Macroeconomic performance of oil price shocks: Outlier evidence from nineteen major oil-related countries/regions

Keyi Ju ^{a,b,c,d,*}, Bin Su^b, Dequn Zhou^{c,d}, Junmin Wu^a, Lifan Liu^a

^a College of Economics and Management, Jiangsu University of Science and Technology, Zhenjiang 212003, China

^b Energy Studies Institute, National University of Singapore, No. 29 Heng Mui Keng Terrace, 119620, Singapore

^c College of Economics and Management, Nanjing University of Aeronautics and Astronautics, 29 Jiangjun Road, Nanjing 211106, China

^d Research Centre for Soft Energy Science, Nanjing University of Aeronautics and Astronautics, 29 Yudao Street, Nanjing 210016, China

Abstract:

As a special energy commodity, oil price shocks can affect not only the energy market but also the performance of the macroeconomy. This research provides complementary explanations for nineteen major oil-related countries/regions' macroeconomic effects caused by unexpected oil price changes. It focuses on the macroeconomic performance of oil price shocks from outlier perspective, investigating the inner hidden factors of co-movements between oil price shocks and macroeconomy. Three methods called Empirical Covariance (EC) method, Robust Covariance (RC) method and one-class Support Vector Machine (SVM) method are used in the outlier detection. Empirical results show that: (a) one-class SVM method has the best environmental adaptability for detecting the outlier performance of the co-movements between oil price shocks and macroeconomy, followed by EC and RC methods; (b) according to the time axis, the outlier performances of gross domestic production, consumer price index and unemployment rate are all concentrated in 2005-2014, which is highly consistent with the oil price shock process; (c) according to the spatial axis (countries), four categories with similar outlier performances are obtained. Outlier performance of macroeconomy of oil-related country/region in lower levels seems to be worse than that of the country/region in higher levels, which can be attributed to more attentions paid by higher-leveled oil-related countries on the oil price shock response system to relieve the severe macroeconomic impacts caused by oil price shocks.

Key words: Oil price shock; Macroeconomy; Outlier detection; Oil-related countries/regions

* Corresponding author. Tel: +86 5118262880695.
E-mail address: jky1983@163.com (K.Y. Ju).

1. Introduction

Oil, an important energy source in energy markets and for the overall economy, has been well studied in the literatures. Since the first oil price shock in 1973, there have been many studies on the empirical link between oil prices and macroeconomic fundamentals. Although such empirical relationship has been confirmed in previous researches (e.g. Hamilton, 1983^[1]; Burbridge, 1984^[2]), it has been argued that this relationship has been weakened after the collapse of oil price in 1986 (Chang and Wong, 2003^[3]). After that, many studies investigate the effects of oil price shocks on macroeconomic indicators (e.g. Darby, 1982^[4]; Hamilton, 1983^[1]; Barsky and Kilian, 2004^[5]; Farzanegan and Markwardt, 2009^[6]; Kilian, 2009^[7]). However, the results are still inconclusive.

As a special energy resource, oil price can be strongly influenced by exogenous factors such as irregular events (Zhang et al., 2009^[8]), global economic status (Chen and Hsu, 2012^[9]), speculation activities (He et al., 2012^[10]) and political and social attitudes (Ozdemir et al., 2013^[11]), whose effects on macroeconomy are sometimes hard to quantify (Yu et al, 2015^[12]). Therefore, further research is to understand how macroeconomy reacts to oil price uncertainty, which can provide complementary explanations for macroeconomic effects caused by unexpected oil price changes that cannot be attributed to fundamental factors.

The latest oil price shock happened since 2002 to 2008, during which an abnormal oil price surge happened. As suggested by Dogrul and Soytas (2010)^[29], oil price shocks lead to increase in the cost of production in many sectors, which might reduce production and increase unemployment while also result in inflation (Cavalcanti and Jalles, 2013^[13]). Therefore, the most three important macroeconomic domains, *Gross Domestic Production (GDP)*, *Consumer Price Index (CPI)* and *Unemployment rate*, are employed in this paper to detect the inner-effects caused by oil price shocks.

Table 1 shows the Top 10 oil consumption, production and importing countries/ regions in 2014, which are called “major oil-related countries/regions” for short. Considering data availability, we choose 19 countries as research targets¹, which are called oil-related countries/regions. Similar studies have been carried out in the case of OECD countries collections (e.g. Mork, 1994^[14]; Katircioglu et al., 2015^[15]). However, there still a room to investigate such relationship using new approaches and in the case of other important major oil-related countries as well.

