LMU Munich
Machine Learning: Applications in Economic Research

Using machine learning methods to evaluate global uncertainty and optimism

Submitted by

Moritz Dillmann
Economics (M.Sc.)

John Doncy
Economics (M.Sc.)

Lukas Franz
Economics (M.Sc.)

Bennet Niederhoefer
Economics (M.Sc.)

March 15, 2021

Advisor
Dr. Thomas Fackler
Department of Economics, ifo Institute

Helene Strandt
Department of Economics

Abstract

We exploit a corpus of 1,844 Article IV consultation reports published by the IMF between 2004 and 2020 for 190 countries to quantify global economic uncertainty and optimism following the methodology of the IMF's World Uncertainty Index (WUI). Our results suggest that both global uncertainty and optimism have overall increased over the last 15 years. We extend upon the IMF measure to construct indices on the country-level and find considerable regional concentration of uncertainty and optimism, respectively. Besides, we employ a simple rule-based sentiment analysis to measure the tone of global economic outlook and note several improvements for future work.

Table of Content

1. Introduction

At the beginning of the COVID-19 pandemic, global uncertainty reached unprecedented levels and remains elevated yet today, having seen a more general increase over recent years (IMF, 2021). In response to the pandemic, governments and central banks around the world engage in notable efforts to re-establish economic confidence and optimism. For example, the Biden administration recently passed a 1.9 trillion US dollar stimulus package - worth 10% of the country's GDP - to pave its return to pre-crisis growth trajectories (Giles, 2021).

In this context, the paper at hand constructs two benchmark indices to measure global uncertainty and optimism over time and across countries, following the methodology of the International Monetary Fund's *World Uncertainty Index* (Ahir et al., 2018). The *World Uncertainty Index* (WUI) is a quarterly measure for global economic and policy uncertainty and is computed by counting the frequency of the word *uncertain* (and its variant) in *Economist Intelligence Unit* reports for 143 countries starting in the first quarter of 1990.

To the best of our knowledge, our optimism index is the first cross–country measure of optimism for a large range of developing and advanced countries. More specifically, our indices capture uncertainty and optimism related to both economic and political developments, combining short- and longer-term outlook. To make our approach operational, we exploit a large text corpus comprising 1,844 Article IV consultation reports published by the International Monetary Fund (IMF) between 2004 and 2020 for 190 countries. Constructing simple frequency measures that exploit the occurrence of the terms *uncertainty* and *optimism* (and their variants) in these reports, we find that both global uncertainty and optimism have overall increased over the last 15 years. Besides, we extend our measures to the country–level and find considerable regional concentration of uncertainty and optimism, respectively. Moreover, our results lend support to the existence of uncertainty spillovers from systemic countries to the rest of the world, while we find little evidence for optimism spillovers from these countries. We extend upon our frequency optimism index using a simple rule-based sentiment analysis and find evidence for similar although lagged patterns of our sentiment index and our benchmark optimism index. Finally, we note several improvements for future work.

Similar to the *World Uncertainty Index*, our approach benefits from the advantage of using one single and standardized source with a pre-defined topic coverage. A notable downside of our time series measures is their annual frequency, as we expect missing out on substantial variation of uncertainty and optimism *within* years to dilute our results.

This paper proceeds in six more sections: Section 2 briefly reviews the related literature, while Section 3 introduces our text corpus. Section 4 summarizes our methodology and results. Section 5 includes robustness checks, followed by a discussion of potential caveats in Section 6. Lastly, Section 7 concludes.

2. Literature

Overall, our study contributes to the empirical literature that aims to measure economic uncertainty and optimism. In methodical terms, the paper at hand relates to a body of literature that lays out the groundwork of Natural Language Processing (NLP) to make our approach operational (see, e.g., Gentzkow et al., 2019).

In thematic terms, our study speaks to the literature that aims to empirically measure macroeconomic uncertainty. Uncertainty in this context is typically defined as the conditional volatility of disturbance that is unforecastable for economic agents. Next to the frequency of particular *uncertainty*-related keywords, popular uncertainty measures employed in previous literature include the implied or realized volatility of stock market returns, or the cross-sectional dispersion of subjective (survey-based) forecasts. Jurado et al. (2015) extend upon the above mentioned measures providing direct econometric estimates of time-varying macroeconomic uncertainty, aiming to detach their measure from theoretical priors and dependencies on any single (or small number of) observable economic indicators. Their results suggest that much of the variation in the above measures of uncertainty is in fact not driven by actual variation in uncertainty.

Exploiting text data, Alexopoulos and Cohen (2009) construct an uncertainty index based on the number of *New York Times* articles containing keywords related to uncertainty and economic activity. In similar fashion, Baker et al. (2016) construct an Economic Policy Uncertainty (EPU) index covering an extended range of news articles and keywords. Furthermore, Larsen (2017) uses a Latent Dirichlet Allocation (LDA) model to disentangle different types of uncertainty in Norwegian business newspaper articles. As indicated above, our approach besides relates closely to Ahir et al. (2018) who develop an uncertainty measure based on term frequency in *Economist Intelligence Unit* reports.

Besides, our study is related to the literature aiming to identify the effects of increased uncertainty on aggregate macroeconomic fluctuations. Several studies find that an increase in uncertainty foreshadows worsening economic conditions such as declines in investment, output, and employment (Bloom, 2009; Jurado et al., 2015). Besides, Bloom et al. (2012) find that elevated uncertainty can alter the relative impact of government policies, making them initially less effective while then subsequently more effective. Basu and Bundick (2012) provide evidence suggesting that households tend to make precautionary savings in response to increasing uncertainty, in turn affecting aggregate demand. The study at hand contributes to this literature by testing the (descriptive) relationship between changes in economic uncertainty and changes in optimism, which we define as overall confidence regarding economic outlook.

In contrast to uncertainty, empirical measures of macroeconomic optimism are rarely explored in the existing literature. Bhandari et al. (2019), for example, create "belief wedges", measuring differences between households' expectations and statistical forecasts on macroeconomic outcomes. Using these wedges as notions of pessimism and optimism, they find a significant effect of beliefs on macroeconomic fluctuations. The paper at hand contributes to this literature by constructing a new measure of economy-wide optimism based on text data and simple machine learning techniques.

3. Data

This section introduces the text corpus we exploit to measure cross-country uncertainty and optimism. Following a brief classification of the IMF Article IV reports, we describe the employed scraping procedure and the main properties of our dataset. We use Article IV consultation reports as a publicly available alternative to the reports disclosed within the Economist Intelligence Unit that are used to construct the WUI.

3.1 Article IV country reports

As one of the most influential international financial institutions, the IMF offers technical assistance on economic affairs, provides loans to countries in need, and engages in monitoring of economic and financial policies. In this context, an essential part of the IMF's monitoring is carried out as part of the consultations based on the Article IV of its *Articles of Agreement*.

The results of the surveillance process are summarized in IMF country reports prepared by IMF staff teams after discussion with officials of the country, serving the purpose of surfacing any risk to domestic and global stability. In particular, a team of IMF economists visits the country in-person to gather data and information, and to hold discussions with mainly government and central bank officials (referred to as the *Authority's View* in the reports, Figure 1), but also labor representatives, members of parliament, civil society organizations, and private investors, which then serve as the basis of the prepared staff report. In principle, Article IV consultations with member countries are carried out annually.[1] The staff reports touch on a variety of key topics, including monetary, fiscal, and regulatory policies, the exchange rate, and macro-critical structural reforms. As such, the reports provide a window into the IMF's assessment of the most important macroeconomic issues and economic risks. (Mihalyi and Mate, 2018)

Many of the key topics are reflected in the below word cloud of the 40 most frequent words across all Article IV reports in our sample (Figure 1). Examples include the terms *exchange* and *inflation* for exchange rate and monetary policy, respectively, while *labor market*, *wage* and *(un-)employment* for regulatory policy, and *wage*, *corporate* and *reserve* for fiscal policy.

[1] Extended consultation cycles longer than 12 months can apply if certain criteria are met, e.g. if the country is *not* of systemic or regional importance. Countries under IMF program may also be placed on a 24-month consultation cycle, but will generally report more frequently (semi-annually or quarterly) within particular program review reports. (Mihalyi and Mate, 2018)

Figure 1: *Word cloud with the 40 most frequent words in our Article IV corpus*.
Note: A bigger font indicates a higher probability for the term to occur in a given report.

3.2 Scraping, cleaning and construction of the dataset

We scraped 8,168 observations from the IMF website[2] which were tagged as country reports, referring to the *entire* database, i.e. reports published between January 1, 1994, and January 31, 2021.[3] Following Mihalyi and Mate (2018) this was done using *Web Scraper*[4]. Within the scraping procedure we stored each document alongside the corresponding meta data as displayed on the IMF website, including the title and publication date of the document, its unique series ID, and the URL. We then converted searchable PDF documents into plain text using the *PDFtotext*[5] tool, while using *textract*[6] for scanned (i.e. non-searchable) documents, and then combined the plain text of the reports together with the meta data in one CSV-file. Besides, we took further steps to clean the metadata and complement the dataset with information on countries' geographic region and income group.

For our machine learning applications, we focus on the subset of Article IV reports in the corpus.[7] In short, this particular choice aims to ensure that the reports feature a similar structure and depth of analysis. Moreover, we expect the text body from these reports to be more balanced across major topics of relevance from the perspective of macroeconomic risks, as opposed to *thematic* reports which tend to focus on fewer, more specific topics.

We select the Article IV reports from the scraped documents by searching "Article IV" in their title in the stored metadata. Coinciding with the year when Article IV reports became published by default, we select the reports published between 2004 and 2020.[8] [9] We also drop 21 reports that do not refer to a single country but to a country group such as currency unions.

As is common practice in the NLP literature, the raw text is cleaned prior to further application (Gentzkow et al., 2019). In this context, we remove numbers, punctuation, line spacing, and unidentified apostrophes. Moreover, we remove stop-words and conduct stemming of the word tokens, offering a clean version of the report text for further application.[10]

---

2 https://www.imf.org/en/publications/search?when=After&series=IMF+Staff+Country+Reports.

3 The IMF website does not report a PDF document for 402 of these report entries (mainly before 1998). We did scrape a PDF document for the remaining 7,766 observations.

