

South Korea EPU Memo

Jessica Yu Kyung Koh

Current version: Friday 16th September, 2016

1 Overview

The purpose of this memo is to document how we construct South Korea Economic Policy Uncertainty (EPU) Index, which is based on the monthly counts of news articles that are likely to contain information about the economic policy uncertainty in South Korea.¹ In this document, we present (i) the term set used in our search, (ii) newspapers used, (iii) web-scraping strategy (iv) methods used in developing the index, (v) results, and (vi) future steps.

2 EPU Term Set

We select the combination of terms that show the economic policy uncertainty in South Korea. The selected terms are divided into the following three subsets that show uncertainty, economy, and policy, respectively:

1. **Uncertainty-relevant terms:** “uncertain,” “uncertainty”
2. **Economy-relevant terms:** “economic,” “economy,” “commerce”
3. **Policy-relevant terms:** “government,” “Blue House,” “congress,” “authorities,” “legislation,” “tax,” “regulation,” “Bank of Korea,” “central bank,” “deficit,” “WTO,” “law/bill,” “ministry of finance”

¹For more information about the EPU, visit <http://www.policyuncertainty.com>.

Our assumption is that an article containing at least one term from each of the above subsets implies that economic policy uncertainty existed during the time when the article was published. In order to construct the index, we count the number of newspaper articles containing at least one term from each of the above subsets. We perform all searches in Korean, as shown in Table 1.

Table 1: Term Sets for South Korean EPU Index, with Translations to English

Category	English Terms	Korean Terms	In Korean Characters
U	uncertainty OR uncertain	bulhwaksilsung OR bulhwaksil	불확실성 OR 불확실
E	economic OR economy commerce	gyeongje OR gyeongjeui sangup OR muyeok	경제 OR 경제의 상업 OR 무역
P	government “Blue House” congress authorities legislation tax regulation “Bank of Korea” “central bank” deficit WTO law/bill “ministry of finance”	jeongbu Chungwade gukhoe dangguk jejeong OR jejeongbu OR ibbub se gyuje OR tongje OR gyejeong Hankukeunhaeng OR Haneun jungangeunhaeng jukja OR bujok WTO OR Segye muyeok gigu bub OR buban gihwaekjaejungbu OR gijaebu	정부 청와대 국회 당국 제정 OR 제정법 OR 입법 세 규제 OR 통제 OR 규정 한국은행 OR 한은 중앙은행 적자 OR 부족 WTO OR 세계 무역 기구 법 OR 법안 기획재정부 OR 기재부

3 South Korean Newspaper Sources

3.1 Major Newspapers

Table 2 shows information on major newspapers in South Korea. If articles of a newspaper are available in multiple sources, we list all of them in the “Archive” column and show the time period available in each archive. Further information about archives are described in the next section.

For constructing South Korean EPU index, we currently use six newspapers: Donga Ilbo, Kyunghyang, Maeil Economic, Hankyoreh, Hankook Ilbo, and Korea Economic Daily. We do not use Chosun Ilbo and Chungang Ilbo at this stage, as the boolean operators do not function well in the online search engines for those newspapers. Once the websites are fixed, we aim to use articles from these two sources in the future.

Table 2: Major Newspapers in South Korea

Newspaper Name	Political Orientation	Type	Archive	Date Coverage	Boolean Operators
Chosun Ilbo	Conservative	Daily News	Chosun Archive	1920-Present	No
Chungang Ilbo	Conservative	Daily News	Chungang Website	2000-Present	No
Donga Ilbo	Conservative	Daily News	1. Donga Archive 2. Naver News Library	1920-Present 1920-1999	Yes Yes
Hankook Ilbo	Moderate	Daily News	Big Kinds	2000-Present	Yes
Kyunghyang	Liberal	Daily News	1. Naver News Library 2. Big Kinds	1946-1999 1990-Present	Yes Yes
Hankyoreh	Liberal	Daily News	1. Naver News Library 2. Big Kinds	1988-1999 1990-Present	Yes Yes
Maeil Economic	Conservative	Economic	1. Naver News Library 2. Big Kinds	1966-1999 1995-Present	Yes Yes
Korea Economic	Conservative	Economic	Big Kinds	1995-Present	Yes

3.2 Newspaper Search Engines

Articles from some South Korean newspapers we use are available in comprehensive news archives. There are two major online news archives (Big Kinds and Naver News Library) that collect historical and current articles from various newspaper sources. These archives are the major sources from which we scrape our data, and we utilize these archives as following:

1. **Big Kinds:** contains articles that have been published by 36 national and local news sources from 1990 to the current date. We use 5 major newspapers from this archive, which are Kyunghyang News, Maeil Economics, Hankyoreh, Hankook Ilbo, and Korea Economic Daily.
2. **Naver News Library:** contains news articles from 4 different sources - Donga Ilbo, Kyunghyang News, Maeil Economics, and Hankyoreh - from 1920s to 1999. Since years of the first publication are different for all 4 newspapers, starting dates of the availability of data are also all different. Donga Ilbo goes furtherest back as this newspaper started publishing from April 1920. Other newspapers started publishing in later years.

Since the post-2000 articles of Donga Ilbo are not available in the above archives, we extract the EPU counts of those articles from Donga Ilbo's own online search engine. Moreover, although the pre-2000 articles of Donga Ilbo are stored in Naver News Library, we use Donga's own archive as it contains more articles over that time period. Except for Donga Ilbo, EPU counts of all other newspapers are web-scraped from either Big Kinds or Naver News Library.

4 Web-Scraping Strategy

The first step in constructing the index is to scrape data from news archives. We scrape the monthly counts of EPU articles and the monthly counts of total articles for each newspaper

using Perl from search engines mentioned in the previous section. Since the HTML codes that are used to construct the websites are different across websites, we use different Perl scripts to scrape the data from different archives. All the scripts used for web-scraping are stored in the South Korean EPU repository.

5 Constructing the Index

5.1 Plotting Counts

Figures 1 - 6 show two sets of information: (1) log of monthly total article counts (the left vertical axis) and (2) ratio of monthly raw EPU article counts over monthly total article counts for each newspaper (the right vertical axis). If article counts of a newspaper are extracted from two archive sources, information of both sources are plotted on the same chart. Log of monthly total article counts help us track the months when a newspaper has abnormally low number of articles.