¹ Target countries/regions are: U.S., E.U., China, Russia, Japan, India, South Korea, Singapore, Germany, Canada, Brazil, Australia, United Arab Emirates, Iran, Iraq, Kuwait, Mexico, Saudi, and Venezuela.

Table 1 Target countries selection (Thousand barrels daily)

No.	Consumption	Production	Importing
1	U.S. (19,035)	U.S. (11,644)	E.U. (12,601)
2	E.U. (12,527) ²	Saudi (11,505)	U.S. (9,221)
3	China(11,056)	Russia (10,838)	China (7,540)
4	Japan(4,298)	Canada (4,292)	Japan (4,333)
5	India (3,846)	China (4,246)	India (4,225)
6	Brazil (3,229)	United Arab Emirates (3,712)	Singapore (3,057)
7	Russia (3,196)	Iran (3,614)	Canada (1,163)
8	Saudi (3,185)	Iraq (3,285)	Middle Est (1,133) ^{*3}
9	South Korea (2,456)	Kuwait (3,123)	Australasia (1,023)
10	Germany (2,371) Canada (2,371)	Mexico (2,784) Venezuela (2,719)	Mexico (641)

Source: BP statistical review of world energy, 2015

The contribution of this paper is two-fold. Firstly, we study the relationship between oil price and macroeconomy in major oil-related countries/regions from outlier perspective. The concept of outlier detection rooted in the work of Hawkins (1980) ^[16] was initially used in data mining for fault detection. This approach has a competitive advantage in that it measures the effects of oil price shocks with original data, instead of considering the restriction of certain assumptions or endogenous/exogenous variables, maximized retains the inherent characteristics of data sets and avoids the loss and distortion of data information. As far as we know, this is the first study using the outlier detection approach for investigating the relationship between oil price and macroeconomy. The other contribution is that three different outlier detection methods are employed to show the relationship between oil price and macroeconomy. This is enhanced by the fact that new methods will generate more confident and robust results.

The rest of the paper is organized as follows: Section 2 reviews the literature on the macroeconomic effects of oil price shocks. Section 3 describes the methodology and data definitions. Section 4 presents an empirical study using the outlier detection of oil price and macroeconomy indexes for nineteen major oil-related countries/regions. Section 5 concludes the whole study.

² According to BP statistics, the calculation of E.U. data excludes Estonia, Latvia and Lithuania prior to 1985 and Slovenia prior to 1991. Additionally, in order to consistent with BP statistics, we calculate the data for both E.U. and Germany together.

³ Economic data are not available for Middle East region, so it will not be considered in this paper.

2. Literature Review

The literature on the macroeconomic effects of oil price shocks are abundant. For example, Hamilton (1983)^[1], the most influential article in this field, demonstrates that all but one of the U.S.'s recessions since the end of World War II was caused by a dramatic oil price increase. After that, a large number of researches were carried out in various branches, e.g. Huang (2008)^[17], Ghosh (2011)^[18], Bai et al. (2012)^[19], Zhang and Wang (2013)^[20]. Recent studies devote more attention to investigate the oil price-macroeconomy relationship in certain countries, e.g. U.S. (Oladosu, 2009^[21]; Tiago and Tovar, 2013^[22]), China (Huang and Guo, 2007^[23]; Cong et al., 2008^[24]; Ju et al., 2015^[25]; Ju et al., 2016^[26]), Japan (Lee et al., 2001^[27]), Singapore (Chang and Wong, 2003^[3]), South Korea (Rumi et al., 2011^[28]), Turkey (Dogrul and Soytas, 2010^[29]), Canada (Sajjadur and Apostolos, 2012^[30]), and other world countries (Farzanegan et al. 2009^[6]; Ramaprasad and Malliaris, 2011^[31]). However, few of them investigate the oil price-macroeconomy relationship for major oil production, consumption and importing countries, let alone giving an international comparison. It is highly likely that oil price changes have significant effects on the macroeconomy of these major oil-related areas.

