MEASURING CYBER RISK

STÉPHANE LHUISSIER* AND FABIEN TRIPIER**

ABSTRACT. We present a Twitter-based indicator of cyber risk by scraping all tweets worldwide that contain both "cyber" and "risk" (including variants) since 2011. We document new facts on the rise of cyber risk and its composition by sectors on a real-time and daily basis. We show that the indicator is unrelated to other measures of uncertainty and risk. Using regression analysis, we test the 'hack effect': a positive impact of the rise of cybercrimes on the performance of cybersecurity companies. We show that the effect is very much in evidence, and has become even stronger since the COVID-19 pandemic.

Date: August 4, 2021.

Key words and phrases. Cyber Risk; Textual Analysis; Twitter; Hack Effect; COVID-19.
JEL Classification. D80, E58, G12.

*Banque de France, 31, Rue Croix des Petits Champs, DGSEI-DEMFI-POMONE 41-1422, 75049 Paris Cedex 01, FRANCE (Email: stephane.lhuissier@hotmail.com; URL: http://www.stephanelhuissier.eu). The views expressed in this paper are those of the authors and should under no circumstances be interpreted as reflecting those of the Banque de France or the Eurosystem. **Université Paris-Saclay, Univ Evry, EPEE, 91025, Evry-Courcouronnes, France & CEPII. (Email: fabien.tripier@univ-evry.fr).

The indicators introduced in this paper are available at http://www.stephanelhuissier.eu/assets/cyber.xlsx and will be periodically updated by the authors.

"*The chances that we would have a breakdown that looked anything like [in 2008] that where you had banks making terrible loans and investment decisions and needing and having low levels of liquidity and weak capital positions, and thus needing a government bailout, the chances of that are very, very low. Very low. But the world changes. The world evolves. And the risks change as well. And I would say that the risk that we keep our eyes on the most now is cyber risk. That's really where the risk I would say is now, rather than something that looked like the global financial crisis. (...) There are cyber-attacks every day on all major institutions now. That's a big part of the threat picture in today's world.*"

—Jerome Powell, Federal Reserve Chairman, CBS News on April 12, 2021.

I. Introduction

Concerns about cyber risk have intensified as the economy and financial system have become more digitalised. For example, in the Bank of England 2019 systemic risk survey of market participants, 61 percent of respondents expressed worries about the impact on the financial system of cyber attacks if they were to materialise, ranking cyber risk ahead geopolitical risk and the risk of a global economic downturn.[1] More recently, the European Systemic Risk Board has identified cyber risk as a source of systemic risk to the financial system that could have significant negative economic consequences (ESRB, 2020). Cyber risk considerations have been progressively integrated into the governance and risk management frameworks of financial and monetary authorities (Kashyap and Wetherilt, 2019).[2] However, an indicator of cyber risk that is consistent over time, and that measures in real time, high-frequency cyber risk as perceived by the public, risk managers, and policy makers is still missing in the literature. This paper attempts to fill this gap.

Using Twitter, we construct daily, weekly, and monthly indexes of cyber risk and examine their evolution since 2011. Our index reflects the frequency of tweets that contain the

[1]Cyber risk has been also regularly identified as the number one risk (or at least one of the top five) for the global financial system by the DTCC systemic risk barometer since the first edition in 2013. See https://www.dtcc.com/systemic-risk/systemic-risk-barometer-surveys.

[2]For example, the G7 Finance Ministers and Central Bank Governors has created its Cyber Expert Group in 2015, the European Central Bank has established the Euro Cyber Resilience Board for pan-European Financial Infrastructures in 2018, and the Basel Committee on Banking Supervision published in 2018 its report on "Cyber-resilience: range of practices".

following duo of terms: "cyber"; and one or more of "risk," "attack," or "threat". Our Twitter-based cyber risk index is characterized by several spikes corresponding to key cyber risk events such as DDoS attackon Spamhaus, WannaCry ransomware attack, and U.S. federal government data breach. Beyond these spikes, our indicator shows an upward trend in cyber risk, which rebounded with the COVID-19 pandemic outbreak after a downturn during the years 2018-2019. We extend our twitter-based approach to measuring cyber risk according to journalists' perception and sensitive sectors from cyber risks.

To capture the perception of journalists about cyber risk, we rely exclusively on Twitter accounts held by newspapers. The resulting media-based index of cyber risk is remarkably well correlated with our benchmark index, suggesting that Twitter users and journalists have similar perceptions about the evolution of cyber risk. Nevertheless, the media-based index suffers from an insufficient number of occurrences to allow a relevant monitoring procedure of cyber risk on a real-time and daily basis. This is precisely where social media such as twitter show their superior decisive advantage: their massive and uninterrupted traffic allows a daily monitoring of the cyber risk.[3]

We develop indicators of cyber risk at the level of sensitive sectors from cyber attacks by specifying more restrictive criteria for tweets that contain terms about cyber and risk. These indexes are particularly helpful to understand the different types of cyber risk. For example, we develop indexes for central banks and governments exposed to cyber risk based on the presence of additional terms like "Federal Reserve", "central bank", "ECB", or "Bank of England", and "government", "national security", "congress", or "white house", respectively. Sector-specific cyber risk events are very much in evidence.

We provide a period update of our indicators, which we make freely available at different frequencies (daily, weekly, and monthly).[4] We believe that our index as well as its sub-components provide useful real-time measures of cyber risk both for policymakers and researchers, and will foster more research on cyber risk and its effects on the economy.

As an illustration, we turn our analysis to the economic effects of cyber risk as measured by our twitter-based indicator. To analyze the economic impact of cyber risk for both the

[3]According to Twitter, there were about 192 million of daily active users in the fourth quarter of 2020, around 500 million tweets are sent each day (e.g., Mention, 2018), and 20% of U.S. internet users access Twitter on a monthly basis (e.g., eMarketer, 2018).

[4]Our indicators are available at via the following link: http://www.stephanelhuissier.eu/assets/cyber.xlsx.

private sector and public entities, for which stock prices do not exist, we investigate the supply side of cyber risk: the stock market return for companies of the cybersecurity industry. The market return for this industry has outperformed the overall S&P 500 stock index since 2011 with an excess return of 9% per year in average. Nasdaq MarketInsite (2019) explains this performance by the 'hack effect': a positive impact of the rise of cybercrimes on the performance of cybersecurity companies. Our regression analysis shows a robust and significant link between cyber risk and the excess return of the cybersecurity industry. In particular, days of extremely high cyber risk are associated with an excess return of 0.247%, five times the average value for all the period. This demonstrates that our cyber risk index captures well the economic impact of cyber risk and highlights the gains of cyber risk for cybersecurity companies. Interestingly, we also find greater sensitivity to cyber risk since the recent period marked by the COVID-19 pandemic outbreak.