4 https://webscraper.io.

5 https://github.com/jalan/pdftotext.

6 https://textract.readthedocs.io/en/stable/.

7 This implies that we do not consider *Program Review Reports* and all thematic country reports, including *ROSC* reports on the compliance with various international standards and codes, *Financial Sector Assessments*, or *Selected Issue Reports*.

8 Prior to selecting reports we correct the given *publication* year to the actual report year based on the title where possible and applicable (i.e. for 548 reports).

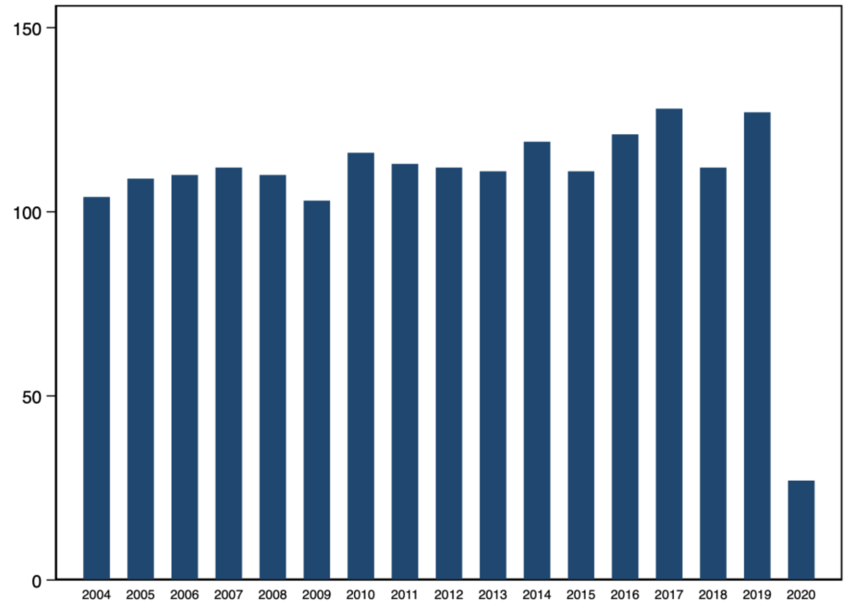
9 Moreover, there tend to be much fewer country reports from before 2004, and among these the majority are scanned documents which makes text recognition difficult and less precise.

10 In particular, we employ the *NLTK* stop-word list (English) and use the Porter stemming algorithm.

3.3 Main properties

The reduced panel dataset comprises 1,844 Article IV reports for 190 countries for the years between 2004 and 2020. Besides the sharp drop in available reports for the year 2020, the number of available reports is relatively stable, although tends to slightly increase over time. A given year sees 113 reports on average, with a minimum of 27 (2020) and maximum of 128 (2017) (Figure 2).

Figure 2: *Number of Article IV reports per year, 2004-2020*.



The distribution of reports per country is rather uneven, presumably as a result of some countries receiving more IMF staff visits than others (Figure 3). Both, the median and average number of reports per country refer to 11, with values ranging between 1 and 18. Considering outliers in the report numbers, there are 12 countries that have 4 or less reports, and 9 countries with 16 or more reports.[11]

Our dataset covers most of the countries worldwide (Figure 4). There are some notable exceptions such as Venezuela, and our panel tends to contain more reports from Europe, North America, Asia, and certain African and Latin American countries.[12]

11 Countries with 4 or less reports: Anguilla, Argentina, Brunei Darussalam, Ecuador, Honduras, Macao, Montserrat, Nauru, Netherlands Antilles, Somalia, South Sudan, Tajikistan. Countries with 16 or more reports: China, France, Germany, Japan, Mexico, Netherlands, Russian Federation, United Kingdom, United States.

12 In particular, the most reports come from *Europe* by a considerable margin. *Asia and the Pacific, Africa*, *America and the Caribbean* all report around two thirds of the number from Europe, while the *Middle East and Central Asia* around one half of the European number (Figure 20 in the appendix). The reports are rather evenly distributed across income groups, with the most reports coming from lower middle income economies (Figure 21).

Figure 3: *Distribution of Article IV reports per country.*

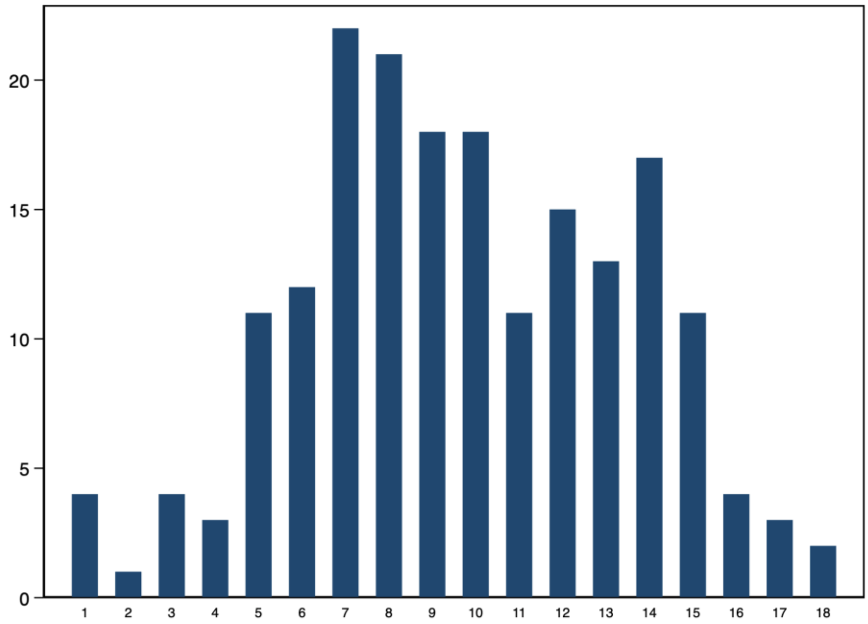
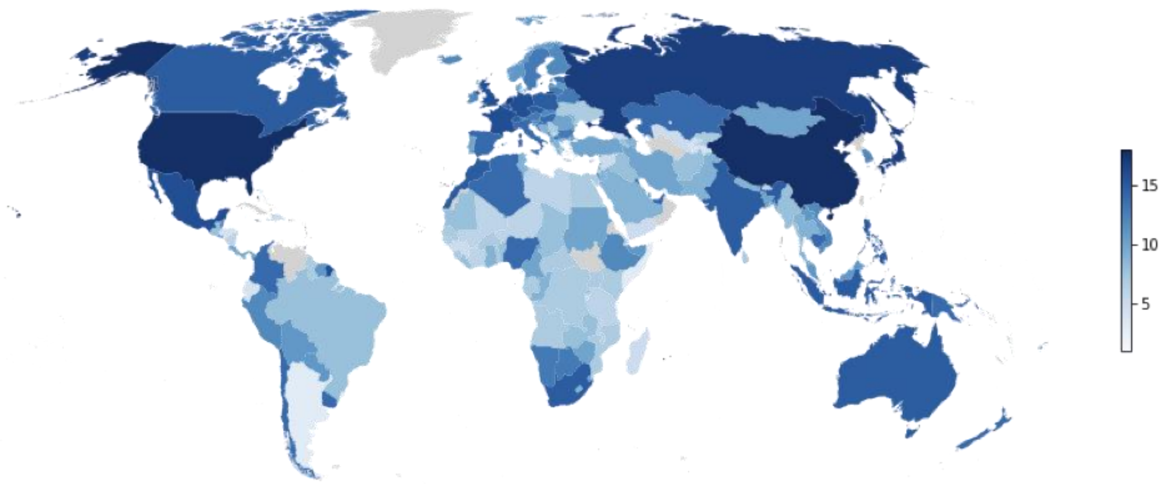


Figure 4: *Country coverage of the Article IV subset*.
Note: Darker blue color indicates a higher number of reports for a given country over the entire sample period.

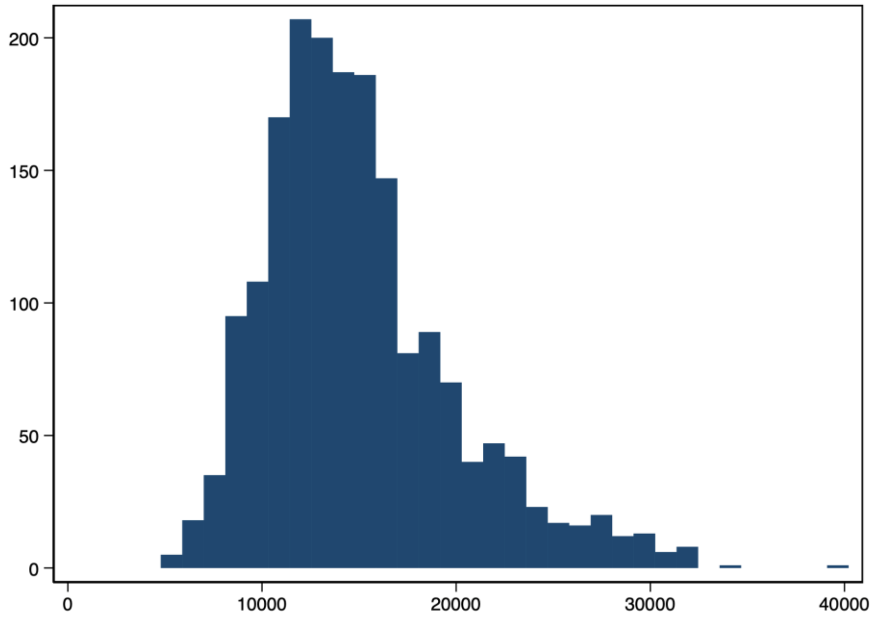


The descriptive statistics for our corpus in Table 1 show that the length of the Article IV reports increases over time, from around 14,000 words on average in 2004 to around 19,000 words in 2020. The standard deviation of the mean word count per report tends to be stable. Moreover, the minimum number of words per report tends to increase over time, while the maximum number of words per report is more or less stable over time. Figure 5 illustrates the distribution of word counts in the corpus, showing that the distribution tends to be skewed towards the right side.

Table 1: *Descriptive statistics for the Article IV corpus*.

