-step Fama-Macbeth regression where we regress stock returns on economic uncertainty betas. A detailed description of our regression model is provided in section 5.1, results can be found in section 5.2.

## 4. Uncertainty perception indicator: measurement and development

### 4.1. Construction and functionality of UPI

Based on Shiller (2020)'s rationale that the study of media narratives can be used as an alternative method to forecast macroeconomic changes and potential future shocks, UPI uses the frequency counts of economic uncertainty-related newspaper articles to measure economic uncertainty perception. It is inspired by the economic policy uncertainty index, which is also based on news coverage and was originally developed by Baker, Bloom, and Davis (2016). In contrast to EPU, UPI is developed by means of computational *topic* analysis in the form of Latent Dirichlet Allocation, while EPU partly relies on manual content analysis. UPI can be decomposed into three subcategories, which are called *uncertainty factors*. The *uncertainty factors* enable us to identify the drivers of perceived economic uncertainty. The following two sections provide further details into the measurement of UPI.

#### 4.1.1. Topic modeling with latent dirichlet allocation

Latent Dirichlet Allocation is an unsupervised *topic* modelling approach that can be used to examine large amounts of text data and detect hidden thematic structures within collections of documents (DiMaggio, Nag, and Blei 2013). Based on the probability distributions of words and patterns of co-occurrence, the underlying algorithm generates clusters of words that are more likely to appear together than they would by chance. These clusters are called *topics* and reflect the latent main themes in a document collection (Blei, Ng, and Jordan 2003; Jacobi, van Atteveldt, and Welbers 2016). One basic assumption of an LDA *topic* model is the bag-of-words premise, which means that the algorithm ignores aspects like semantics and word order. As a result, words can appear across multiple *topics* depending on their respective context (Blei, Ng, and Jordan 2003). LDA is thus able to detect the presence of multiple meanings within one word (polysemy) and multiple meanings within one document (heteroglossia) (DiMaggio, Nag, and Blei 2013).

Prior to starting the analysis, the granularity of the model, i.e. its level of detail, must be determined by setting the parameter $k$, the number of *topics* generated during the LDA process, to an adequate value. The main goal here is to describe the data with fewer dimensions (topics) than are actually present, but with enough dimensions so that as little relevant information as possible is lost (Jacobi, van Atteveldt, and Welbers 2016, 93). In other words, the optimal value for $k$ marks a balance between granularity and conflation of *topics*: If there are many *topics* that are thematically associated to one another, this usually indicates that granularity is too high and $k$ should be reduced. Contrastingly, if there are *topics* that seem to cover more than just one subject, this indicates a high level of conflation and it might be more suitable to set $k$ to a higher value in order to make the model split these *topics* into separate ones according to their content. While statistical methods for this purpose do exist, such as the perplexity measure (Blei, Ng, and Jordan 2003), the literature highlights that the interpretability of *topics* is a more important factor in social sciences (Jacobi, van Atteveldt, and Welbers 2016). The optimal balance between granularity and conflation is also related to the underlying research question and must be chosen such that the resulting model is suitable for investigating that question accordingly. The general guiding principle is that on the one hand, $k$ should be high enough to ensure that *topics* can be clearly distinguished from one another in terms of their content. On the other hand, $k$ must be low enough in order to prevent the formation of *topics* that are too similar content-wise (von Nordheim, Müller, and Gerret 2019). With the aid of the respective lists of topwords (i.e. the statistically most frequent words per *topic*) and toptexts (i.e. articles containing a high proportion of vocabulary from a certain *topic* an can thus be considered the most representative articles per *topic*), a human coder is able to draw conclusions regarding the content of the *topics* and can therefore determine which value for $k$ is the best. In this study, we compared the results for different values of $k$ and eventually set the parameter to $k = 18$, because this value provides the optimal balance of granularity and conflation of *topics* with respect to our research objective.[5]

As mentioned above, a human coder is able to interpret the *topics* using the list of topwords. These topwords are thematically associated to one another and usually activate a coherent cognitive scheme with which it is possible to find headlines (labels) that reflect the underlying subject in each word cluster. Sometimes, the meaning of *topics* is not intuitive (Puschmann and Scheffler 2016) and can neither be derived from the topwords, nor in reference to other *topics*. In such cases it is necessary to investigate the toptexts, which are a collection of the most representative articles in each *topic*. Regarding UPI, both *topic* structure and the toptexts are important tools to estimate percieved levels of economic uncertainty. The contents of the toptexts do not only shed light on the context in which economic uncertainty is mentioned in each *topic*, but they are also useful for detecting certain events, actors or places associated with uncertainty. Furthermore, the contents of the toptexts and the distribution of *topics* within the corpus can provide information about the sources of economic uncertainty.

Validity and interpretability are two of the greatest concerns in the context of *topic* modeling with LDA. An overview of approaches to deal with these challenges – for instance statistical methods for *topic* selection or strategies for testing the semantic validity of *topics* – is provided by Maier et al. (2018). In this study, a heuristic analysis of topwords and toptexts was combined with intruder tests. To assess intratopic validity we ran word intrusion tests, scanning a list of words for the one term that is not associated with a given *topic*. For instance, among the words china, beijing, trump, washington, war, us, tariffs, trade, said and cars, cars would be marked as the intruder (more examples are provided in the online Appendix). *Topic* intrusion was tested analogously following the approach of Chang et al. (2009).[6] Additionally, we tested intertopic validity, i.e. the statistical proximity of *topics*, by means of hierarchical clustering (see Figure 2).

In this study, LDA is conducted using the R packages tosca (Koppers et al. 2020) and ldaPrototype (Rieger 2020). The purpose of the latter package in particular is to increase the interpretability as well as the reproducibility of our results as ldaPrototype has been developed to mitigate a major disadvantage of the LDA method, namely the aspect of instability, by determining a prototype model from multiple LDA runs (Rieger, Jörg, and Carsten 2020).

#### 4.1.2. Decomposing economic uncertainty perception

As mentioned earlier, one major advantage in comparison to Baker, Bloom and Davis's EPU index is the fact that UPI can be decomposed into subcategories or, as we call them, *uncertainty factors*. In reference to the three kinds of economic narratives identified by Shiller (2020), we distinguish the following *uncertainty factors*:

(1) *UPI real economy*
(2) *UPI financial markets*
(3) *UPI politics*

This decomposition is possible because the *topics* that are generated during the LDA process can be grouped together not only according to their respective labels, but also with respect to the predominant actors mentioned in each word cluster. For instance, the *topic* T8 (business) from our analysis deals with subjects like risk management, entrepreneurship or business activity and the toptexts commonly mention managers, businessmen, employers, entrepreneurs etc. T8 is therefore assigned to *UPI real economy*.