5.1.1 Donga Ilbo

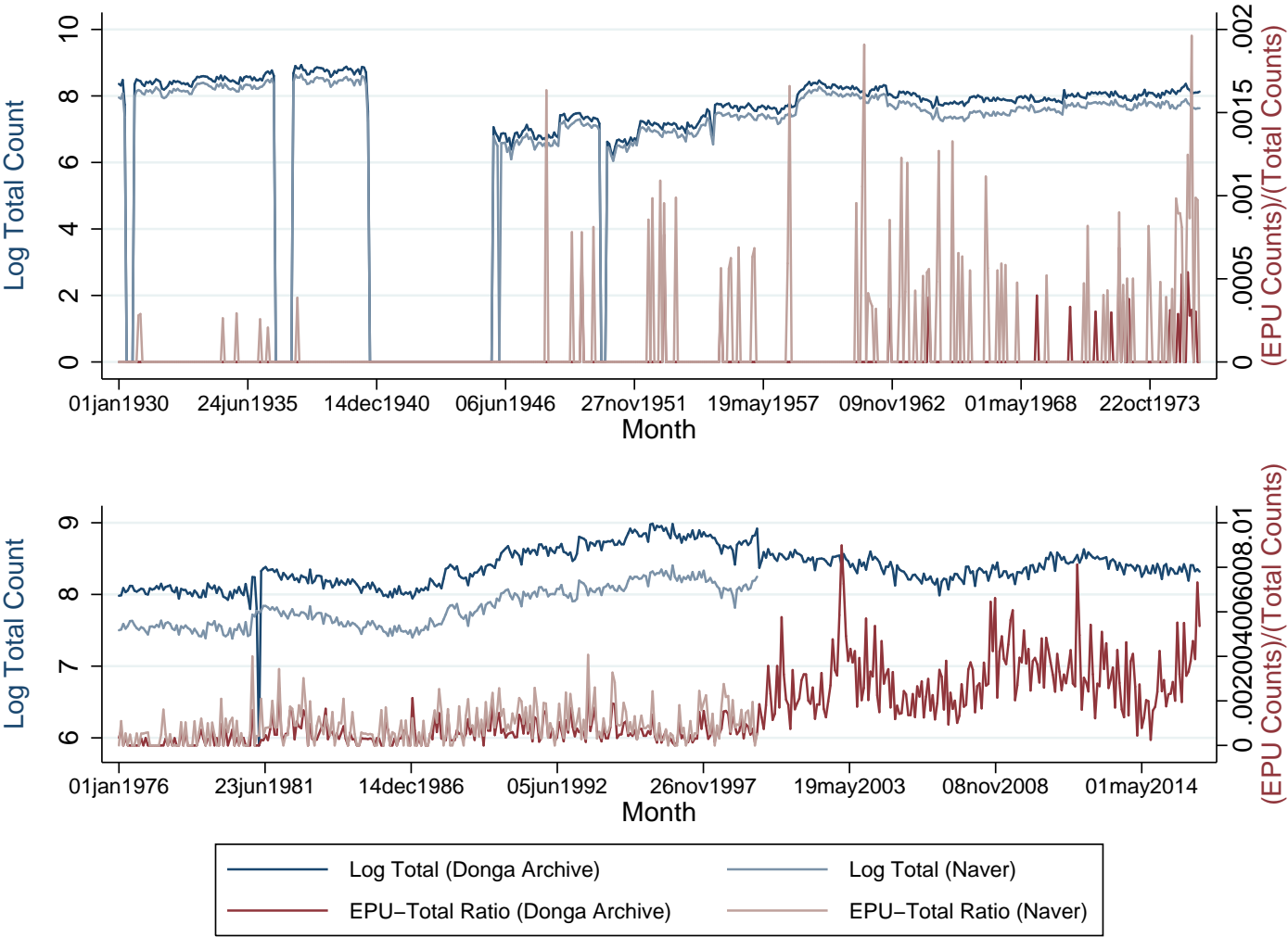
Figure 1 plots information regarding Donga Ilbo, which goes back furtherest in date among the newspapers we currently use for South Korea. The figure shows some abnormalities regarding the log total counts.

The first two sets of abnormally low total counts are due to forced suspension of publication. From October 1920 to January 1921, Donga Ilbo was forced to stop publishing when it published articles criticizing Japan's forced ruling of Korea. Similarly, Donga Ilbo temporarily suspended publishing from May 1930 and August 1930 due to the censorship.

The longer period of abnormally low counts of articles is shown in the 1940s. This is because Donga Ilbo discontinued publishing from the period from August 1940 to December 1945 due to the pressures from the Japanese General Government, which continued to colonialize Korea during those periods.

The figure also shows that the ratio of the EPU articles over total articles significantly differ across archives before the 1980s. This is because the Korean news articles contained a big proportion of Chinese characters in the text. This gradually changed, and articles published after the 1980s mainly contain Korean characters. Naver News Library translates the Chinese characters to Korean characters and allow searches with Korean keywords even for the articles published before the 1980s. However, Donga Archive does not translate Chinese characters, and the search performed in Korean does not capture articles that are mainly written in Chinese characters. This is why the ratio of EPU over total is much higher for Naver News Library than Donga Archive for articles published before the 1980s.