Year	No. of reports	Mean word count	Std. deviation	Minimum	Maximum
2004	104	13,602.38	4,468.76	7,304	29,990
2005	109	13,733.31	4,453.61	6,646	29,301
2006	110	13,086.78	4,152.34	5,271	24,682
2007	112	12,278.99	4,108.20	5,205	25,698
2008	110	12,398.15	3,840.24	6,055	22,773
2009	103	13,009.64	4,378.29	4,790	26,937
2010	116	13,159.97	4,218.43	6,770	29,713
2011	113	13,797.22	4,484.18	6,376	31,736
2012	112	15,481.79	5,005.33	8,080	32,101
2013	111	15,934.78	5,177.71	8,391	30,686
2014	119	15,349.51	4,304.39	7,837	30,020
2015	111	15,666.34	4,582.35	7,859	31,961
2016	121	1,7295.36	4,882.04	8,705	40,210
2017	128	17,204.35	4,492.02	10,543	31,696
2018	111	17,165.71	4,849.35	10,134	32,281
2019	127	18,354.80	4,850.98	8,775	34,102
2020	27	19,228.22	4,193.96	13,127	27,947

Figure 5: *Distribution of word counts per report in the corpus*.



4. Methodology and results

This section describes the construction of our benchmark indices and discusses their properties both over time and across countries. Besides, we extend our measures to quantify the role of spillovers in the formation of global uncertainty and optimism.

4.1 Uncertainty index

Our methodology to quantify global uncertainty over time closely follows the construction of the World Uncertainty Index (WUI) (Ahir et al., 2018). The WUI is a quarterly measure of global economic and policy uncertainty covering 143 countries over the period from 1996 to 2020. Essentially, the measure relies on a simple term frequency procedure that returns how often the word *uncertain* (and its variant) occurs in the *Economist Intelligence Unit* (EIU) country reports. To make the measure comparable across countries, the raw word counts are scaled by the total number of words in each report. The relative measure is then rescaled by 1,000,000 such that, for example, an index value of 400 corresponds to the word *uncertainty* accounting for 0.04% of all words, which refers to 4 words per report, given that the EIU reports are on average 10,000 words long. Moreover, the series is weighted by real GDP to account for heterogeneity in the contribution to global uncertainty across countries. Taken together, a higher value of the index thus indicates higher levels of uncertainty and vice versa.

In particular, our procedure counts the occurrence of the stems *uncertainti* and *uncertain* in every report to capture the variants of the word *uncertain*, and then follows the above procedure.[13] Figure 6 shows our benchmark uncertainty index for the period from 2004 to 2019.[14] Overall, the figure suggests that global uncertainty has increased over time. In particular, the level of uncertainty reached an overall high in 2013, being around three times higher than the period-time low in 2007. Similarly, the level of uncertainty tends to range below its sample average in the first half of the period, while above its mean in the second half. Our index besides suggests uncertainty to have increased in the three years after 2007, a pattern overall consistent with the narrative of the financial crisis. The level of uncertainty again notably rises after 2011. In principle, this increase could at least in part be related to the US fiscal cliff and sovereign debt crisis in Europe.

Figure 6: *Benchmark uncertainty index, 2004-2019*.

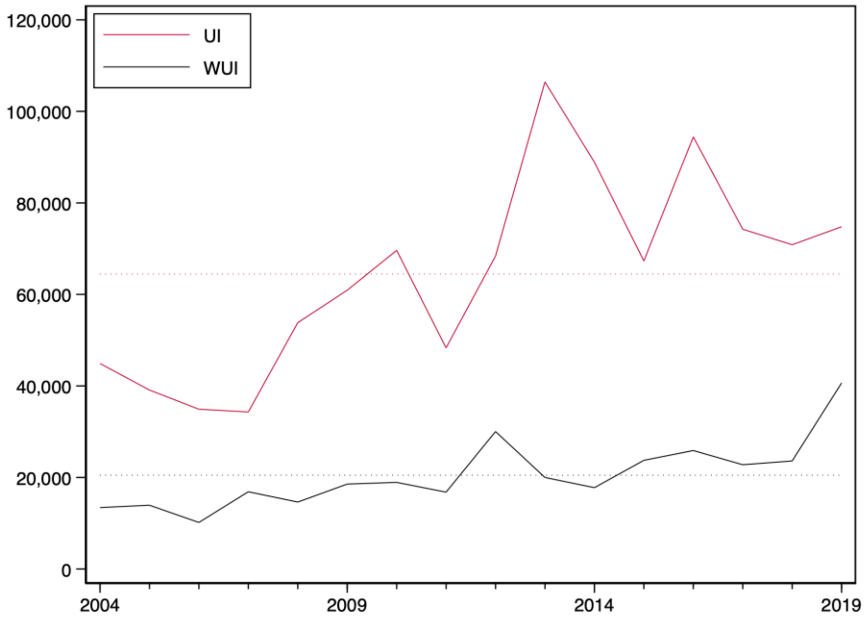
Note: The index is computed by counting the frequency of the stems *uncertainiti* and *uncertain* in Article IV consultation reports, and is then normalized by the total number of words in each report, and rescaled by multiplying with 1,000,000. The horizontal line indicates the sample average of the index.

[13] The former stemm refers to the words *uncertainty* and *uncertainties*, while the latter stemm refers to *uncertain* and *uncertainness*.

[14] The real GDP (constant 2010 USD) data comes from the World Bank (https://data.worldbank.org/indicator/NY.GDP.MKTP.KD?view=map). As this source does not report any data for 2020, we report our time series until 2019.

Figure 7 compares our index against the WUI.[15] Our uncertainty index (UI) features higher index *levels* as compared to the WUI, which indicates that the word *uncertain* (and its variant) tends to generally occur more often in the IMF reports as compared to the EIU reports. In terms of correlation between the two series, we find a positive coefficient of 0.51 which is significant on the 5% level. This indicates that our replication index tends to mimic movements in the WUI.

Figure 7: *Benchmark uncertainty index vs. World Uncertainty Index, 2004-2019*.
Note: The quarterly WUI series has been converted to annual frequency by taking the mean of the quarterly index values in a given year. The horizontal lines indicate the sample average of the annual indices.



Furthermore, we extend upon WUI methodology and construct an aggregate *country*-level measure of uncertainty. In particular, we count all words in all reports for a given country (i.e. over the entire sample period) and then count the occurrence of the word *uncertain* (and its variant) in all reports for a given country. We then calculate the relative word frequency as percentage share and rescale the index as above by multiplying with 1,000,000. This procedure returns an index value for every country in our sample, measuring the average level of uncertainty in a country over the entire sample period.

Figure 8 illustrates our country-level uncertainty index for the entire set of countries in our sample. Our calculations suggest that the North America region is characterized by comparably high levels of uncertainty, with Central Europe featuring similar levels. Moreover, South America also sees a considerable degree of uncertainty on average, just as the Middle East. Uncertainty tends to be rather less pronounced in Asia, with some exceptions such as Japan and Thailand. Besides, uncertainty in South Africa is notably higher than in the remainder of Africa, where uncertainty levels tend to be low compared to the rest of the world. The five countries featuring the lowest level of uncertainty are (in ascending order): Montserrat, Somalia, Yemen, Senegal, and Benin.[16] By contrast, the five countries in our sample characterized by the highest levels of uncertainty in the reporting period are (in descending order): United Kingdom, Thailand, Japan, Germany, and Mexico.

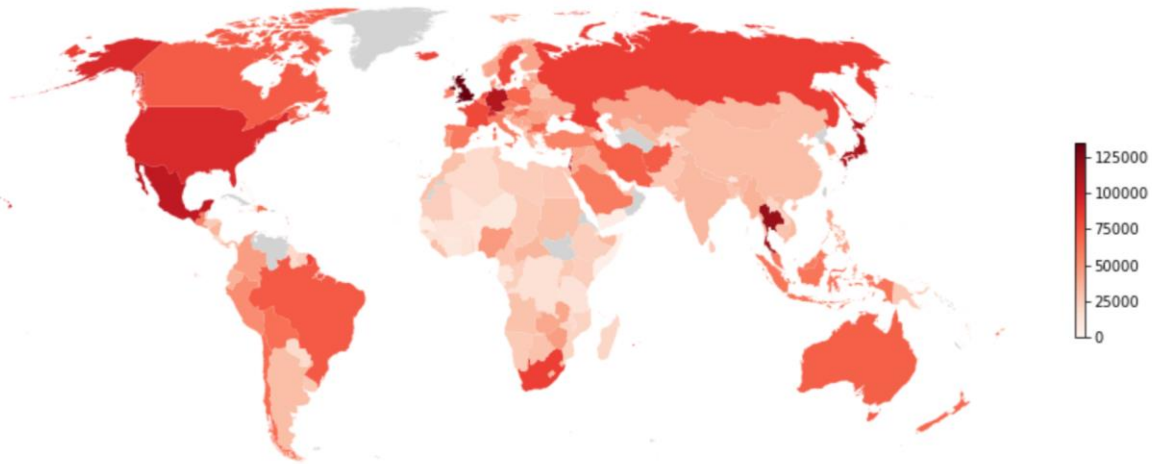
We leave open for future work to quantify the extent to which our country-level indices are correlated with dimensions like population or real GDP (growth).

15 The data on the WUI comes from https://worlduncertaintyindex.com. Note that we are taking out some of the variation of the quarterly WUI series by converting to annual frequency (see Figure 22 in the appendix).

16 It may be surprising that our measure suggests Yemen, a country under ongoing civil war since several years, to be among the countries with the (globally) lowest level of uncertainty. Besides, a variety of armed conflicts currently ongoing globally are taking place in the regions that our index suggests to be characterized by low uncertainty, e.g. in Africa (see https://en.wikipedia.org/wiki/List_of_ongoing_armed_conflicts). In case the situation (e.g. war) in a country in some way affects the Article IV consultation in this country, potential measurement errors in our index are unlikely to be random, which would negatively affect the internal validity of our approach.

Figure 8: *Country-level uncertainty index*.

Note: The index is computed by counting the frequency of the stems *uncertainiti* and *uncertain* in all reports for each country, and is then normalized by the total number of words over all reports for each country, and rescaled by multiplying with 1,000,000.