Another point of reference for UPI decomposition is the specific type of perceived uncertainty a *topic* deals with. Müller and Hornig (2020) differentiate between three varieties of perceived economic uncertainty: market-based uncertainty, economic policy uncertainty and truly exogenous uncertainty. The first type of perceived uncertainty refers to sudden changes in market sentiment and expectations. Respective *topics* are categorized as *UPI real economy* or *UPI financial markets*, because the perception of market-based uncertainty originates from the market itself and is thus directly related to subjects like business development, banking or trade. In contrast, the perception of economic policy uncertainty deals with unanticipated political events such as surprising outcomes of an election. Moreover, the implementation and consequences of certain policy instruments would also be classified as economic policy uncertainty. *Topics* that deal with this type of uncertainty are thus categorized as *UPI politics*. The third type of perceived uncertainty is less clearly defined (Müller and Hornig 2020, 4), refer to it as truly exogenous uncertainty (Müller and Hornig 2020, 4), which originates from developments taking place outside the market or political system. It can take various shapes: For instance, natural disasters or pandemics as well as unforeseen technological developments. *Topics* may be classified as either one of the three *uncertainty factors* depending on the area that a particular uncertainty shock has affected the most. Getting back to the case of T8, the European debt crisis is mentioned as one source of economic uncertainty. Even though financial markets are certainly affected here, too, T8 is categorized as *UPI real economy*, because market-based uncertainty originating from the crisis and its outcomes is featured most prominently in the respective toptexts.

For the task of matching the LDA *topics* to one of the three *uncertainty factors*, a close examination and thorough reading of the respective toptexts is necessary. To reduce the corpus to a manageable and yet representative number of articles, we sampled 50 toptexts for each *topic* (18 *topics* à 50 topics = 900 articles).

### 4.2. Development of UPI china

#### 4.2.1. General observations

Before going into detail about how the uncertainty perception indicator and its subcategories have developed over time in the case of China, we start with some general observations. As Figure 1 illustrates, the monthly share of economic uncertainty-related newspaper articles has increased during the 20-year observation period from under one percent of overall publication volume in the year 2000 to almost four percent in 2020. An all-time maximum of 4.6 percent was reached in August 2019. Interpreting media coverage as a reflection of public sentiment, we observe a surge in economic uncertainty perception. Also, the curve in Figure 1 peaks at certain points in time, for instance in late 2001, 2008/2009, 2012, 2016 and particularly 2019/2020. These maxima represent moments with particularly high levels of perceived economic uncertainty and thus point towards uncertainty shocks. In accordance with

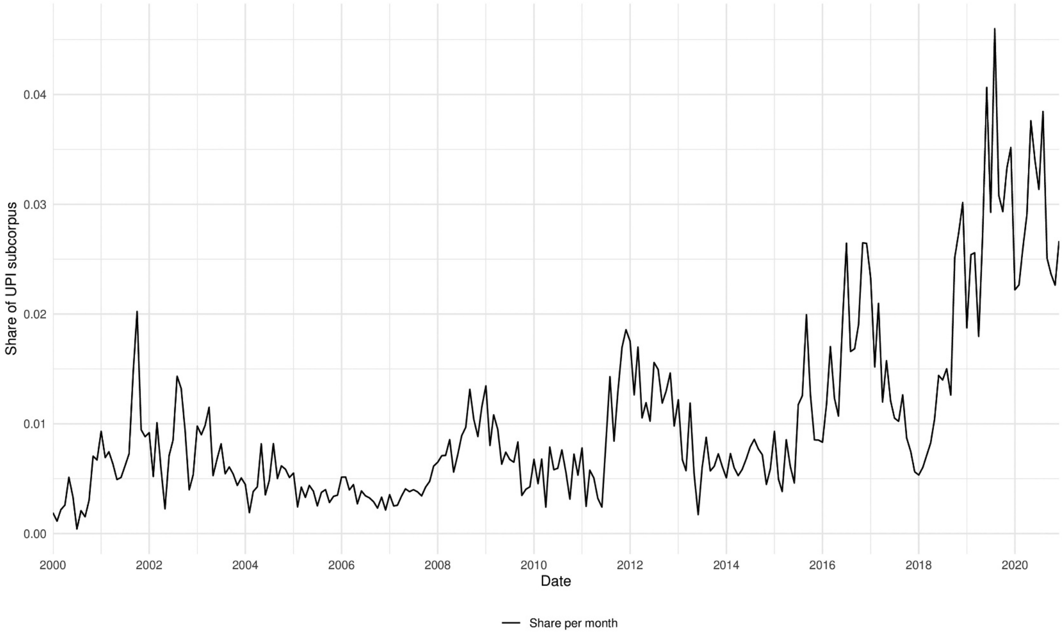


**Figure 1.** Monthly share of uncertainty-related newspaper articles in relation to overall publication volume in south China morning post between January 1, 2000 and December 31, 2020. Source: authors' own calculations.

existing studies on UPI and EPU (Baker, Bloom, and Davis 2016; Müller and Hornig 2020), events like 9/11, the global financial crisis, the Euro crisis, China's political transition, Trump's presidential election, the US-China trade war as well as the COVID-19 pandemic are likely the causes of these shocks. A closer look at the contents of the LDA *topics* and the respective toptexts confirms this assumption, as we will explore in more detail in the following subsection.