The rest of the paper is organized as follows: Section II explains our contributions to the literature; Section III describes the definition of cyber risk that we adopt in our paper; Section IV discusses how we construct the index and describes its key features; Section V presents indexes for specific vulnerable sectors from cyber threats; Section VI shows the reaction of the stock market to cyber risk for the cybersecurity industry. Finally, VII provides concluding remarks.

II. Related literature

Our paper contributes to several strands of literature. First, it contributes to the literature on the economics of the cyber risk by providing a new measure of cyber risk. Empirical contributions to this field usually exploited historical databases of effective cyber attacks which can be related to the specific attacked firms as in the seminal investigations of Campbell, Gordon, Loeb, and Zhou (2003) and Cavusoglu, Mishra, and Raghunathan (2004), and in the most recent contributions of Corbet and Gurdgiev (2019), Kamiya, Kang, Kim, Milidonis, and Stulz (2021), and Tosun (2021). Bouveret (2018, 2019) presents evidence of cyber attacks on financial institutions around the world and develop a quantitative framework to assess cyber risk for the financial sector. Jamilov, Rey, and Tahoun (2021) have recently expanded the literature by incorporating additional measures of cyber risk based on textual analysis of quarterly earnings calls of publicly listed firms. Textual analysis allows for the assessment

of perceived cyber attack threats to companies, not just cyber attacks experienced by companies. Taken together, these methods provide very rich databases that allow for a granular analysis of cyber risks for companies, taking into account their microeconomic characteristics such as their geographic location or their business sector. Our twitter-based indicator clearly does not have an advantage over these databases in providing accurate information at the micro level.

Nevertheless, it has several specific advantages over these indicators. First, our twitter-based indicator is a high-frequency indicator and allows for daily monitoring of cyber risk in a way that financial reports published on a quarterly basis does not. Second, our twitter-based indicator is a real-time indicator that allows for the tracking of risks as they occur, unlike event databases that are by definition constructed ex-post. Third, our twitter-based indicator covers the entire economy (households, small and large businesses, and public entities such as governments and central banks) and not just large publicly traded companies.[5]

Second, our paper contributes to the literature on the economics of the cyber risk by studying its effects on the supplier of cybersecurity services. The vast majority of the literature on cyber risk investigates the costs on financial markets of cyber attacks on the attacked firms and the potential transmission to peer firms trough contagion effects which may give rise to systematic risk, see Corbet and Gurdgiev (2019); Kamiya, Kang, Kim, Milidonis, and Stulz (2021); Jamilov, Rey, and Tahoun (2021); Tosun (2021). On the contrary, we focus on the potential benefits for the cybersecurity industry.[6] Besides the direct interest of studying one of the most profitable industries at the present time, it allows to gauge the economic impact of cyber risk more generally, not only for publicly listed companies, but also for public entities.

Third, our paper contributes to the literature on the use of social media data in economics. In particular, we rely on the methodology developed by Baker, Bloom, Davis, and Renault (2021) that build a Twitter-based index of economic uncertainty. They review the pro and

[5]It may worth to mention that in the field of cyber risk, as observed by Caldara and Iacoviello (2018) in the field of geopolitical risk, many business companies publish various indicators of cyber risk which suffer from several shortcomings (lack of transparent methodology, wide-ranging definition of cyber risk, hard to replicate indexes, available only for a few years) that make them poorly suited for empirical analysis.

[6]Cavusoglu, Mishra, and Raghunathan (2004) is one of the few studies investigating this supply side of the market for cybersecurity

cons of twitter data for economic analysis and show that the twitter-based index of economic policy uncertainty is remarkably similar to the newspaper-based index developed in Baker, Bloom, and Davis (2016).[7] Among the advantages of social media data, Baker, Bloom, Davis, and Renault (2021) point out that "*tweets come with a precise timestamp and cannot be revised or edited. This last feature is especially useful in constructing high-frequency indicators and in relating tweet-based measures to near contemporaneous developments and financial market responses.*" This is of great important for twitter-based measured of cyber risk which may be tracked by public agencies and regulators.

Fourth, our paper contributes to the literature on the COVID-19 crisis. The changes in the organization of work and production implemented to respond to the development of COVID-19 have had major consequences in terms of cyber security as explained by Lallie, Shepherd, Nurse, Erola, Epiphaniou, Maple, and Bellekens (2021), all the more important as these changes are expected to be long term according to Barrero, Bloom, and Davis (2021). Our twitter-based cyber risk indicator and our analysis of the financial markets confirm the increase in cyber risk during the COVID-19 crisis but also that the economic impact of this risk has been multiplied by the pandemic outbreak.

Finally, our paper is related to the literature on fluctuations in uncertainty. Financial time-series data were first used to isolate changes in uncertainty and its effects on aggregate activity. Notable examples include Bloom (2009), Leduc and Liu (2016), Basu and Bundick (2017), and Lhuissier and Tripier (forthcoming). Then, this literature opened up to other types of data sources, such as survey data or textual analysis, to enrich the measurement of uncertainty and to characterize its different facets like economic policy uncertainty (e.g., Baker, Bloom, and Davis, 2016), macroeconomic forecasting errors (e.g., Jurado, Ludvigson, and Ng, 2015) or goepolitical risks (e.g., Caldara and Iacoviello, 2018) – see Ferrara, Lhuissier, and Tripier (2018) for a literature review. Our article supplements this literature by adding to it the variations over time of cyber-related risk, another facet of uncertainty that threatens the economy. We show in particular that our indicator is unrelated to other measures of uncertainty and risk widely used in the literature.

[7] Altig, Baker, Barrero, Bloom, Bunn, Chen, Davis, Leather, Meyer, Mihaylov, Mizen, Parker, Renault, Smietanka, and Thwaites (2020) show the behavior of the twitter-based index of economic uncertainty before and during the COVID-19 crisis.

III. Definition: Cyber Risk, Cyber Threat, and Cyber Attack

According to the Financial Stability Board (FSB) Cyber Lexicon[8], cyber risk refers to "the combination of the probability of cyber incidents occurring and their impact". A "cyber incident", in turn, is "any observable occurrence in an information system that: (i) jeopardises the cyber security of an information system or the information the system processes, stores or transmits; or (ii) violates the security policies, security procedures or acceptable use policies, whether resulting from malicious activity or not".

In defining cyber risk, we want to emphasize both the risk that attacks materialize, and the new risks associated with an escalation of existing attack. The former is more complex. Recognising the risks from cyber incidents, authorities, financial firms, and businesses could be encouraged to take real actions even when threats never materialize.