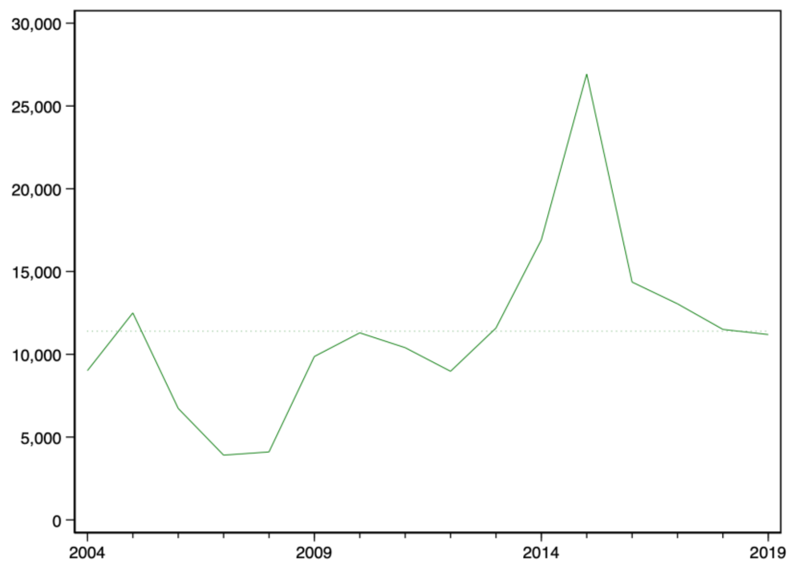


4.2 Optimism index

To obtain a measure for global economic optimism over time, we transfer the methodology from above and exploit the term frequency of the stems *optim* and *optimist* to capture the variants of the term *optimism*.[17] Figure 9 shows our benchmark optimism index for the period from 2004 to 2019. The figure illustrates that optimism overall increased over our sample period, with notable variation over time, however. The level of optimism overall ranges below its sample average in the first half of the period (where uncertainty was below average), while above its mean in the second half (where uncertainty was above average). Moreover, we find a period-low level of optimism around 2007 and 2008, and a consistent increase after 2012 mounting into a period-high in 2015, to then return to sample average levels in 2018. In particular, our index suggests the world was around six times as optimistic at peak in 2015 compared to the low in 2007.

Figure 9: *Benchmark optimism index, 2004-2019*.

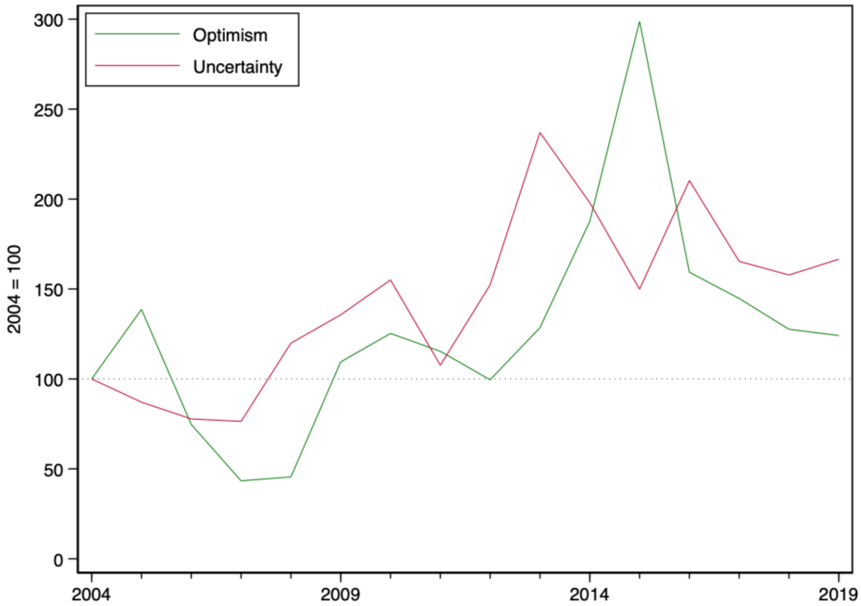
Note: The index is computed by counting the frequency of the stems *optim* and *optimist* in Article IV consultation reports, and is then normalized by the total number of words in each report, and rescaled by multiplying with 1,000,000. The horizontal line indicates the sample average of the index.



[17] The former stem refers to the word *optimism*, while the latter to *optimistic*. We note that the stem *optim* also picks up words such as *optimal*, and leave a robustness check of the indices without stemming for future work.

Figure 10 plots our benchmark optimism index against our benchmark uncertainty index. Overall, the figure illustrates that global optimism and uncertainty follow similar trends, with the relationship being reversed in the years around the optimism peak in 2015, however. In line with this pattern, we find a positive correlation coefficient of 0.45 for the two series over the entire sample period, being significant on the 10% level. This suggests that (changes in) optimism and uncertainty overall tend to go hand in hand, i.e. increasing uncertainty tends to be associated with increasing optimism and vice versa. By contrast, we find the correlation between our benchmark optimism index and the WUI to be insignificant (see Figure 23 in the Appendix).

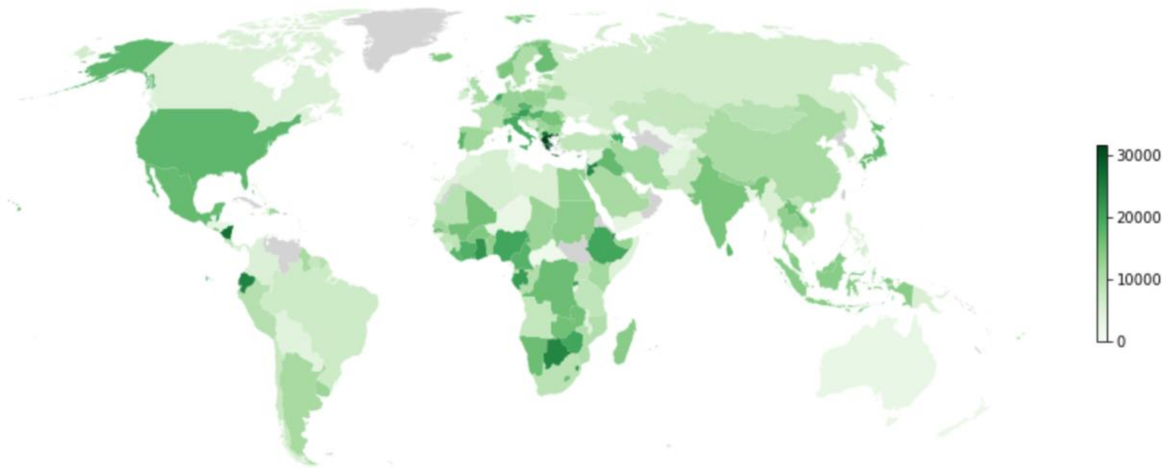
Figure 10: *Benchmark optimism index vs. benchmark uncertainty index, 2004-2019*.



Furthermore, we construct the equivalent of our above country-level index for uncertainty also for optimism. Figure 11 illustrates our country-level optimism index for the entire set of countries in our sample. Our index suggests that the United States together with Mexico and Europe feature comparably high levels of optimism. Moreover, Africa is also suggested to be characterized by comparably high optimism overall. The five most optimistic countries in our sample are (in descending order): Greece, Gambia, Nicaragua, North Macedonia, and Ecuador. By contrast, the five least optimistic countries are (in ascending order): Uzbekistan, Micronesia, Anguilla, Netherlands Antilles, and Montserrat. At the same time, we do not find a significant relationship between our country-level uncertainty and optimism indices.

Figure 11: *Country-level optimism index*.

Note: The index is computed by counting the frequency of the stems *optim* and *optimist* in all reports for each country, and is then normalized by the total number of words over all reports for each country, and rescaled by multiplying with 1,000,000.



4.3 Spillovers

Economic growth in key systemic economies like the United States or the European Union are a key driver of economic activity in the rest of the world. Given increasing levels of interconnectedness of countries worldwide, one could expect that uncertainty arising from, for example, presidential elections in the United States, Brexit, or China-US trade tensions spills over and affects uncertainty in other countries. (IMF, 2021)

To test whether uncertainty in systemic economies matters for uncertainty around the world, we follow Ahir et al. (2018) and construct an index that measures the extent of uncertainty spillovers from the G7 countries plus China to the rest of the world. Essentially, our measure exploits the frequency of the word *uncertainty* (and its variants) in the reports in proximity to a term related to the respective systemic country. Our word list closely follows the keywords as applied within the WUI, and includes the country's name, name of presidents, name of the central bank, name of central bank governors, and selected country's major events (such as Brexit).[18] [19]

In particular, we search for the country-specific keywords in every report in a given year (other than from the systemic country of interest) and count the occurrence of *uncertainty* (and its variants) in the two words before and after a given keyword (Table 2).[20] Besides, we count the occurrence of *uncertainty* (and its variants) in all reports in a given year and then calculate the share of the yearly aggregated neighborhood counts in the entire uncertainty count as measure for uncertainty spillovers from this particular systematic country.[21] Put differently, our spillover index measures the ratio of uncertainty related to the systemic country to overall uncertainty.

Table 2: *Country-specific keywords for the G7 plus China*.

Country	Keywords
Canada	Canada, Bank of Canada, David A. Dodge, Gordon Thiessen, Jean Chretien, Justin Trudeau, Mark Carney, Paul Martin, Stephen Harper, Stephen S. Poloz, Tiff Macklem
France	Banque de France, Christian Noyer, Emmanuel Macron, France, Francois Hollande, Francois Villeroy de Galhau, Jacques Chirac, Jean-Claude Trichet, Nicolas Sarkozy
Germany	Germany, Bundesbank, Jens Weidmann, Axel A. Weber, Ernst Welteke, Hans Tietmeyer, Helmut Kohl, Gerhard Schröder, Angela Merkel
Italy	Antonio Fazio, Bank of Italy, Enrico Letta, Giuliano Amato, Ignazio Visco, Italy, Lamberto Dini, Mario Draghi,

18 We thank Hites Ahir from the IMF for providing us with the entire list of country-specific keywords.

19 For reasons of simplicity, we take the exact list (augmented with *Great Britain* for the UK) as the WUI does, despite Bill Clinton, for example, not being president of the United States during our sample period.

