AMERICAN UNIVERSITY OF BEIRUT

TWITTER-BASED ECONOMIC POLICY UNCERTAINTY INDEX FOR LEBANON

by
MAJDY MOFID BARAKAT

A thesis
submitted in partial fulfillment of the requirements
for the degree of Master of Arts
to the Department of Economics
of the Faculty of Arts and Sciences
at the American University of Beirut

Beirut, Lebanon
April 2023

AMERICAN UNIVERSITY OF BEIRUT

TWITTER-BASED ECONOMIC POLICY UNCERTAINTY INDEX FOR LEBANON

by
MAJDY MOFID BARAKAT

Approved by:

Sumru Altug

Dr. Sumru Guler Altug, Professor | Advisor
Economics

Dr. Simon Neaime, Professor | Member of Committee
Economics

Ali Abboud

Dr. Ali Abboud, Assistant Professor | Member of Committee
Economics

Date of thesis defense: April 25, 2023

AMERICAN UNIVERSITY OF BEIRUT

THESIS RELEASE FORM

Student Name: Barakat Majdy Mofid
Last First Middle

I authorize the American University of Beirut, to: (a) reproduce hard or electronic copies of my thesis; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes:

☒ As of the date of submission

☐ One year from the date of submission of my thesis.

☐ Two years from the date of submission of my thesis.

☐ Three years from the date of submission of my thesis.

May 5, 2023

Signature Date

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to the members of my thesis committee, Dr. Sumru Guler Altug, Dr. Simon Neaime, and Dr. Ali Abboud, for their guidance throughout this research project. I am especially grateful to my advisor, Dr. Sumru Guler Altug, for her exceptional mentorship, unwavering support, and assistance in securing the necessary funds to ensure the success of this project.

I would also like to thank Anthony J. Saab for his contribution to the creation and development of the Python code for our TEPU index. His research assistance was crucial in making this project possible.

Last but not least, I wish to express my deepest gratitude to my family, my mother Roudaina, my father Mofid, and my brother Marwan, for their unconditional love and encouragement throughout my academic journey. Their motivation has been my guiding light, and this work would not have been possible without them.

ABSTRACT
OF THE THESIS OF

Majdy Mofid Barakat for Master of Arts
Major: Economics

Title: Twitter-based Economic Policy Uncertainty Index for Lebanon

This paper presents a novel approach to construct a weekly and monthly Twitter-based Economic Policy Uncertainty (TEPU) index for Lebanon, covering the period from January 1, 2011, till January 18, 2023. We have developed a unique and distinctive methodology that was specifically designed to overcome the challenges posed by the unavailability and lack of reliability of data in Lebanon. Our methodology enabled us to create a reliable and effective TEPU index for Lebanon. In this paper, we created and employed Python scripts that interact with the official Twitter API to fetch and transform tweets into actionable data. This curated data was the basis for generating a TEPU index for Lebanon.

Moreover, we employed two scaling methods to construct our TEPU index. The first was proposed by Baker et al., (2021), and the second was also proposed by Baker et al., (2021) as a variant of their main index and adopted by Lee et al., (2023) as their primary scaling method. We showed that despite some notable differences, the two TEPU indexes shared many common points, such as the general trend of economic policy uncertainty over time in Lebanon. We concluded that our TEPU index using the first scaling method was more reliable than the index using the second scaling method since the latter index tends to be inaccurate when outliers are present.

Finally, we conducted an event analysis to demonstrate how our TEPU index links to significant political and economic events that occurred in Lebanon throughout our sample period. We observed that our TEPU index significantly spikes during major political and economic events that occurred in Lebanon, effectively tracking the evolution of economic policy uncertainty.

Keywords
Twitter-based data, Economic Policy Uncertainty, TEPU index, Python scripts, Twitter API, TEPU scaling methods, event analysis

TABLE OF CONTENTS

ACKNOWLEDGEMENTS 1

ABSTRACT 2

INTRODUCTION 5

LITERATURE REVIEW 11

2.1 Navigating the Impact of Uncertainty on Economic Outcomes 11

2.2 Developing Empirical Measures for Uncertainty 12

THE TWITTER-BASED ECONOMIC POLICY UNCERTAINTY INDEX 15

3.1 Methodology 16

3.1.1 Initial Set 20
3.1.2 Intermediate Set 21
3.1.3 Final Set 23

3.2 Twitter Developer Platform 26

3.2.1 Twitter API 26
3.2.2 Twitter API with Python 30
3.2.3 Twitter API: Challenges and Limitations 31

3.3 Constructing the Weekly and Monthly TEPU Index 36

3.3.1 Scaling and Results 40

3.4 Event Analysis 46

CONCLUSION 50

APPENDIX 1 52

APPENDIX 2 57

REFERENCES 62

CHAPTER 1

INTRODUCTION

The financial crisis in Lebanon has been ongoing since 2019 and is considered one of the worst in Lebanon's history. This crisis was triggered by a combination of factors that took place in Lebanon, including high current account deficits (21.7 % of GDP in 2019), high debt-to-GDP ratios (155.1% of GDP in 2019), and a fixed exchange rate system (1505 LL in exchange to one US dollar since 1997).[1] The Lebanese Lira, which has been pegged to the US dollar for over two decades, began to experience severe pressure in late 2019, and the situation worsened in 2020 with the outbreak of the COVID-19 pandemic and the political instability in Lebanon. Given that Lebanon was and remains a politically and economically crisis-ridden country, we were interested in creating a Twitter-based Economic Policy Uncertainty (TEPU) index for Lebanon similar to the indexes developed by other countries such as the US, China, Turkey, Chile, etc.

This paper presents a novel approach to construct a weekly and monthly Twitter-based Economic Policy Uncertainty (TEPU) index for Lebanon, covering the period from January 1, 2011, till January 18, 2023, that aims to monitor and track developments in Lebanon's economic and political landscape. To the best of our knowledge, this paper represents the first attempt to construct an Economic Policy Uncertainty index for Lebanon, a country known for its highly volatile economic policy uncertainty environment.

[1] Statistics are provided on Lebanon by The International Monetary Fund (IMF), which can be accessed by following this link: https://www.imf.org/en/Countries/LBN

In this paper, we created and employed Python scripts that interact with the official Twitter API to fetch and transform tweets into actionable data. This curated data was the basis for generating a TEPU index for Lebanon. Our primary objective was to develop a TEPU index that best fits Lebanon and can effectively identify and uncover significant economic policy uncertainty events that have taken place in the country since 2011. However, collecting reliable Twitter data for Lebanon was a significant challenge. Thus, we developed a distinctive methodology to overcome issues related to the unavailability and lack of reliability of data in Lebanon. This methodology played a pivotal role in ensuring the success and accuracy of our index. During the development of our distinctive methodology to create a TEPU index for Lebanon, we encountered some sensitive issues that we addressed carefully to ensure the reliability of our index. These sensitive issues are as follows: First, Lebanon is a relatively small country with a limited number of well-known Twitter influencers, and Twitter is not the most commonly used social media platform in Lebanon. Second, the issue of plagiarism and automated accounts, or "bots," posed a challenge as it made Twitter data very noisy. Third, there are several places around the world called "Lebanon," which created ambiguity in identifying relevant tweets related to the country of Lebanon in the Middle East.

To generate our weekly and monthly Twitter-based economic policy uncertainty index for Lebanon, our methodology was heavily influenced by the work of Baker et al., (2016) and Baker et al., (2021), in which they counted the frequency of newspaper articles and tweets, respectively, that are related to economic policy uncertainty. In this paper, we searched for all global tweets from various locations worldwide containing at least one keyword from the following four categories: (1) "Lebanon," (2) "Economic," (3) "Policy," and (4) "Uncertainty." The keywords in categories 1 and 4 were straightforward.

However, constructing the keyword sets for categories 2 and 3 proved challenging, as we adopted a data-driven approach that involved manual annotation, similar to the "Let the data speak for themselves" approach. So, instead of randomly selecting words based on their frequency or importance, we opted for a manual annotation process. This allowed us to select existing keywords in tweets and ensure that we extracted the precise English terms commonly used by Lebanese people, Arabs, and people from other parts of the world when discussing specific economic policy events related to Lebanon. Although this approach was time-consuming, we were certain our manual annotation was necessary to ensure the accuracy of our keyword sets for the different categories, especially since we were dealing with a politically and economically crisis-ridden country.

Our Economic Policy Uncertainty index for Lebanon was constructed using exclusively Twitter-collected data. While we didn't use data from traditional media platforms such as newspapers, as Baker et al., (2016) did, we found several practical reasons why Twitter data can serve as a viable substitute. In recent years, Twitter has emerged as a direct, real-time, and network-enhanced platform for economic policy experts to communicate their opinions on economic and financial issues with a large social network. This enables a diverse range of viewpoints to be captured almost immediately, without the filtering and significant time lag that occurs in newspaper articles. Moreover, Twitter Usage Statistics indicate that approximately 500 million tweets are sent daily. Therefore, Twitter data provides an abundance of information that may contain crucial insights. Twitter also captures the beliefs and opinions of a broad cross-section of social media users worldwide, rather than just academics, policymakers, journalists, or experts. In the construction of our TEPU index for Lebanon, we make use of the latter point in two ways: First, we gathered global data from Twitter to broaden the search space beyond Lebanon, allowing for a more comprehensive analysis. Second, we expanded our data collection beyond certain populations previously defined in the existing literature, such as expert opinion tweets utilized by Yeşiltaş et al. (2022). Furthermore, it is worth mentioning that each tweet comes with a precise timestamp and cannot be revised or edited, adding to the credibility and reliability of Twitter data.[2]

Further in this paper, we employed two scaling methods to construct the TEPU index for Lebanon, one proposed by Baker et al., (2021) and the other was also proposed by (Baker et al., 2021)as a variant of their main index and adopted by Lee et al., (2023) as their primary scaling method. Our objective was to determine the most suitable scaling method for identifying and uncovering major economic policy uncertainty events that took place in Lebanon during the last decade. We found that both scaling methods were effective in detecting periods of high economic policy uncertainty in Lebanon, but there were some real differences between the two graphs presented in Figures 3 and 4 that reflect the use of both scaling methods. In addition, we found that these two graphs share many common points, such as the general trend of Lebanon's economic policy uncertainty index over time. This means that while the choice of the scaling method may have some impact on the interpretation of the data, it does not fundamentally change the overall message conveyed by our indexes. Furthermore, we showed that our index which was constructed using the second or alternative scaling method was not accurate when dealing with outliers, such as the total number of tweets that contained "Lebanon" or "Lebanese" in the month of August 2020. Our findings suggest that the TEPU index based on the first scaling method, as shown in Figure 3, is the most reliable for Lebanon. This

[2] In September 2022, Twitter released a new feature allowing users to edit their tweets, but only within 30 minutes of sending the tweet. This feature is only available for users who have a Twitter Blue subscription.

is because the first scaling method was not impacted by outliers and directly relates to our data sample.

Our paper is related to other papers in the literature that has developed Twitter-based Economic Policy Uncertainty (TEPU) indexes for countries such as the US (Baker et al., 2021), Turkey (Yeşiltaş et al., 2022), Chile (Becerra & Sagner, 2020) and China (Lee et al., 2023). However, in this paper, we have developed a unique and distinctive methodology that was specifically designed to overcome the challenges posed by the unavailability and lack of reliability of data in Lebanon. This methodology enabled us to create a reliable and effective TEPU index for Lebanon. To evaluate the effectiveness of our TEPU index, we conducted an event analysis and found that our TEPU index is significant, as it links with major political and economic events in Lebanon that occurred between 2011 and 2022. We observed that our TEPU index significantly spikes during these major political and economic events, effectively tracking the evolution of economic policy uncertainty.

The remainder of this paper is organized as follows. Section 2 provides a literature review. Section 3 is divided into three subsections. Section 3.1 explains the distinctive methodology we developed to construct the TEPU index for Lebanon. In this section, we give an overview of how we created and employed Python scripts that interact with the official Twitter API to fetch and transform tweets into actionable data. In Section 3.2, we discuss our reasons for using the Twitter API and the challenges and limitations we encountered while using this service from Twitter. In Section 3.3, we describe how we constructed our TEPU index using two scaling methods and determine which method best fits Lebanon. In this section, we also conducted an event analysis to show how our TEPU index is linked to major political and economic events in Lebanon between 2011 and

2022. Section 4 concludes the paper. Finally, Section 6 contains all our figures and tables for this paper.

CHAPTER 2

LITERATURE REVIEW

The concept of uncertainty has been a prominent theme in economics, particularly in macroeconomics where it was referred to as macroeconomic uncertainty. Recently, the literature has specifically focused on the concept of economic policy uncertainty, as opposed to general macroeconomic uncertainty. The renowned economist Frank Knight's definition of uncertainty, in 1921, has become the modern understanding of this concept. Knight (1921) starts by defining risk, which is a related concept of uncertainty, as a known probability distribution over a specific set of events, while uncertainty pertains to the inability of individuals to forecast the likelihood of specific events occurring. This distinction between risk and uncertainty has been explored across several fields, including finance, psychology, philosophy, and economics. Bloom (2014) explores the concept of uncertainty, which he describes as a vague and undefined notion. In addition, he investigates the frequently asked question of how economic agents such as households, policymakers, and businesses adjust their economic behaviors in response to fluctuations in uncertainty. Uncertainty plays a significant role in shaping investment and consumption decisions at the micro level. In fact, when uncertainty is high, economic agents tend to delay making these decisions (Pindyck, 1988; Rodrik, 1991).