#### *4.2.2. Topic structure and contents*

For the present study, we generated a total of 18 *topics* using LDA. All *topics* were labelled according to the respective list of topwords (see Table 1 in the online Appendix). The dendrogram in Figure 2 illustrates the relative distance (Y-axis as Hellinger distance) between *topics*, which can be interpreted as thematic relatedness. For instance, the relative distance between T3, T13, T2, T15, T5 and T9 ranges between 0.6 and 0.7, which means that these *topics* are closely related in terms of their content. A look into the toptexts reveals that all these *topics* do indeed deal with different aspects of the economy, which is also reflected in their respective labels (see Table 2 in the online Appendix). Moreover, the dendrogram indicates that T1, T7, T14 and T17 can be considered so-called outlier *topics* meaning that they diverge more from the other *topics* in the same cluster. The contents of the toptexts of these four *topics* confirm this observation: T1 contains a broad mixture of articles dealing with all kinds of different subjects such as crime and terrorism, nature and wildlife or sports. Because these articles are not useful in terms of our research goal, T1 is excluded from further analysis. The same applies to T7 (education) and T14 (arts & leisure). While the exclusion of T14 is self-explanatory, the case of T7 requires some further explanation. T7 predominantly covers subjects like Hong Kong's educational system or studying in general. Even though it also contains some articles about recruitment and the job market, uncertainty is not mentioned in this particular context. Also, the actors appearing in T7 do not belong to either one of the three *uncertainty factors* so that it is impossible to assign this *topic* to any of the three UPI subcategories. In order to investigate the sources of uncertainty perception in China, all remaining *topics* were categorized as either *UPI real economy* or *UPI financial markets* or *UPI politics* (see subsection 4.1.2 for details).

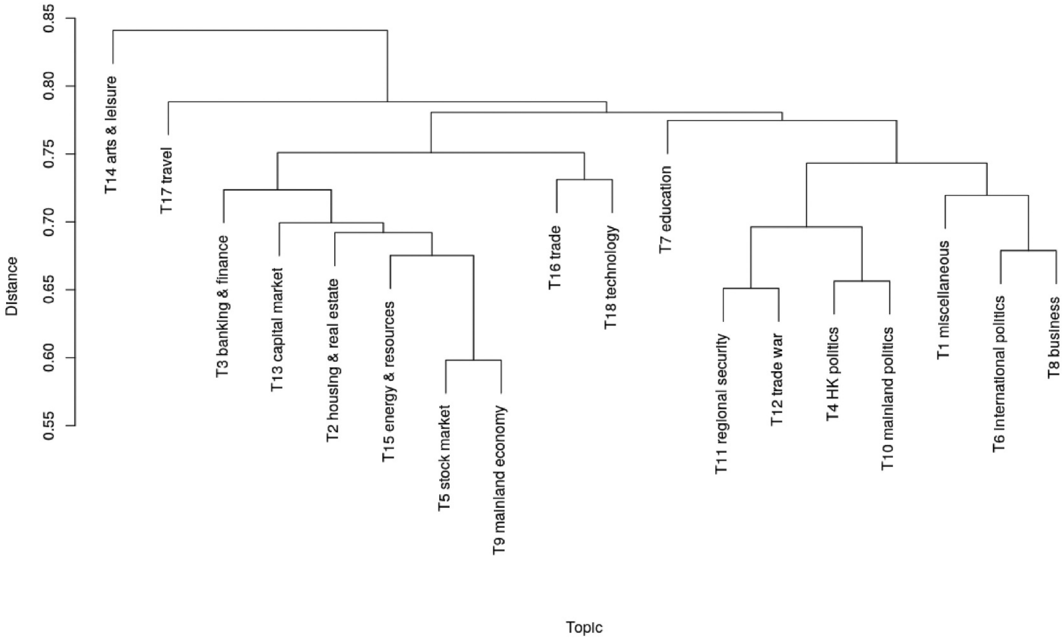


**Figure 2.** Cluster dendrogram for $k$ = 18 illustrating the relative distance (hellinger distance) between *topics*.

### *4.2.3. The sources of perceived uncertainty in China*

Table 1 lists all labelled *topics* that are relevant in terms of our research goal. It gives an overview of *topic* shares within the filtered UPI subcorpus ($n$ = 5,600) and provides a brief summary of contents for each *topic*. Depending on the main actors and the type of perceived uncertainty mentioned in the toptexts (see section 4.1.2), the *topics* are assigned to one of the three *uncertainty factors*. As Table 1 makes evident, most *topics* deal with market-based uncertainty and are thus assigned to *UPI real economy*. This makes *UPI real economy* the most influential *uncertainty factor* accouting for an accumulated share of 42.8 percent of the entire corpus between 2000 and 2020. Three *topics* are categorized as *UPI financial markets*, summing up to 16.66 percent of the corpus, while four *topics* cover economic policy uncertainty and are thus grouped together as *UPI politics* constituting 24.91 percent of the corpus. When comparing these findings to the observations made by Müller and Hornig (2020) in the case of economic uncertainty perception in Germany, it becomes clear that our results seem to reflect a peculiarity of economic uncertainty perception in China. While Müller and Hornig argue that in the case of Germany *UPI politics* is the dominant source of overall perceived economic uncertainty, the analysis for China reveals that political actors or politics in general seem to play a less important role for perceived levels of economic uncertainty.

The circumstance that *UPI real economy* is more pronounced than the other two *uncertainty factors* is illustrated in Figure 3: As the dynamics of the curves make clear, perceived levels of uncertainty in the real economy are particularly high after 2018, while *UPI politics* and *UPI financial markets* lose relative importance compared to *UPI real economy*. A closer look at the individual *topics* in each of the three *uncertainty factors* is helpful for finding explanations for this development and for identifying the sources of perceived economic uncertainty. For instance, with 10.88 percent, T9 (mainland economy) has the largest share of all *topics* within the UPI subcorpus (see Table 1). T9 was also the main source of perceived uncertainty in *UPI real economy*, but the role of *topics* within this *uncertainty factor* has changed over time. Shortly after Donald Trump was elected president of the United States, T12 (trade war) started to become a dominant *topic* for perceived uncertainty in the real economy (see Figures 1