20 Technically, we search for the (unstemmed) keywords, extract and then stemm the neighborhood tokens, and then count the frequency of our respective uncertainty stems.

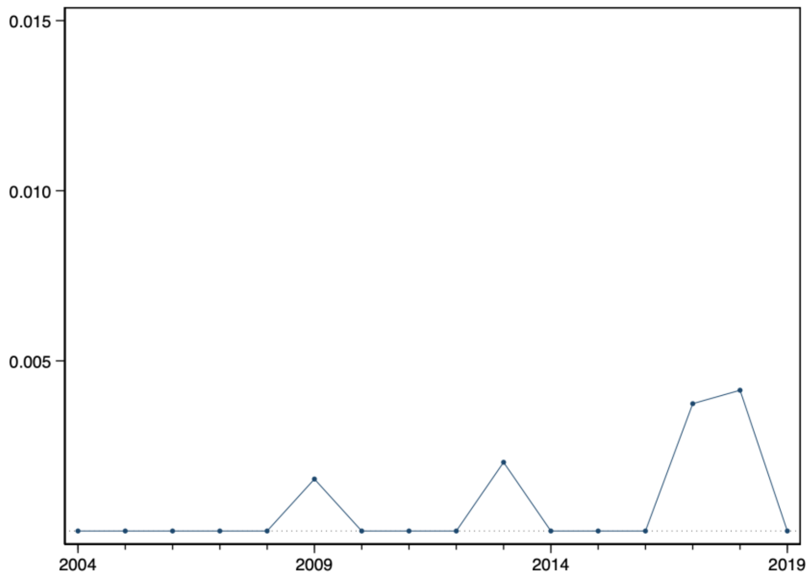
21 Note that our quantification of spillover effects is directly subject to the defined range around a given keyword. Deciding on a particular range sees a trade-off between the *accuracy* to capture *uncertainty* only around a respective keyword and the *possibility* to pick up a sufficient number of words around a keyword (e.g. for semantic or grammatical reasons). Given that we remove stopwords before extracting neighborhood words, we are confident to address this trade-off rather well. We find very similar patterns and magnitudes when extending our set of *optimism* terms to the third random sample (i.e. benchmark optimism terms plus five randomly selected words). We leave a robustness check with a broader range (e.g. four words in either direction) for future work. In this sense, our results serve as a lower bound of spillover effects.

	Mario Monti, Massimo D'Alema, Matteo Renzi, Paolo Gentiloni, Romano Prodi, Silvio Berlusconi
Japan	Bank of Japan, Haruhiko Kuroda, Japan, Junichiro Koizumi, Keizo Obuchi, Masaaki Shirakawa, Masaru Hayami, Naoto Kan, Ryutaro Hashimoto, Shinzo Abe, Taro Aso, Toshihiko Fukui, Yasuo Fukuda, Yasuo Matsushita, Yoshihiko Noda, Yoshiro Mori, Yukio Hatoyama
United Kingdom	Andrew Bailey, Bank of England, Boris Johnson, Brexit, Britain, David Cameron, Edward George, Gordon Brown, Great Britain, John Major, Mark Carney, Mervin King, Theresa May, Tony Blair, United Kingdom
United States	Alan Greenspan, America, Barack Obama, Ben Bernanke, Bill Clinton, Donald Trump, Federal Reserve, George H. W. Bush, George W. Bush, Janet Yellen, Jerome Powell, NAFTA, North America, United States
China	China, Dai Xianglong, Hu Jintao, Jiang Zemin, People's Bank of China, Xi Jinping, Yi Gang, Zhou Xiaochuan

Figure 12 shows our index for spillovers from the United States. In particular, we find marginal spillovers of US uncertainty in the years 2009, 2013, 2017, and 2018. By contrast, the WUI more generally reports higher levels of US spillovers than our calculations do, with US uncertainty at peak in 2017 accounting for up to 30% of uncertainty in other countries. (IMF, 2021)

Figure 12: *Uncertainty spillovers from the United States*, 2004-2019.

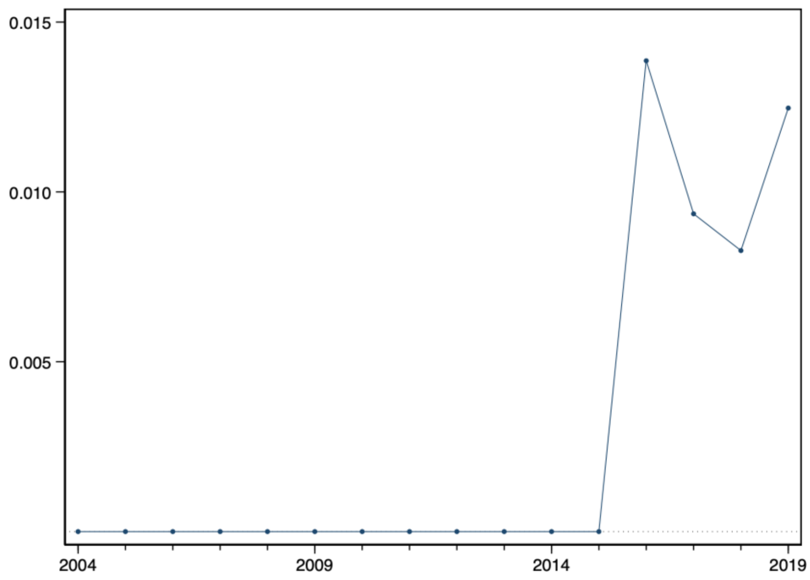
Note: The figure shows the ratio of uncertainty related to the United States to global uncertainty in a given year.



Moreover, Figure 13 shows our index for uncertainty spillovers from the United Kingdom. Interestingly, our results suggest that the United Kingdom became a contributor to uncertainty in other countries after the Brexit referendum in 2016, with spillovers of around 1.2% at peak. Besides, spillovers from the United Kingdom remain on elevated levels in the years after the referendum. Overall, these results line up with what WUI spillovers suggest, however again with significantly lower *levels* of spillovers in our case.

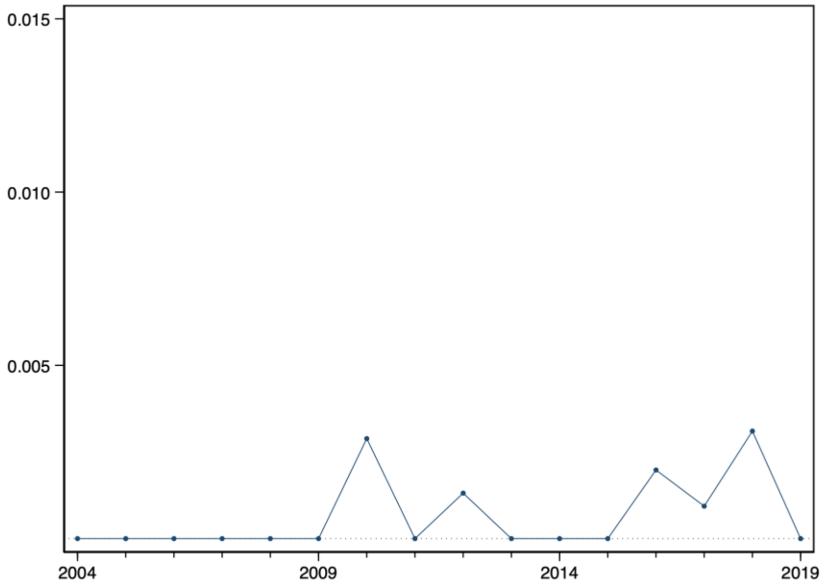
Figure 13: *Uncertainty spillovers from the United Kingdom*, 2004-2019.

Note: The figure shows the ratio of uncertainty related to the United Kingdom to global uncertainty in a given year.


0.015
0.010
0.005
2004
2009
2014
2019

Finally, Figure 14 shows our index for (combined) uncertainty spillovers from the remaining five G7 countries (i.e. Canada, France, Germany, Italy, and Japan) and China. Overall, we find that these countries combined have little spillover effects on the rest of the world, lining up with what WUI spillovers suggest. However, while these systemic economies might have limited spillovers on *global* uncertainty, they could still have more pronounced *regional* uncertainty effects, like Germany for the other European economies (IMF, 2021).

Figure 14: *Uncertainty spillovers from the other G7 countries and China*, 2004-2019.
Note: The figure shows the ratio of uncertainty related to Canada, France, Germany, Italy, Japan (other G7) and China to global uncertainty in a given year.



We also extend our analysis to an equivalent index of optimism spillovers using the respective words stems as introduced above. Essentially, our results do not lend support for the existence of optimism spillovers from systemic economies to other countries.

Taken together, our analysis lends support for the existence of uncertainty spillovers from systemic economies on the rest of the world. By contrast, we find little evidence for optimism spillovers from these countries.

5. Robustness

To address the concern that terms other than our above keywords *uncertainty* and *optimism* (and their variants) could be used to capture respective sentiment, we augment our frequency measure with further terms related to *uncertainty* and *optimism*. In this sense, our frequency measures see a trade-off between *accuracy* (i.e. suitable words) and the *size* of the set of selected words to reliably capture sentiment.

In a first step, we randomly select a certain number of "other words" that are labelled as "most relevant" in the *Thesaurus* online dictionary for the terms *uncertainty* and *optimism*, respectively.[22] In particular, we generate three random sets for each of our two keywords and include them in our index measure: Next to *uncertainty* or *optimism* we include the word frequency of (1) two randomly selected words out of the set of synonyms, (2) two randomly selected words not selected in the first set, and (3) five randomly selected words from the entire set.

Figure 15 shows our benchmark uncertainty index from above together with the respective indices resulting from our random selection experiment. The figure shows that our selection picks up one or more words that considerably increase the levels of our index. This should not be surprising per se, given the *frequency* nature of our index. By contrast, the remaining two sets closely mimic the path of our benchmark index, featuring a correlation coefficient of 0.995 and 0.999, respectively, significant on the 1% level. For the other (higher-level) series we find a correlation coefficient of around 0.52 with the benchmark index, yet significant on the 5% level.

Figure 15: *Benchmark uncertainty index with random samples, 2004-2019*.

Note: Horizontal line indicates average level of benchmark index over the entire period.