Different methods have been adopted in measuring the relationship between oil price shocks and macroeconomic performance in the literature. In general, they can be grouped into two categories: linear and nonlinear statistical methods. The most commonly used approach in early literature was to specify a linear or log-linear relationship between oil price and macroeconomy (Darby, 1982^[4]; Gisser et al., 1986^[32]). As time flies, scientists began to aware that, from a theoretical point of view, the reasons for why oil price shock and to what extent will the shocks affect macroeconomy variable, and they called for non-linear specifications for oil price-macroeconomy relationship, most of which are statistical methods (Lee et al., 1995^[33]; Hamilton, 1996^[34]; Jimenez et al., 2004^[35]). A number of other recent references on this topic includes Gronwald (2008^[36]), Huang (2008^[37]), Hamilton (2011^[38]), Ji and Fan (2015^[39]), and Bal and Rath (2015^[40]). Despite the popularity of nonlinear statistical models, these models have to be built upon some strong assumptions, which are hardly satisfied with due to the uncertain nature of oil prices and macroeconomy. In addition, the macroeconomic effects of oil price shocks are usually measured with dummy variables, so that information loss is inevitable.

In order to fill these gaps, various powerful artificial intelligence (AI) models have been applied to overcome the difficulties in modeling nonlinear hidden patterns of international crude oil price and macroeconomy (Zhang et al., 2009^[41]; Yu et al., 2008^[42]; Jammazi and Aloui, 2012^[43]; Yu et al., 2014^[44]; Ju et al., 2014^[45]; Zhang and Zhang (2015)^[46]; Zhang and Wang (2015)^[47]; Zhang

and Yao (2016)^[48]. Unlike traditional techniques, AI methods are self-adaptive, nonlinear and data driven models. They have been proved to be more effective in modeling complex, nonlinear, irregular and uncertain data than traditional models. As a typical AI branch, outlier detection is widely used in data mining, aiming to discover useful exceptional and unusual patterns hidden in data sets and provide interesting insights about them. The outlier detection algorithms can be classified into four aspects (Huang et al., 2006^[49]): statistical-based methods (Barnett and Lewis, 1994^[50]), distance-based methods (Knorr and Ng, 1998^[51]), density-based methods (Spiros and Hiroiyuki, 2003^[52]) and deviation-based methods (Jagadish and Koudas, 1999^[53]; Zheng and Jiao, 2001^[54]). Outlier detection has been used in several fields, such as fraud analyzing and detecting in financial market (Hübler, 2014^[55]), network intrusion detection (Kershaw et al., 2011^[56]), the outlier detection in business surveys in the presence of missing data (Todorov et al, 2011^[57]), etc. However, it has not been used to study the relationship between oil price and macroeconomy, which is the aim of this study.

3. Methodology and Data

3.1 Overview of outlier detection methodology

Outlier detection (also known as anomaly detection) and clustering analysis are two highly related tasks. Clustering finds the majority patterns in the data set and organizes the data accordingly, whereas outlier detection tries to capture those exceptional cases that deviate substantially from the majority patterns. Hawkins (1980)^[16] generally exposes the key points of outlier: “An outlier is an observation that deviates so much from other observations as to arouse suspicions that it was generated by different mechanism”.

Generally speaking, outliers can be generated by mis-inputs and exceptional events. The first one does not have research value. However, the outliers resulting from the second one are the truth of actuality and can provide additional meaningful information which the majority data sets can’t provide. These outliers are worthy of more attentions. Outliers detected belongs to the latter.

3.2 Three methods for outlier detection

In this paper, we choose three outlier detection methods, called *Empirical Covariance (EC) method*, *Robust Covariance (RC) method* and *one-class Support Vector Machine (SVM) method*

to detect the outlier relationship between oil price and the macroeconomy of nineteen major oil-related countries/regions. Both of the EC and RC methods are based on robust mahalanobis distances computed with the minimum covariance determinant estimator, which was introduced by Rousseeuw (1984)^[58], and one-class SVM method is an AI method. There are two reasons for choosing them in this study: (a) these three methods are commonly used in outlier detection (see Filzmoser, 2004^[59]; Hubert et al., 2008^[60]; Rousseeuw and Hubert, 2011^[61]; Todorov et al., 2011^[57]; Filzmoser and Todorov, 2013^[62]; Gschwandtner and Filzmoser, 2013^[63]); (b) As mentioned in Filzmoser and Todorov (2013)^[62], these three methods are also useful for outlier detection in high dimensional dataset, which is an important target for our further research.

3.2.1 Empirical covariance (EC) method

The Minimum Covariance Determinant (MCD) estimator is a robust, high-breakdown point, which can be used to estimate the covariance matrix of highly contaminated datasets, up to $(n_{\text{samples}} - n_{\text{features}} - 1)/2$ outliers) estimator of covariance. The idea is to find $(n_{\text{samples}} + n_{\text{features}} + 1)/2$ observations whose empirical covariance has the smallest determinant, yielding a “pure” subset of observations from which to compute standard estimates of location and covariance. After a correction step aiming at compensating the fact that the estimates were learned from only a portion of the initial data, we end up with robust estimates of the data set location and covariance.