2.1 Navigating the Impact of Uncertainty on Economic Outcomes

A significant portion of the literature focuses on identifying the mechanisms by which uncertainty affects economic outcomes. For instance, high levels of uncertainty often lead households to postpone their purchases of durable goods (Eberly, 1994). Early studies indicate that investing in projects with positive net present value is a rational decision, and greater uncertainty may potentially increase investment returns if it leads to an upside risk for the project (Hartman, 1972; Abel, 1983; Paddock et al., 1988). More comprehensive investment models have been developed to address the interplay between uncertainty, the choice of timing, and irreversibility (Bernanke, 1983; Pindyck, 1988) Additionally, Dixit (1989) and Dixit & Pindyck (1994) demonstrate that increased uncertainty impacts investment behavior by generating an option value of waiting. During uncertain times, firms choose to postpone capital expenditures that involve sunk costs until more information becomes available. Uncertainty can also have real effects on employment. Schaal (2017) argues that uncertainty has limited effects on real options when search frictions are the only costs associated with labor reassignment, provided that firms can easily reverse their employment decisions. Although Schaal's model can explain approximately 40% of the overall increase in unemployment, it seems that uncertainty alone cannot fully account for the magnitude and persistence of unemployment during the 2007-2009 recession. Furthermore, rising uncertainty can prompt companies to adapt to flexible margins, such as part-time employment, due to the fixed expenses associated with hiring and firing (Valletta et al., 2018).

2.2 Developing Empirical Measures for Uncertainty

The recent literature focuses explicitly on the impact of economic policy uncertainty, as opposed to general macroeconomic uncertainty (Rodrik, 1991; Hassett & Metcalf, 1999; Altug et al., 2009; Baker et al., 2016; Baker et al., 2021; Lee et al., 2023). In recent years many empirical studies have placed great emphasis on developing accurate measures for uncertainty. Understanding people's perceptions, attitudes, and emotions across various online platforms has become crucial with the increasing use of social media. The reason for the growing interest in understanding such behaviors is the ease of accessing digital archives of opinionated data, as well as the application of natural language processing (NLP) techniques that enable researchers to save memory. The field of measuring economic policy uncertainty has been expanding rapidly, with researchers utilizing text search methods to create new indexes. A prominent example is the economic policy uncertainty (EPU) index created by Baker et al. (2016). This EPU index is based on the frequency of articles in ten leading U.S. newspapers since 1985 that contain specific economic policy uncertainty keywords. In recent years, various other indexes have emerged to measure economic uncertainty. For example, the World Uncertainty Index (WUI), created by Ahir et al. (2018), measures uncertainty by tracking the frequency of the word "uncertainty" in country reports released by the Economist Intelligence Unit for 143 countries quarterly since 1996. Moreover, Baker et al. (2021) develop daily, weekly, and monthly Twitter-based Economic Uncertainty indicators. Yeşiltaş et al. (2022) create a high-frequency Twitter-based Economic Policy Uncertainty (TEPU) index for Turkey using a select group of Twitter user accounts deemed to reflect an expert opinion on the subject. Similarly, Becerra & Sagner (2020) develop a daily-frequency measure of economic uncertainty for Chile by scraping Twitter data. Additionally, Lee et al. (2023) develop novel daily and monthly frequency "censorship-free" Twitter-based indices to measure Chinese economic policy uncertainty from 2010 onwards. Other studies also built economic policy uncertainty indexes based on news coverage and examined their impact on domestic firms (Jirasavetakul & Spilimbergo, 2018). It is worth noting that all these indexes were developed using text search methods that depend on specific words. This is similar to the technique used in the newspaper-based economic policy uncertainty index developed by Baker et al. (2016), which is based on the frequency of newspaper articles and served as the foundation for this keyword search approach. This novel approach guarantees the dependability and significance of the economic policy uncertainty index and has been shown to be effective.

CHAPTER 3

THE TWITTER-BASED ECONOMIC POLICY UNCERTAINTY INDEX

Understanding people's perceptions, attitudes, and emotions across various online platforms has become crucial with the increasing use of social media. Fortunately, researchers can access the data from many online platforms, including Twitter, which offers historical archives of opinionated data from public conversations. We obtained Twitter data from the digitalized Twitter archives through Twitter API for Academic Research. To analyze this massive amount of data, researchers have utilized sentiment analysis or opinion mining, which involves using computational methods to study people's opinions, sentiments, emotions, and attitudes toward various topics, such as products, organizations, services, and their attributes (Liu, 2015). Although sentiment analysis may pose research challenges, according to Liu (2015) it is a promising method for anyone interested in opinion and social media analysis, and its importance is growing in society and businesses.

Sentiment analysis has been one of the most active and important research fields in natural language processing (NLP) since the early 2000s, with applications in data mining, text mining, web mining, and information extraction (Zhang et al., 2018). Given the significant impact of sentiment analysis, it has spread into management and social sciences, including marketing, finance, economics, political science, and even history. This spread is due to the fact that opinions have a crucial impact on our beliefs, behaviors, and perceptions of reality, affecting almost all human actions. Consequently, whenever individuals or organizations need to make a decision, they often seek the opinions of others.

In this study, we develop a measure of economic policy uncertainty based on historical Twitter public conversations that we believe reflect a useful and important general sentiment on economic policy uncertainty events in Lebanon.

3.1 Methodology

The Twitter-based Economic Policy Uncertainty (TEPU) Index for Lebanon was the new variable we aimed to construct from scratch by gathering data from Twitter. To gather this data in an efficient and quick way, we utilized the Twitter application programming interface, or Twitter API for short. However, accessing the Twitter API was not without its challenges and limitations, which we will discuss in detail in the next section. In this paper, we created and employed Python scripts that interact with the official Twitter API to fetch and transform tweets into actionable data. This curated data is the basis for generating a Twitter-based Lebanese Economic Policy Uncertainty index. The unavailability and lack of reliability of data in Lebanon were taken into account. To gain access to global public Twitter data using Twitter API, new developers are required to complete an Academic Research form on Twitter's website. Having met the requirements, we were granted access to the Twitter API for Academic Research, allowing us to collect and analyze up to 10 million tweets per month as one of our special access limitations.[3] Twitter has granted us access to its real-time and historical public data, as well as additional features and functionality that helped us in the collection of more accurate and unbiased datasets. We will elaborate on our experience using the Twitter API in section 3.2.

---

[3] More information about the data that developers can access through the Twitter API for Academic Research can be found by following this link: https://developer.twitter.com/en/products/twitter-api/academic-research

Many countries have developed their own Twitter-based Economic Policy Uncertainty (TEPU) indexes such as the US (Baker et al., 2021), Turkey (Yeşiltaş et al., 2022), Chile (Becerra & Sagner, 2020), and China (Lee et al., 2023). Given that Lebanon was and remains a politically and economically crisis-ridden country, we were interested in creating a Twitter-based Economic Policy Uncertainty (TEPU) index for Lebanon similar to the indexes developed by the above countries. Our primary objective was to develop a TEPU index that best fits Lebanon and can effectively identify and uncover significant economic policy uncertainty events that have taken place in the country since 2010. However, collecting reliable data from Twitter for Lebanon can be a challenging task. Therefore, we recognized that developing an optimal methodology in this context would be crucial for ensuring the success and reliability of our index. We were convinced that creating a unique methodology, which utilizes Twitter data and is tailored to Lebanon's data availability and characteristics, would result in a reliable TEPU index for Lebanon. We believe that creating a distinctive methodology for Lebanon would be an excellent starting point for further research in Lebanon and the MENA. This section outlines the challenges we faced during the construction of the Twitter-based Economic Policy Uncertainty (TEPU) index for Lebanon and describes the methodology developed to overcome these challenges.

During the development of our distinctive methodology, which utilizes Twitter as the primary social media platform, we encountered some sensitive issues that we addressed carefully to ensure the reliability of our index. These sensitive issues are as follows: Firstly, Lebanon is a relatively small country with a limited number of well-known Twitter influencers, and Twitter is not the most commonly used social media platform in Lebanon. Therefore, it was essential to consider the representativeness of the data. Secondly, the issue of plagiarism and automated accounts, or "bots," posed a challenge as it made Twitter data very noisy. It was curial to filter out the noise and maintain a high level of accuracy in our data collection. Thirdly, there are several places around the world called "Lebanon," which created ambiguity in identifying relevant tweets related to the country of Lebanon in the Middle East. Hence, it was important for us to limit our searches to events that were relevant to the country of Lebanon.

First and foremost, let us delve into the sensitive issues we encountered during the development of our methodology. We began drafting our distinctive methodology with the proposed solutions we had in mind for these sensitive issues. The first sensitive point was that Lebanon is a relatively small country with a limited number of well-known Twitter influencers, and Twitter is not the most commonly used social media platform. For instance, Lebanese economists on Twitter have significantly fewer followers than economists from other countries, and they occasionally post on Twitter. The same is true for many other Lebanese academics, parliamentarians, journalists, private sector professionals, and policymakers who have very few followers and therefore limited social influence in Lebanon, especially on Twitter. According to recent data from Internet World Stats, Facebook and WhatsApp are the most popular social media apps in Lebanon, followed by Twitter and Instagram.[4] In addition, the data from Twitter's advertising resources shows that there were only 531,000 Twitter users in Lebanon at the beginning of 2023, which is approximately 9.8% of the total population.[5] It's important to note that the number of users reached by Twitter's ads is different from the number of monthly active users. Therefore, there may be differences between the total number of active users

[4] https://medialandscapes.org/country/lebanon/media/social-networks

[5] https://datareportal.com/reports/digital-2023-lebanon

on Twitter and the number of users reached by Twitter's ads. Consequently, these numbers were not large enough to base our sentiment analysis on. The second sensitive issue was the challenge created by plagiarism and automated accounts, or "bots," which made Twitter data very noisy. We observed that many people copy-paste another person's tweet on their page without retweeting it, while some bots and fake accounts upload the same tweet multiple times on the same day to create distractions. It was challenging to control this behavior because the Twitter API only allows us to control retweets, capitalization i.e., capital-small letters, the language used, and a limited set of other options that we will discuss in the following section. This behavior presented a problem because we did not want to count a Tweet more than once. Counting a Tweet multiple times would create a bias and give it more significance than other Tweets that were only posted once. Finally, the existence of several places around the world called "Lebanon," was another sensitive issue. There are several cities in the United States of America called Lebanon, including the city of Lebanon in Boone County, Indiana. Additionally, there is a county named Lebanon in Pennsylvania. Furthermore, there are other places around the world with the name "Lebanon." This issue could bias our results by counting unrelated tweets containing the word "Lebanon" but with no relation to the country of Lebanon in the Middle East which our study is based.

We have successfully addressed the three sensitive issues mentioned earlier. Shortly, we will provide a detailed discussion on how we resolved each of them. Now, let's delve into the distinctive methodology that we have developed to overcome these issues. To start with, our methodology was heavily influenced by the work of (Baker et al., 2016) and (Baker et al., 2021), in which they counted the frequency of newspaper articles and tweets, respectively, that are related to economic policy uncertainty. In order to develop a practical methodology for Lebanon, our approach focused on two key components. First, we gathered global data from Twitter to broaden the search space beyond Lebanon, allowing for a more comprehensive analysis. Second, we expanded our data collection beyond certain populations previously defined in the existing literature, such as expert opinion tweets utilized by Yeşiltaş et al. (2022), to ensure a more holistic perspective. To generate our weekly and monthly Twitter-based economic policy uncertainty index for Lebanon, we searched for all global tweets from various locations worldwide containing at least one keyword from the following quartet of terms: (1) "Lebanon," (2) "Economic," (3) "Policy," and (4) "Uncertainty." By doing so, we successfully addressed the first sensitive issue caused by Lebanon's relatively small size, particularly the issue of the limited number of well-known Twitter influencers and the relatively low usage of Twitter compared to other social media platforms in Lebanon. Hence, we addressed this regional issue by widening our search to include global tweets, including those from the MENA region, Gulf countries, Egypt, Europe, The US, and others that provide jobs and investment funds for small, open economies like Lebanon. This approach ensured an inclusive way of capturing economic policy uncertainty in Lebanon through the lens of global public tweets.

3.1.1 Initial Set

To elaborate on our distinctive methodology for Lebanon, we began by retrieving public global Tweets from Twitter. First, we extracted all English tweets related to Lebanese uncertainty (categories 1 and 4) using the keywords listed in Table 1. Particularly, we retrieved all global Tweets from various locations worldwide containing at least one keyword from the following two categories: (1) "Lebanon," and (4)

"Uncertainty." This step led to the formation of our largest collection of global Tweets, which we named the "Initial Set". Our Initial set contained 7,451 original global tweets sent by a total of 4,933 distinct users from January 1, 2010, to August 1, 2022. The Tweets in this set were raw data obtained from Twitter API. We applied controls for retweets, capitalization, and language used to ensure a high-quality dataset. Therefore, we excluded all retweets from our search to ensure that we only selected the original tweets. Additionally, we standardize capitalization to avoid distinguishing between words based on capital-small letters. For example, "Lebanon" was treated the same as "LEBanon," "lebanoN," or "LeBaNoN." Lastly, we restricted our search to English tweets and excluded Arabic Tweets from our analysis. This was because extracting and analyzing Arabic Tweets was extremely difficult due to the complex nature of the Arabic alphabet and the use of additional characters when writing Arabic words.