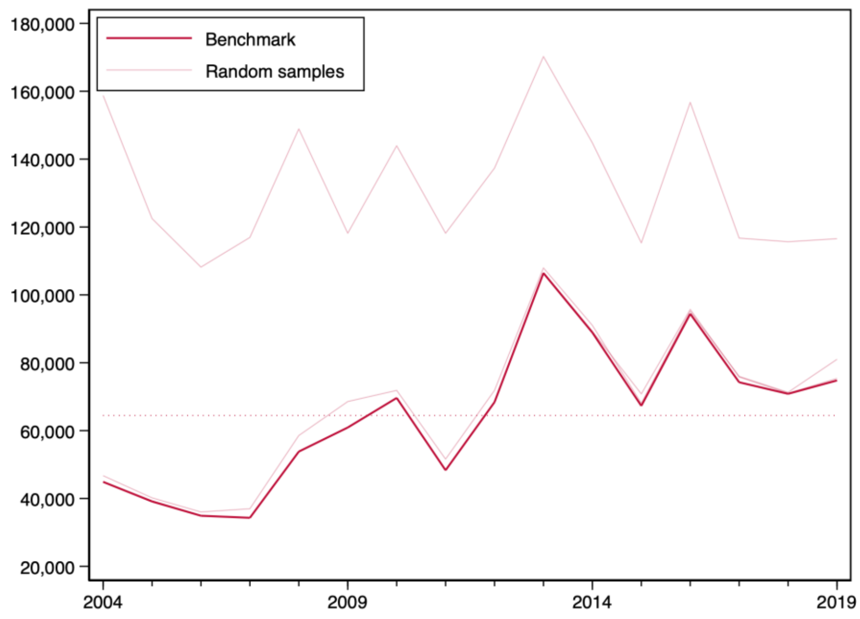


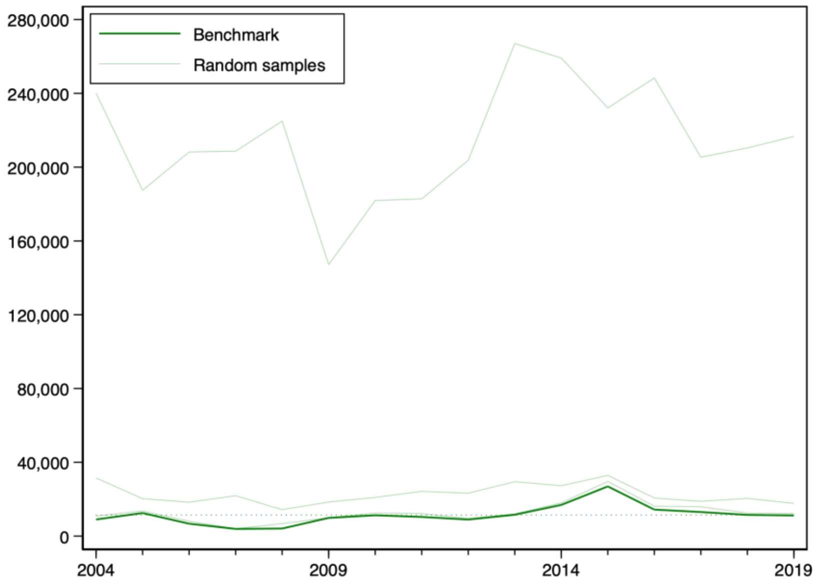
Figure 16 shows the equivalent results for our optimism index. Again, the random selection picks up one or more words that generally tend to increase the level of the index, while the two remaining sets again closely follow the path of the benchmark index. We find a correlation coefficient between the benchmark index and the latter two samples of around 0.99 and 0.57, significant on the 1% percent level, while an insignificant correlation for the remaining (high-level) series.

As such, we emphasize that a comparison of levels *across* indices is directly subject to the (aggregated) occurrence of the *selected* terms and therefore should be treated with caution. However, a consistent representation using the same set of words over time allows to compare index levels *within* a particular series. Besides, our experiment illustrates that the *variation* of the index (i.e. the main property we are interested in) can, in principle, be affected by including additional terms in our frequency measure. We leave open for future work to discuss the relevance of the employed word list in measuring uncertainty and optimism in economic terms.

Figure 16: *Benchmark optimism index with random samples, 2004-2019*.

22 For *uncertainty*, these (13) "most relevant" synonyms are: ambiguity, ambivalence, anxiety, concern, confusion, distrust, mistrust, skepticism, suspicion, trouble, uneasiness, unpredictability, and worry (https://www.thesaurus.com/browse/uncertainty?s=t). The equivalent set for *optimism* (9) refers to: anticipation, certainty, confidence, elation, enthusiasm, expectation, happiness, idealism, trust (https://www.thesaurus.com/browse/optimism?s=t).

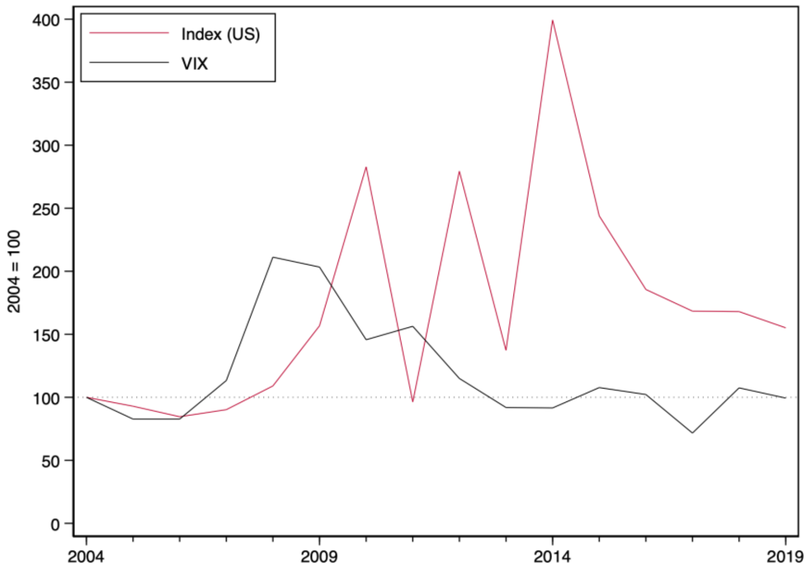
Note: Horizontal line indicates average level of benchmark index over the entire period.



Next, we compare our benchmark uncertainty index for the US against the CBOE Volatility Index (VIX), a popular measure of the stock market's expectation of volatility based on S&P 500 index options (Whaley, 2009). Specifically, we expect higher values of the VIX (and thus higher volatility) to capture higher uncertainty among investors. Figure 17 shows that our index and the VIX tend to follow similar movements until the start of the financial crisis in 2008, while the correlation for the entire sample period turns out to be insignificant.[23] [24]

We refer to earlier literature regarding uncertainty in financial markets (Arellano et al., 2010; Slovik, 2010) and leave a discussion of differences to overall economic uncertainty aside for future work.

Figure 17: *Benchmark uncertainty index for the US vs. CBOE volatility index, 2004-2019*.
Note: The VIX series has been converted to annual frequency by taking the mean of the daily index values in a given year.



[23] Note that we are taking out a substantial part of the variation in the VIX by converting the series to annual frequency (Figure 24 in the appendix). The data for the VIX comes from: https://ww2.cboe.com/products/vix-index-volatility/vix-options-and-futures/vix-index/vix-historical-data.

[24] By contrast, we find the WUI to be slightly positive and significantly correlated with the VIX (Figure 25 in the appendix).

Moreover, a major point of criticism of our approach is that we create our indices independent of context. For example, context is crucial when the word *neither* or *not* appears in front of our key terms. Another example would be the phrase “uncertainty has little to do with it”.

The goal of sentiment analysis is to incorporate context into the evaluation and understand the meaning behind the spoken or written word. This can be achieved by two variants of sentiment analysis: automatic approaches and rule-based approaches. The former approach is usually used when labeled data is available to train the model on different sentiments. In our case, this would mean manually classifying the reports into *optimistic* or *pessimistic* reports. Given the time constraints, we opted for the rule-based variant.

This approach uses a predefined lexicon and classifies words into categories such as positive, negative or neutral. Based on this classification, a picture of the overall sentiment of the report then emerges. A widely used field-specific dictionary in this context is the financial dictionary by Loughran and McDonald (2011). While this dictionary weights each word equally, providing a very simple method of sentiment analysis, we use the field-specific dictionary of Picault and Renault (2017). Using labeled phrases from European Central Bank (ECB) press conferences, the authors create a term-weighted dictionary that assigns a value between 0 and 1 to phrases (n-grams) for the categories of negative, positive, and neutral economic outlook.

In our pre-processing of the raw dataset in this context, we shrink our usual set of stop-words to include only words that are not found in the lexicon. This leaves us with a much shorter list, since Picault and Renault (2017) use a list of only 32 stop-words. Using the cleaned data, we split each report into tokens of two words each (bigram) and tokens of single words. For each n-gram that occurs in at least 1 percent of all reports, we count the frequency of their occurrence in each report. Bigrams are chosen for computational reasons. In the next step, we assign the corresponding value from the lexicon (positive, negative, or neutral) to each n-gram found in both the lexicon and the report and weigh them accordingly. The degree of positive outlook of a report is then the sum of the frequency of occurrence of all n-grams in a report weighted by the positive outlook. The negative outlook accordingly is the sum of the frequencies of the n-grams weighted by negativity. Finally, we obtain our sentiment index from the difference between the positive and negative economic outlook of all reports in one year. In line with our other indices, the sentiment index is weighted by the reporting countries' share in world real GDP.

Figure 18 shows the sentiment index for the period from 2004 to 2019. Starting with a positive outlook in the country reports in 2004, the positive outlook steadily declines, with one exception in 2010, and turns into a negative economic outlook in 2011. However, the aftermath of the global economic crisis thus appears to be over for the IMF's country experts, and the index evolves over a long period of a growing positive outlook with a short period of stagnation from 2014-2016, ending at an all-time high in 2017.

Figure 19 plots our sentiment index against our benchmark optimism from above. Taken together, we find that, in the wake of the global economic crisis, the optimism index has already picked up a positive trend in the reports somewhat earlier than the sentiment index. Similarly, the global maximum in the optimism index is already two years earlier than in the sentiment index. For the entire sample period, we find an insignificant correlation coefficient of 0.25, although the extremes in the two indices are very similar, albeit with a time lag.