3.2.2 Robust covariance (RC) method

The estimation of multivariate location and shape is one of the most difficult problems in robust statistics (Donoho, 1982^[64]), because many known methods will break down if the fraction of outlier is larger than $1/(p+1)$, where p is the dimension of the data. Hardin and Rocke (2005)^[65] tackled this problem with Mahalanobis Squared Distance (MSD). Given n data points, the MCD of those data is the mean and covariance matrix based on the sample of size $h(h \leq n)$ that minimizes the determinant of the covariance matrix.

$$MCD = (\bar{X}_J^*, S_J^*) \quad (1)$$

where

$$J = \{\text{set of } h \text{ points: } |S_J^*| \leq |S_K^*| \forall \text{ sets } K \text{ s.t. } |K| = h\}$$

$$\bar{X}_J^* = \frac{1}{h} \sum_{i \in J} x_i$$

$$S_J^* = \frac{1}{h} \sum_{i \in J} (x_i - \bar{X}_J^*)(x_i - \bar{X}_J^*)^T$$

The value h is the minimum number of points that cannot be outlying. The MCD has its highest possible breakdown at $h = \lfloor (n+p+1)/2 \rfloor$ where $\lfloor \cdot \rfloor$ is the greatest integer function (Lopuhaä and Rousseeuw, 1991^[66]). In outlier detection, we use h at its highest possible breakdown $h = \lfloor (n+p+1)/2 \rfloor$ in our calculations, and we refer to a sample of size h as a “half

sample”. The MCD is computed from the “closest” half sample, and therefore, the outlying points will have little effect on the MCD location of shape estimate. The calculation of MCD is quite computationally intensive with Python codes⁴.

3.2.3 One-class Support Vector Machine (SVM) method

The one-class SVM introduced by Schölkopf et al. (2001)^[67] is an unsupervised algorithm that learns a decision function for novelty detection: classifying new data as similar or different to the training set. Strictly-speaking, the one-class SVM is not an outlier-detection method but a novelty-detection method. Outlier detection in high-dimension, or without any assumptions on the distribution of the inlying data is very challenging, and one-class SVM method can give useful results in these situations. Furthermore, one-class SVM can create a non-linear decision boundary by projecting the data through a non-linear function. By providing the normal training data, one-class SVM algorithm can create a representational model of the original data. When newly encountered data is very different from original dataset according to certain measurements, it will be labeled as “outliers” by one-class SVM detection.

The one-class SVM algorithm maps input data into a space (via a kernel) and iteratively finds the maximal margin hyperplane which best separates the training data from the origin. The one-class SVM may be viewed as a regular two-class SVM where all the training data lies in the first class, and the origin is taken as the only member of the second class. Thus, the hyperplane (or linear decision boundary) corresponds to the classification rule:

$$f(x) = [w, x] + b \quad (2)$$

where w is the normal vector and b is a bias term. The one-class SVM solves an optimization problem to find the rule f with maximal geometric margin. We use this classification rule to assign a label to a test example x . If $f(x) < 0$, we label x as an anomaly; otherwise, it is labeled as normal. In practice there is a trade-off between maximizing the distance of the hyperplane from the origin and the number of training data points contained in the region separated from the origin by the hyperplane.

Solving the one-class SVM optimization problem is equivalent to solving its dual quadratic programming problem:

$$\min_{\alpha} \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j K(x_i, x_j) \quad (3)$$

subject to the constraints

⁴ The source for Empirical & Robust covariance outlier detection methods can be seen in the following linkage:
<https://github.com/scikit-learn/scikit-learn/blob/master/examples/covariance>

$$0 \leq \alpha_i \leq \frac{1}{\nu l} \quad \text{and} \quad \sum_i \alpha_i = 1 \quad (4)$$

where α_i is a Lagrange multiplier (or “weight” on example i such that vectors associated with one-zero weights are called “support vectors” and solely determine the optimal hyperplane), ν is a parameter that controls the trade-off between maximizing the distance of the hyperplane from the origin and the number of data points contained by the hyperplane, l is the number of points in the training dataset, and $K(x_i, x_j)$ is the kernel function. by using the kernel function to project input vectors into a feature space that allows for nonlinear decision boundaries. Given a feature map:

$$\phi: X \rightarrow R^N \quad (5)$$

where ϕ maps training vectors from input space X to a high-dimensional feature space, we can define the kernel function as:

$$K(x, y) = \langle \phi(x), \phi(y) \rangle \quad (6)$$

Feature vectors need not be computed explicitly, and in fact it greatly improves computational efficiency to directly compute kernel values $K(x, y)$. For more details, please see Schölkopf et al. (2001) ^[67].