3.1.2 Intermediate Set

To address our second sensitive issue regarding plagiarism and bots, we developed a Python code to ensure the accuracy of our dataset and eliminate any repeated tweets by bots or users. This step led to the formation of our "Intermediate Set," which contained 5806 tweets. Figure 1 provides an overview of our Python algorithm and outlines the methodology we followed, starting from obtaining the raw tweets from the Twitter API (7451 Tweets). Initially, we inputted the raw tweets into our Python code and applied various layers of checks to ensure their accuracy. The first step involved reapplying Twitter API filters using our Python code, just like what Twitter API did, to ensure that Twitter API retrieved the relevant tweets that aligned with our requirements. This step resulted in the deletion of 1343 tweets from our dataset that did not match our filters (categories 1 and 4), producing a new set of 6108 tweets. We will elaborate in the next section on why certain Tweets were identified as relevant by Twitter API but did not align with our Python code that used the same filters. Despite the Twitter API identifying these Tweets as relevant, we removed them from our dataset as they did not meet our requirements. Our priority was to ensure accuracy and generate a reliable index, which we achieved by only including relevant tweets in our dataset that contained the specific keywords we were searching for. The second step involves implementing our "Remove Parrots" algorithm to eliminate repeated tweets by the same user, whether posted at the same or different times, from our previous set of 6108 tweets. We only selected the original relevant tweet posted by the user based on the tweet timestamp. This step resulted in the removal of 83 tweets, leaving us with a new set of 6025 tweets that contained no repeated tweets by the same users, whether posted at the same or different times. The second step involves implementing our “Remove Plagiarism" algorithm to the latter set of 6025 tweets to detect any plagiarized tweets with identical characters posted by different users (repeated tweets posted by the same user were already detected from step 2). The algorithm retained the original tweet as relevant, resulting in the elimination of 219 tweets. It is worth noting that our "Remove Parrots" algorithm detected 83 duplicate tweets, which could have been identified by our "Remove Plagiarism" algorithm in the third step of our methodology (repeated tweets by the same user, whether posted at the same or different times are also considered plagiarized tweets since these tweets include identical characters). However, we intentionally removed the repeated tweets by the same user first, whether posted at the same or different times. We then proceeded to detect any plagiarized tweets with identical characters posted by different users. As a result, this process allowed us to obtain accurate statistics on each step of our methodology and better understand how bots and automated accounts work. We have now reached our

"Intermediate Set", which contained 5806 tweets that have passed through a second round of filtering using Python Algorithms and then carefully screened for both Parrots and Plagiarism.

3.1.3 Final Set

After we reached our Intermediate Set of 5806 Tweets, we proceeded to the important part of our methodology, which involved building Economic and Policy categories, which we called categories two and three, respectively. Several papers, including Yeşiltaş et al. (2022), have developed a Twitter-based Economic Policy Uncertainty index by gathering the most commonly used terms related to "Economic" and "Policy" and then categorizing them based on their relevance to each category in a particular country. For instance, words that are commonly associated with Economics such as Banks, Money, Capital, Cost, and so on, were classified under the Economic category. Similarly, words associated with Policy such as Government, Leaders, Ministers, Politicians, Regimes, and so on, were classified under the Policy category. Hence, this approach to developing the Economic Policy Uncertainty index involved completing all categories based on plausible predictions of relevant terms in each category. However, in our paper, we adopted a slightly different approach, which was inspired by Lee et al. (2023). Our primary methodology for creating the Economic and Policy categories, as shown in Table 1, was to manually annotate all the tweets from the Intermediate Set. This process involved identifying the most frequent and important keywords in each tweet of our Intermediate Set that could be used for building these categories. The manual annotation process was time-consuming because we carefully scrutinized each tweet and assessed the type of English words used in relation to Lebanon.[6] The manual annotation process involved two different human annotators who were experts in the field of economics and finance in Lebanon. We then manually selected the relevant and significant terms to distribute them into the Economic and Policy categories. It is important to note that these keywords were also selected based on their actual meaning as found in the tweets, which was helpful in assigning them to their appropriate categories. In summary, we manually annotated all the tweets from the Intermediate Set, then we identified and extracted the most frequently used and important keywords from the 5806 tweets, and finally, we distributed these keywords to the Economic and Policy categories. Therefore, we completed the first draft of each of the four categories shown in Table 1.

Following that, we copied all the keywords from the four categories into our Python code and let Python do the remaining work. The Python code generated a subset from the Intermediate Set, which we named our “Final Set.” This Final Set was filtered to contain relevant tweets that had at least one keyword in each of the four following categories: (1) "Lebanon," (2) "Economic," (3) "Policy," and (4) "Uncertainty." The set of tweets that met all four criteria was named the Final Set, which contained 1034 Tweets. In summary, we used a data-driven approach in building Categories 2 and 3 by using manual annotation on the Intermediate Set, similar to the "Let the data speak for themselves" approach. Rather than randomly selecting words based on the most frequent and important terms that could be associated with categories 2 and 3, we chose to use a manual annotation process. Although this approach was time-consuming, we were certain our manual annotation was necessary to ensure the accuracy of our four categories, especially since we were dealing with a politically and economically crisis-ridden

[6] The manual annotation process took approximately 1.5 months.

country. Furthermore, we opted for the manual annotation process since we wanted to ensure that we were using the precise English terms used by the Lebanese people as well as Arabs, and people from other parts of the world when discussing specific economic policy events related to Lebanon. Since Arabic is the primary language of Lebanon, we were concerned that selecting keywords randomly for our categories might not cover all the relevant terms used in public Twitter conversations. Hence, we were convinced that a manual annotation approach for building categories 2 and 3 was necessary to build an accurate TEPU index for Lebanon.

After we reached our Final Set of 1034 Tweets, we proceeded to the last step in creating the TEPU index which was to scale our data. The standard scaling method was to compute M, which represented the mean of our entire sample, and then multiply each weekly or monthly series by its own 100/M. In this paper, we employed two scaling methods, one proposed by Baker et al. (2021) and the other was also proposed by Baker et al. (2021) as a variant of their main index and adopted by Lee et al. (2023) as their primary scaling method. At the end of this Chapter, we will go over both scaling methods in detail, presenting both TEPU index graphs generated using the same updated Final Set but with different scaling methods. Then we proceed to compare how each TEPU index behaves under these scaling methods. Our aim is to determine which scaling method is preferable in the case of Lebanon in terms of identifying major economic policy uncertainty events that occurred over the last decade.

Furthermore, Table 2. shows the frequency of each of our keywords across all four categories in our Final Set. This frequency table displays the frequency of specific English terms used by the Lebanese people as well as Arabs, and people from other parts of the world when discussing specific economic policy events related to Lebanon.

Additionally, this table will be useful in any future updates to our keyword table (Table 1). By analyzing the frequency of each keyword, we can gain a better understanding of the significance of each keyword within its context. Hence, this information can help determine which keywords to keep or remove from our keyword table in future reviews.

3.2 Twitter Developer Platform

The Twitter Developer Platform offered us the necessary tools, data, and API features for our data collection and analysis of public Twitter conversations, which are dynamic exchanges of information and ideas that occur freely among users on Twitter. This platform is officially provided, managed, and maintained by Twitter itself, and not by any third party. Therefore, it was well-structured and suitable for our usage. The Twitter Developer Platform consists of three distinct products provided to Twitter-approved developers: Twitter for Websites, Twitter Advertising API, and Twitter API. In our study, our primary focus was on using the Twitter API for retrieving data from Twitter. There were many purposes for each of these three distinct services provided under the Twitter Developer Platform. For this paper, we will only discuss our reasons for using the Twitter API, as well as the challenges and limitations we encountered while using this service from Twitter.

3.2.1 Twitter API

Let us begin by asking ourselves some important questions. What does the term API mean? What is the Twitter API? When was it launched? And which type of Twitter API access did we use? It is essential that we start by thoroughly answering these questions since they will serve as the sole means for collecting relevant data for our research study.

To start with, API is an abbreviation for Application Programming Interface. In the context of APIs, the term "Application" refers to any program that performs a certain function. The term "Interface" can be viewed as a service contract between two applications. This contract specifies how the two applications will communicate with one another via requests and responses. APIs, in other words, provide methods allowing two software components to interact with one another through the use of a set of definitions and protocols. APIs are essential for software systems to interact with each other. For example, when we use a weather app on our phone, the app reaches out to a weather bureau's software system via APIs to display daily weather updates. Another way of thinking about APIs is to imagine what a waiter's role is in a restaurant. Without the waiter, you would need to go all the way into the kitchen and grab something to eat (assuming you are allowed into the kitchen in the first place!). In this case, the waiter acts as a link between two entities: the customer and the chef. The same analogy applies to APIs, which serve as a link between two pieces of software, enabling them to interact and exchange information. Furthermore, anytime we use social media platforms like Facebook, Instagram, Twitter, TikTok, etc., we are communicating with an API. APIs have transformed the way we use technology by simplifying communication across software systems and improving user experience.

Specifically, the Twitter API is a collection of programmatic endpoints that are employed for fetching and posting data to the Twitter conversation. The Twitter API documentation, which is offered to the public by Twitter, includes detailed information on how developers should arrange and structure their requests and how they should parse responses from the API.[7] With this API, we can retrieve and interact with a wide range of resources, such as Tweets, Users, Spaces, Places, Direct Messages, Trends, Lists, and Media. The latest version of the Twitter API available to all Twitter-approved developers is Twitter API v2.

In June 2020, Twitter granted developers early access to its Twitter API v2 to provide them with a preview of what this program may offer. Then, in November 2021, Twitter released a production-ready, stable version of v2 to all Twitter-approved developers. The Twitter API has evolved over time by granting different levels of access, allowing developers and academic researchers to analyze the public conversation on Twitter in diverse ways, gaining important insights into user behavior, sentiment analysis, and more.

In our study, we utilized the Twitter API for Academic Research, which is only granted to approved academic researchers by Twitter. It is important to note that in order to be authorized as an academic researcher by Twitter and recognized as a Twitter-approved developer, you must first complete an Academic Research Form.[8] We submitted our Academic Research Form through the Twitter Developer Platform website. For no cost at all, Twitter accepted our application and granted us the status of approved academic researchers and Twitter-approved developers. As part of the application process, Twitter required a comprehensive research proposal to assess our eligibility to access the Twitter API for Academic Research. A research proposal must adhere to the terms and conditions set forth by Twitter, as well as show how the research study would

---

[7] The Twitter API documentation can be found by following this link: https://developer.twitter.com/en/docs/twitter-api

[8] Twitter API v2 Academic Research Access is available by following this link: https://developer.twitter.com/en/products/twitter-api/academic-research. You may check your eligibility for this access as well as details on the data you have access to.

be useful by obtaining data from Twitter. To gain access to the Twitter API for Academic Research, we had to answer a series of questions through a form provided on the Developer Platform. Namely, “Describe your research project." "Describe how Twitter data and/or Twitter APIs will be used in your research project." “Describe your methodology for analyzing Twitter data, Tweets, and/or Twitter users”. Finally, “Describe how you will share the outcomes of your research (include tools, data, and/or resources)." The Twitter API for Academic Research provided us with data from the public Twitter conversation at no cost. This specialized access offered us the ability to use all Twitter API v2 endpoints. It is worth mentioning that our access to the Twitter API for Academic Research is primarily restricted to university academics with specific research and study objectives, known as Twitter-approved academic researchers. Practically, any of the following may be applicable for a free version of the Twitter API for Academic Research: Graduate students preparing for their theses, Ph.D. candidates preparing for their dissertations, or research professionals affiliated with an academic institution in some capacity. Researchers who are interested in applying for such academic free access must have a well-defined research proposal and a detailed strategy for using, analyzing, and sharing Twitter data as we previously discussed. Once granted access to Twitter API for Academic Research, academic researchers benefit from downloading a total of 10 million tweets per month, longer query length than other types of access, more filters to be applied when searching tweets, access to the Full archive of tweets since 2005, and advanced tweet search operators like a place, country, geographic location, etc. Furthermore, there are some limitations to take into account like query criteria being restricted to 1024 characters and rate limits on full archive search. The full-archive search has a request limit of 300 queries every 15-minute window, which includes the pagination of returned Tweets.

3.2.2 Twitter API with Python

Social media platforms are becoming increasingly recognized as rich sources of real-world data. Data is freely available on platforms such as Twitter and Facebook. This data may be utilized in a variety of beneficial ways, such as sentiment analysis. In our study, we focused on Twitter as our sole social media platform since we believe that it offers rich and useful data that can lead to fruitful research outcomes for Lebanon.

As mentioned earlier, we were granted access to the Twitter API for Academic Research after Twitter approved our application for a Twitter Developer Account. This account was essential because it was the only way for the Twitter API to accept our requests. We made sure to set up our Twitter credentials correctly to ensure smooth and easy access to the Twitter API. The very first step we did after creating our Twitter Developer Account was to create a new project in our account that we called "Thesis Data". Next, we created a new app specific to our project. At this point, we reached a page with our keys and tokens, which we used in each request sent to the Twitter API.