Cyber crimes are committed using a large range of methods. Though interconnected, most common attack types are malware, phishing attacks, and Denial-of-service to name just a few. Malware is defined as "software designed with malicious intent containing features or capabilities that can potentially cause harm directly or indirectly to entities or their information systems" according to FSB. The concept of malware itself has many names and variants including, for example, ransomware, adware, and spyware. Ransomware is designed to lock down the system and deny access to data until a ransom is paid. Adware is software that produces malicious advertisements such as pop-up windows on web pages. Spyware is software that aims to gather information about a person or organization, and transmit such information without permission. Phishing attacks are social engineering attacks where the attacker sends a fraudulent message designed to trick a victim into revealing confidential information to the attacker. Many forms of phishing exists like Email, Voice, and SMS phishing. Denial-of-service aims to prevent a computer system from functioning properly by overwhelming the networks and servers with traffic.

IV. Measurement

In January 2021, Twitter launched the Academic Research product track on the new Twitter API, allowing researchers to have free access to the full history of public conversation via the full-archive search endpoint. We make extensive use of this new Twitter API to

[8] https://www.fsb.org/wp-content/uploads/P121118-1.pdf

construct our index.[9] More specifically, we extract all tweets from January 2011 containing the following duo of terms: "cyber"; and "risk" or "attack" or "threat". For each tweet, we extract the tweet content, the date of the tweet, and the unique identifier of the tweet. We end up with a database of 2,147,126 original tweets (without retweets) from January 1st, 2011 to March 31st, 2021.

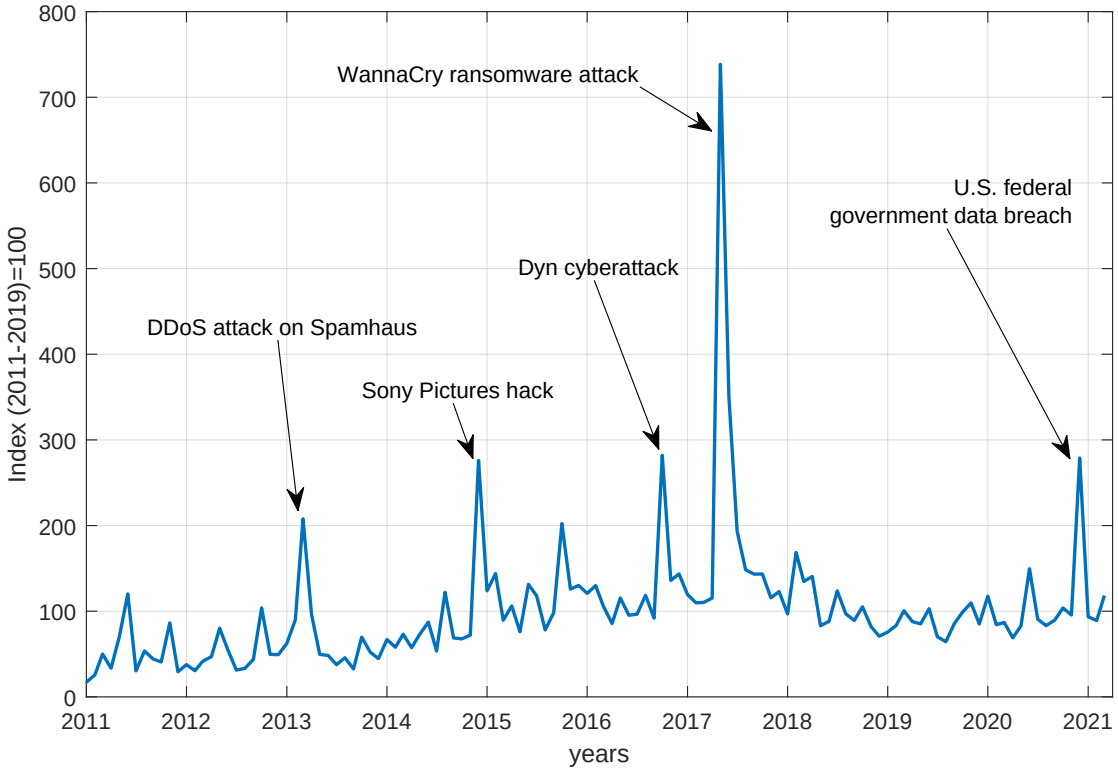
The index reflects, in each day, the number of tweets discussing rising cyber risks. We compute monthly tweet frequency and normalize the index to average a value of 100 in the 2011-2019 period, which we plot in Figure 1. By construction, the index captures cyber risks as perceived by English-speaking Twitter users, including English language Twitter users outside the United States like Canada, U.K. and Australia. Our definition encompasses events that has occurred worldwide.

Our Twitter-based cyber risk index shows several spikes corresponding to key cyber risk events. The first spike is recorded in March 2013 and corresponds to the distributed denial-of-service (DDoS) attack of unprecedented scale that targeted the Spamhaus Project, an international spam-fighting organization. The second spike happens in December 2014 and is caused by a malicious hacker group that leaked a release of confidential data from the film studio Sony Pictures. The index spikes again at the late of October 2016, during a series of distributed DDoS attacks, targeting systems operated by Domain Name System (DNS) provider Dyn. The index reaches its maximum during the Wannacry ransomware attack in May 2017, a worldwide cyber attack that targeted computers running the Microsoft Windows operating system and demanding a ransom in Bitcoin to decrypt the encrypted files. The WannaCry attack infected more than 250,000 computer systems in 150 countries. The index also rises in December 2020, during a major cyber attack that targeted multiple parts of the United States federal government, leading to a series of data breaches.

In addition to our benchmark index, we produce a cyber risk index using exclusively Twitter accounts held by newspapers. The index captures cyber risk as perceived and chronicled by the press in English-speaking countries, particularly in the United States. Indeed, to construct the index, we use fifteen U.S., and two British newspapers. More specifically, the index relies on the following leading newspapers: *New York Times*, *Los Angeles Times*, *USA*

[9]We employ the Python wrapper `searchtweets-v2` available at https://pypi.org/project/searchtweets-v2/ to extract tweets.

FIGURE 1. Twitter-based Cyber Risk Index



Note: The line plots the monthly cyber risk index from January 2011 to March 2021. The index is normalized to average a value of 100 in the 2011-2019 period.

Today, *Chicago Tribune*, *Washington Post*, *Boston Globe*, *Wall Street Journal*, *Miami Herald*, *Dallas Morning News*, *Houston Chronicle*, *San Francisco Chronicle*, *The Times*, *The Guardian*, *Financial Times*, *New York Post*, *Newsday*, and *Star Tribune*.[10] Taking monthly averages of our daily index, it correlates at 0.92 with our all-users monthly index, indicating a high degree of similarity. This suggests that Twitter users and journalists have similar perceptions about the evolution of cyber risk.