Figure 1: Monthly Information on Raw EPU and Total Counts (Donga Ilbo)



5.1.2 Kyunghyang News

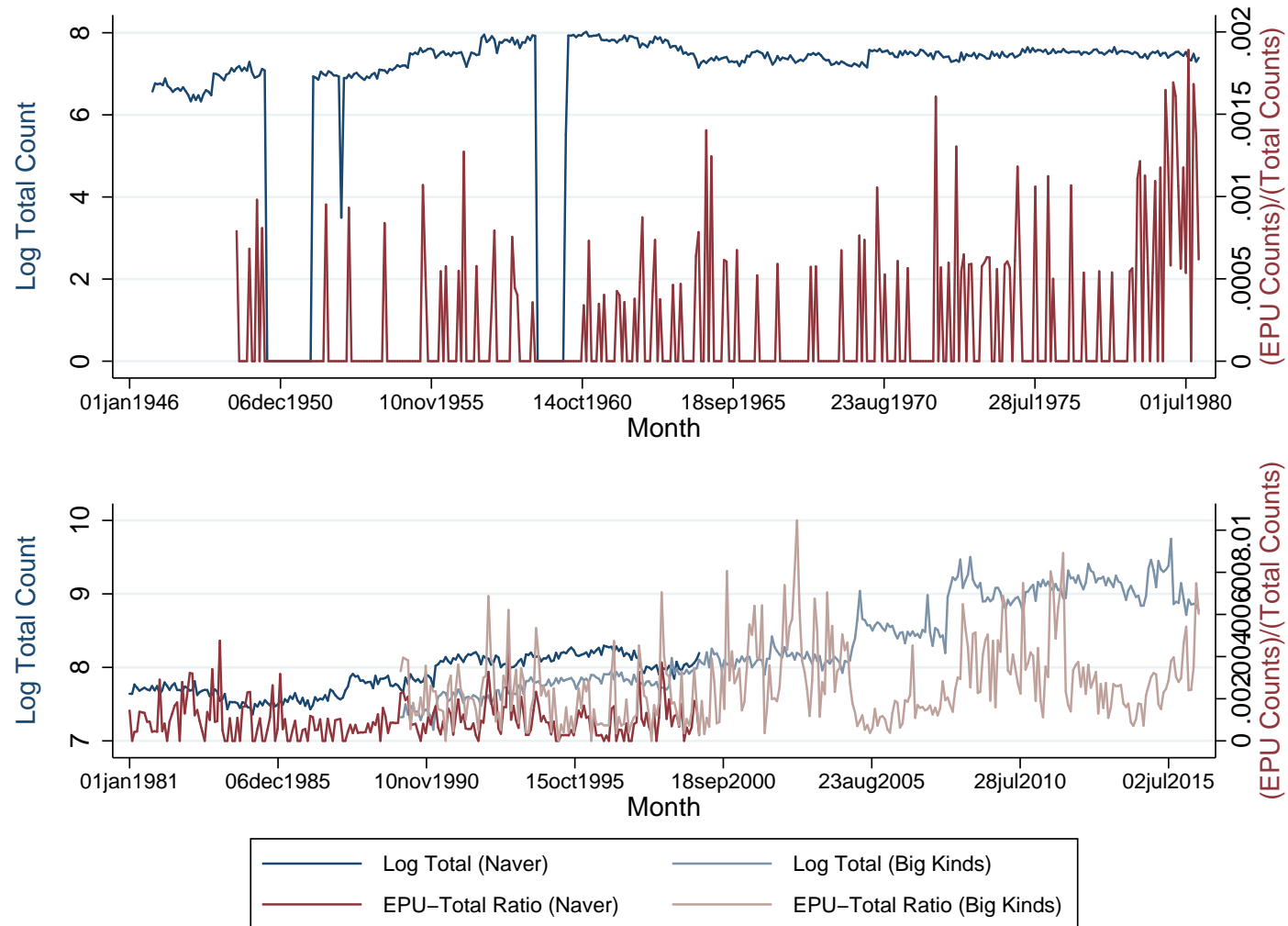
Figure 2 plots information regarding Kyunghyang News, which goes back to 1946. There are a few abnormalities regarding the log total counts of the data before the 1970s due to historical events.

The first abnormal months is the period between July 1950 and December 1951. According to the written history of Kyunghyang News, there is no evidence that Kyunghyang News discontinued publishing during this period. However, since this period was when the Korean War occurred, it is possible that Kyunghyang News might not have published articles during this period.

The second abnormal month is December 1952. This is due to the attacks made by the right-wing gang when Kyunghyang News published articles criticizing the the selected amendment bill to the Constitution of South Korea.

The last set of abnormal months, which is the period between May 1959 to March 1960, occurred because the South Korean government forced Kyunghyang News to stop publishing. This was due to the articles published by Kyunghyang that criticized the fraudulent election that allowed for a prolonged one-man rule by President Seung-man Lee.

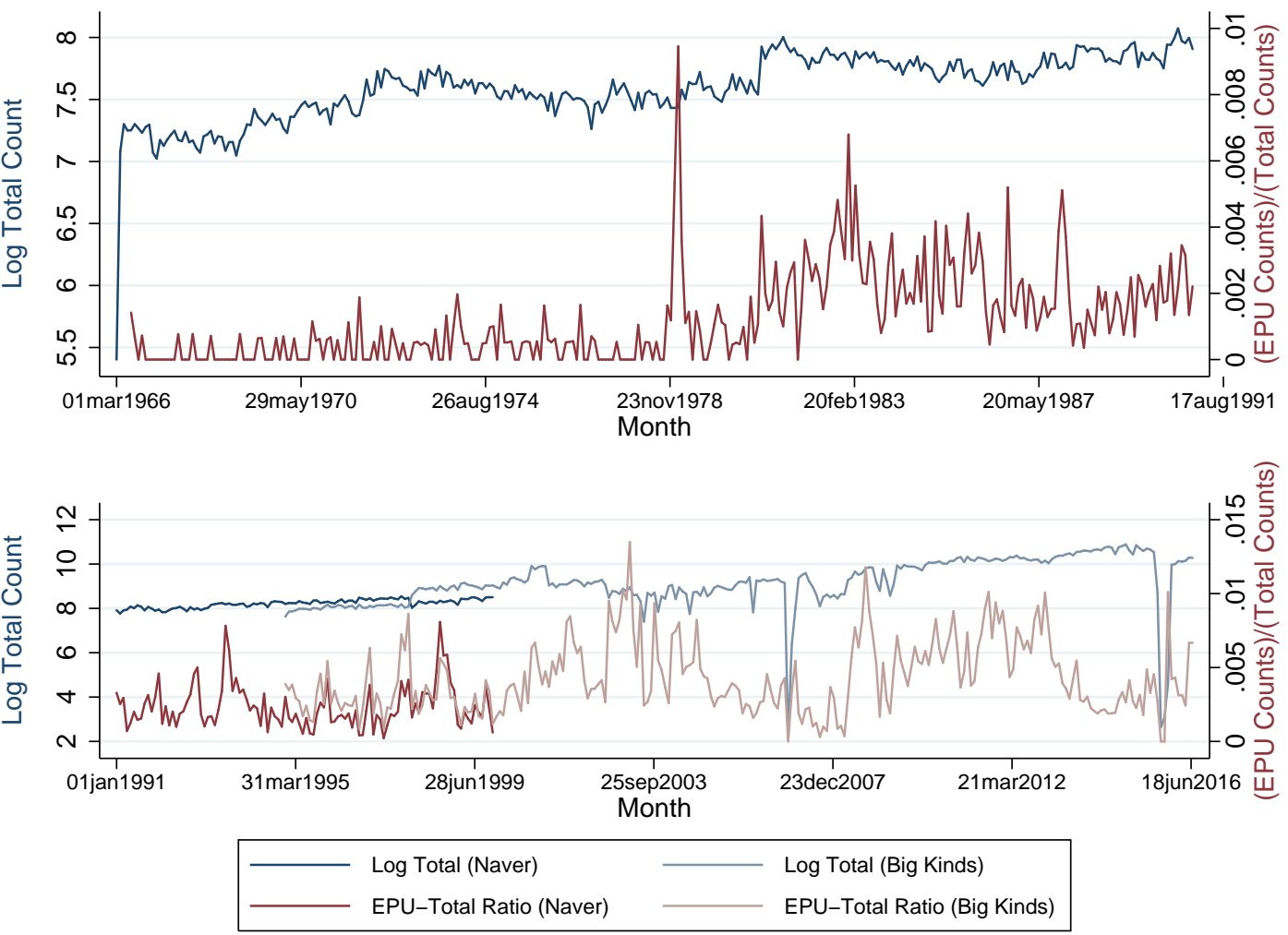
Figure 2: Monthly Information on Raw EPU and Total Counts (Kyunghyang)