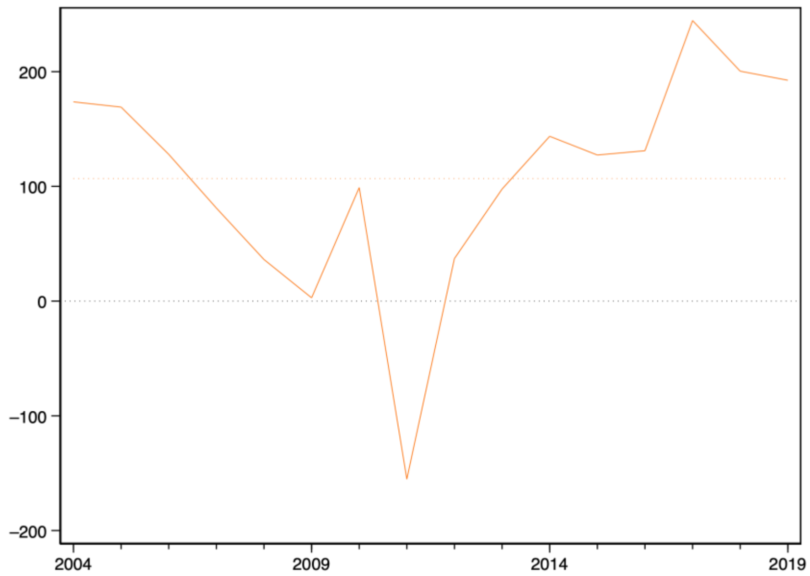
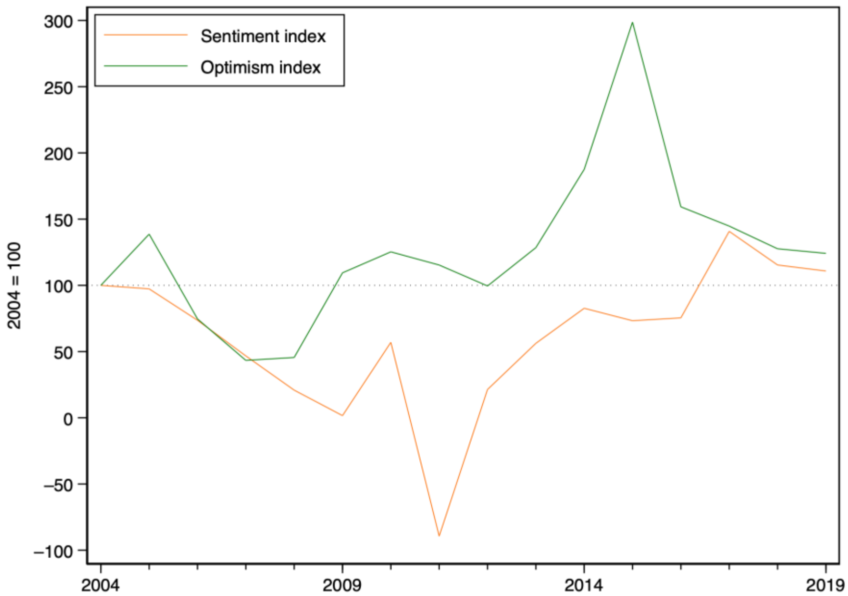


Figure 18: *Sentiment index, 2004-2019*.

Figure 19: *Sentiment index vs. benchmark optimism index, 2004-2019*.



6. Caveats

Next to potential limitations already addressed in the previous sections, this section briefly discusses further caveats to be aware of, together with potential improvements. First and foremost, as opposed to the *quarterly* WUI, the IMF's Article IV consultations usually take place *annually*, hence priming the frequency of our time indices.[25] In this sense, we tend to capture longer-term rather than *within*-year changes in uncertainty and optimism. We

25 For an *Economist Intelligence Unit* sample report as used for the WUI, see: https://store.eiu.com/article.aspx?productid=50000205. Overall, we expect the EIU reports to better capture *recent* policy developments in a given country (in the sense of news type reporting), which is likely to improve the measurement of short-term uncertainty that adds up to longer-term trends.

therefore expect missing out on substantial variation of uncertainty and optimism *within* years to dilute our results to a considerable degree.

Besides, Article IV reports are unlikely to be missing at random in our dataset for several reasons. First, not all IMF member countries get the same amount of surveillance. Given resource constraints of the IMF, for example, smaller or less accessible countries (e.g. through geography or language) may be monitored in less detail than others. By contrast, countries undergoing IMF program or countries with higher macroeconomic vulnerabilities may be more likely to get more intense monitoring. Besides, country authorities are per se able to block the publication of respective Article IV reports, which happened in around 20% of cases in 2004-2005, and then gradually declined to 5% of cases in 2014-2016. (Mihalyi and Mate, 2018)

Similarly, Edwards et al. (2011) find that more democratic governments are more likely to release their country reports, next to notable variation in regional patterns. Taken together, this could negatively affect the internal validity of our approach to the extent that the likelihood of a keyword appearing at all in reports will depend on the likelihood of the report being published in the first place, and its depth (Mihalyi and Mate, 2018). Put together with additional findings of Mihalyi and Mate (2018), our frequency measures could be somewhat biased towards larger, less wealthy, and more democratic countries as well as years of financial difficulty. The number of available reports per country in particular relates to our country-level measures. Here, we expect our indices to be more precise in the number of available reports for a given country.

Moreover, the Article IV consultation reports do notably vary in length (Table 1), which we do take into account by normalizing our index measures. With particular regard to the comprehensiveness of appendices, however, the reports can notably differ from each other. For example, the appendices often contain statements from persons such as government officials. Therefore, when taking the whole report as a basis (as we do above), the content of the reports does not necessarily come exclusively from the IMF staff, which could be another source of measurement noise. To achieve a higher degree of standardization in the sample, we started to extract the body of the reports and leave open for future work to refine our extraction algorithm. To give an example, we could use only relevant paragraphs covering the economic outlook section to avoid picking up uncertainty or optimism counts that refer to the past and therefore do not reflect the current (or expected future) situation.[26]

Finally, referring to our sentiment extension, alternatives to our rule-based method could likely refine the measurement of sentiment in the context of (standardized) country reports, which tend to strive for neutrality in tone in the first place. In particular, one could manually classify exemplary reports (or mostly relevant extracts from the reports) according to uncertainty or optimism and then train the model. However, the classification tends to be time-intensive since a sufficiently large number of reports has to be classified in order to serve as training data for the validation of the model in the sample, and as test data for the out-of-sample validation.

7. Conclusions

In this paper, we have used machine learning methods to exploit a large text corpus of Article IV consultation reports published by the IMF between the years 2004 and 2020 to quantify patterns in global economic uncertainty and optimism. Constructing simple frequency measures, we find that both global uncertainty and optimism have overall increased over the last 15 years, and our benchmark uncertainty index tends to mimic movements in the World Uncertainty Index. Moreover, our results lend support to the existence of uncertainty spillovers from systemic countries on the rest of the world, while we find little evidence for optimism spillovers from these countries. We extend upon our frequency index using a simple rule-based sentiment analysis and find a similar although lagged pattern between our sentiment index and our benchmark optimism index. Finally, we note a variety of improvements for future work, emphasizing more general methodical refinements and a higher index frequency.

References

Ahir, H., Bloom, N. & Furceri, D. (2018). The World Uncertainty Index. *Mimeo*, 1-33.

Alexopoulos, M. & Cohen, J. (2009). Uncertain times, uncertain measures. *University of Toronto Working Paper*,

26 Some reports separately include individual *chapters*, while the majority of reports covers information on the economic outlook within *paragraphs*.

No. 352, 1-34.

Arellano, C., Bai, Y. & Kehoe, P. (2010). Financial Markets and Fluctuations in Uncertainty. *Federal Reserve Bank of Minneapolis Research Department Staff Report*, 1-41.

Baker, S. R., Bloom, N. & Davis, S. J. (2016). Measuring Economic Policy Uncertainty. *Mimeo*, 1-57.

Basu, S. & Bundick, B. (2017). Uncertainty shocks in a model of effective demand. *NBER Working Paper*, No. 18420, 1-28.

Bhandari, A., Borovicka, J. & Ho, P. (2019). Survey Data and Subjective Beliefs in Business Cycle Models. *Federal Reserve Bank of Richmond Working Paper*, No. 19-14, 1-60.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica*, 77(3), 623-685.

Bloom, N., Floetotto, M., Jaimovich, N., Saporta-Eksten, I. & Terry, S. J. (2012). Really uncertain business cycles. *NBER Working Paper*, No. 18245, 1-39.

Edwards, M. S., Coolidge, K. A. & Preston, D. A. (2011). Who reveals? Transparency and the IMF's Article IV Consultations, *Mimeo*, 1-44.

Giles, C. (2021). Joe Biden's huge bet: the economic consequences of 'acting big'. https://www.ft.com/content/49ca176d-8fa4-45a9-8c77-c837d1ad8e39. Accessed March 13, 2021.

Gentzkow, M., Kelly, B. & Taddy, M. (2019). Text as Data. *Journal of Economic Literature*, 57(3), 535-574.

IMF (2020). Transparency at the IMF. https://www.imf.org/en/About/Factsheets/Sheets/2016/07/27/15/35/Transparency-at-the-IMF. Accessed March 12, 2021.

IMF(2021). What the Continued Global Uncertainty Means for You. https://blogs.imf.org/2021/01/19/what-the-continued-global-uncertainty-means-for-you/. Accessed March 14, 2021.

Jurado, K., Ludvigson, S. C. & Ng, S. (2015). Measuring Uncertainty. *American Economic Review*, 105(3), 1177-1216.

Larsen, V. H. (2017). Components of Uncertainty. *Norges Bank Working Paper*, No. 5/2017, 1-39.

Loughran, T. & McDonald, B. (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *Journal of Finance*, 66, 35-65.

Mihalyi, D. & Mate, A. (2018). Text-mining IMF country reports - an original dataset. *Mimeo*, 1-24.

Picault, M. & Renault, T. (2017). Words are not all created equal: A new measure of ECB communication. *Journal of International Money and Finance*, 79(C), 136-156.