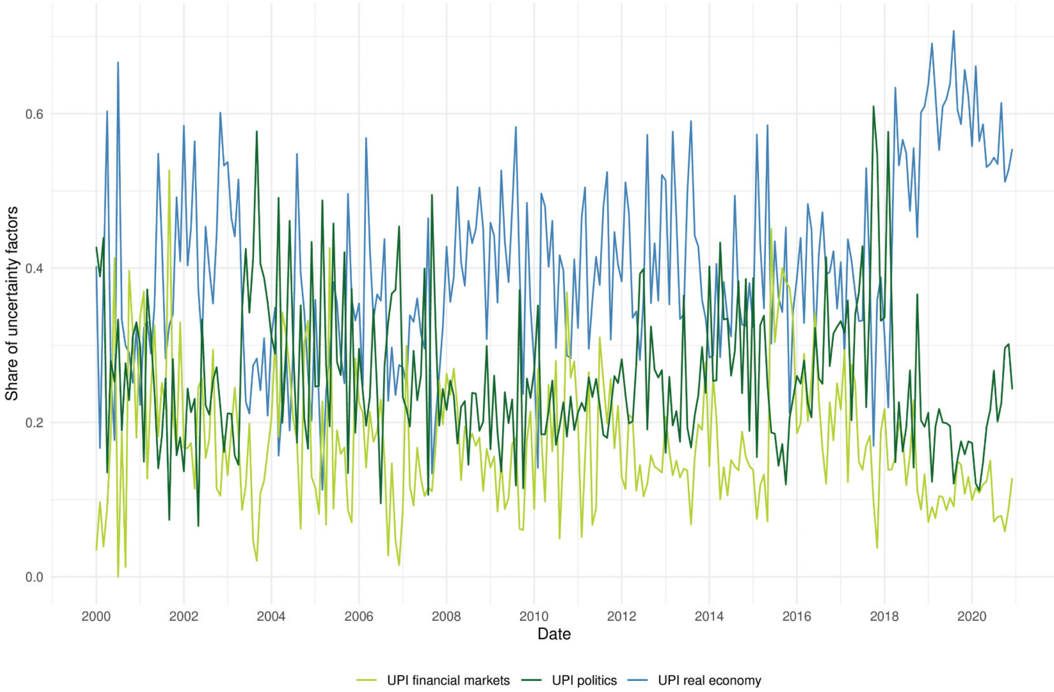


**Figure 3.** Aggregated share of *uncertainty factors* in UPI subcorpus ($n$ = 5,600) over time (2000–2020). Source: authors' own calculations.

and 2 in the online Appendix). We can assume from this observation as well as the contents of the respective toptexts that the lengthy negotiations between the US and China have not only had a direct economic impact on both countries, but they have also induced uncertainty in the Chinese economy.

In contrast to rising levels of *UPI real economy*, the sources of uncertainty on financial markets have been relatively stable (see Figure 3 in the online Appendix). As Table 1 shows, T3 (banking & finance) and T5 (stock market) are the main drivers for the levels of perceived economic uncertainty in the context of financial markets. A look into the toptexts reveals that among other events, the terrorist attack on the World Trade Center on 1 September 2001 had an impact on perceived financial market uncertainty. Moreover, China's accession to the World Trade Organisation (WTO) in 2001 as well as the SARS pandemic, the Global Financial Crisis, US elections or oil price fluctuations have had an impact on financial market uncertainty in the early 2000s. Later on, Brexit is a focal issue in many toptexts, which explains the brief increase of *UPI financial markets* in 2016 (see Figure 3). Interestingly, our analysis also indicates that the role of the stock market seems to have diminished in recent years: While news articles about stock price volatility or trader sentiment accounted for more than a quarter of all coverage in 2001, the overall share of T5 has declined until 2020 and articles about banking and finance play a more important role now (see Figures 3, and 4 in the online Appendix).

As for *UPI politics*, all *topics* combined in this *uncertainty factor* seem equally decisive for overall levels of perceived policy-related uncertainty, as T4, T6, T10 and T11 all have a share of around six percent of the analyzed UPI subcorpus (see Table 1). However, similar to *UPI financial markets*, the role of individual *topics* within *UPI politics* has changed throughout the observation period: Until around 2010 to 2012, the perception of policy-related uncertainty was for the most part influenced by T4 (HK politics). But the share of this *topic* has declined over time, while the share of others has increased, particularly T10 (mainland politics), which reached an all-time high around 2018 (see Figures 5 and 6 in the online Appendix). According to the toptexts, this development can be seen as a reflection of Xi Jinping's rise to power. Elected General Secretary of the Central Committee of the Communist Party of China (CPC) in 2012, Xi has put particular emphasis on foreign affairs and an expansive international strategy, a development which is not only alluded to by the toptexts, but which is also being discussed in current scientific literature (Drinhausen et al. 2021). More importantly, the toptexts in T10 discuss rising levels of political uncertainty in relation to subjects like the CPC's inner party struggles, social instability as well as the government's anti-corruption campaign and increased media control.

Considering the background of Alibaba's acquisition of the South China Morning Post in 2016, the increase in coverage of mainland politics is quite interesting, because the buyout had caused widespread concerns among observers that mainland Chinese influence over the free press in Hong Kong would increase and that coverage of politically sensitive issues would become more difficult (Council on Foreign Relations 2015). Accordingly, we should have seen a decline in T10 after 2016 rather than an increase. Moreover, we should be able to observe a change in content structure of *topics* T4 (HK politics) and T10 (mainland politics), as these are the *topics* that are most likely to include articles covering issues that would be affected by a potential editorial shift. However, a thorough reading of the respective toptexts did not provide any evidence for such assumptions.

While there's a variety of studies looking into the deterioration of press freedom and associated practices of self-censorship in Hong Kong post-handover (e.g. Chan 2020; Frisch, Belair-Gagnon, and Agur 2018; S. Wu 2023), the literature on editorial shifts of particular media is quite scarce. Wiebrecht (2018) investigates the potential editorial shift that SCMP may have undergone after 2016. The author ascertains that the editorial stance of the paper had already begun shifting towards mainland China before Alibaba's buyout and that there's no causal link between the two. In addition to the importance of domestic mainland politics for overall uncertainty perception, T11 (regional security) exhibits some interesting peaks, for instance in 2018 (see Figure 5 in the online Appendix). The reading of the respective toptexts for this *topic* reveals that China's growing assertiveness in regional security affairs such as the South China Sea or its role in regional initiatives are among the possible explanations for these dynamics.