5.1.3 Mael Economic Daily

Figure 3 plots information regarding Mael Economic Daily, which goes back to 1966. There are no abnormal log counts shown before the 2000s except for the beginning of the publication month. However, there are some abnormally low total counts after the 2000s. However, according to the current evidence, there has not been any suspension of the publication for Mael Economic Daily after the 2000s. We are currently investigating if this is a problem with the Big Kinds archive.

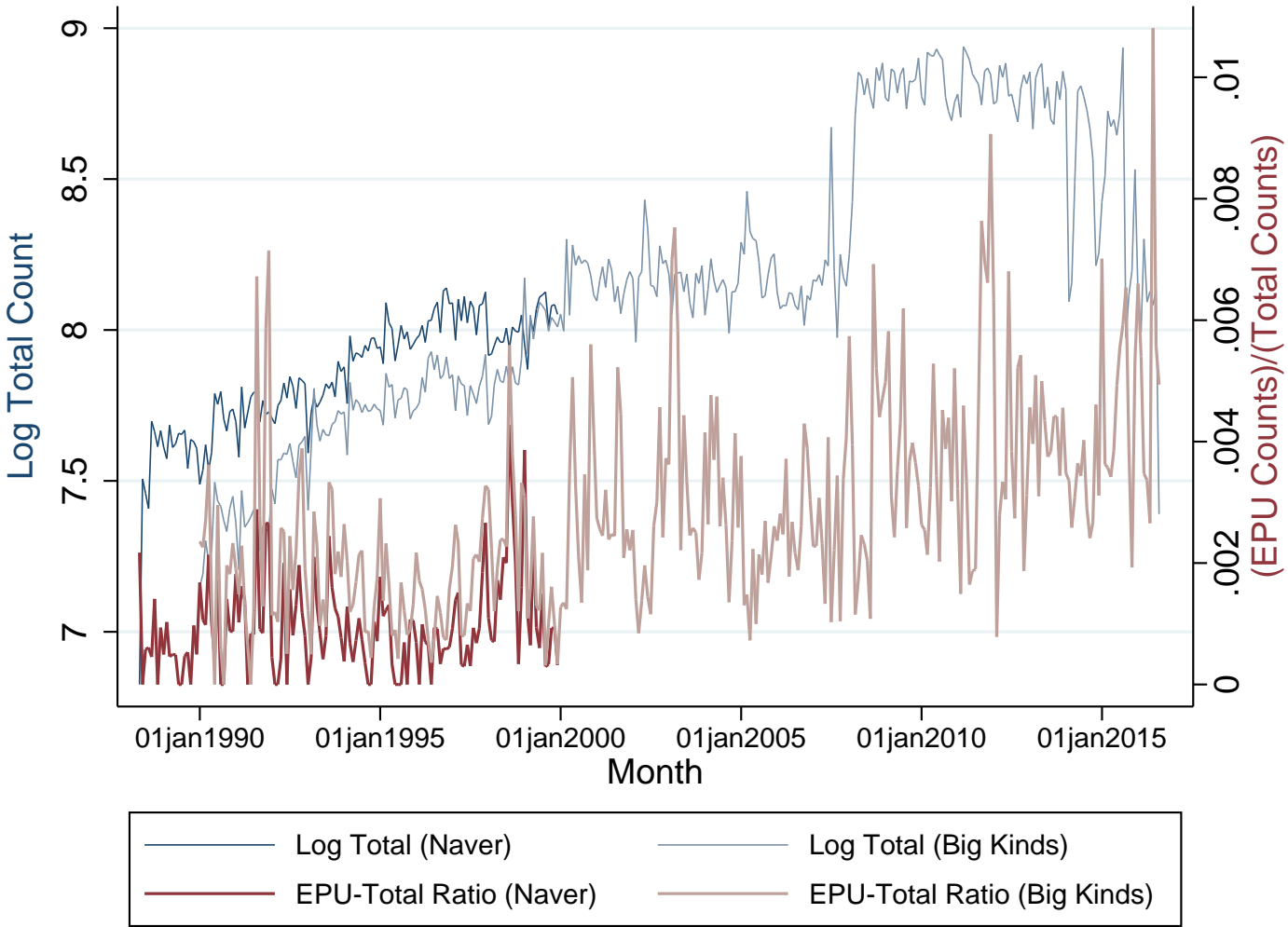
Figure 3: Monthly Information on Raw EPU and Total Counts (Maeil Economic)



5.1.4 Hankyoreh

Figure 4 plots information regarding Hankyoreh, which goes back to 1988. According to the figure, there is no noticeable abnormal log total counts for this newspaper.