Now that we were ready to go, we proceeded to submit our first request through the API. We kept in mind that the v2 version of the API was still relatively new, which meant that there were fewer community resources available to help us if we encountered any issues while collecting data for our research. The next step was to design and implement robust and deterministic Python scripts that would communicate with the Twitter API and process the downloaded data. Since we wanted to achieve accurate and stable results, we did not use any third-party Twitter library to communicate with

Twitter's API. This is because third-party libraries tend to be slow in keeping up with the Twitter API's updates, and they are often not reliable. Searching through Twitter third-party libraries and reviewing their code to determine their quality is a lot more time-consuming than developing our own API communicator from scratch. That is why we only used state-of-the-art Python libraries that provide basic boiler-plate functions, mainly: "requests" helped us in crafting and sending custom HTTP requests to the API, "re", short for regex (Regular Expressions) allowed us to parse the downloaded Tweets by recognizing patterns we specified, "os" allowed us to get the secret API token from the system's environment, "pickle" backed up all API downloaded data in a RAW format in case our quota ran out, etc.

This approach enabled us to fully utilize the Twitter API's capabilities and gain complete access to its functions. We utilized the Twitter API in different ways throughout our study such as retrieving locations when applicable, gathering tweet statistics, and more.

3.2.3 Twitter API: Challenges and Limitations

Although the Twitter API for Academic Research enabled us to gather data easily for our research project, we encountered some challenges and limitations while working with this Twitter service. We will discuss the main challenges and limitations that we faced throughout our experience while working with Twitter API for Academic Research.

The first challenge we encountered was the lack of choices for retrieving public data from Twitter. We had to rely on Twitter API since there was no other efficient way to retrieve public data from Twitter (There was no way to scrape the data since Twitter only shows recent tweets to normal users, not the full archive. Additionally, Twitter's bot-fighting mechanisms are sophisticated and very hard to bypass. Furthermore, no third party could provide us with the data from Twitter that we wanted.) While the Twitter API documentation was crucial in our work, we observed that it requires additional improvement and illustration in several key components. In other words, the Twitter API documentation was vague and hard to comprehend, but it was still workable with some additional effort. Therefore, in order to properly use Twitter API and address any issues encountered, we had to search for solutions outside of the Twitter API documentation. So, we examined several sites and studied numerous reviews from Twitter developer's forums to pinpoint the particular issues that we couldn't explain from the Twitter API documentation.[9] Sometimes when Google was unable to provide us with the answers we needed, we had to do our own black box testing to infer the mechanisms employed by the Twitter API. This approach helped us sometimes to pinpoint wrong assertions listed on the Twitter API Documentation website. Furthermore, we witnessed inconsistency and discrepancy between the expected behavior described in the Twitter API documentation and the actual behavior observed while we utilized Twitter API for the purpose of data retrieval from Twitter. For example, the Twitter API should only search for keywords in the body of the Tweet and not anywhere else since a relevant keyword is defined in the Twitter API documentation as a "relevant keyword within the body of a Tweet". However, we observed that the Twitter API matched the relevant keywords we provided with tweet data fields beyond the tweet body itself. Specifically, the Twitter API was identifying tweets as relevant if it found any of the relevant keywords in the body of the tweet, the account's username, the bio, or even the user's location, which was not what

[9] For example, an extensive guide can be found by following this link: https://towardsdatascience.com/an-extensive-guide-to-collecting-tweets-from-twitter-api-v2-for-academic-research-using-python-3-518fcb71df2a

we wanted; we only wanted tweets that had our relevant keywords in its body. As a result, we faced a real problem because many tweets retrieved by Twitter API and classified as relevant were not actually relevant to our study, so we had no choice but to reapply our own filters on our "Initial Set" to ensure that we fixed the issue that we could not solve directly using Twitter API. In the second step of our methodology, depicted in Figure 1, we reapplied the filters from categories one and four again to our initial set. This step was crucial in ensuring the reliability of our index. In our distinctive methodology, we placed great importance on ensuring that relevant tweets contained relevant keywords in the tweet's body and not in other tweet data fields.

The second challenge we faced was the meticulous and time-consuming process of applying for the Twitter API for Academic Research. Twitter is extremely cautious in approving access to its API and in classifying new developers as Twitter-approved developers. We did not acquire access as a result of our first application submission, therefore we had to submit a second application with a better research proposal and projected outcomes. So, once we provided Twitter with the necessary information and documentation in the second application indicating that I am a full-time Master's student in the economics department at the American University of Beirut (AUB) conducting research for my thesis, access was granted to my account and we officially had a Twitter-approved developer account.

Our Third challenge revolved around the need for a query builder webpage, which was sometimes inaccessible. We made an effort to use Twitter API only when necessary, as we continuously searched for appropriate methods to deal with delays, such as the aforementioned issue with the inaccessibility of the query builder webpage. This page helps the developer craft valid queries that the Twitter API will comprehend. The search endpoints take a single query and return a list of historical tweets that match our query. These queries are composed of operators that match various tweet properties, but it is important to note that the queries are restricted based on the developer's access level. Since we were granted Academic Research access, our query was 1024 characters long.

Now, let us discuss some of the limitations that we encountered when utilizing Twitter API for Academic Research. The first limitation we faced while using the Twitter API was the rate limit, which prevented us from retrieving all tweets at once. We could send only 300 requests per 15-minute window to the search endpoint of the API, and each request would return a result set containing a maximum of 100 tweets. As a result, this rate limit was quite restrictive, requiring us to spend additional time to get all our relevant tweets. In addition, when attempting to extract all tweets containing only the keywords "Lebanon" or "Lebanese" for our scaling approach, our monthly retrieval limit of 10,000,000 tweets was insufficient. We identified nearly 40,000,000 tweets using the keywords from Category 1 shown in Table 1. Given the difficulty in obtaining all of these tweets and applying our filters, we turned to utilize the tweet count-only service provider by Twitter API. It is important to note that counting tweets does not consume from the 10 million quotas; only fetching the content of tweets does consume the quota)

The second limitation we encountered was the lack of confidence in the reliability of the Twitter API. Although it provided us with all relevant tweets, it also generated some noise within the data. Furthermore, the Twitter API for Academic Research was a new service from Twitter that required further development and clarification in the documentation for new developers. Because this service was new, the Twitter API documentation was somewhat ambiguous as we discussed earlier. As a result, we questioned ourselves, "How can we determine if Twitter API hides some of the relevant tweets?" This quandary arose when we downloaded the raw data from Twitter and discovered some irrelevant tweets in our Initial Set. It is essential to highlight that raw data must only contain relevant tweets from Twitter based on our set of relevant keywords shown in Table 1. with no noise or extraneous data, but that was not the case while working with the data retrieved by Twitter API. As a result, we created a unique methodology to address these issues effectively.

The third and final limitation we faced was the inability of Twitter API to retrieve publicly relevant tweets that have been deleted by users or existing relevant tweets from private accounts. Deleted tweets are simply inaccessible through Twitter API, and private tweets cannot be detected by the search endpoints of Twitter API. This is a significant problem that we can't solve, as these tweets may be relevant and important to our study or other studies. Therefore, if certain accounts become private at any point, all of their tweets will become inaccessible to us. As a result, we are unable to access all relevant data on Twitter using the Twitter API.

Finally, it is worth mentioning that in February 2023, Twitter announced that it will begin charging third parties for access to its platform's public data. The price for this API service with "basic" level access (10,000 tweet retrievals per month) starts at $100 per month. However, this amount of data is not enough for many users, while the high-end pricing ranges from $42,000 per month for the small package (50,000,000 tweet retrievals per month) to $210,000 per month for the large package (200,000,000 tweet retrievals per month), making it unaffordable for most academic researchers. Prior to this change, researchers could pay around $500 per month to access up to 10% of the approximately 1 billion tweets that are generated on Twitter each month.[10]

[10] More information could be found by following this link: https://edition.cnn.com/2023/04/05/tech/academic-researchers-blast-twitter-paywall/index.html

3.3 Constructing the Weekly and Monthly TEPU Index

To construct our Twitter-based Lebanese Economic Policy Uncertainty (TEPU) index, we needed our Final Set of tweets which contained 1034 tweets from January 1, 2010, to August 1, 2022. It is worth noting that our TEPU index is instantly updated in real-time via the Twitter API for Academic Research, which we were granted access to by Twitter. Consequently, whenever we receive updated raw data from Twitter, referred to as the Initial Set in Figure 1, and once we apply our methodology and scaling technique to the data, we can obtain an updated Lebanese TEPU index for our sample. The last update for our raw data from Twitter was from January 1, 2010, to January 18, 2023, and we obtained an updated Final Set of tweets that contained 1011 tweets. It was not surprising that we got a lower number of relevant tweets in our updated Final Set than our original Final Set from the first sample period, which was shorter in time duration, from January 1, 2010, to August 1, 2022 (This original Final Set contained 1034 tweets, see Figure 1). We have many explanations for why the number of relevant tweets decreased. Although Twenty-two relevant tweets were detected from August 1, 2022, to January 18, 2023, this slight decrease in the total number of relevant tweets between the updated Final Set and the original Final Set was mainly because some accounts turned to private or deleted some relevant tweets in their history of tweets. Hence, some tweets no longer appeared in the updated Final Set since Twitter API was unable to retrieve publicly relevant tweets that have been deleted by some users or existing relevant tweets from accounts that turn private. Deleted tweets are simply inaccessible through Twitter API, and tweets in private accounts cannot be detected by the search endpoints of Twitter API.

For a more in-depth look at the challenges and limitations we faced when using the Twitter API, please refer to Section 3.2 of this paper.

We constructed our TEPU index for Lebanon based on our updated Final Set that contained 1010 tweets, covering the period from January 1, 2011, to January 18, 2023. Our Methodology to reach the Final Set of relevant tweets was clear and presented in detail in section 3.1. We removed the year 2010 while constructing our TEPU index for Lebanon since it only contained one relevant tweet. Therefore, we decided to construct our index from 2011 onward, as this period in Lebanon coincides with the onset of many economic policy uncertainty events that we are trying to identify and uncover. To construct our TEPU index, we followed a similar procedure to the one developed by Baker et al. (2021). We treated the Twitter data-based series as if they had been extracted from a single newspaper. As already discussed, we created and employed Python scripts that interact with the official Twitter API to fetch and transform tweets into actionable data. This curated data is the basis for generating a Twitter-based Lebanese Economic Policy Uncertainty index for our specified time interval from January 1, 2011, to January 18, 2023.

Using our Final Set of relevant tweets for the updated sample period, we first constructed a simple graph to display the evolution of the total number of relevant tweets sorted monthly (see Figure 2). Given that the Twitter API provided us with the precise posting date of each relevant tweet, we adopted two coinciding approaches to show our results, one was a monthly approach and the other was a weekly approach. Figure 5 also displays the evolution of the total number of relevant tweets in our sample period; however, this graph is sorted weekly instead of monthly. The month with the highest number of relevant tweets was August 2020 with a total of 167 relevant tweets as well as the week with the highest number of relevant tweets was “W10820” which refers to the first week (W1) of August (08) of the year 2020 (20). Figures 2 and 5 provide a general overview of how many relevant tweets our methodology captured over the past decade in Lebanon and plot the numbers of these relevant tweets as detected in each month or week in graphs that display the evolution of the total number of relevant tweets in our updated sample period. This step involves gaining an understanding of how the relevant tweets were distributed between months and, more precisely, between weeks in each month in our updated sample period.

After plotting the numbers of the relevant tweets that comprised our updated Final Set on a monthly (Figure 2) and weekly (Figures 5) basis, the final step in creating our TEPU index was to scale the data for our chosen time interval, which we determined to be primarily from January 1, 2011, to January 18, 2023. In this paper, we employed two scaling approaches, one proposed by Baker et al. (2021) and the other was also proposed by Baker et al. (2021)as a variant of their main index and adopted by Lee et al. (2023) as their primary scaling method.

Figures 3 and 6 present our finalized monthly and weekly TEPU indexes, respectively, for Lebanon. These indexes have enabled us to identify and uncover major economic policy events in Lebanon in the last decade. In addition, Figures 4 and 7 also present our finalized monthly and weekly TEPU indexes, respectively, for Lebanon but using a different scaling method from Figures 3 and 6. The monthly and weekly indexes were the exact same indexes that revealed the same information in terms of capturing relevant economic policy uncertainty events in Lebanon, but the only difference was that they presented the number of relevant Tweets on a weekly basis in the weekly TEPU index and on a monthly basis in the monthly TEPU index. We will only shed light on our monthly TEPU index while analyzing how our index identified and uncovered the important economic policy events that occurred in Lebanon in the past decade. Our focus is on Figures 3 and 4 which present our two monthly TEPU indexes for the updated sample period from January 1, 2011, to January 18, 2023, using different scaling methods. Our main goal was to determine which scaling method would provide the best accuracy in interpreting the data in order to reveal significant economic policy uncertainty events in Lebanon. As mentioned before, our TEPU indexes shown in Figures 3 and 4 were constructed using the same updated Final Set of relevant tweets, but the only difference was the scaling method used to construct these indexes. In the context of the TEPU index, scaling involves normalizing the keywords count data to account for differences in measurement units or time. In order to conduct meaningful analysis and interpretation, it is essential that all the data be on a comparable scale. We selected these two main scaling methods because they have been shown to be useful and accurate in the literature. The aim was to determine which method would provide the most accurate interpretation of our data and reveal important economic policy uncertainty events in Lebanon. Overall, our TEPU index is an important variable for uncovering Lebanon's political and economic uncertainty events, and the choice of which scaling method we adopt can significantly impact our index's accuracy. Further interpretation is necessary to fully understand both methods and the results presented in Figures 3 and 4. A conclusion about the best scaling method for constructing the TEPU index for Lebanon will be discussed in the following section.