## 5. Stock market reactions to economic uncertainty perception

### 5.1. Measuring the stock market reaction to economic uncertainty perception

In examining the reaction of monthly excess stock returns to economic uncertainty we apply standard asset pricing methodology. Our approach relates closely to studies investigating the asset pricing implications of economic uncertainty like Brogaard and Detzel (2015) and especially Bali, Brown, and Tang (2017). Specifically, we estimate the first stage of a Fama-Macbeth regression (Fama and MacBeth 1973):

$$\begin{aligned} R_{i,t} \quad &= \alpha_{i,t} + \beta_{i,t}^{\mathrm{UNC}}.\mathrm{UNC}_t + \beta_{i,t}^{\mathrm{MKT}}.\mathrm{MKT}_t + \beta_{i,t}^{\mathrm{SMB}}.\mathrm{SMB}_t + \beta_{i,t}^{\mathrm{HML}}.\mathrm{HML}_t \\ &\quad + \beta_{i,t}^{\mathrm{MOM}}.\mathrm{MOM}_t + \beta_{i,t}^{\mathrm{RMW}}.\mathrm{RMW}_t + \beta_{i,t}^{\mathrm{CMA}}.\mathrm{CMA}_t + \grave{o}_{i,t} \end{aligned} \tag{1}$$

We regress monthly excess returns for a month t on our three uncertainty factors, where our uncertainty factors are measured relative to the overall news coverage of the SCMP in the respective month. We control for the Fama-French five factors, namely market (MKT), size (SMB), book-to-market (HML), profitability (RMW) and investment (CMA). We also include a momentum factor (MOM). Similar to Bali, Brown, and Tang we estimate this regression equation for every stock $i$ via 60-month fixed-window estimation. Since our sample starts in January 2000, we use all monthly return data from January 2000 to December 2004 to obtain our first set of uncertainty betas. This first set of uncertainty betas is later used to predict the excess returns of month $t+1$, in this case January 2005. We then keep rolling the sample one month into the future until the sample is exhausted in December 2020. In total, we obtain $i \times (t-59)$ uncertainty betas for each of our three uncertainty factors described in subsection 4.1.2.[7] For each month we sort the cross-section of stocks into decile portfolios by their uncertainty beta. The first decile portfolio contains stocks with the lowest uncertainty betas, while the tenth decile portfolio contains stocks with the highest uncertainty betas. We compute the average uncertainty beta for each decile portfolio by averaging over all stocks included. Then, we average the uncertainty betas of each of the 10 decile portfolios over time.

Next, we investigate how investors price the exposure to the three UPI uncertainty factors. We compute the average excess returns of stocks contained in each monthly decile portfolio for the subsequent month. We then average these portfolio excess returns over all months. In order to adjust returns to risk, we again follow Bali, Brown and Tang and regress the time-series of decile portfolio excess returns on the Fama-French five factors and momentum as well as a constant $\alpha$:

$$\begin{aligned} R_{PF,t} \quad &= \alpha_{PF,t} + \beta^{\mathrm{MKT}}_{PF,t}.\mathrm{MKT}_t + \beta^{\mathrm{SMB}}_{PF,t}.\mathrm{SMB}_t + \beta^{\mathrm{HML}}_{PF,t}.\mathrm{HML}_t \\ &+\beta^{\mathrm{MOM}}_{PF,t}.\mathrm{MOM}_t + \beta^{\mathrm{RMW}}_{PF,t}.\mathrm{RMW}_t + \beta^{\mathrm{CMA}}_{PF,t}.\mathrm{CMA}_t + \grave{o}_{PF,t} \end{aligned} \tag{2}$$

### 5.2. For which uncertainty factor do investors require the largest risk premium?

Having established that developments in the real economy are the most important source of perceived economic uncertainty in China, we now turn to look at how stock market investors react to uncertainty shocks. More precisely, we investigate the asset pricing implications of economic uncertainty by estimating the first stage of a Fama-Macbeth regression for each of the three uncertainty factors and then sorting stocks into decile portfolios by their exposure to the respective factor.

Columns 1 to 3 of Table 2 report the average beta exposure to the uncertainty factors for each decile portfolio, where the first row contains stocks with the lowest uncertainty betas. Confirming our ex ante expectation, the exposure of average stock prices to *UPI real economy* is considerably higher than exposure to *UPI politics* and *UPI financial markets*. We standardized the three UPI uncertainty factors by their corresponding standard deviation, meaning that a one sd increase in *UPI real economy* is associated with a decrease in monthly excess returns of 12.5 percentage points for the average stock contained in the first decile portfolio. This reaction is much more pronounced compared to a one sd increase in *UPI financial markets*, which only leads to a decrease of 5.7 percentage points in monthly excess returns for those stocks most negatively exposed to it. We further observe that a greater number of stocks are negatively exposed to *UPI real economy* than *UPI financial markets* and respectively *UPI politics*, since the average exposure to financial market and politics uncertainty is already slightly positive in the 6th decile, while the average exposure to perceived real economic uncertainty is not.

Columns 4 to 6 of Table 2 report the average excess returns of the decile portfolios in the subsequent month. For all three *uncertainty factors*, the lowest excess returns are earned by stocks in decile portfolio 10 (1.12 to 1.25 percent). Investors therefore prefer to hold stocks that are positively exposed to economic uncertainty, even at the drawback of lower returns. Moreover, we would expect stocks in the lower decile portfolios to have considerably higher returns in order for investors to hold these negatively exposed stocks. For *UPI real economy* and *UPI financial markets* this is mostly the case, as the highest excess return are achieved in decile portfolio 2. However, investors do not require a significant risk premium for stocks negatively exposed to political uncertainty, as all excess returns in portfolios 1 through 9 are very similar.

In columns 7 to 9 of Table 2 we report the average alphas of the decile portfolios as estimated in Equation 2. If this implied six factor model fully explains the time variation in the decile portfolio excess returns $R_{PF,t}$, we would expect $\alpha$ to be zero. However, risk-adjusted portfolio excess returns are still considerably larger than zero, implying that the six factor model cannot fully explain the variation in $R_{PF,t}$. The distribution of the $\alpha$-values over the decile portfolios confirms our results from the raw excess return analysis, with investors requiring the lowest risk premium for holding stocks that are positively exposed to financial market uncertainty (0.58) and the highest risk premium for holding stocks that are negatively exposed to financial market uncertainty (1.16).