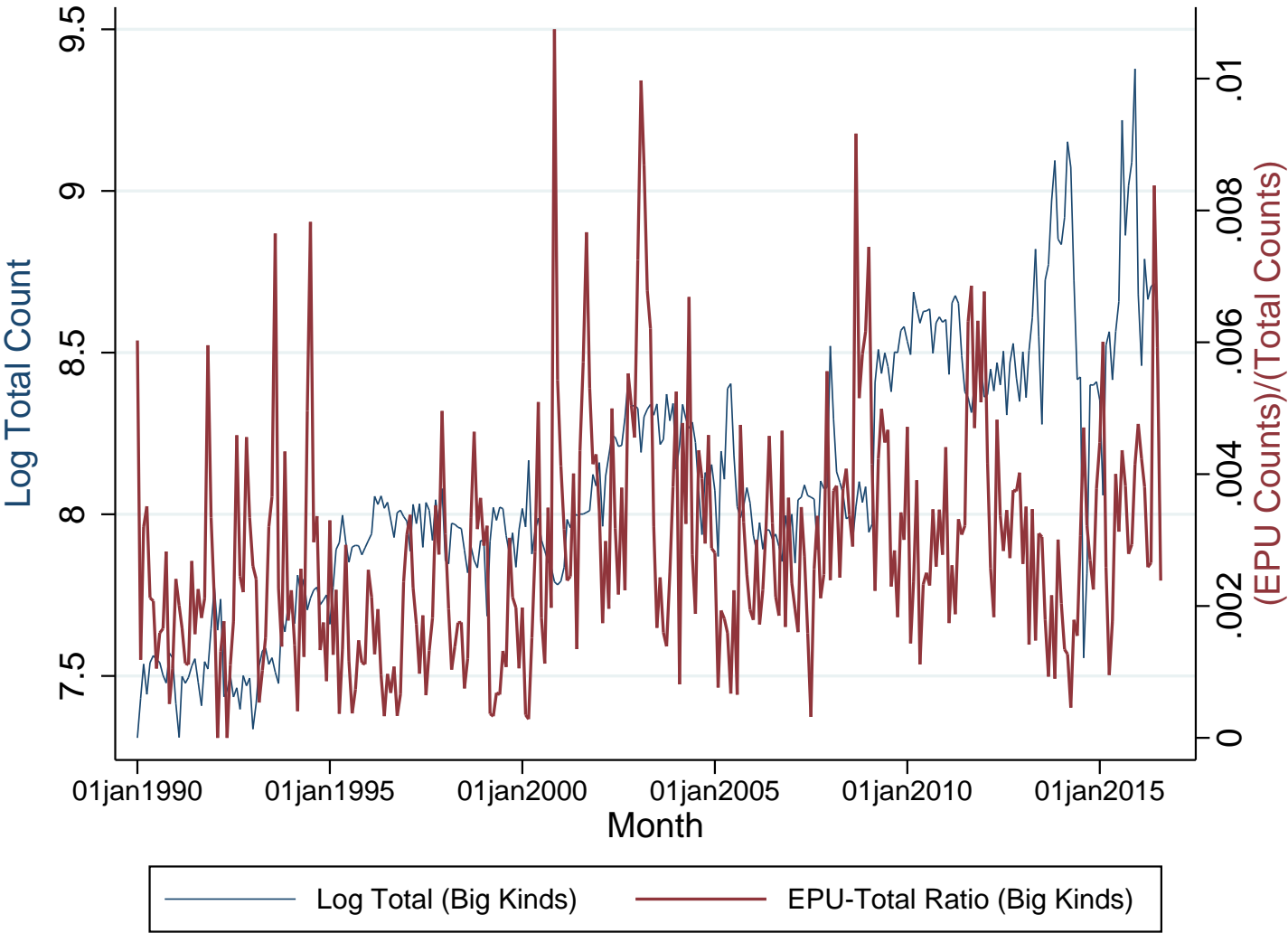
Figure 4: Monthly Information on Raw EPU and Total Counts (Hankyoreh)



5.1.5 Hankook Ilbo

Figure 5 plots information regarding Hankook Ilbo. According to the figure, there is no strong evidence of suspension of the publication for this newspaper.