3.3.1 *Scaling and Results*

We adopted two scaling methods in constructing our TEPU index for Lebanon. The first was proposed by Baker et al. (2021), which involves computing the mean, M, of our entire sample from January 1, 2011, to January 18, 2023, and then multiplying each weekly or monthly series by its own $\frac{100}{M}$ in order to renormalize our index to a mean of 100. The second scaling method was also proposed by Baker et al., (2021) as a variant of their main index and adopted by Lee et al. (2023) as their primary scaling method. It is based on the evolution of the total number of tweets worldwide containing the words "Lebanon" or "Lebanese" each month or week. [11] After using the "Twitter API count" service, we determined that there was a total of 39,506,906 tweets containing the words "Lebanon" or "Lebanese" between January 1, 2011, and January 18, 2023. Hence, our second scaling method was as follows: We scaled the number of tweets in each month by the total number of tweets that mention "Lebanon" or "Lebanese" (we denoted the resulting ratio by $R_{0,t}$). Then, we standardized the ratio $R_{0,t}$ to a unit standard deviation to obtain what we called $R_{1,t}$ using the standard deviation from January 1, 2011, to January 18, 2023. Finally, we computed the mean M of $R_{1,t}$ from January 1, 2011, to January 18, 2023, and multiply $R_{1,t}$ by $\frac{100}{M}$ to renormalize our index to a mean of 100.

After applying the two scaling methods to our data, we plotted the resulting graphs to compare their effectiveness in identifying major economic policy uncertainty events that occurred in Lebanon over the last decade. Figure 3 shows the monthly TEPU index based on the first scaling method proposed by Baker et al., (2021). Figure 4, on the other hand, shows the monthly TEPU index based on the second or alternative scaling method proposed also by Baker et al. (2021) as a variant of their main index and adopted by Lee

[11] Baker et al., (2021) used the total number of tweets containing the word "have"

et al. (2023) as their primary scaling method. While both scaling methods were effective in identifying periods of high economic policy uncertainty in Lebanon over the past decade, there were some real differences between the two graphs presented in Figures 3 and 4, which are worth exploring in more detail.

One of the most noticeable differences between the two graphs is the magnitudes of the spikes in each index. For example, in Figure 3, the spike that resulted from the August 4, Beirut port massive explosion and the resignation of the Lebanese government is the highest in the graph, with a significantly higher magnitude compared to other spikes in the same graph. The same event was identified by the TEPU index in Figure 4, but the spike is not with the highest magnitudes within Figure 4. Therefore, it is crucial to consider the scaling method employed since it can significantly impact how specific events are represented in our TEPU index. Additionally, Figure 4 exhibits a higher number of significant spikes compared to Figure 3, without a significant dominant spike like that of Figure 3 in August 2020. This suggests that the scaling method used in Figure 4 is more sensitive to smaller changes in economic policy uncertainty, whereas the scaling method used in Figure 3 is more effective at highlighting major events that have a significant impact on economic policy uncertainty in Lebanon.

Another important difference between Figures 3 and 4 is the varying levels of importance that each index assigned to specific economic policy uncertainty events, in comparison to other events that occurred in Lebanon since 2011. This distinction is visually apparent in the magnitude of the spikes. For instance, while the monthly TEPU index shown in Figure 3 highlights the August 4, 2020, Beirut catastrophic port explosion event as the most important economic policy uncertainty event in Lebanon in the past decade, the TEPU index shown in Figure 4 highlights Hariri stepping down as the

Lebanese prime minister-designate event that occurred on July 15, 2021, as the most important economic policy uncertainty event in Lebanon in the past decade. In addition, it should be noted that our TEPU index shown in Figure 4 assigns greater weight to certain events compared to our TEPU index shown in Figure 3. Specifically, events such as the collapse of Hariri's unity government in Lebanon on 12 January 2011, uncertainty about the third trial of already postponed general elections in Lebanon in May 2017, the political uncertainty that began in September 2018 and harms investor confidence in Lebanon, the confirmation of the first case of COVID-19 in Lebanon on February 21, 2020, the defaulting of Lebanon on its foreign debt for the first time on March 9, 2020, and the general elections that result in no clear majority in Lebanon on May 15, 2022, were all significantly highlighted in Figure 4 compared to other events in the same figure. While these events are also highlighted by our index in Figure 3, they are not as significant as they are in Figure 4 compared to other spikes in the same index in Figure 3. In order to make a valid comparison, we compared the magnitude of each spike with other spikes in the same index that had the same scaling method. Therefore, the choice of the scaling approach can have a big impact on how the data is interpreted and how well our index can identify events that have the biggest influence on Lebanon's economic policy uncertainty.

Despite these differences, the two graphs share many common points, such as the general trend of the economic policy uncertainty index over time. This means that while the choice of the scaling method can have some impact on the interpretation of the data, it does not fundamentally change the overall message conveyed by the graphs. It is important to mention that the correlation between the two graphs, presented in Figures 3 and 4, is 0.539, which indicates a moderate positive relationship between them.

While understanding the differences and similarities discussed between the monthly indexes shown in Figures 3 and 4 is crucial, choosing the appropriate scaling method for the Lebanese TEPU index is equally important. It is imperative to consider the most relevant scaling method when interpreting our data for Lebanon. In this regard, we would first like to shed light on the main reason for the observed differences between our two indexes. As noted before, Figure 3 displays the monthly TEPU index based on the first scaling method, which involves calculating the mean of our entire sample. On the other hand, Figure 4 shows the monthly TEPU index based on the second or alternative scaling method, which takes into account a new out-of-sample set referred to as the New Set. This New Set comprised the total number of tweets worldwide containing the words “Lebanon” or “Lebanese” in each month or week in our sample period. The major difference between the two scaling methods lies in the New Set that we used in the alternative scaling method. As noted earlier, we created the New Set using the Twitter API's count feature and discovered that it contained 39,506,906 tweets spanning from January 1, 2011, to January 18, 2023. The month with the lowest number of tweets in the New Set was February 2011, with a total of 62,562 tweets, while the month with the highest number of tweets was August 2020, with a total of 5,097,894 tweets. It is important to shed light on the month with the second-highest number of tweets in our New Set, which was October 2019 and it contained 770,593 tweets, revealing the extreme difference in the number of tweets from August 2020. This significant difference between the month with the highest number of tweets and the month with the second-highest number of tweets highlighted the month of August 2020 as an outlier in the total number of tweets within the New Set. It is worth mentioning that we recorded the highest number of relevant tweets in August 2020, with a total of 167 tweets.[12] The month of August 2020 had the most significant spike in Figure 3. On the other hand, the most significant spike in Figure 4 was in the month of July 2021, with a total of 73 relevant tweets and 330,713 detected tweets in our New Set, and not in the month of August 2020, which had the highest number of relevant tweets (167 tweets). Understanding why this happened and identifying the issue with the second scaling approach was crucial. In the second scaling method, we scaled the number of tweets in each month or week by the number of tweets that mentioned “Lebanon" or “Lebanese" (we denoted the resulting ratio by $R_{0,t}$). Therefore, $R_{0,t}$ equals $\frac{167}{5,097,894}$ for the month of August 2020 and $\frac{73}{330,713}$ for the month of July 2021. Then, we standardized the ratio $R_{0,t}$ to a unit standard deviation to obtain what we called $R_{1,t}$, using the standard deviation from January 1, 2011, to January 18, 2023. Finally, we computed the mean M of $R_{1,t}$ from January 1, 2011, to January 18, 2023, and multiply $R_{1,t}$ by $\frac{100}{M}$ to ensure an average value of 100 in our sample. Therefore, A larger value of $R_{0,t}$ will lead to a higher value of $R_{1,t}$ resulting in a higher spike magnitude. So, the highest spike magnitude should be in the month of July 2021, and similarly, that logic explains why the August 2020’s spike is much shorter than that of the month of July 2021 shown in Figure 4. As a result, our index that was constructed using the second or alternative scaling method was not accurate when dealing with outliers, such as the total number of tweets in the month of August 2020 in our New Set. Usually, such outliers are observed in a crisis-ridden country like Lebanon, which is why it's crucial to take measures to control them.

[12] Recall, relevant tweets are the tweets that that had at least one keyword in each of the four following terms: (1) "Lebanon," (2) "Economic," (3) "Policy," and (4) "Uncertainty." The set of tweets that met all four criteria was named the Final Set.

The economic and political environment in Lebanon has been unstable over the past 10 years and beyond. Several significant economic policy uncertainty events have occurred, which have caused a direct impact on the country's political stability as well as its economy. In summary, to better understand how the two different scaling methods that we adopted dealt with these events, we examined the month of August 2020 as a direct clear example. Through this examination, we gained insights into each scaling method's strengths and weaknesses. Furthermore, by looking at our index in Figure 4, we can view that it highlighted some events in a more significant way than our index in Figure 3. This was explained by the fact that the number of tweets in our New Set for a specific month could be the trigger to highlight an event or less highlight it. This sheds light on the importance of considering specific criteria when determining which scaling method better represents the data of Lebanon. Since both scaling methods highlight the same economic policy uncertainty events in Lebanon, it was difficult to determine which scaling method best fits Lebanon. We showed that the second scaling method can be inaccurate when outliers are present, such as in the case of the month of August 2020. This resulted that the second scaling method of our TEPU index, which is shown in Figure 4, highlighted August 2020 as a very low month of economic policy uncertainty, which is clearly not true based on the relevant tweets data and events that occurred during that month. Namely, August 4, 2020, Beirut’s catastrophic Port explosion and the resignation of the Lebanese government events, were among the most significant economic policy uncertainty events in the last decade. Therefore, we believe that our TEPU index shown in Figure 3 is the most reliable Lebanese TEPU index since the scaling method was not affected by outliers and directly relates to our data sample. This underscores the importance of developing robust and reliable scaling methods when analyzing economic policy uncertainty events in Lebanon and other similar countries. By doing so, we can better understand the economic policy landscape in Lebanon and make an informed analysis that can help drive positive change.

3.4 Event Analysis

In what follows, we conduct an event analysis to demonstrate how our TEPU index links to significant political and economic events that occurred in Lebanon throughout the sample period from 2011 to 2022. Figures 8 and 9 display the monthly TEPU index for Lebanon from January 1, 2011, to January 18, 2023, and track the evolution of the economic policy uncertainty events in Lebanon. Earlier we discussed how we developed our TEPU index for Lebanon from scratch, using two different scaling techniques. In this section, we will discuss the meaning behind the spikes depicted in Figure 8 and Figure 9 to gain a better understanding of the economic policy uncertainty events that occurred in Lebanon in the past decade. It is worth noting that we have included Figure 9 in our analysis as a reference to support our previous idea that although there are some differences in our monthly TEPU index shown in Figures 3 and 4, it is still true that the two graphs share many common points, such as the general trend of the economic policy uncertainty index over time. This means that while the choice of the scaling method may have some impact on the interpretation of the data, it does not fundamentally change the overall message conveyed by the graphs. Table 3 provides a comprehensive timeline of the most significant economic policy uncertainty events that happened in Lebanon in the past decade. Notably, our TEPU index effectively captures the impact of these events. For instance, the first indexed spike in our TEPU index (See Figure 8) corresponds to the resignations of the opposition parties from Hariri's unity government, which led to its collapse on 12 January 2011.[13] This event caused a significant period of economic policy uncertainty in Lebanon, which our TEPU index accurately reflects. Our TEPU index experienced another spike in February 2012 due to the warning issued by the International Monetary Fund (IMF) regarding the severe risk posed to Lebanon's economy by the internal political uncertainty in the country and the growing crisis in Syria.[14] As we move towards November 2015, we can observe a notable increase in the magnitude of the spike in our TEPU index. This spike was higher than before and can be attributed to the emergence of a significant Syrian refugee crisis in Lebanon. The arrival of over 1.1 million Syrian refugees to Lebanon strained its resources, causing a period of significant economic policy uncertainty, which our TEPU index accurately reflects.[15] Moving ahead to the year 2017, Lebanon faced several significant economic policy uncertainty events. In May 2017, our TEPU index recorded a spike due to the uncertainty surrounding the third trial of already postponed general elections in Lebanon. Additionally, ongoing election uncertainty also put oil and gas tenders in Lebanon at risk, contributing to the country's high level of economic policy uncertainty, which our TEPU index accurately captures.[16] At the end of 2017, specifically on 4 November 2017, our TEPU index experienced a spike in response to Saad Al Hariri's sudden announcement of his resignation as the Prime Minister of Lebanon from Saudi Arabia. This event caused an economic policy uncertainty shock in Lebanon. In September 2018, our TEPU index also showed a spike, which was mainly attributed to

---

[13] https://www.theguardian.com/world/2011/jan/12/hezbollah-quits-lebanon-unity-government

[14] https://www.elibrary.imf.org/view/journals/002/2012/040/article-A003-en.xml

[15] https://edition.cnn.com/2015/09/09/world/welcome-syrian-refugees-countries/index.html

[16] https://www.reuters.com/article/lebanon-economy-oil-idUKL8N1IB5BN

the harmful effects of political uncertainty on investor confidence.[17] This event coincided with warnings from authorities, bankers, and economists that Lebanon was facing a growing economic crisis.[18] The year 2019 was important for Lebanon, as reflected by a significant spike in our TEPU index. This spike was a result of the beginning of Lebanon's October 2019 revolution, which started on 17 October 2019. This revolution was marked by widespread protests in all regions of Lebanon, where people demanded the resignation of the political regime. The magnitude of this spike in our TEPU index highlights the significant impact of this event on economic policy uncertainty in Lebanon. In 2020, Lebanon faced a series of devastating economic and policy events, both internal and external, making it the most challenging year for the country. It all started with the confirmation of Lebanon's first case of COVID-19, as a 45-year-old lady returning from Iran tested positive for coronavirus and was placed in quarantine at Beirut's Rafik Hariri Hospital on 21 February 2020.[19] The Syrian refugee crisis reached its peak in Lebanon in February 2020, as the ongoing health and economic uncertainties left the refugees in a devastating condition.[20] On 9 March 2020, Lebanon defaulted on its foreign debt for the first time, as it failed to repay a $1.2 billion Eurobond, marking the country's first sovereign default.[21] Furthermore, the most devastating event that occurred in 2020 and resulted in the most significant spike in our index, as shown in Figure 8, was the

[17] The September 2018 spike, among others, is more prominently highlighted in Figure 9, as it corresponds to the alternative scaling method we adopted for our TEPU index.