Another way to analyze our cyber risk index is by comparison with other measures of uncertainty or cyber risk, as illustrated in Figure 2. The first column compares our index to the VIX (an index of 30-day option-implied volatility in the S&P500 index), the Economic

[10]More specifically, we use the following twitter accounts: @nytimes, @latimes, @USATODAY, @chicagotribune, @washingtonpost, @BostonGlobe, @WSJ, @MiamiHerald, @dallasnews, @HoustonChron, @sfchronicle, @thetimes, @guardian, @FT, @nypost, @newsday, @StarTribune.

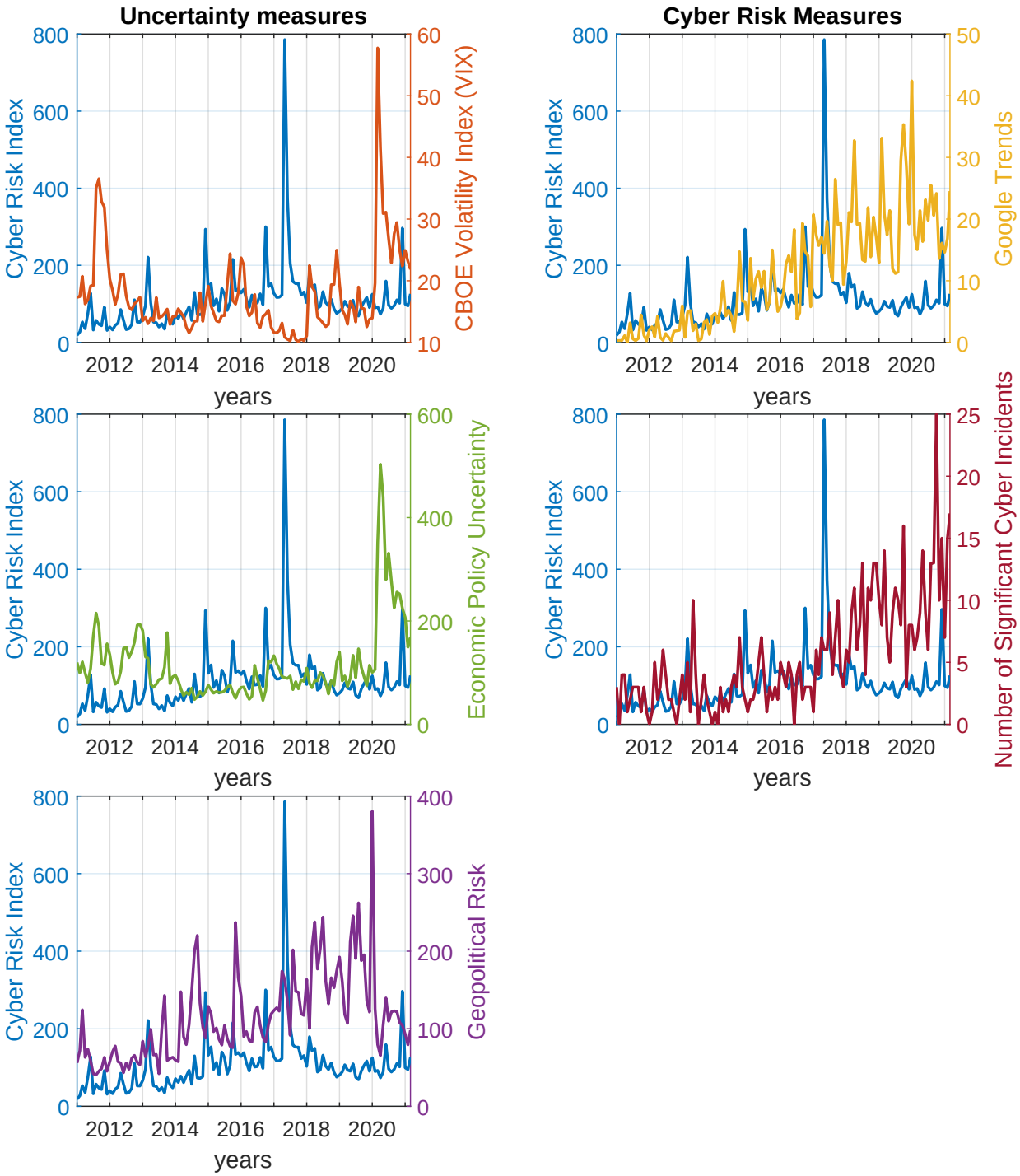
Policy Uncertainty (EPU) index of Baker, Bloom, and Davis (2016), and the geopolitical risk (GPR) index constructed by Caldara and Iacoviello (2018). The VIX and EPU indexes feature a large amount of independent variation and do not share common spikes with respect to the cyber risk index. The correlation of our index is -0.17 and -0.10 with the VIX and EPU indexes, respectively. Clearly, the cyber risk index captures events that are more likely to be exogenous to economic policy and financial uncertainty. The GPR index and our cyber risk index often do not move together, but they share a common peak; the GPR index spikes in 2017 during the Wannacry cyber attack, a period during which U.S. blamed North Korea to be behind this attack. This suggests that the correlation may run sometimes from cyber-attack events to geopolitical events. Overall, the correlation between the two measures is 0.21, a modest but non-negligible value as the GPR index appears to capture partly risk related to cyber attacks.

Several studies have constructed quantitative proxies of cyber incidents. For example, Aldasoro, Gambacorta, Giudici, and Leach (2020) construct a proxy for time-varying cyber risk by making use of Google Trends data. The authors count the daily number of online searches for "cyber risk" over the last decade. We replicate the proxy, extend it until March 2021, and normalize the index such that the worldwide search interest is relative to the highest point (=100). The second column of Figure 2 presents the series at a monthly frequency alongside our cyber risk index. The proxy appears to be relatively independent of overall movements in our cyber risk index.

We also consider a cyber risk indicator based on the center for strategic and international studies (CSIS) database. CSIS provides a reliable monthly timeline of significant cyber incidents that occurred since 2006. CSIS narrows down their definition of "significant attacks" by focusing exclusively on "cyber attacks on government agencies, defense and high tech companies, or economic crimes with losses of more than a million dollars".[11] This database has been largely used in many applied sciences. The measure, which counts the number of significant cyber incidents per month, is plotted alongside the cyber index in the second column of Figure 2. Both indexes display some degree of comovement in several historical periods, such as the Wannacry ransomware attack, the DDos attack on Spamhaus, and the

[11]The detailed list of significant cyber incidents recorded by CSIS is available at https://www.csis.org/programs/strategic-technologies-program/significant-cyber-incidents.