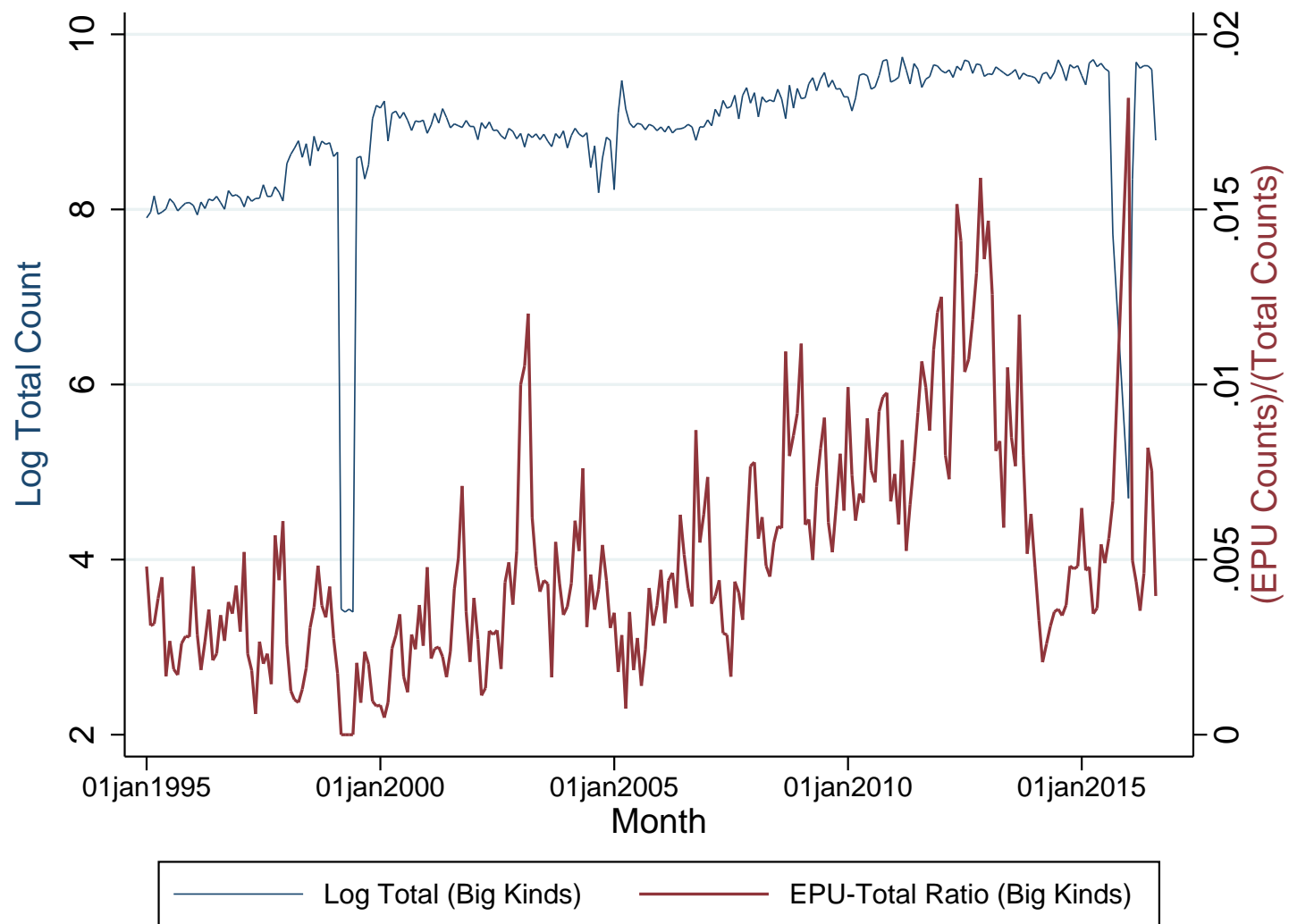
Figure 5: Monthly Information on Raw EPU and Total Counts (Hankook Ilbo)



5.1.6 Korea Economic Daily

Figure 6 plots information regarding Korea Economic Daily. There are some abnormally low total counts shown in the figure. However, according to the current evidence, there has not been any suspension of the publication for Korea Economic Daily during this period. We are currently investigating if this is a problem with the Big Kinds archive.

Figure 6: Monthly Information on Raw EPU and Total Counts (Korea Economic)



5.2 Identification of Abnormal Months

We use the following rule to identify abnormal months for each newspaper in the post-1990 period: Let $N(t)$ be the total newspaper article count in month t , and let $\bar{N}(t-15, t-4)$ denote the average monthly article in the same paper from 15 months earlier to 4 months earlier. If $\frac{N(t)}{\bar{N}(t-15, t-4)} < 0.3$, we identify the month t as an “abnormal month.”

The following list shows abnormal months for each newspaper over the years 1990-2015:

- **Hankook Ilbo:** August 2014
- **Maeil Economic Daily:** July 2003, December 2006, January - February 2007, September - December 2015
- **Korea Economic Daily:** March - June 1999, September - December 2015, January 2016

5.3 Imputation Method

Each newspaper’s raw EPU rate is estimated by dividing monthly EPU article counts over monthly total article counts. However, due to the abnormal months shown in Section 5.2, we use the regression method to deal with the problem and adjust each newspaper’s EPU rate for the post-1990 period.

Let E_t^i be the EPU rate for newspaper i at month t . Let T^i be the set of abnormal months for newspaper i . In order to impute values for the EPU rate during the abnormal months, we first regress EPU rates of a newspaper with abnormal month on EPU rates of newspapers without any abnormal months (which are Donga Ilbo, Kyunghyang, and Hankyoreh). That is,

$$E_t^i = \alpha + \beta_1 E_t^{Donga} + \beta_2 E_t^{Kyunghyang} + \beta_3 E_t^{Hankyoreh} + \varepsilon_t^i, \forall t \notin T^i \quad (1)$$

Note that we do not include the abnormal months when estimating the above regression equation. Once Equation 1 is estimated, we impute the EPU rates for abnormal months

using the following equation:

$$\tilde{E}_t^i = \hat{\alpha} + \hat{\beta}_1 E_t^{Donga} + \hat{\beta}_2 E_t^{Kyunghyang} + \hat{\beta}_3 E_t^{Hankyoreh}, \forall t \in T^i \quad (2)$$

where $\hat{\alpha}$, $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$ are the estimated coefficients from Equation 1.

5.4 Standardization Method

In order to construct an overall South Korean EPU index, we use the adjusted EPU rates that have imputed numbers for abnormal months. We use the following procedures to construct the index:

1. Multiplicatively standardize each newspaper's adjusted EPU rate so that it has unit standard deviation from January 1995 to December 2014. That is, divide each newspaper's adjusted EPU rate by its raw standard deviation. Our choice of January 1995 as the starting month is because it is the starting month of two of the newspapers used.
2. Using the normalized EPU rates, average across newspapers by month to obtain an overall South Korean EPU index from January 1990 onward.
3. Multiplicatively standardize the overall South Korean EPU index to have a mean of 100 from January 1990 to December 2014.

6 Result

In this section, we present final outputs of newspaper-level EPU indices and the overall South Korea EPU index. Figure 7, which is divided into three time periods, shows the standardized newspaper-level EPU indices over time. Figure 8 plots the overall South Korean EPU index over time. Since our resources for news articles have changed from last year's South Korea EPU index, we compare the new version of the South Korea index with the old version in Figure 9, which shows that there is no noticeable change.

Figure 10 shows the annotated South Korean EPU index over time. Major events that might have contributed to South Korean economic policy uncertainty are marked on the spikes. Those events are selected according to (i) the review of the articles that contains the EPU terms, and (ii) the investigation of the major events occurred during the periods of spikes.