The RC estimate, as a non-robust one, is highly influenced by the heterogeneous structure of the observations. Although the RC estimate is able to focus on the main mode of the data distribution, it sticks to the assumption that the data should be Gaussian distributed, yielding some biased estimation of the data structure, but yet accurate to some extent. The one-class SVM is able to capture the real data structure, but the difficulty is to adjust its kernel bandwidth parameter so as to obtain a good compromise between the shape of the data scatter matrix and the risk of over-fitting the data⁵.

3.3 Data description

Nineteen oil-related countries are selected in this study, and their samples cover the period 1980-2014⁶ (i.e. 35 years). Data series included oil prices and three macroeconomic variables, i.e. GDP (billion U.S. dollars), CPI (Year 2000=100), and unemployment rate (as a percentage of the total labor force). The data are obtained from the BP Statistical Review of World Energy and IMF

⁵ The three methods mentioned here are available from Scikit-Learn, a simple and efficient tool for data mining and data analysis. [The source code of the algorithm is given in the attached data and code file.](#)

⁶ Authors choose the starting time of 1980 with the consideration that the data sets can cover the total four oil price crisis happened in 1973, 1979, 1990, and 2008 (which is called “the fourth oil price crisis” in recent studies by [Yu et al., 2015^{\[12\]}](#) and [Ju et al., 2014^{\[45\]}](#)). Another reason is that the data before 1980s are not available in some target countries.

database. Table 2 shows the summary of the datasets used are shown in Table 2.

Table 2 Data descriptions

No.	Country/Region	Years, data amounts		
		GDP	CPI	Unemployment rate
1	Australia	1980-2014, 35	1980-2014, 35	1980-2014, 35
2	Brazil	1980-2014, 35	1980-2014, 35	1980-2014, 35
3	Canada	1980-2014, 35	1980-2014, 35	1980-2014, 35
4	China	1980-2014, 35	1980-2014, 35	1980-2014, 35
5	E.U.	1980-2014, 35	1980-2014, 35	1980-2014, 35
6	Germany	1980-2014, 35	1980-2014, 35	1980-2014, 35
7	India	1980-2014, 35	1980-2014, 35	<i>No data</i>
8	Iran	1980-2014, 35	1980-2014, 35	<i>1990-2014, 25</i>
9	Iraq	<i>2004-2014, 11</i>	<i>2004-2014, 11</i>	<i>No data</i>
10	Japan	1980-2014, 35	1980-2014, 35	1980-2014, 35
11	Kuwait	1980-2014, 35	1980-2014, 35	<i>1981-2014, 34</i>
12	Mexico	1980-2014, 35	1980-2014, 35	1980-2014, 35
13	Russia	<i>1992-2014, 23</i>	<i>1992-2014, 23</i>	<i>1992-2014, 23</i>
14	Saudi	1980-2014, 35	1980-2014, 35	<i>1999-2014, 16</i>
15	Singapore	1980-2014, 35	1980-2014, 35	1980-2014, 35
16	South Korea	1980-2014, 35	1980-2014, 35	1980-2014, 35
17	United Arab Emirates	1980-2014, 35	1980-2014, 35	<i>No data</i>
18	U.S.	1980-2014, 35	1980-2014, 35	1980-2014, 35
19	Venezuela	1980-2014, 35	1980-2014, 35	<i>1999-2014, 16</i>

Note: Data of E.U. are calculated by authors from all E.U. members.

4. Outlier Detection for the Relationship between Oil Price Shocks and Macroeconomy

4.1 Outlier detection results for major oil-related countries/regions

In this paper, we investigate the relationship between oil price shocks and macroeconomy from outlier perspective. After the analysis of the three outlier detection methodologies, we can get some “normal patterns” in either countries or indicators, those which are not included in the

1 “normal patterns” and called “outliers”⁷. The outlier detection results of nineteen oil-related
2 countries/regions are shown in Appendix A. In order to obtain more valuable results, we discuss
3 the outliers from two aspects: macroeconomic indexes (time axis) and countries/regions (spatial
4 axis) in the following subsections.