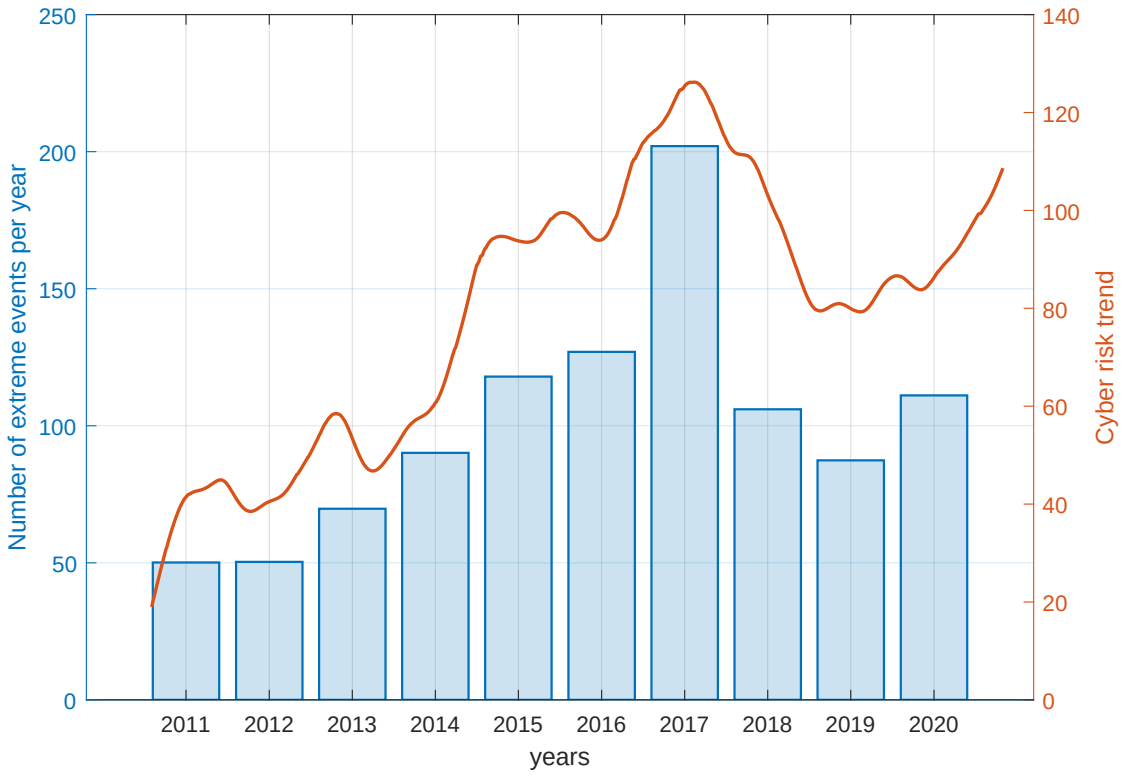
FIGURE 2. Comparison with Other Measures of Uncertainty and Cyber Risk



Note: The lines plot the monthly measures from January 2011 to March 2021.

U.S. federal government data breach. But they also show distinct variation, such as during the 2017-2019 period, when the CSIS measure rose sharply, while the cyber risk index declined in steady manner. However, the two alternative indicators of cyber risk differ conceptually in several respects, which may explain some differences. For example, the CSIS index counts cyber incidents, including those that might receive little attention to the public.

FIGURE 3. Distribution of extreme events per year and the trend of the cyber risk



Note: The bar displays the yearly number of cyber risk extreme events from 2011 to 2020 (left scale). The red line shows the daily trend of the cyber risk index from January 2011 to March 2021 (right scale).

Overall, our analysis shows that the proposed method provides a robust and reliable measure of cyber risk. However, It does not easily provide a clear view of the long-term pattern of cyber risk as well as its short-term fluctuations. To assess the fluctuations in cyber risk, Figure 3 shows the number of extreme events per year between 2011 and 2020, defined here as the highest 5% of the daily index values. From 2011 to 2017, the number of extreme events increases steadily, before dropping in 2018 and 2019 and then recovering in 2020.

This evolution of cyber risk is not only driven by extreme events. Figure 3 also shows the trend component of the cyber risk index excluding days with extreme events. The trend is estimated by applying the Hodrick-Prescott Filter (the smoothing parameter is equal to 10e6 to the daily cyber risk index after removing days with extreme events). Again we observe a gradual increase in cyber risk until 2017, then a slowdown, followed by a rebound in 2020. The slowdown of cyber risk in 2018-2019 as well as the rebound in 2020 (generally associated to the Covid-19 crisis and the general lockdown of economies) are consistent with other measures of cyber risk provided in the literature; see, for example, Figure 2, or Figure 4 in Jamilov, Rey, and Tahoun (2021), and Figure 1(b) in Tosun (2021).

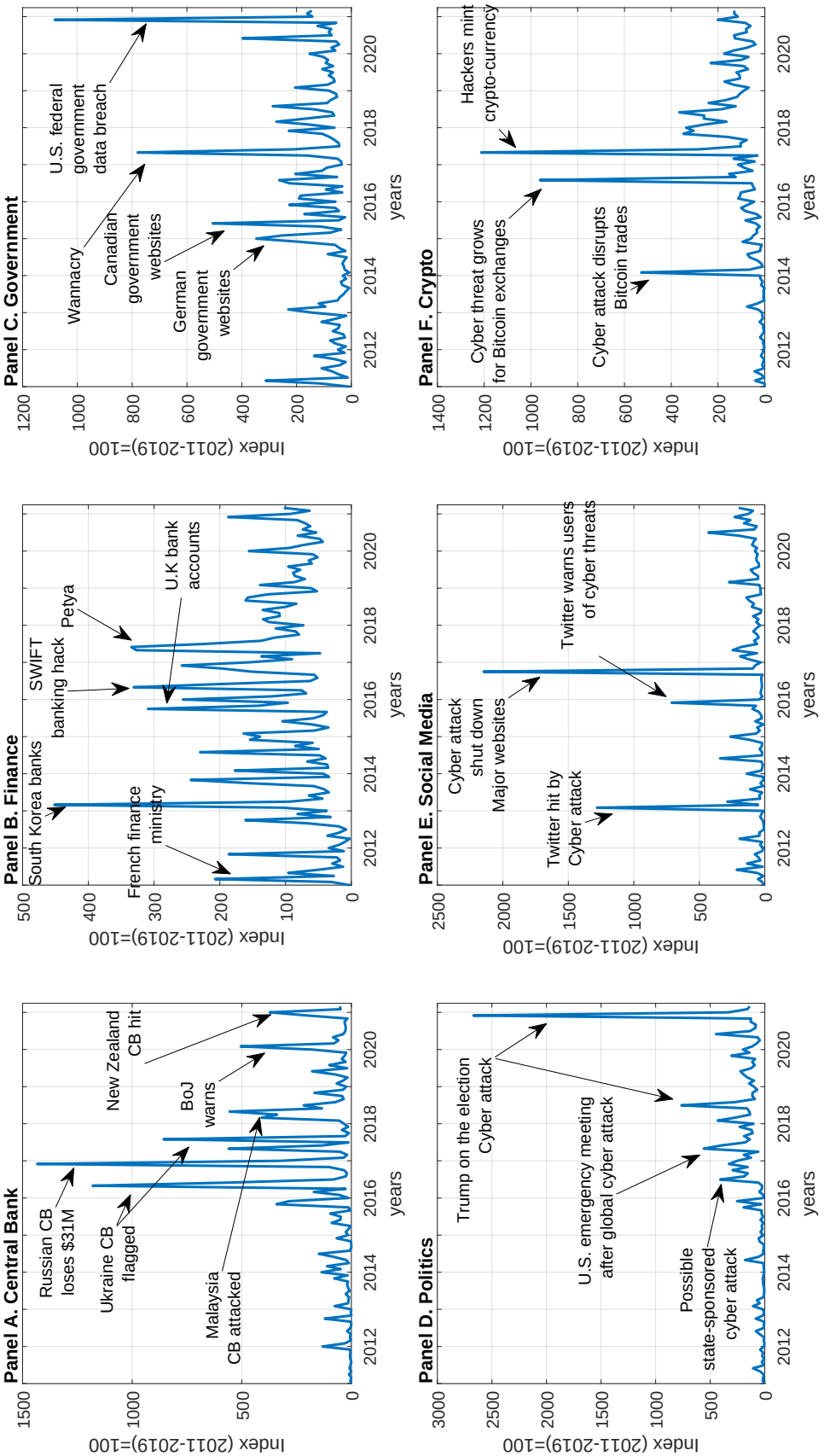
V. Cyber Risk Indexes for Sub-categories