Slovik, P. (2010). Market uncertainty and market instability. *IFC Bulletin*, No. 34, 430-435.

Whaley, R. E. (2009). Understanding the VIX. *The Journal of Portfolio Management*, 35(3), 98-105.

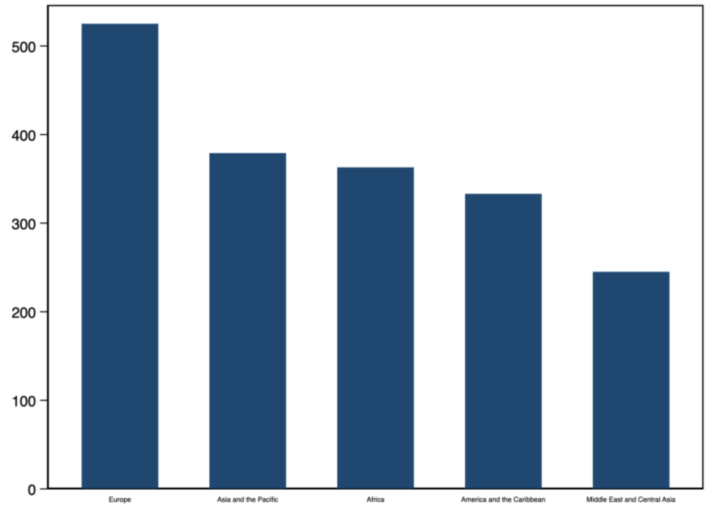
Appendix

This appendix contains complementary results referenced in the main text.

Section 3.3 - Number of reports by geographic region

Figure 20 shows the number of Article IV reports in our sample by geographic region. *Europe* contributes most to the number of contracts, followed by *Asia and the Pacific, Africa*, *America and the Caribbean*, and the *Middle East and Central Asia*.

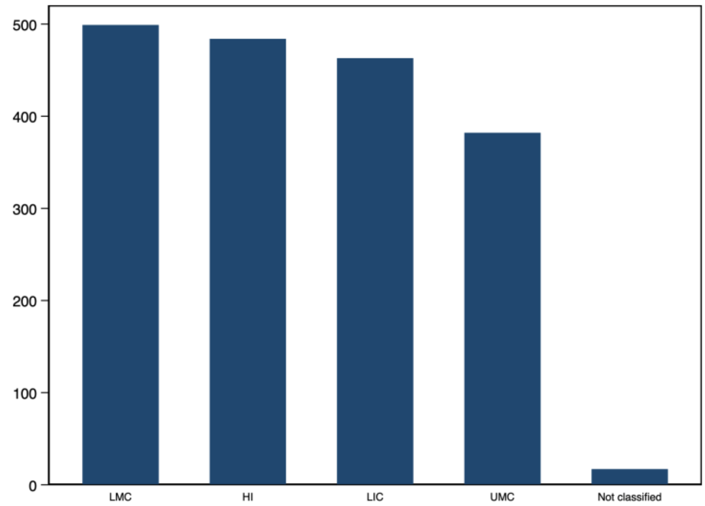
Figure 20: *Number of Article IV reports by geographic region, 2004-2020*.



Section 3.3 - Number of reports by income group

Figure 21 shows the number of Article IV reports by income group.[27] Lower middle income countries (LMC) contribute most to the number of reports, closely followed by high income countries (HI) and low income countries (LIC), and then upper middle income countries (UMC).

Figure 21: *Number of Article IV reports by income group, 2004-2020*.

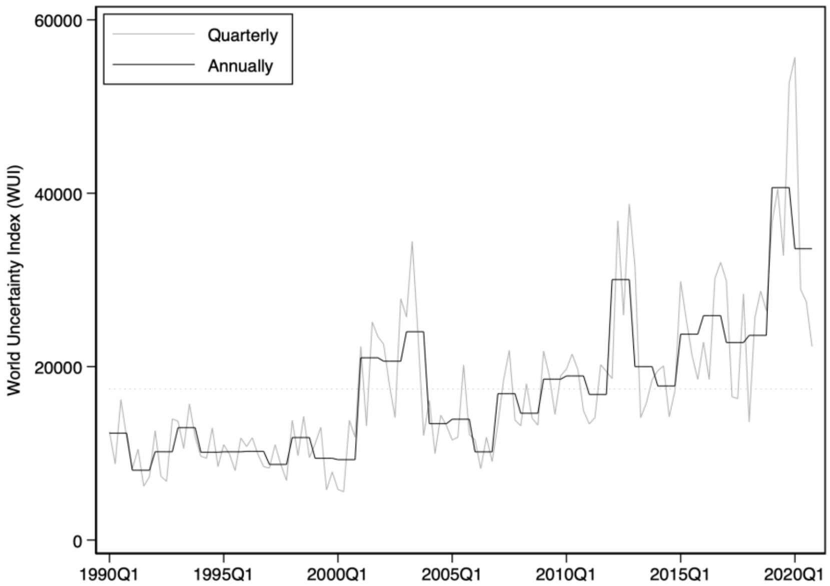


27 The classification follows the World Bank's 2006 World Development Report (http://documents1.worldbank.org/curated/en/435331468127174418/pdf/322040World0Development0Report02006.pdf).

Section 4.1 - World Uncertainty Index over time

Figure 22 shows the path of the WUI on quarterly and annual frequency between 1990Q1 and 2020Q4. The quarterly series has been converted to annual frequency by taking the unweighted mean of the index values in a given year.

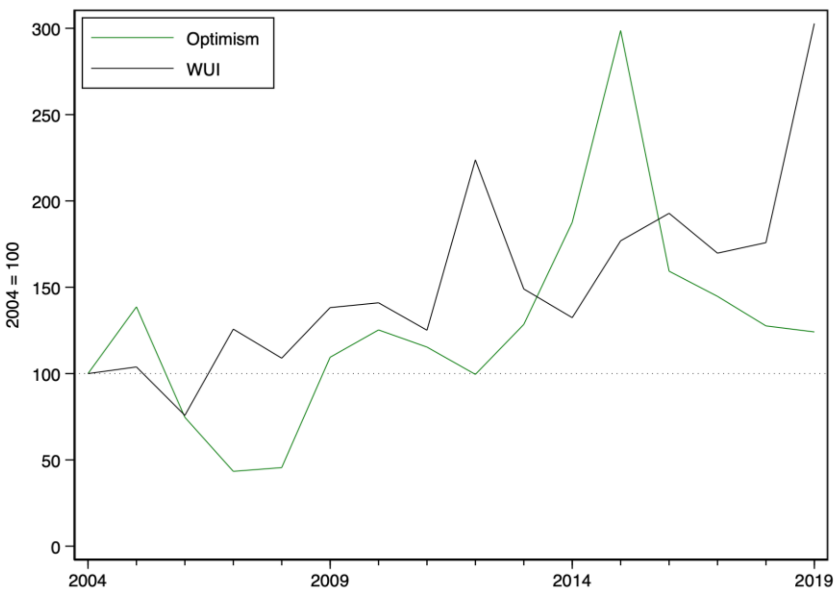
Figure 22: *World Uncertainty Index on quarterly and annual frequency, 1990-2020*.
Note: The quarterly series has been converted to annual frequency by taking the mean of the quarterly index values in a given year. The horizontal line indicates the quarterly sample average of the index.



Section 4.2 - Benchmark optimism index vs. World Uncertainty Index over time

Figure 23 plots our benchmark optimism index against the WUI for the period 2004 to 2019. By contrast to the relationship between our optimism and uncertainty indices, the correlation between our benchmark optimism index and the WUI is insignificant.

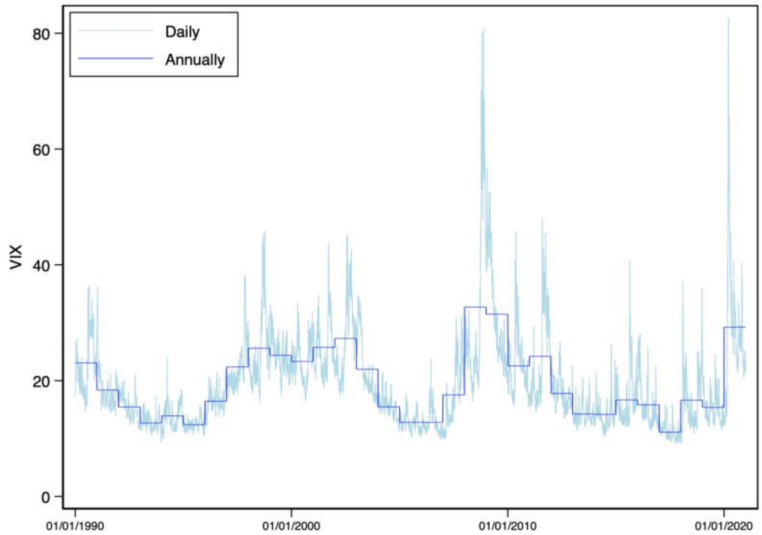
Figure 23: *Benchmark optimism index vs. World Uncertainty Index, 2004-2019*.
Note: The quarterly WUI series has been converted to annual frequency by taking the mean of the quarterly index values in a given year.



Section 5 - Conversion of VIX series to annual frequency

Figure 24 shows the level of the CBOE volatility index (VIX) for the period from January 1, 1990 to December 31, 2020, on daily and annual frequency, respectively. The daily series has been converted to annual frequency by taking the unweighted mean of the index values in a given year. The figure shows that this reduction takes out substantial variation in the index, particularly during (temporary) peaks of market volatility.

Figure 24: *CBOE volatility index on daily and annual frequency, 1990-2020*.
Note: The daily series has been converted to annual frequency by taking the mean of the daily index values in a given year.



Section 5 - World Uncertainty Index vs. CBOE volatility index

Figure 25 plots the World Uncertainty Index together with the VIX for the period from 1990Q1 to 2020Q4. By contrast to our uncertainty index, we find the WUI to be marginally positive and significantly correlated with the VIX on the 5% level, with a correlation coefficient of 0.15.

Figure 25: *World Uncertainty Index vs. CBOE volatility index, 1990Q1-2020Q4*.
Note: The VIX series has been converted to quarterly frequency by taking the mean of the daily index values in a given quarter.