5 **4.2 Outlier analysis—comparisons across macroeconomic indexes**

6
7 **Fig. 1 Outlier detection for the relationship between oil price and GDP**

⁷ “Outlier performance” mentioned in this paper is just used to represent some unique expressions in the macroeconomic indicators when oil price shocks.

Fig. 2 Outlier detection for the relationship between oil price and CPI

Fig. 3 Outlier detection for the relationship between oil price and unemployment rate

Fig. 1~Fig. 3 show the outlier detection results of the relationship between oil price and GDP, CPI, and unemployment rate with three outlier detection methods, from the time perspective. Taking Fig.1 for example to explain the meanings of such three radar figures. Fig.1 shows the outlier results of the relationship between oil price shocks and GDP since 1980 to 2014. In this figure, the yellow dot line represents the oil price (It is shown in the secondary axis. The

coordinate of the secondary axis is invisible in the figure in order to avoid ambiguity). The other three solid lines in the primary axis represent the outliers of GDP detected by the three algorithms respectively, and the number in the axis stands for the outliers of GDP occurred during the year 1980 to 2014.

It is obvious that the former oil price crisis (1990-1991) had greater impacts on unemployment rate while the later oil price crisis (2002-2008) brought more impacts on both GDP and CPI to target countries/regions. Outlier evidences from the three figures also show that, period 2005-2014 is the most specific time span among the whole research horizon because it has the largest amount of outliers in the macroeconomic indexes of GDP, CPI, and unemployment rate. Such results are highly consistent with the oil price shock trend, during which international oil price hovers at a historical high level. In that case, we can conclude that oil price shock indeed have significant impacts on macroeconomy, and it is feasible and reasonable to describe such inner-relationship from outlier perspective.

According to three outlier detection methods, the outliers found by RC method are similar among the three macroeconomic indicators, which are concentrated in 2006-2014 years' region, so it shows that RC method has less distinguish ability in detecting the outliers from oil price and macroeconomic data sets.

4.3 Outliers analysis—comparisons across countries

Fig. 4 Outlier detection for the relationship between oil price and GDP (by Country)

Fig. 5 Outlier detection for the relationship between oil price and CPI (by Country)

Fig. 6 Outlier detection for the relationship between oil price and unemployment rate (by Country)

We can get some other interesting results by comparing the outliers among the nineteen major oil-related countries/regions. Fig. 4~Fig. 6 show the total outliers occurred in the 19 countries in three macroeconomic domains. Taking Fig.4 for example, the three solid lines in the primary axis represent the total macroeconomic outliers of the major oil-related countries detected by the three algorithms respectively. Generally speaking, major oil production countries show less outliers than major oil consumption or importing countries. Those who are both major oil consumption and

1 importing countries, such as China, Japan, and U.S, show more serious outliers than others. Such
2 results also confirm that major oil importing or consumption countries are more sensitive to oil
3 price shocks than oil production ones.

4 In addition, from the orange line (the results of RC method) in Fig.4 to Fig.6, outliers occurred
5 in majority countries/regions during 2006 to 2014. In other words, majority of the target countries
6 show similar outlier distributions in RC method. Combined with the results discussed in Section
7 4.2, we can conclude that RC method is not so suitable for detecting the outlier performance
8 between oil price shock and macroeconomy, and one-class SVM method might be the most
9 appropriate.

11 ***4.4 Classification of countries/regions according to outliers***

12 As analyzed above, one-class SVM method is the most effective method in describing the
13 outlier evidence for macroeconomic performance of oil price shocks. In the following section, the
14 results of one-class SVM will be used for countries/regions classification. We intend to find out
15 which countries/regions have the most similar outlier co-movement features between oil price
16 shocks and macroeconomy. Fig.7 shows the distributions of outliers of nineteen major oil-related
17 countries/regions since 1980 to 2014 with one-class SVM method. The solid lines are the same as
18 the results of one-class SVM occurred in Fig.4~Fig.6, and the dotted lines represent the average
19 level of outliers of GDP, CPI, and unemployment rate, respectively. According to those whose
20 amount of outliers are greater than the average level of total outliers, we can divide the countries
21 into four categories⁸ as shown in Table 3.

⁸ For fairness, the countries whose data are incomplete are not mentioned here. Such countries includes: India, Iran, Iraq, Kuwait, Saudi, Arab, and Venezuela.