Figure 7: Standardized Newspaper-level EPU Indices



Figure 8: South Korea EPU Index

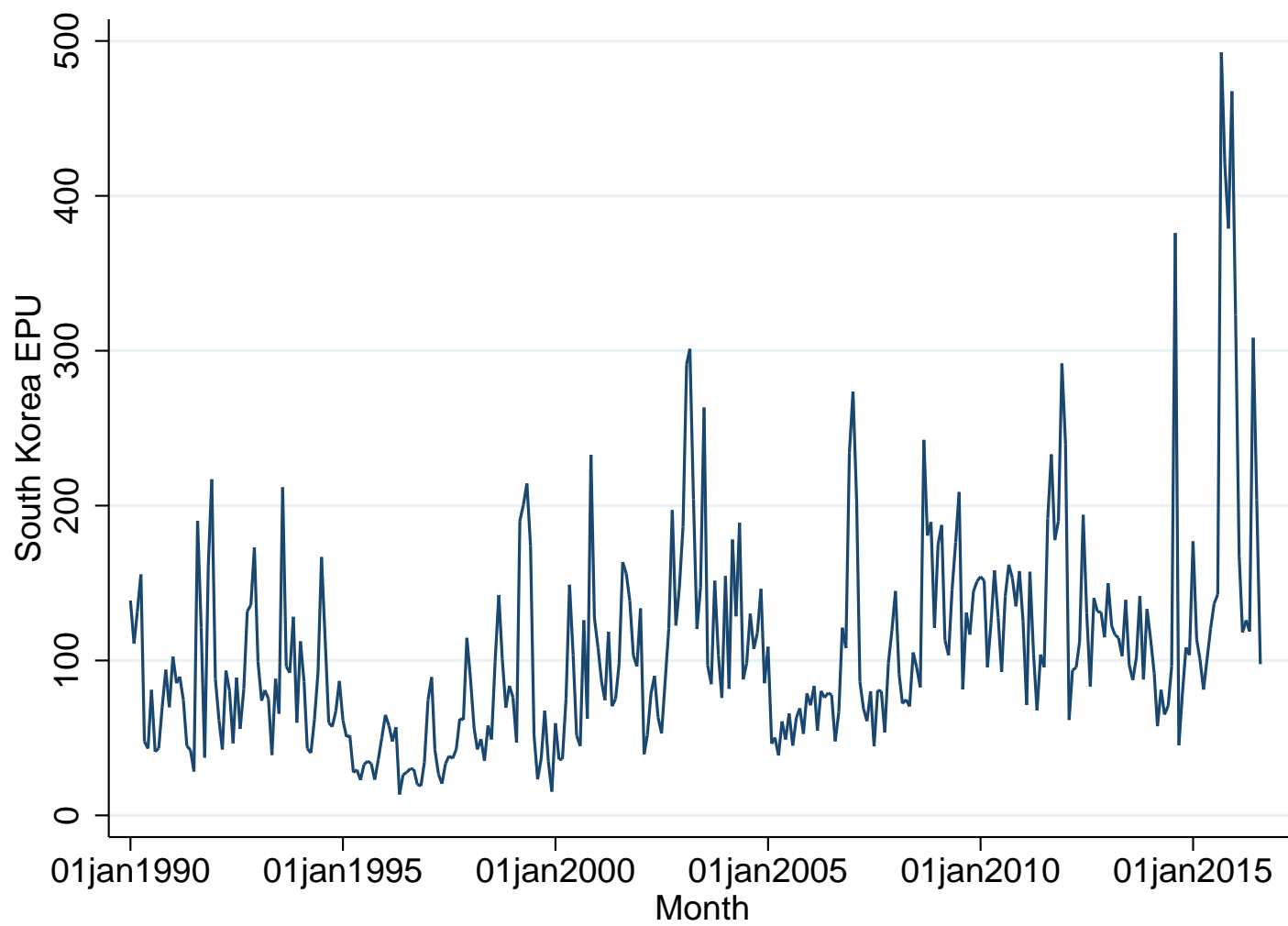


Figure 9: Comparison of South Korea EPU Index with the 2015 Version

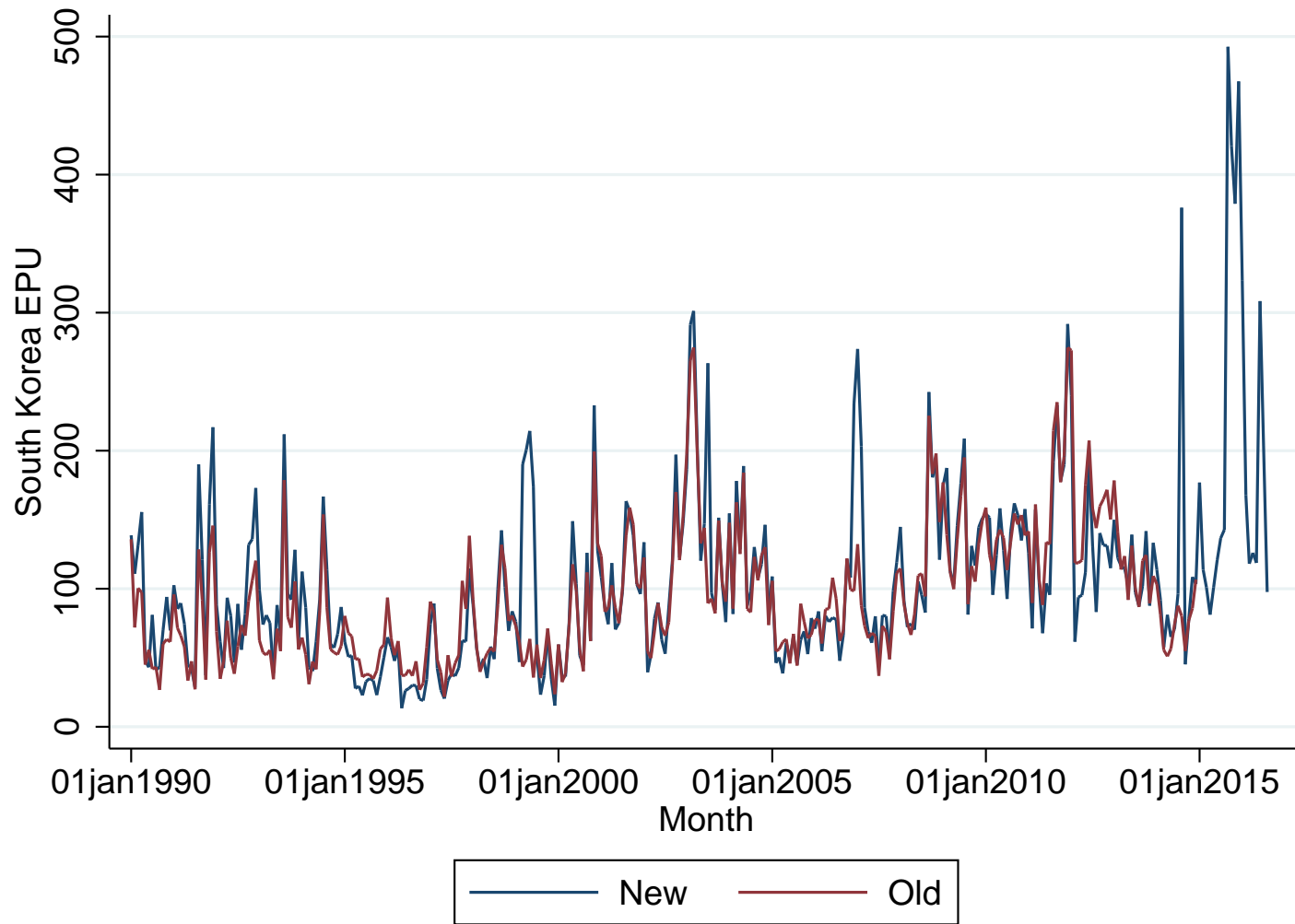
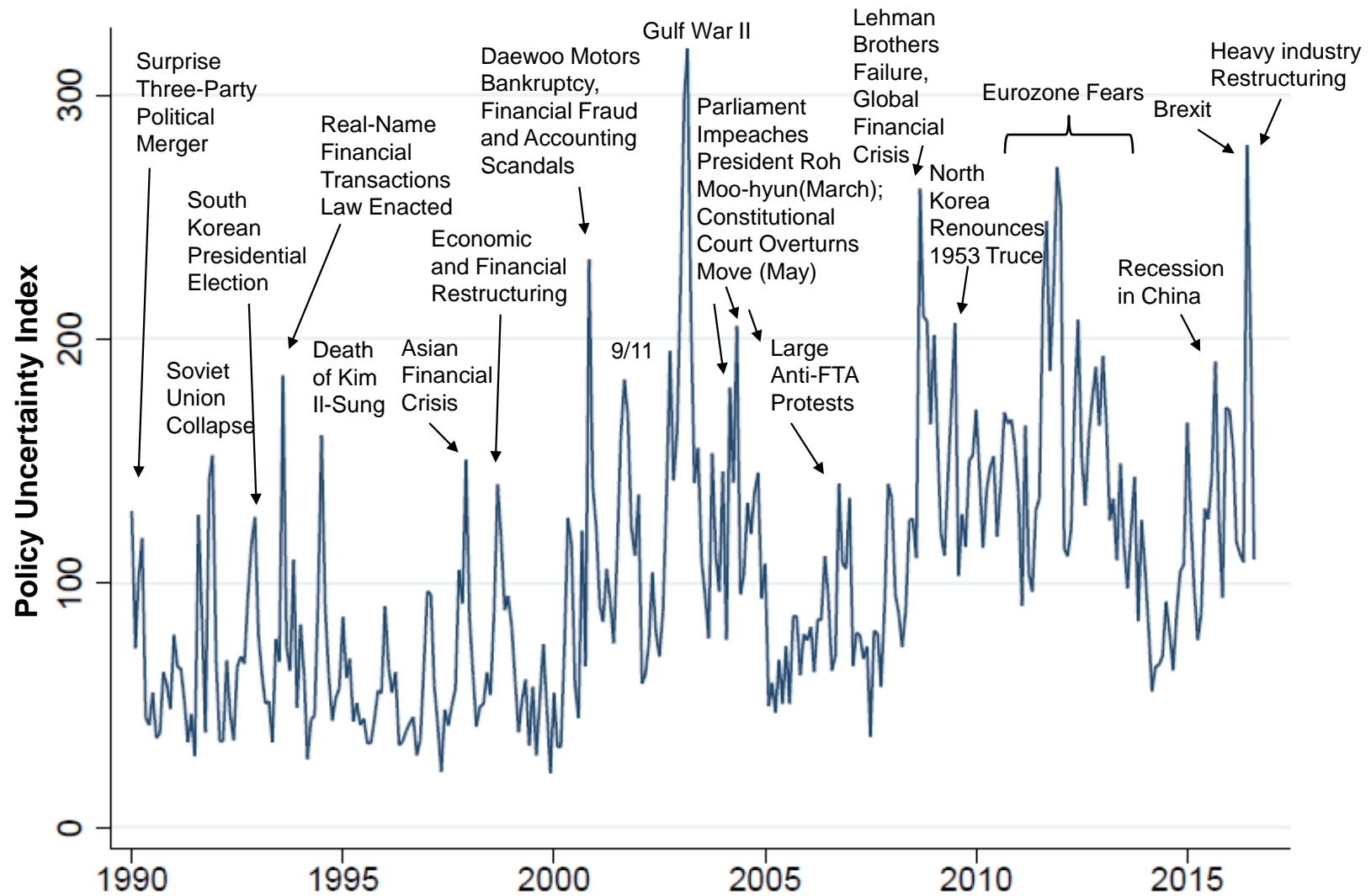


Figure 10: Annotated South Korean EPU Index



Notes: Index reflects scaled monthly counts of articles in six South Korean newspapers containing ‘uncertain’ or ‘uncertainty’, ‘economic’ or ‘economy’ or ‘commerce’, and one or more of the policy terms specified for South Korea in Appendix A. The series is normalized to mean 100 from 1995 to 2014.

7 Next Steps

1. Consider furthering the construction of the South Korean EPU index back to the older dates. In order to do so, we need to **(1)** investigate more into the selection of appropriate terms that capture the economic policy uncertainty and **(2)** include more newspaper sources, especially Chosun Ilbo, once the archive problems are fixed.
2. Investigate why different news archive show different EPU ratio for same newspaper at some periods. Perform day-level analysis for those periods.
3. Further document the causes of abnormal months for each newspaper.