To create indexes for sensitive sectors from cyber attacks, we apply additional criteria to those tweets that contain our duo of terms about cyber and risk. We consider six sectors: central bank, finance, government agencies, politics, social media, and cryptocurrency. Each of them are popular targets for cyber criminals since they hold a large trove of confidential and personal information that can naturally and potentially result in money takeovers.[12]

As seen in Figure 4, the central bank cyber risk index spiked sharply in connection with major events, and reached its highest value with the cyber attack on Russian central bank in 2016, where malicious hackers stole more than 2 billion rubles ($31 million) from correspondent accounts. However, spikes are not necessarily linked to the realization of cyber

[12]For each sector, the additional criteria involve the presence of one or more category-relevant terms: "Federal Reserve," "central bank," "European Central Bank," "Bank of England," "Bank of Japan," "Bank of China," "Bundesbank," "Bank of France," or "Bank of Italy" (including variants like "BoJ," "ECB," or "the Fed") for central bank category; "bank," "bond," "capital market," "credit crunch," "currency," "debt," "dividend," "equities," or "finance" for finance category; "government," "national security," "congress," "white house," or "authorities" for government category; "election", "state", "sponsor", "state sponsored", "state-sponsored", "espionage", "democratic national committee", "DNC", "Trump", "Clinton", "Assange" for social media category; "Zoom", "webex", "hangouts", "Facebook", "Google", "Twitter", "Bing", "Snapchat", "Linkedin", "MailChip", "Baidu", "Tencent", "Weibo", "Yandex", "Rambler", "Line", "whatsapp" for social media category; and "crypto", "crypto currency", "cryptocurrency", "ledger", "cryptography", "blockchain", "bitcoin", "altcoin", "token", "ethereum", "rupple", "litecoin", "tether", "libra", "monero", "diem" for crypto-currency category. For the categories social media, politics, and crypto, we use the terms employed in Jamilov, Rey, and Tahoun (2021).

FIGURE 4. Twitter-based Cyber Risk Index: Sub-indexes



Note: The line plots the monthly cyber risk index from January 2011 to March 2021. The index is normalized to average a value of 100 in the 2011-2019 period.

attacks. Cyber risk has become, from time to time, a topic of concern in central banks' speeches. For example, the central bank cyber risk index rocketed in January 2020 when Bank of Japan warned of cyber-attack vulnerability ahead of Olympic Games. Exploring different forms of communication (speech, statements, or reports) of central banks using for example textual analysis techniques to extract the information about cyber risk would seem to be very relevant for better understanding the interest of central banks to cyber risk.

The finance cyber risk index rose sharply during big events like the cyber attack of French finance ministry targeting G20 documents, SWIFT banking hack hitting member banks, Petya malware in 2017, and U.K bank account hack. The central bank and finance indexes present a high degree of comovement, with a correlation of about 0.42. More specifically, major spikes in the central bank index coincide with spikes in the finance index. However, there is still a high amount of independent variation between both indexes. For example, the finance cyber risk rose sharply in the months of major events like the attack on the French finance ministry, while the central bank cyber risk index remains unchanged during the same period. The government cyber risk index reacts to attacks targeting German and Canadian government websites, Wannacry in 2017, and the U.S. federal government data breach.

The political cyber risk index exhibits a sharp break in 2016, year marked by Trump election; the average value of post-2016 index remains essentially above that of the pre-2016 index, and is characterized by several peaks related to concerns about cyber attacks targeting U.S. elections. Cyber risk related to social media is also characterized by major peaks such as the shut down of major social media websites (Twitter, Facebook, etc...) in 2017 and the alert of cyber threats by Twitter in 2016. Finally, cyber risk concerns related to crypto-currency are important and appears to be higher since 2017 with the unprecedented boom in Bitcoin. The digital currency was marked by a 2000% increase in value from January 1, 2017, to December 16, 2017.

VI. Cybersecurity Industry Returns and Cyber Risk

We then turn our analysis to the role that cyber risks play for the pricing of financial assets. In the empirical literature (e.g., Corbet and Gurdgiev, 2019; Kamiya, Kang, Kim, Milidonis, and Stulz, 2021; Jamilov, Rey, and Tahoun, 2021; Tosun, 2021), the stock market reaction is used to measure the cost of cyber attacks for victim firms and to assess the potential

contagion to peer firms which may give rise to a systemic cyber risk. This methodology is however limited for cyber attacks that concern publicly listed companies and cannot be applied to cyber attacks on other sectors as public entities for which stock market data do not exist.

To circumvent this difficulty, we investigate the supply side of cyber security services to gauge the economic impact of cyber risk. The companies of the cyber security industry, without being the originators of the cyber attacks, could benefit from them because of the increased demand for their services generated which may in turn induced a rise in the stock market valuation of these companies. In the lack of accurate data on which cyber security companies were involved in each specific attack, we use the sectoral stock index ISE Cyber Security UCITS Index (HUR) provided by the NASDAQ.[13] In doing so, we follow the usual approach in finance to explore the reactions of industrial stock indexes of the industries most affected by shocks; see Kilian and Park (2009) for industries impacted by oil shocks and Scholtens and Boersen (2011) by energy accidents.