Fig.7 Distributions of outliers for major oil-related countries/regions (1980-2014)

Table 3 Four country categories according to outliers

Categories	Countries/Regions
0	Russia
1	E.U., Japan, U.S.
2	Brazil, China, Singapore, South Korea
3	Australia, Canada, Germany, Mexico

Category 0-3 stands for that: there are 0(-3) macroeconomic indicator(s) whose amount of outliers are above the average level of corresponding indicator(s).

Evidences in Fig.7 and Table 3 show that Russia is the most stable country in oil price shocks, followed by E.U., Japan, and U.S. (Category 1), China, Singapore, South Korea and Brazil are at the third level (Category 2), which are consistent with their oil consumption and macroeconomic situations in the international oil market. Additionally, Australia, Canada, Germany and Mexico are at the fourth level (Category 3), which indicates that there are the most outliers in the co-movement process of their macroeconomy and oil price shocks.

Russia and U.S. are two major oil consumption and production countries in the world. However, Russia is one of the major oil exporters, while U.S. is one of the major oil importers. Existed conclusions support an asymmetric relationship between oil price shocks and macroeconomy in oil-importing and exporting countries (Mork, 1989^[68] and 1994^[69]; Hooker, 1996^[70] and 1999^[71]). It explains why U.S.'s macroeconomy has more outliers than Russia. Japan's dependence of

importing oil is similar as Singapore and South Korea whose oil importing dependence is above 90%. However, the outlier performance of Japan is better than the other two countries, partially because that Japan has the most effective Strategy Petroleum Reserve (SPR) system in the world. China is also building its national SPR, and have completed its first phase of the SPR project, which will enhance the oil supply chain as well as oil price security situation of China.

The countries in the fourth category, including Australia, Canada, Germany and Mexico, have the most common outliers performance of the co-movements between oil price shocks and macroeconomy. These countries are all located at bottom of Table 1. Compared with other major or higher-leveled oil-related countries/regions, outliers of macroeconomy of those lower-leveled oil-related countries/regions perform worse. That is to say, macroeconomy of countries in lower level seems more sensitive to oil price shocks. It may due to that, after oil price shocks bring more severe effects on their macroeconomy, higher-leveled oil-related countries have paid more attentions on their oil price shock response system in order to ensure steady and healthy development of their macroeconomy.

5. Conclusion and Future Work

Oil is the most traded commodity in the world and its price volatility has great influences on macroeconomic activities, which has been widely accepted in the previous literatures. In this paper, we investigate the impact of oil price shocks on nineteen major oil-related countries/regions' macroeconomy from outlier perspective. The outlier detection method allows us to study the inner effects of oil price shocks on these countries' major macroeconomic indicators, including GDP, CPI and unemployment rate. Several major conclusions are drawn in this paper. First of all, according to the time axis (macroeconomic outlier performance of each year), the outlier performances of GDP, CPI and unemployment rate are all concentrated in 2005-2014, which is highly consistent with the oil price shock process. Second, according to the spatial axis (countries), we divided the nineteen major oil-related countries/regions into four categories, countries in each category have similar outlier performances. It shows that macroeconomic outlier performance of those oil-related countries/regions in lower levels seems worse than the countries in higher levels, due to the reason that higher-leveled oil-related countries have paid more attentions on the oil price shock response system to relieve the severe macroeconomic impacts caused by oil price shocks. Lastly, among the three outlier detection methods used in this paper, RC method is not so suitable for detecting the outlier performance between oil price shock and macroeconomy, and one-class SVM method might be the most appropriate.

Results show that Russia is the most stable country in oil price shocks, followed by EU, Japan,

U.S., and China, etc., which is consistent with their oil consumption and macroeconomic situations in the international oil market. Meanwhile, conclusions from outlier perspective support the asymmetric relationship between oil price shocks and macroeconomy in oil-importing and exporting countries. In order to improve the performance of macroeconomy when oil price shock, there are some policy recommendations. For one thing, countries should give more attention on oil import independence and diversification of oil import sources oil import, the pioneers include U.S., E.U., Singapore. For another, countries should give more efforts on the construction of SPR, the pioneers include U.S. and Japan.

Although the above analyses have provided some interesting insights, it has a few limitations. First, emphasis may also be given to increasing energy resources, such as coal, natural gas and renewable and biomass energy resources, which would be primary energy consumption factors in the near future. It will be interesting and also valuable to know the market rules of outlier performances caused by different energy factor prices, then the research may involve the outlier detection in high dimensional data sets. Meanwhile, oil price shocks may affect other macroeconomic indicators not mentioned in this paper, different indicators may have different and inter-related results between each other. In that case, composite index seems to be more reasonable for detecting the outlier performance of macroeconomy caused by oil price shocks. Thirdly, spot oil prices of each countries may be more suitable to explain the internal behavior of the macroeconomy-oil price shock co-movement outlier performance. These issues will be included in our future research.