[18] https://www.aljazeera.com/videos/2018/9/19/political-uncertainty-plunges-lebanon-into-economic-crisis

[19] https://www.reuters.com/article/uk-china-health-lebanon-minister-idUKKBN20F1XH

[20] https://www.aljazeera.com/economy/2020/2/23/syrian-refugees-in-lebanon-suffer-amid-economic-uncertainty

[21] https://www.economist.com/middle-east-and-africa/2020/03/12/for-the-first-time-lebanon-defaults-on-its-debts

catastrophic explosion at Beirut's Port on August 4, 2020. This explosion caused severe damage to a significant part of Lebanon's capital, leaving 218 deaths, 7,000 injuries, and 300,000 displaced.[22] Furthermore, in alignment with the same massive spike in our TEPU index, Hassan Diab announced his resignation as the prime minister of Lebanon on August 10, 2020, following the devastating blast in Beirut. After the resignation of Hassan Diab, Hariri was elected as prime minister-designate to form a Lebanese unity government. However, he was unable to form a government after eight months of trying and stepped down on 15 July 2021, resulting in a significant spike in Figure 8 (and the most significant spike in Figure 9). Finally, on 15 May 2022, the parliamentary elections were held. Independent candidates made a breakthrough, resulting in no single party or division gaining the majority in the Lebanese parliament. The exceptional economic policy uncertainty surrounding what this parliament can accomplish is reflected in a significant spike in our index as shown in Figure 8. Hence, our TEPU index clearly links to significant political and economic events that occurred in Lebanon throughout the sample period from 2011 to 2022 and track the evolution of economic policy uncertainty events in Lebanon.

[22] https://www.hrw.org/video-photos/interactive/2021/08/02/lebanon-evidence-implicates-officials-beirut-blast-targeted

CHAPTER 5

CONCLUSION

In this paper, we construct a weekly and monthly Twitter-based Economic Policy Uncertainty (TEPU) index for Lebanon, covering the period from January 1, 2011, till January 18, 2023. We have developed a unique and distinctive methodology that was specifically designed to overcome the challenges posed by the unavailability and lack of reliability of data in Lebanon. Our methodology enabled us to create a reliable and effective TEPU index for Lebanon. Our approach involved creating and employing Python scripts that interact with the official Twitter API to fetch and transform tweets into reliable and actionable data, which was crucial for generating a TEPU index for Lebanon. Our primary objective was to develop a TEPU index that best fits Lebanon and can effectively identify and uncover significant economic policy uncertainty events that have taken place in the country since 2011. Furthermore, we employed and test two scaling methods to construct our TEPU index, the first was proposed by Baker et al. (2021), and the second was also proposed by Baker et al. (2021) as a variant of their main index and adopted by Lee et al. (2023) as their primary scaling method. We also observed that our TEPU index significantly spikes during major political and economic events that occurred in Lebanon, effectively tracking the evolution of economic policy uncertainty. Finally, we believe that our development of a distinctive methodology to construct the TEPU index for Lebanon would be an excellent starting point for future research on tracking economic policy uncertainty events in Lebanon and the MENA region.

In recent years, Twitter has become a popular platform for both economic policy experts and non-experts to share their thoughts on financial and economic matters with a broad audience in real-time. However, as highlighted by Baker et al., (2021), it is important to note that Twitter data does not provide a perfect method to track economic policy uncertainty. Our index primarily reflects the views of individuals who favor digital platforms, mainly Twitter, which may not be representative of the wider population's perceptions of economic policy uncertainty in Lebanon. Moreover, it is crucial to keep in mind that the tweets we used in constructing our TEPU index may have some limitations. For instance, tweets generally tend to be shorter and more informal than newspaper articles, meaning valuable information could be omitted from tweets.

APPENDIX 1

FIGURES

Figure 1. An Outline of Our Distinctive Methodology for Developing the TEPU Index for Lebanon

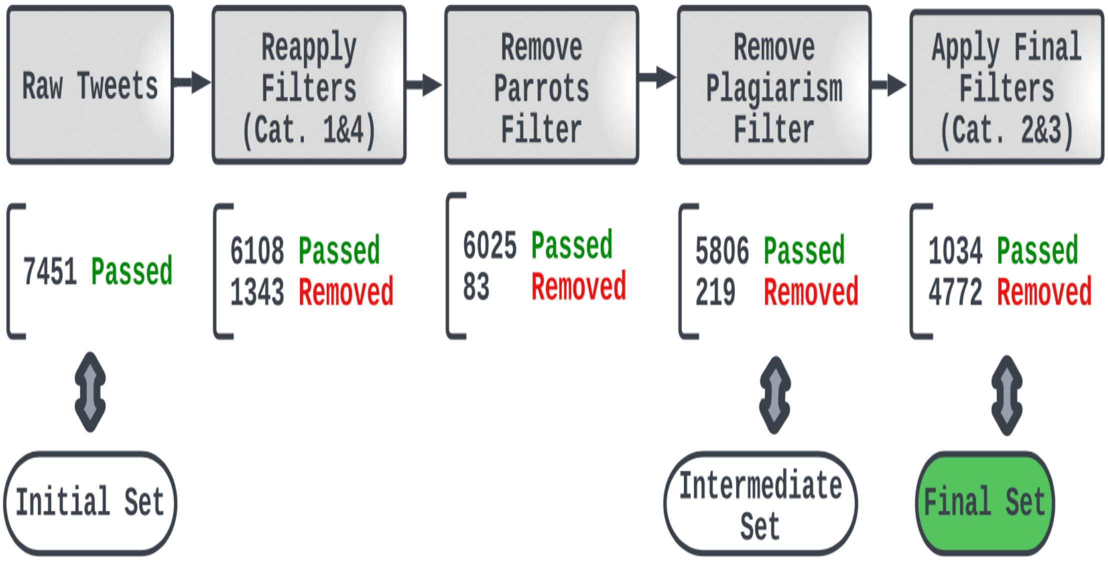


Figure 2. Monthly Evolution of the Total Number of Relevant Tweets

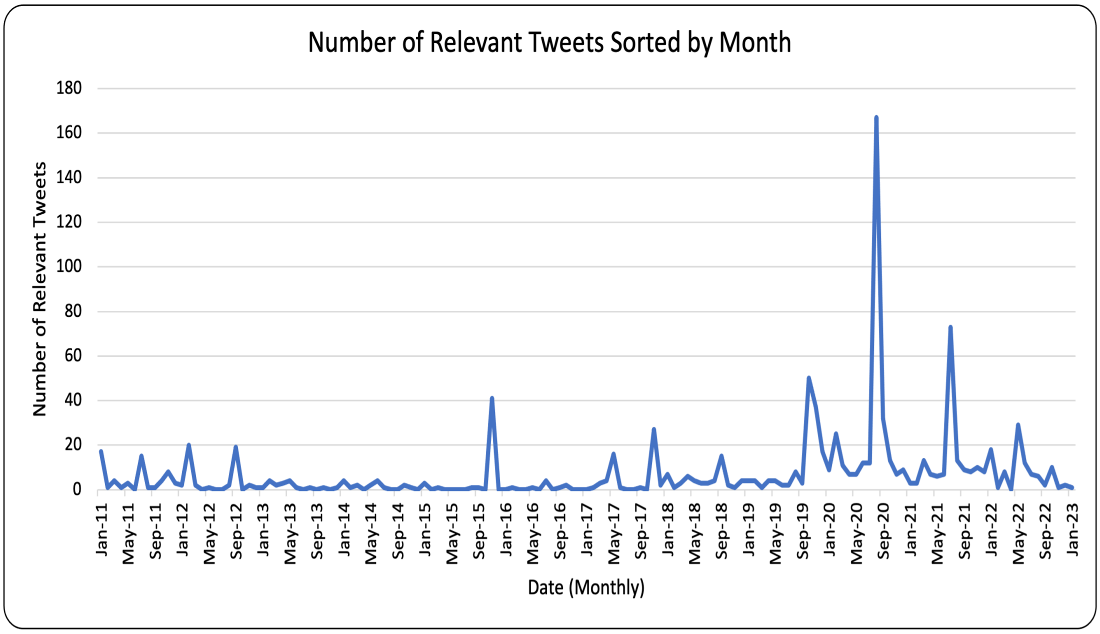


Figure 3. Monthly TEPU Index for Lebanon

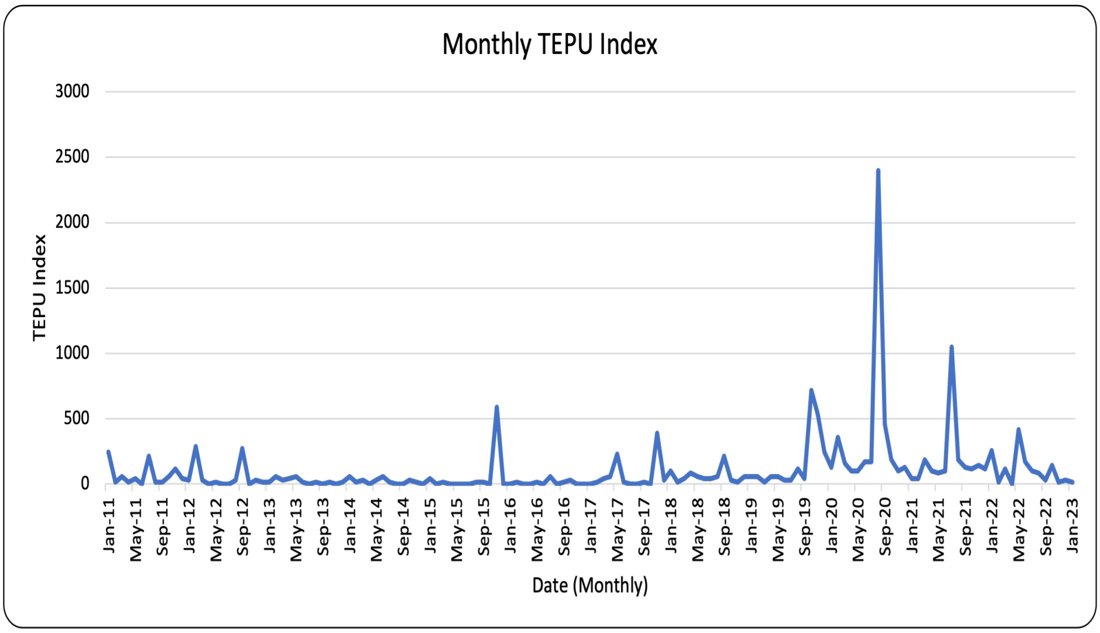


Figure 4. Monthly TEPU Index for Lebanon (Alternative Scaling Method)