The stock market index for cyber security industry has clearly outperformed the overall S&P 500 stock index since 2011. The return for cyber security industry is equal to 32% per year against 21% for the S&P 500 index, hence an excess return of 9% in average. Nasdaq MarketInsite (2019) explains this performance by the 'hack effect': the rise of cybercrimes has had a positive impact on the performance of cybersecurity companies. We assess herein the importance of this 'hack effect' by analyzing the stock market reaction to our twitter-based indicator of cyber risk following the traditional methodology developed to study stock markets' reaction to news (e.g., Bernanke and Kuttner, 2005) extended to social media data notably by Bollen, Mao, and Zeng (2011).

To gauge the economic impact of cyber risk, the benchmark regression is as follows:

$$s_t = c + \alpha \times \text{CyberRisk}_t + \epsilon_t \tag{1}$$

where t is the business day, s_t the stock market excess return for the cyber security industry at the daily frequency, c is the constant, CyberRisk_t is the cyber risk index, and ϵ_t denotes the residual. To take into account the closing time of the markets, the cyber risk twitter

[13] "The ISE Cyber Security UCITS Index is designed to track public companies that are actively involved in providing cyber security technology and services." Source: https://indexes.nasdaqomx.com/Index/Overview/HUR.

index at date t measures the traffic on the social network between date $(t-1)$ at 4pm and date t at 4pm. The coefficient of interest is α which measures the effect of cyber risk on the stock market excess return for the cyber security industry. Table 1 provides the estimation outcome.

The estimation in column (1) suggests that the impact of cyber risk on cybersecurity returns is not of great importance since the estimated coefficient is not significantly different from zero. However, this result conceals marked differences between the various components of cyber risk as shown by the columns (2)-(11).

In columns (2) and (3), we decompose the cyber risk index into two components. Columns (2) reports the estimation for the extreme events of cyber risk using a dummy which is equal to one if $\text{CyberRisk}_t > P_{95}(\text{CyberRisk})$, that is the threshold used to define extreme cyber risk events. Days of extreme events are associated with an excess return of 0.247%, significantly different from zero at the 1% level, ten times higher than the average excess return for the full period (0.024%). The impact of extreme events of cyber risk is then not only huge but also highly significant. This is not the case for the non-extreme values of cyber risk. Column (3) reports the estimation of equation (1) for days without extreme events. The estimated coefficient is no longer significant, even at the 10% level, as for the benchmark cyber risk index in column (1). This suggests that the impact of cyber risk on cybersecurity industry returns is mainly driven by the extreme variations of the index.

Columns (4)-(5) compare the recent period, marked by the COVID-19 crisis, with previous years. We introduce D_t a dummy variable which is equal to one before the 1st of January, 2020, and zero otherwise. Then, equation (1) is estimated with the following two explanatory variables: $D_t \times \text{CyberRisk}_t$ and $(1-D_t) \times \text{CyberRisk}_t$. Columns (4) reports the results for the cyber risk index. While the estimated coefficient for the cyber risk index is not significant in column (1) for the overall period, it is slightly significant (at the 10% level) since 2020 contrary to the 2011-2019 period. The value of 0.101 for the estimated coefficient indicates an important effect of cyber risk on the return for the cyber security industry. A typical value of 300 for the cyber risk index (as observed during the Sony Pictures hack and, more recently, the U.S. federal governement data breach) is associated to a daily excess return

TABLE 1. Cyber risk and excess return in the cybersecurity industry

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Cyber risk	0.021										
	(0.016)										
Days with extreme events		0.247***									
		(0.084)									
Cyber risk without extreme events			-0.016								
			(0.036)								
Cyber risk: 2020-2021				0.101*							
				(0.060)							
Cyber risk: 2011-2019				0.014							
				(0.012)							
Days with extreme events: 2020-2021					0.965*						
					(0.512)						
Days with extreme events: 2011-2019					0.186**						
					(0.077)						
Cyber risk: Finance						0.009					
						(0.005)					
Cyber risk: Government							0.014**				
							(0.007)				
Cyber risk: Central banks								0.004**			
								(0.002)			
Cyber risk: Politics									0.014***		
									(0.002)		
Cyber risk: Social media										0.004**	
										(0.001)	
Cyber risk: Crypto											0.000
											(0.004)
Constant	0.000	0.010	0.026	-0.005	0.010	0.014	0.007	0.020	0.005	0.021	0.024
	(0.027)	(0.020)	(0.041)	(0.025)	(0.020)	(0.020)	(0.021)	(0.020)	(0.020)	(0.020)	(0.020)
Observations	2028	2028	1912	2028	2028	2028	2028	2028	2028	2028	2028

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table reports the coefficients α introduced in the regression equation (1). The dependent variable is the excess return for the cyber security industry defined as $100 \times (\Delta \log \mathrm{HUR}_t - \Delta \log \mathrm{SP}_t)$ where HUR_t is the ISE Cyber Security UCITS Index and SP the S&P 500 Index. Source: Eikon Datastream. Cyber risk is our twitter-based measure of cyber risk. Extreme events are the 5% higher values for the benchmark index. All cyber risk indexes are divided by 100 in the regressions. Robust standard errors in parentheses.

for the cyber security industry of $(300/100) \times 0.101 = 0.303\%$.[14] This suggests a great magnitude of the cyber risk impact on the Cyber security industry. The daily average of the excess return for this industry is 0.024%, therefore the realization of a 300-value for the Cyber risk index multiplies by almost 12 of the average daily return in this industry in the recent period. The column (5) confirms this pattern using the extreme events of cyber risk. In both cases, the coefficients are positive and significantly different from zero (even if it is at the 5% and 10% levels, against the 1% level in the second column). However, they differ in terms of magnitude. The estimated coefficient for the recent period is equal to 0.965 five time higher than the value of 0.186 estimated for the previous period. Estimation results reported in columns (4)-(5) are consistent with the analysis of cyber security in the age of COVID-19 done by Lallie, Shepherd, Nurse, Erola, Epiphaniou, Maple, and Bellekens (2021). The lockdown and home working context not only made the attacks more likely but also increased their consequences given the particular organization of work at that time. Our estimate indicates that cyber attacks during the COVID-19 crisis generates more demand for financial services, and profits for their suppliers, than in the period before.

Columns (6)-(11) compare various cyber risk subindexes concerning the affected sectors by the cyber risk: finance, governments, central banks, politics, crypto, and social media. The impact turns out to be significant for governement, central banks, politics, and social media subindexes (at the 5% or 1% level), but not for the finance and crypto subindexes. However the magnitude shows sizeable differences between them. The effect impact much higher higher for the subindexes on governments and politics than for the other subindexes, highlighting the exposition of public entities to cyber risk. The interest of our methodology, based on stock market capitalization of cyber security companies, is to identify this exposition and supplements the evidence provided in the literature on the exposition of the publicly listed corporations.