**Table 2.** Decile portfolios sorted by uncertainty betas. This table reports average uncertainty betas, average excess returns, and average alphas for all 10 decile portfolios sorted by their respective uncertainty betas. Uncertainty betas were estimated by first stage fama-macbeth regression: for each stock, we regress monthly excess returns on an uncertainty factor and a market (MKT), size (SMB), book-to-market (HML), profitability (RMW), investment (CMA) and momentum (MOM) factor. The uncertainty factors we use are *UPI financial markets* (UPI fin), *UPI politics* (UPI pol) and *UPI real economy* (UPI real). Newey-West adjusted *t*-statistics are reported in parentheses.

| | Betas | | | Returns | | | Alphas | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) UPI fin | (2) UPI pol | (3) UPI real | (4) UPI fin | (5) UPI pol | (6) UPI real | (7) UPI fin | (8) UPI pol | (9) UPI real |
| 1 | −5.66 | −7.35 | −12.48 | 1.44 | 1.48 | 1.44 | 0.98 | 0.97 | 0.98 |
| | | | | (6.05) | (6.22) | (6.12) | (2.35) | (2.26) | (2.30) |
| 2 | −2.82 | −3.59 | −6.36 | 1.70 | 1.51 | 1.63 | 1.16 | 0.97 | 1.11 |
| | | | | (7.59) | (6.86) | (7.43) | (3.16) | (2.78) | (2.97) |
| 3 | −1.76 | −2.29 | −4.23 | 1.65 | 1.49 | 1.49 | 1.09 | 0.97 | 0.92 |
| | | | | (7.44) | (6.64) | (6.67) | (3.21) | (2.78) | (2.65) |
| 4 | −1.01 | −1.35 | −2.75 | 1.44 | 1.56 | 1.47 | 0.87 | 0.99 | 0.90 |
| | | | | (6.62) | (7.32) | (6.91) | (2.56) | (2.94) | (2.89) |
| 5 | −0.37 | −0.54 | −1.47 | 1.58 | 1.54 | 1.44 | 1.00 | 0.92 | 0.86 |
| | | | | (7.35) | (7.14) | (6.65) | (2.99) | (2.68) | (2.65) |
| 6 | 0.24 | 0.22 | −0.27 | 1.40 | 1.51 | 1.53 | 0.78 | 1.01 | 1.00 |
| | | | | (6.57) | (6.91) | (7.16) | (2.41) | (3.00) | (2.87) |
| 7 | 0.88 | 1.02 | 1.00 | 1.62 | 1.56 | 1.53 | 1.06 | 0.99 | 0.92 |
| | | | | (7.82) | (7.61) | (7.24) | (3.64) | (3.07) | (3.01) |
| 8 | 1.62 | 1.95 | 2.45 | 1.58 | 1.49 | 1.47 | 1.03 | 0.90 | 0.86 |
| | | | | (7.45) | (7.20) | (6.81) | (3.09) | (2.87) | (2.77) |
| 9 | 2.65 | 3.20 | 4.43 | 1.26 | 1.43 | 1.54 | 0.70 | 0.88 | 1.03 |
| | | | | (6.09) | (6.96) | (7.40) | (2.30) | (2.91) | (3.16) |
| 10 | 5.81 | 6.90 | 10.27 | 1.12 | 1.21 | 1.25 | 0.58 | 0.64 | 0.69 |
| | | | | (5.27) | (5.56) | (5.98) | (1.64) | (1.84) | (2.00) |

## 6. Discussion

In this paper, we utilize the measure of UPI to identify the drivers of perceived economic uncertainty in China based on the frequency counts of economic uncertainty-related newspaper articles. For this task, we rely on the computer-assisted *topic* modeling approach LDA. The *topics* generated during the LDA process allow us to investigate the sources of economic uncertainty perception in more detail, because we are able to distinguish between perceived uncertainty in three different areas of the economy, forming so-called *uncertainty factors*: *UPI real economy*, *UPI financial markets* and *UPI politics*. In a second step, we investigate stock market reactions to these three types of perceived economic uncertainty.

Our results can be summarized as follows: First, we find that perceived economic uncertainty in China has risen to unprecedented levels in recent years and particularly after 2018. We also observe that developments in the real economy rather than politics or financial markets are the most important source of perceived economic uncertainty. Even though regional security issues and increasing domestic policy concerns in terms of Xi Jinping's unrestrained consolidation of power have recently led to a surge in *UPI politics*, *UPI real economy* remains the most pronounced of the three *uncertainty factors* throughout the entire observation period. Despite our hope that media censorship would not be an issue for our analysis given that South China Morning Post is based in Hong Kong, we assume that the main reason for the dominance of *UPI real economy* could be due to the Chinese media system: While newspapers in Western democracies are free to report on any subject, Chinese media organisations are subject to strict censorship rules. And because political coverage is considered particularly sensitive (Stockmann 2012), it seems logical that uncertainty-related newspaper articles will deal with economic facts rather than putting political decision-makers up for discussion. Conversely, this also means that UPI China might probably be less useful for making assumptions about unexpected economic uncertainty shocks originating from political decisions.

Secondly, we find that Chinese stock prices react most strongly to uncertainty shocks in the real economy. To be precise, the exposure of stock prices to *UPI real economy* is considerably higher than exposure to *UPI politics* and *UPI financial markets*. However, when pricing stocks investors require the largest (smallest) risk premium for stocks that are negatively (positively) exposed to uncertainty shocks related to financial markets.

To conclude, it should be noted that our analysis suffers from several limitations: For instance, our corpus was derived from the South China Morning Post. Because SCMP is based in Hong Kong, the editorial team is probably more prone to choose news items that are relevant for Hong Kong rather than the mainland. Even though we have tried to manage this circumstance by adjusting our query accordingly, it contradicts our goal of investigating the drivers of economic uncertainty perception in *the People's Republic* of China, not Greater China. We suggest future researchers to conduct a UPI analysis based on a corpus of mainland Chinese newspapers – either English language ones or, following Huang and Luk (2020)'s work, Mandarin Chinese papers. Another shortcoming of our analysis is connected to media censorship. As mentioned before, the production and publication of news in China is meticulously monitored. We assume that this might hamper the validity of UPI, because the concept is based on the notion that newspaper articles can be interpreted as a reflection of public sentiment. However, in a restricted media environment such as China, this argument becomes fragile. Although SCMP should not be affected by censorship, our findings let us assume otherwise, as noted above. Further research is required to investigate this proposition more thoroughly.