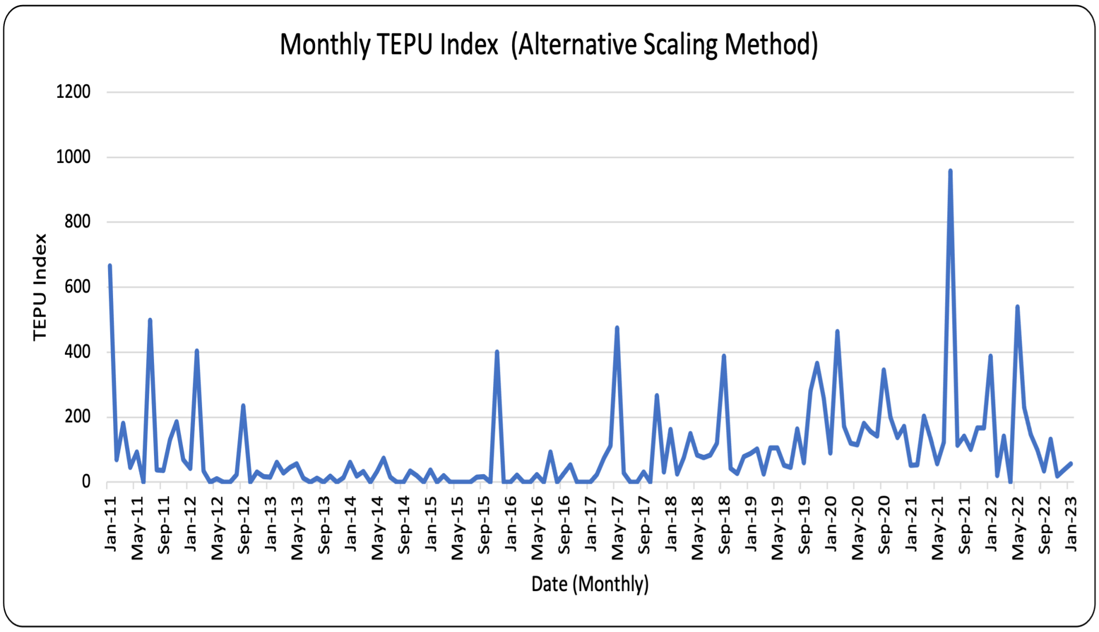


Figure 5. Weekly Evolution of the Total Number of Relevant Tweets

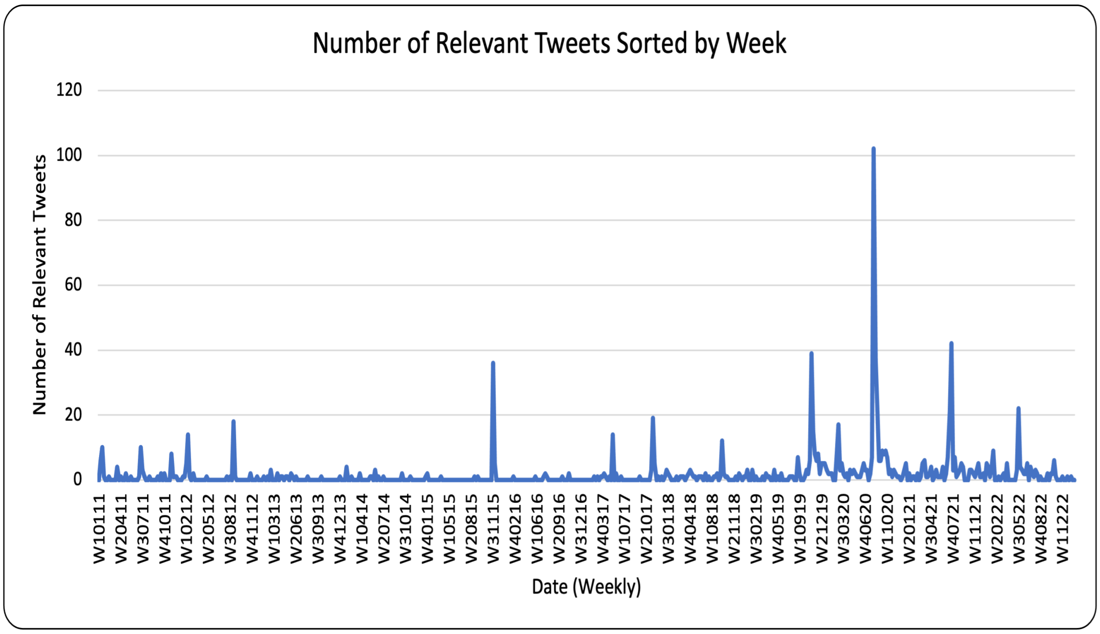


Figure 6. Weekly TEPU Index for Lebanon

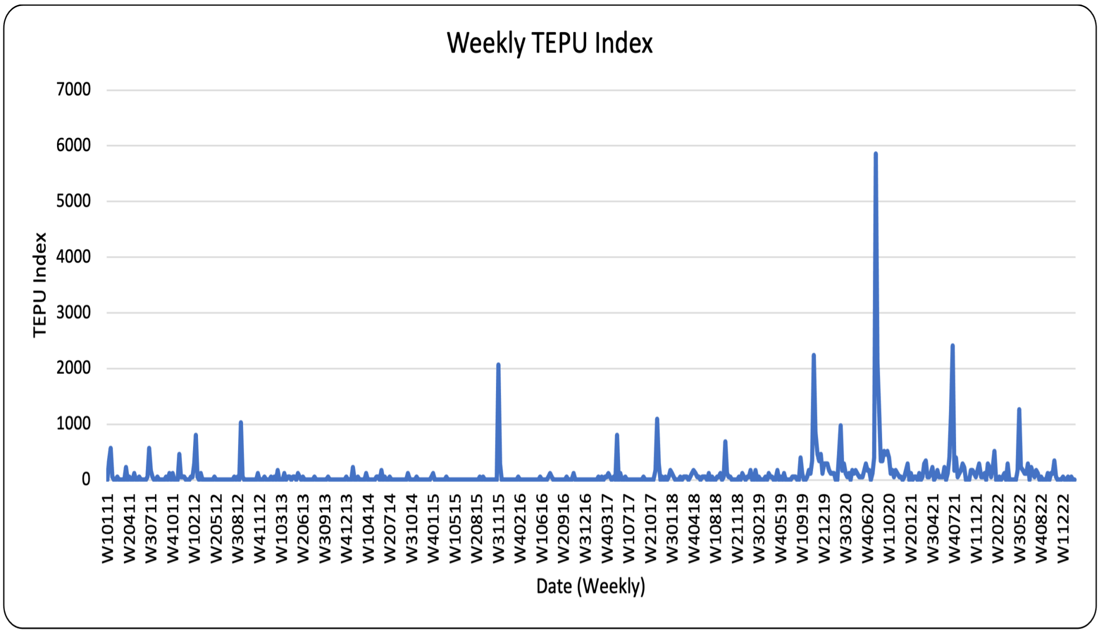


Figure 7. Weekly TEPU Index for Lebanon (Alternative Scaling Method)

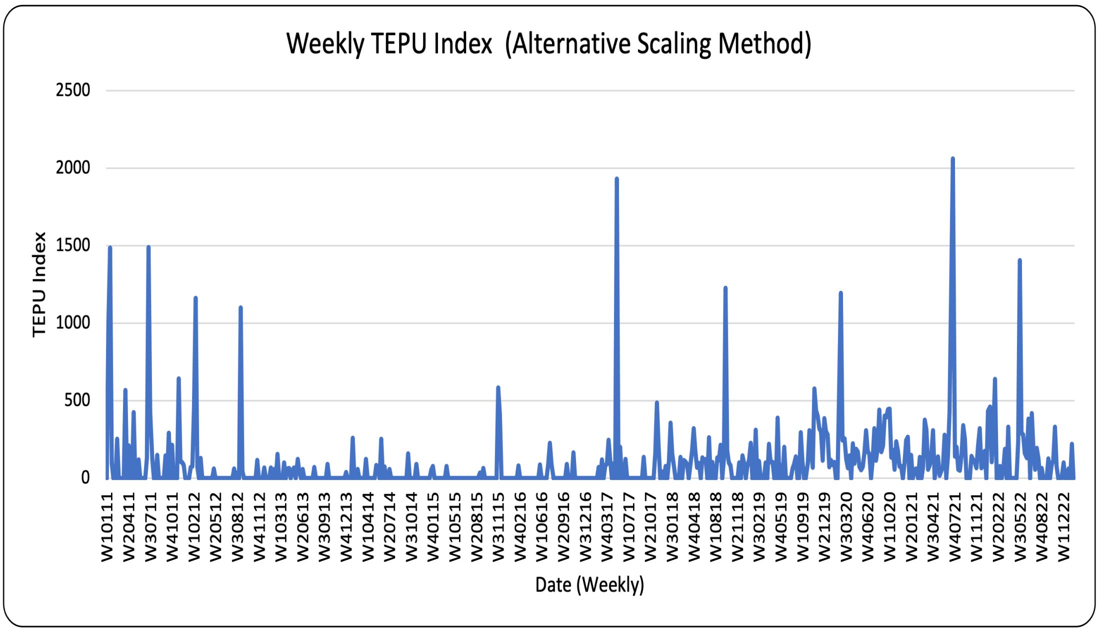


Figure 8. Monthly TEPU Index for Lebanon from January 2011 to January 2023 with Major Economic Policy Uncertainty Events

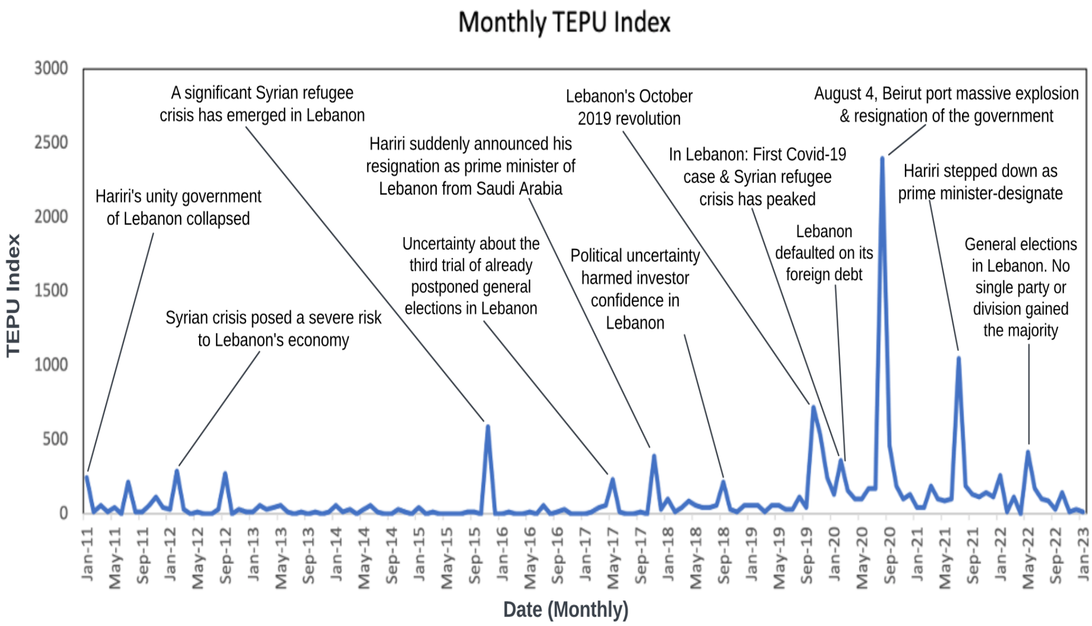
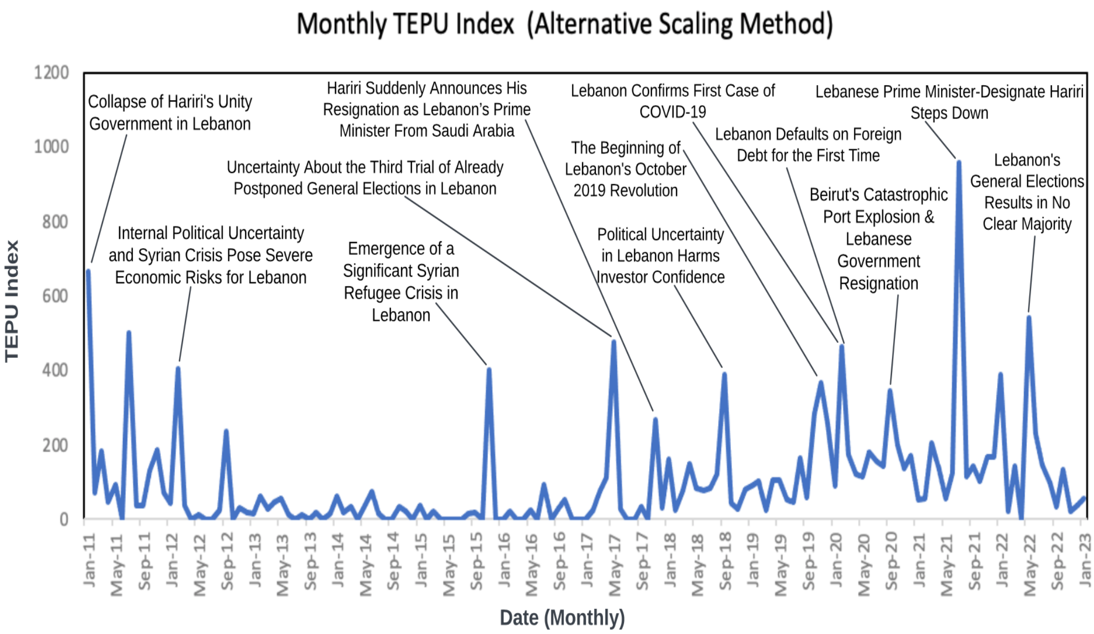


Figure 9. Monthly TEPU Index (Alternative Scaling Method) for Lebanon from January 2011 to January 2023 with Major Economic Policy Uncertainty Events