VII. Conclusion

We construct in this paper a time-series index of cyber risk by tapping a large collection of tweets since 2011. The two key advantages of this indicator are to cover all economic actors

[14]The 300-value is divided by 100 because the cyber index risk index is divided by 100 in regressions to facilitate the reading of values reported in the the table.

exposed to cyber risk and to be available in real time for monitoring high frequency cyber risk. Narrative evidence and stock market reactions show the relevance of the fluctuations in cyber risk described by our indicator.

Our analysis is in line with the various warnings issued by regulators and specialists in the digital economy about the growing cyber risk. It is important to underline that the cyber risk, whose evolution and economic impact are measured herein, is still contained to limited events and that there has not yet been a cyber disaster — by analogy with the rare disasters *à la* Barro (2006) — leading to a complete shutdown of a country's economy (for example, due to a paralysis of transport or payment system). From this perspective, the recent theoretical works by Duffie and Younger (2019) and Eisenbach, Kovner, and Lee (2020) provide some insight into our understanding of the risk presented by cyber attacks to the U.S. financial system and their financial and economic consequences.

References

Aldasoro, I., L. Gambacorta, P. Giudici, and T. Leach (2020): "The drivers of cyber risk," BIS Working Papers 865, Bank for International Settlements.

Altig, D., S. Baker, J. M. Barrero, N. Bloom, P. Bunn, S. Chen, S. J. Davis, J. Leather, B. Meyer, E. Mihaylov, P. Mizen, N. Parker, T. Renault, P. Smietanka, and G. Thwaites (2020): "Economic uncertainty before and during the COVID-19 pandemic," *Journal of Public Economics*, 191, 104274.

Baker, S. R., N. Bloom, and S. J. Davis (2016): "Measuring Economic Policy Uncertainty," *The Quarterly Journal of Economics*, 131(4), 1593–1636.

Baker, S. R., N. Bloom, S. J. Davis, and T. Renault (2021): "Twitter-Derived Measures of Economic Uncertainty," .

Barrero, J. M., N. Bloom, and S. J. Davis (2021): "Why Working from Home Will Stick," Working Paper 28731, National Bureau of Economic Research.

Barro, R. J. (2006): "Rare Disasters and Asset Markets in the Twentieth Century," *The Quarterly Journal of Economics*, 121(3), 823–866.

Basu, S., and B. Bundick (2017): "Uncertainty Shocks in a Model of Effective Demand," *Econometrica*, 85, 937–958.

Bernanke, B. S., and K. N. Kuttner (2005): "What explains the stock market's reaction to Federal Reserve policy?," *The Journal of Finance*, 60(3), 1221–1257.

Bloom, N. (2009): "The Impact of Uncertainty Shocks," *Econometrica*, 77(3), 623–685.

Bollen, J., H. Mao, and X. Zeng (2011): "Twitter mood predicts the stock market," *Journal of Computational Science*, 2(1), 1–8.

Bouveret, A. (2018): "Cyber Risk for the Financial Sector: A Framework for Quantitative Assessment," IMF Working Papers 2018/143, International Monetary Fund.

——— (2019): "Estimation of losses due to cyber risk for financial institutions," *Journal of Operational Risk, Forthcoming*.

Caldara, D., and M. Iacoviello (2018): "Measuring Geopolitical Risk," International Finance Discussion Papers 1222, Board of Governors of the Federal Reserve System (U.S.).

Campbell, K., L. A. Gordon, M. P. Loeb, and L. Zhou (2003): "The economic cost of publicly announced information security breaches: empirical evidence from the stock market," *Journal of Computer Security*, 11(3), 431–448.

Cavusoglu, H., B. Mishra, and S. Raghunathan (2004): "The effect of internet security breach announcements on market value: Capital market reactions for breached firms and internet security developers," *International Journal of Electronic Commerce*, 9(1), 70–104.

Corbet, S., and C. Gurdgiev (2019): "What the hack: Systematic risk contagion from cyber events," *International Review of Financial Analysis*, 65, 101386.

Duffie, D., and J. Younger (2019): "Cyber runs," Hutchins Center Working Paper 51, Brookings.

Eisenbach, T. M., A. Kovner, and M. J. Lee (2020): "Cyber Risk and the U.S. Financial System: A Pre-Mortem Analysis," Staff Reports 909, Federal Reserve Bank of New York.

eMarketer (2018): "Snapchat and Twitter 2018," Discussion paper.

ESRB (2020): "Systemic Cyber Risk Report on Systemic Cyber Risk," Discussion paper.

Ferrara, L., S. Lhuissier, and F. Tripier (2018): "Uncertainty fluctuations: Measures, effects and macroeconomic policy challenges," in *International macroeconomics in the wake of the global financial crisis*, pp. 159–181. Springer.

Jamilov, R., H. Rey, and A. Tahoun (2021): "The Anatomy of Cyber Risk," Working Paper 28906, National Bureau of Economic Research.

Jurado, K., S. C. Ludvigson, and S. Ng (2015): "Measuring uncertainty," *American Economic Review*, 105(3), 1177–1216.

Kamiya, S., J.-K. Kang, J. Kim, A. Milidonis, and R. M. Stulz (2021): "Risk management, firm reputation, and the impact of successful cyberattacks on target firms," *Journal of Financial Economics*, 139(3), 719–749.

Kashyap, A. K., and A. Wetherilt (2019): "Some Principles for Regulating Cyber Risk," *AEA Papers and Proceedings*, 109, 482–487.

Kilian, L., and C. Park (2009): "The impact of oil price shocks on the US stock market," *International Economic Review*, 50(4), 1267–1287.

Lallie, H. S., L. A. Shepherd, J. R. Nurse, A. Erola, G. Epiphaniou, C. Maple, and X. Bellekens (2021): "Cyber security in the age of covid-19: A timeline and analysis of cyber-crime and cyber-attacks during the pandemic," *Computers & Security*, 105, 102248.

Leduc, S., and Z. Liu (2016): "Uncertainty shocks are aggregate demand shocks," *Journal of Monetary Economics*, 82(C), 20–35.

Lhuissier, S., and F. Tripier (forthcoming): "Regime-dependent Effects of Uncertainty Shocks: a Structural Interpretation," *Quantitative Economics*.

Mention (2018): "2018 Twitter Report," Discussion paper.

Nasdaq MarketInsite, N. (2019): "The Hack Effect: The Effect of Data Breaches on the Nasdaq CTA Cybersecurity Index," *Nasdaq articles*.

Scholtens, B., and A. Boersen (2011): "Stocks and energy shocks: The impact of energy accidents on stock market value," *Energy*, 36(3), 1698–1702.

Tosun, O. K. (2021): "Cyber-attacks and stock market activity," *International Review of Financial Analysis*, 76, 101795.