## Notes

1. When referring to topics in the context of computational content analysis with LDA, Italics are used in order to emphasize the semantic difference between LDA topics and the word 'topic' meaning 'subject' or 'theme'.
2. The shares don't add up to 100 percent, because we have excluded some of the LDA topics from further investigation (see also Table 1 in section 4.2.3 for further reference).
3. Instead of EPU, Bali, Brown, and Tang (2017) use an alternative measure of economic uncertainty developed by Jurado, Ludvigson, and Ng (2015), which correlates at 0.42 with the original EPU index (Baker, Bloom, and Davis 2016).
4. Due to reasons of data availability, we revert to SCMP data for the present analysis. As Baker, Bloom, and Davis (2016)'s original EPU for China is based on SCMP as well, our results should be equally meaningful. In fact, we find that UPI correlates at 0.94 with EPU.
5. For more detailed information on the choice of k see Table 3 in the Online Appendix.
6. To ensure intratopic validity, we ran word and topic intrusion tests on a randomized sample of 18 sets of ten topwords and 18 toptexts using the respective built-in functions of the R-package tosca (Koppers et al. 2020). The samples used for the intruder tests are provided in the online Appendix. 83.3 percent of intruder words were identified correctly, Krippendorff's alpha was 0.81 (where the alpha value was calculated in reference to the correct solution). As for topic intrusion, 55.6 percent of intruder topics were identified correctly and Krippendorff's alpha was 0.38.
7. We require at least 24 monthly observations within the 60-month estimation window. Otherwise the corresponding uncertainty beta is set to missing.

## Acknowledgments

The authors gratefully acknowledge the computing time provided on the Linux HPC cluster at TU Dortmund University (LiDO3), partially funded in the course of the Large-Scale Equipment Initiative by the German Research Foundation (DFG) as project 271512359.

## Disclosure statement

Ralf Metzler reports a relationship with Ruhr Graduate School in Economics (stipend from 2019-2020). The sponsor had no role in study design, collection, analysis and interpretation of data, in writing of the article, or in the decision to submit the article for publication.

## Data availability statement

The data that support the findings of this study were derived from NexisLexis database (newspaper articles) and Refinitiv Eikon/Datastream (monthly stock returns). Both are not publicly available due to privacy restrictions and were provided by third parties under license.

## References

Baker, S., N. Bloom, and S. Davis. 2016. "Measuring Economic Policy Uncertainty." *Quarterly Journal of Economics* 131 (4): 1593–1636. https://doi.org/10.1093/qje/qjw024.

Bali, T. G., S. J. Brown, and Y. Tang. 2017. "Is Economic Uncertainty Priced in the Cross-Section of Stock Returns?" *Journal of Financial Economics* 126 (3): 471–489. https://doi.org/10.1016/j.jfineco.2017.09.005.

Bali, T. G., and H. Zhou. 2016. "Risk, Uncertainty, and Expected Returns." *The Journal of Financial and Quantitative Analysis* 51 (3): 707–735. https://doi.org/10.1017/S0022109016000417.

Barron, O. E., M. Harris Stanford, and Y. Yu. 2010. "Further Evidence on the Relation Between Analysts' Forecast Dispersion and Stock Returns." *Contemporary Accounting Research* 329–357. https://doi.org/10.1506/car.26.2.1.

Bekaert, G., and M. Hoerova. 2016. "What Do Asset Prices Have to Say About Risk Appetite and Uncertainty?" *Journal of Banking and Finance* 67:103–118. https://doi.org/10.1016/j.jbankfin.2015.06.015.

Bernanke, B. S. 1983. "Irreversibility, Uncertainty and Cyclical Investment." *Quarterly Journal of Economics* 98 (1): 85–106. https://doi.org/10.2307/1885568.

Blei, D., A. Ng, and M. Jordan. 2003. "Latent Dirichlet Allocation." *Journal of Machine Learning Research* 3 (Special Issue): 993–1022. https://doi.org/10.5555/944919.944937.

Brogaard, J., and A. Detzel. 2015. "The Asset-Pricing Implications of Government Economic Policy Uncertainty." *Management Science* 61 (1): 3–18. https://doi.org/10.1287/mnsc.2014.2044.

Carr, P., and L. Wu. 2009. "Variance Risk Premiums." *The Review of Financial Studies* 22 (3): 1311–1341. https://doi.org/10.1093/rfs/hhn038.

Chan, C.-K. 2020. "China's Influence on Hong Kong's Media: Subduing Press Freedom." In *China's Influence and the Center-Periphery Tug of War in Hong Kong, Taiwan and Indo-Pacific*, edited by C. H. Brian, J. Fong, M. Wu, and A. J. Nathan, 121–138. London: Routledge.

Chang, J., J. Boyd-Graber, S. Gerrish, C. Wang, and D. Blei. 2009. "Reading Tea Leaves: How Humans Interpret Topic Models." *Advances in Neural Information Processing Systems 22 (NIPS 2009)* 32:288–296.

Chen, J., F. Jiang, and G. Tong. 2017. "Economic Policy Uncertainty in China and Stock Market Expected Returns." *Accounting & Finance* 57 (5): 1265–1286. https://doi.org/10.1111/acfi.12338.

Clement, M., R. Frankel, and J. Miller. 2003. "Confirming Management Earnings Forecasts, Earnings Uncertainty, and Stock Returns." *Journal of Accounting Research* 41 (4): 653–679. https://doi.org/10.1111/1475-679X.00119.

Council on Foreign Relations. 2015. "With New Ownership, Can Influential Hong Kong Daily Keep Its Soul?" Accessed February 19, 2024. https://foreignpolicy.com/2015/12/11/china-hong-kong-media-free-press-alibaba-jack-ma-censorship-business/.

DiMaggio, P., M. Nag, and D. Blei. 2013. "Exploiting Affinities Between Topic Modeling and the Sociological Perspective on Culture: Application to Newspaper Coverage of U.S. Government Arts Funding." *Poetics* 41 (6): 570–606. https://doi.org/10.1016/j.poetic.2013.08.004.

Drinhausen, K., M. Huotari, J. Lee, and H. Legarda. 2021. "The CCP's Next Century: Expanding Economic Control, Digital Governance and National Security." *MERICS Papers on China* (10).

Du, Y., X. Sui, W. Wei, and J. Du. 2023. "Economic Policy Uncertainty and Stock Price Crash Risk." *Asia-Pacific Journal of Accounting & Economics* 30 (3): 667–689. https://doi.org/10.1080/16081625.2021.2007407.

Fama, E. F., and K. French. 2015. "A Five-Factor Asset Pricing Model." *Journal of Financial Economics* 116 (1): 1–22. https://doi.org/10.1016/j.jfineco.2014.10.010.