APPENDIX 2

TABLES

Table 1. List of Economic Policy Uncertainty Related Keywords

Category	Words
Lebanon	lebanon, lebanese
Economic	ammonium nitrate, airport, assistance, aid, banks, bank, banking, bankers, businesses, business, bankruptcy, billions, billion, bonds, brexit, bse, collapse, crash, company, currency, crisis, crises, commercial, cash, capital, cost, costs, creditors, developing, debt, diesel, demand, dollars, dollar, deposit, economic, economy, economical, economically, economiccrisis, energy, electricity, export, exchange, economist, funding, fund, funds, financialcrisis, forecasts, fresh, fuel, fuels, financial, farmers, firms, fiscal, finance, gas, gdp, goods, hardship, hyperinflation, interest, interests, investing, investment, invest, infrastructure, investors, imf, inflation, imports, import, industrialists, income, incomes, industry, industrial, job, lira, loss, loans, liquidity, maritime arena, markets, market, million, monetary, manufacturing, macroeconomy, ngo, outcome, outcomes, opportunity, oil, planning, plan, prices, price, petrol, port, property, poverty, pound, production, profit, real estate, risks, risk, resources, sustainable, shares, solidere, salary, savings, supply, startup, stock, stocks, shortage, supplies, supplied, sanctions, subsidies, stability, stagnation, trader, trade, supplies, supplied, sanctions, subsidies, tons, thousands, usd, unemployment, violations, volatility, work, warehouse.
Policy	army, allies, amal, authorities, authority, ambassador, ambassadors, august 4th, august 4, 4 august, administration, administer, aoun, assassination, blast, blasts, beirutexplosion, beirutblast, budget, banned, bdl, boom, covid, candidates, curfew, corruption, corrupt, corona, coronavirus, central bank, cabinet, deal, demarcation, defenses, defense, diab, diplomatic, drone, elections, election, enforce, electoral, explosions, explosion, eurobond, forces, fpm, government, governments, gov, governed, governance, geagea, geopolitical, hezbollah, hizbollah, hariri, humanrights, israel, independents, invasion, joumblatt, lockdown, lebanonexplosion, lebanonblast, leaders, leader, leadership, law, legal, lebanonprotests, mikati, minister, ministers, military, missiles, middle class, macron, management, opposition, occupied territories, official, officials, polls, poll, protests, protest, protestors, protestor, political, politics, power, powers, protecting, president, party, parliament, parliamentary, psp, politicians, policy, pandemic, police, pm, premier, refugees, refugee, religions, religion, rockets, reconstruction, resignation, resigns, resign, revolution, restrictions, regime, saudi, shocks, shock, sectarianism, soldiers, security, sociopolitical, telecom, terror, tax, taxes, vote, votes, voter, war, wars, weapons, withdrawal

Uncertainty	uncertain, uncertainly, uncertainties, uncertainty, unpredictable, unpredictably, unclear, unclearly

Table 2. Frequency of Keywords Related to Economic Policy Uncertainty in Our Final Set of Relevant Tweets

Keyword	**Freq.**	**Keyword**	**Freq.**	**Keyword**	**Freq.**
august 4	4	financialcrisis	1	pm	31
4 august	3	firms	10	police	6
administer	1	fiscal	7	policy	13
administration	5	forces	14	political	298
aid	29	forecasts	2	politicians	5
airport	11	fresh	14	politics	16
allies	11	fuel	17	polls	4
amal	3	fuels	14	port	77
ambassador	2	fund	7	pound	12
ammonium nitrate	24	funding	23	poverty	58
aoun	11	funds	3	power	20
army	14	gas	57	powers	2
assassination	2	gdp	8	premier	2
assistance	4	geagea	1	president	19
august 4th	1	geopolitical	4	price	8
authorities	6	goods	6	prices	12
authority	2	gov	8	production	4
bank	35	governance	3	profit	3
bankers	4	governed	1	property	20
banking	26	government	80	protecting	2
bankruptcy	4	governments	3	protest	11
banks	45	hardship	10	protestors	4
bdl	6	hariri	69	protests	45
beirutblast	17	hezbollah	78	real estate	8
beirutexplosion	12	hizbollah	3	refugee	17
billion	6	humanrights	1	refugees	96
billions	4	hyperinflation	2	regime	8
blast	46	imf	48	religion	2
blasts	8	import	5	religions	1
bonds	3	imports	2	resign	2
boom	4	income	3	resignation	22

brexit	8	incomes	1	resigns	8
budget	4	industrial	3	resources	5
business	64	industrialists	6	restrictions	3
businesses	9	industry	17	revolution	3
cabinet	9	inflation	12	risk	68
candidates	1	infrastructure	7	risks	18
capital	89	interest	16	rockets	1
cash	8	interests	8	salary	5
central bank	13	invasion	2	sanctions	44
commercial	7	invest	2	saudi	14
company	4	investing	2	savings	4
corona	3	investment	15	sectarianism	2
coronavirus	4	investors	7	security	26
corrupt	1	israel	32	shares	6
corruption	12	job	17	shock	2
cost	9	law	9	shocks	1
costs	5	leader	4	shortage	6
covid	23	leaders	6	sociopolitical	2
crash	13	leadership	10	soldiers	2
creditors	2	lebanese	1340	solidere	1
crises	23	lebanon	5583	stability	24
crisis	181	lebanonblast	1	stagnation	5
currency	20	lebanonexplosion	5	startup	1
deal	26	lebanonprotests	6	stock	7
debt	12	legal	5	stocks	6
defense	1	liquidity	3	subsidies	6
defenses	1	lira	19	supplied	2
demand	8	loans	3	supplies	8
demarcation	3	lockdown	4	supply	11
deposit	3	loss	14	sustainable	4
depot	11	macroeconomy	1	tax	7
developing	5	macron	5	taxes	3
diab	6	management	1	telecom	1
diesel	2	manufacturing	1	terror	2
diplomatic	3	market	72	thousands	36
dollar	12	markets	12	tons	16
dollars	7	middle class	1	trade	8
drone	2	mikati	4	trader	10
economic	294	military	9	uncertain	1172
economical	5	million	20	uncertainly	3

economically	1	minister	85	uncertainties	160
economiccrisis	2	ministers	6	uncertainty	2765
economist	6	monetary	5	unclear	1583
economy	91	ngo	5	unemployment	6
election	20	official	15	unpredictable	319
elections	14	officials	19	unpredictably	6
electoral	5	oil	36	usd	7
electricity	22	opportunity	11	violations	2
energy	14	outcome	31	volatility	6
eurobond	5	outcomes	3	voter	1
exchange	20	pandemic	11	votes	22
explosion	128	parliament	2	war	22
explosions	14	parliamentary	3	warehouse	19
export	3	party	1	weapons	8
farmers	3	petrol	1	withdrawal	6
finance	8	plan	30	work	43
financial	61	planning	46		

Table 3. Timeline of Events in Lebanon

Year	Date	Event
2011	12 January 2011	Opposition Parties Resignations Lead to Collapse of Hariri's Unity Government in Lebanon
2012	February 2012	IMF Warns of Severe Economic Risks for Lebanon Amid Internal Political Uncertainty and Syrian Crisis
2013	22 March 2013	Lebanese Government Collapses Following Resignation of Prime Minister Najib Mikati
2015	November 2015	Over 1.1 Million Syrian Refugees Strain Lebanon's Resources, Creating Significant Crisis
2017	May 2017	Uncertainty About the Third Trial of Already Postponed General Elections in Lebanon Oil and Gas Tenders in Lebanon at Risk Due to Ongoing Election Uncertainty
	16 June 2017	Lebanese Parliament Extended Its Own Term by 11 Months. Under Various Pretexts, Elections Were Postponed Three times: 2013, 2014, and 2017

	4 November 2017	Saad Al Hariri Suddenly Announces His Resignation as Lebanon's Prime Minister From Saudi Arabia
2018	6 May 2018	Lebanon's General Elections Were Held
	September 2018	Political Uncertainty in Lebanon Harms Investor Confidence, Fueling Warnings of Growing Economic Crisis from Bankers and Economists
2019	17 October 2019	The Beginning of Lebanon's October 2019 Revolution
2020	21 February 2020	Lebanon Confirms First Case of COVID-19: Woman Returning from Iran Tests Positive and Placed in Quarantine at Beirut's Rafik Hariri Hospital
	February 2020	Reaching the Peak of the Syrian Refugee Crisis in Lebanon Amidst Current Health and Economic Challenges
	9 March 2020	Lebanon Defaults on Foreign Debt for the First Time, Failing to Repay $1.2 Billion Eurobond and Marking the First Sovereign Default
	4 August 2020	Beirut's Port Explosion: Catastrophic Blast Damages Large Parts of Lebanon's Capital, Leaving 218 Dead, 7,000 Injured, and 300,000 Displaced
	10 August 2020	Lebanese Prime Minister Hassan Diab Resigns Following Beirut Port Explosion Tragedy
2021	15 July 2021	Lebanese Prime Minister-Designate Saad Al Hariri Steps Down After 8 Months of Failed Attempts to Form Unity Government in Lebanon
2022	15 May 2022	Lebanon's General Elections Result in Independent Candidates' Breakthrough, Creating Uncertainty as No Single Party Gains Majority in Parliament

Source: Authors' elaboration.

REFERENCES

Abel, A. B. (1983). Optimal Investment Under Uncertainty. *American Economic Review*, *73*(1), 228–233. https://ideas.repec.org/a/aea/aecrev/v73y1983i1p228-33.html

Ahir, H., Bloom, N., & Furceri, D. (2018). The World Uncertainty Index. *SSRN Electronic Journal*. https://doi.org/10.2139/SSRN.3275033

Altug, S., Demers, F., & Demers, M. (2009). The Investment Tax Credit and Irreversible Investment. *Journal of Macroeconomics*, *31*(4), 509–522. https://EconPapers.repec.org/RePEc:eee:jmacro:v:31:y:2009:i:4:p:509-522

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

Baker, S. R., Bloom, N., Davis, S. J., & Renault, T. (2021). *Twitter-Derived Measures of Economic Uncertainty*. https://blog.twitter.com/developer/en_us/topics/tools/2021/enabling-the-future-of-academic-research-with-the-

Becerra, J. S., & Sagner, A. (2020). *Twitter-Based Economic Policy Uncertainty Index for Chile*.

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

Bloom, N. (2014). Fluctuations in Uncertainty. *Journal of Economic Perspectives*, *28*, 153–167. https://doi.org/10.1257/jep.28.2.153

Dixit, A. (1989). Entry and Exit Decisions under Uncertainty. *Journal of Political Economy*, *97*(3), 620–638. https://doi.org/https://doi.org/10.1086/261619

Dixit, A., & Pindyck, R. (1994). *Investment Under Uncertainty*. Princeton University Press.

Eberly, J. (1994). Adjustment of Consumers' Durables Stocks: Evidence from Automobile Purchases. *Journal of Political Economy*, *102*(3), 403–436. https://doi.org/10.1086/261940

Hartman, R. (1972). The Effects of Price and Cost Uncertainty on Investment. *Journal of Economic Theory*, *5*(2), 258–266. https://doi.org/10.1007/s10158-003-0022-4

Hassett, K., & Metcalf, G. (1999). Investment with Uncertain Tax Policy: Does Random Tax Policy Discourage Investment? *Economic Journal*, *109*(457), 372–393. https://EconPapers.repec.org/RePEc:ecj:econjl:v:109:y:1999:i:457:p:372-93

International Monetary Fund (IMF), Lebanon. (2019). Retrieved from https://www.imf.org/en/Countries/LBN

Jirasavetakul, L.-B. F., & Spilimbergo, A. (2018). *Economic Policy Uncertainty in Turkey*.

Knight, F. H. (1921). *Risk, Uncertainty and Profit*. https://papers.ssrn.com/abstract=1496192

Lee, K., Choi, E., & Kim, M. (2023). Twitter-based Chinese economic policy uncertainty. *Finance Research Letters*, *53*(103627). https://doi.org/https://doi.org/10.1016/1.1rl.2023.103627

Liu, B. (2015). Sentiment analysis: Mining opinions, sentiments, and emotions. *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*, 1–367. https://doi.org/10.1017/CBO9781139084789

Paddock, J. L., Siegel, D. R., & Smith, J. L. (1988). Option Valuation of Claims on Real Assets: The Case of Offshore Petroleum Leases. *The Quarterly Journal of Economics*, *103*(3), 479–508. https://doi.org/10.2307/1885541

Pindyck, R. (1988). Irreversible Investment, Capacity Choice, and the Value of the Firm. *American Economic Review*, *78*(5), 969–985. https://EconPapers.repec.org/RePEc:aea:aecrev:v:78:y:1988:i:5:p:969-85

Rodrik, D. (1991). Policy uncertainty and private investment in developing countries. *Journal of Development Economics*, *36*(2), 229–242. https://doi.org/10.1007/s00792-002-0304-5

Schaal, E. (2017). Uncertainty and Unemployment. *Econometrica*, *85*(6), 1675–1721. https://doi.org/10.3982/ECTA10557

Valletta, R. G., Bengali, L., & Van Der List, C. (2018). *Cyclical and Market Determinants of Involuntary Part-Time Employment*. https://doi.org/10.24148/wp2015-19

Yeşiltaş, S., Şen, A., Arslan, B., & Altuğ, S. (2022). A Twitter-Based Economic Policy Uncertainty Index: Expert Opinion and Financial Market Dynamics in an Emerging Market Economy. *Frontiers in Physics*, *10*. https://doi.org/10.3389/fphy.2022.864207

Zhang, L., Wang, S., & Liu, B. (2018). *Deep Learning for Sentiment Analysis: A Survey*. https://arxiv.org/pdf/1801.07883.pdf