Fama, E. F., and J. D. MacBeth. 1973. "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy* 81 (3): 607–636. https://doi.org/10.1086/260061.

French, K. 2022. "Data Library." Accessed March 18, 2022. https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/datalibrary. html\#Emerging.

Frisch, N., V. Belair-Gagnon, and C. Agur. 2018. "Media Capture with Chinese Characteristics: Changing Patterns in Hong Kong's News Media System." *Journalism* 19 (8): 1165–1181. https://doi.org/10.1177/1464884917724632.

Gavaldon, A. 2017. "Developing News-Based Economic Policy Uncertainty Index with Unsupervised Machine Learning." *Economics Letters* 158:47–50. https://doi.org/10.1016/j.econlet.2017.06.032.

Huang, Y., and P. Luk. 2020. "Measuring Economic Policy Uncertainty in China." *China Economic Review* 59 (C): 101–367. https://doi.org/10.1016/j.chieco.2019.101367.

International Monetary Fund. 2021. *World Economic Outlook: Managing Divergent Recoveries*. Washington DC: International Monetary Fund.

Jacobi, C., W. van Atteveldt, and K. Welbers. 2016. "Quantitative Analysis of Large Amounts of Journalistic Texts Using Topic Modelling." *Digital Journalism* 4 (1): 89–106. https://doi.org/10.1080/21670811.2015.1093271.

Jurado, K., S. C. Ludvigson, and S. Ng. 2015. "Measuring Uncertainty." *The American Economic Review* 105 (3): 1177–1216. https://doi.org/10.1257/aer.20131193.

Kelly, B., L. Pastor, and P. Veronesi. 2016. "The Price of Political Uncertainty: Theory and Evidence from the Option Market." *The Journal of Finance* 71 (5): 2417–2480. https://doi.org/10.1111/jofi.12406.

Koppers, L., J. Rieger, K. Boczek, and G. von Nordheim. 2020. "Tosca: Tools for Statistical Content Analysis. R Package Version 0.2-0." https://github.com/Docma-TU/tosca.

Leduc, S., and L. Zheng. 2015. "Uncertainty Shocks are Aggregate Demand Shocks." Federal Reserve Bank of San Francisco Working Paper.

Maier, D., A. Waldherr, P. Miltner, G. Wiedemann, A. Niekler, A. Keinert, B. Pfetsch, et al. 2018. "Applying LDA Topic Modeling in Communication Research: Toward a Valid and Reliable Methodology." *Communication Methods and Measures* 12 (2–3): 93–118. https://doi.org/10.1080/19312458.2018.1430754.

Mankiw, N. G., and P. Z. Stephen. 1991. "The Consumption of Stockholders and Nonstockholders." *Journal of Financial Economics* 29 (1): 97–112. https://doi.org/10.1016/0304-405X(91)90015-C.

Mehra, R., and E. C. Prescott. 1985. "The Equity Premium: A Puzzle." *Journal of Monetary Economics* 15 (2): 145–161. https://doi.org/10.1016/0304-3932(85)90061-3.

Mirza, S. S., and T. Ahsan. 2020. "Corporates' Strategic Responses to Economic Policy Uncertainty in China." *Business Strategy and the Environment* 29 (2): 375–389. https://doi.org/10.1002/bse.2370.

Müller, H., and N. Hornig. 2020. "Expecting the Unexpected. A New Uncertainty Perception Indicator (UPI) – Concept and First Results." *Working Paper #1-2020*. Dortmund Center for Data-based Media Analysis (DoCMA).

Müller, H., J. Rieger, and N. Hornig. 2021. ""We're Rolling". Our Uncertainty Perception Indicator (UPI) in Q4 2020: Introducing RollingLDA, a New Method for the Measurement of Evolving Economic Narratives." Working Paper #6. Dortmund Center for Data-based Media Analysis (DoCMA).

Müller, H., G. von Nordheim, K. Boczek, L. Koppers, and J. Rahnenführer. 2018. "Der Wert der Worte – Wie digitale Methoden helfen, Kommunikations- und Wirtschaftswissenschaft zu verknüpfen." *Publizistik* 63 (4): 557–582. https://doi.org/10.1007/s11616-018-0461-x.

Puschmann, C., and T. Scheffler. 2016. "Topic Modeling for Media and Communication Research: A Short Primer." *HIIG Discussion Paper Series*.

Rieger, J. 2020. "ldaPrototype: A Method in R to Get a Prototype of Multiple Latent Dirichlet Allocations." *Journal of Open Source Software* 5 (51): 2181. https://doi.org/10.21105/joss.02181.

Rieger, J., R. Jörg, J. Carsten. 2020. "Improving Latent Dirichlet Allocation: On Reliability of the Novel Method LDAPrototype." In *Natural Language Processing and Information Systems*, edited by E. Métais, F. Meziane, H. Horacek, and C. Philipp Cimiano, 118–125. Cham: Springer International Publishing.

Rossi, B., and T. Sekhposyan. 2015. "Macroeconomic Uncertainty Indices Based on Nowcasting and Forecast Error Distributions." *The American Economic Review* 105 (5): 650–655. https://doi.org/10.1257/aer.p20151124.

Shiller, R. J. 2020. *Narrative Economics*. Princeton & Oxford: Princeton University Press.

Stockmann, D. 2012. *Media Commercialization and Authoritarian Rule in China. Communication, Society and Politics*. Cambridge: Cambridge University Press.

von Nordheim, H. M. Müller, and M. Gerret. 2019. "Young, Free and Biased: A Comparison of Mainstream and Right-Wing Media Coverage of the 2015–16 Refugee Crisis in German Newspapers." *Journal of Alternative and Community Media* 4 (1): 38–56. https://doi.org/10.1386/joacm_00042_1.

Wiebrecht, F. 2018. "Cultural Co-Orientation Revisited: The Case of the South China Morning Post." *Global Media & China* 3 (1): 32–50. https://doi.org/10.1177/2059436418778306.

Wu, J., D. Kong, Y. Wu, and J. Zhang. 2018. "When to Go Abroad: Economic Policy Uncertainty and Chinese Firms' Overseas Investment." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3108128.

Wu, S. 2023. "Evaluating "Exemplary Data Journalism" from Asia: An Exploration into South China Morning Post's Data Stories on China and the World." *Journalism* 24 (9): 2042–2058. https://doi.org/10.1177/14648849221093509.