Acknowledgements

Keyi Ju is Visiting Scholar at Energy Studies Institute, National University of Singapore. This work is partially supported by grants from the National Science Foundation of China (No. 71203081, 71203055, 71403131, and 71573121), the Ph.D. Programs Foundation of Ministry of Education of China (No. 2014M551591), Talent Project of Jiangsu University of Science and Technology, and Graduate Innovation Foundation of Jiangsu University of Science and Technology (No. 104080603).

Appendix A. Outlier detections of selected countries/regions

Australia

Brazil

Canada

China

EU

Germany

India (Unemployment rate data are unavailable)

Iran (Unemployment rate since 1990)

Iraq (Since 2004, Unemployment rate data are unavailable)

Japan

Kuwait (Unemployment rate since 1981)

Mexico

Russia (Since 1992)

Saudi (Unemployment rate since 1999)

Singapore

South Korea

US

United Arab Emirates (Unemployment rate data are unavailable)

Venezuela (Unemployment rate since 1999)

References

- [1] Hamilton, J.D. Oil and the macroeconomy since World War II. *Journal of Political Economy*, 1983, 91 (2): 228-248
- [2] Burbidge, J., Harrison, A. Testing for the effects of oil-price rises using vector auto regressions. *International Economic Review*, 1984, 23 (2): 459-484.
- [3] Chang Y., Wong J.F. Oil price fluctuations and Singapore economy. *Energy Policy*, 2003, 31:1151-1165.
- [4] Darby M.R. The price of oil and world inflation and recession. *American Economic Review*, 1982, 72: 738-751.
- [5] Barsky RB, Kilian L. Oil and the macroeconomy since the 1970s. *Journal of Economic Perspectives*, 2004, 14(4):115-134.
- [6] Farzanegan M.R., Markwardt G. The effects of oil price shocks on the Iranian economy. *Energy Economics*, 2009, 31:134-151.
- [7] Kilian L. Not all oil price shocks are alike: disentangling demand and supply shocks in the crude oil market. *American Economic Review* 2009; 99(3):1053–1069.
- [8] Zhang X, Yu L, Wang S, Lai KK. Estimating the impact of extreme events on crude oil price: an EMD-based event analysis method. *Energy Econ* 2009; 31(5):768–778.
- [9] Chen SS, Hsu KW. Reverse globalization: does high oil price volatility discourage international trade? *Energy Econ* 2012; 34(5):1634–1643
- [10] He KJ, Yu L, Lai KK. Crude oil price analysis and forecasting using wavelet decomposed ensemble model. *Energy* 2012; 46(1):564–574.
- [11] Ozdemir ZA, Gokmenoglu K, Ekinici C. Persistence in crude oil spot and futures prices. *Energy* 2013; 59:29–37.
- [12] Yu L.A., Wang Z.S., Tang L.. A decomposition-ensemble model with data-characteristic-driven reconstruction for crude oil price forecasting. *Applied Energy*, 2015, 156:251-267.
- [13] Cavalcanti T, Jalles JT. Macroeconomic effects of oil price shocks in Brazil and in the United States. *Appl Energy* 2013; 104:475–486
- [14] Mork KA, Mysen HT, Olsen. Macroeconomic responses to oil price increases and decreases in seven OECD countries. *Energy J* 1994; 15(4):19-36
- [15] Katircioglu ST, Sertoglu K, Candemir M, Mercan M. Oil price movements and macroeconomic performance: Evidence from twenty-six OECD countries. *Renewable and Sustainable Energy Reviews*, 2015, 44: 257-270
- [16] Hawkins D. Identification of Outliers. London: Chapman and Hall, 1980
- [17] Huang BN. Factors affecting an economy's tolerance and delay of response to the impact of a positive oil price shock. *The Energy Journal*, 2008, 29:1-34
- [18] Ghosh S. Examining crude oil price- exchange rate nexus for India during the period of extreme oil price volatility. *Applied Energy* 2011; 88: 1886-1889.
- [19] Bai Y, Zhou DQ, Zhou P. Modelling and analysis of oil import tariff and stockpile policies for coping with supply disruptions. *Applied Energy* 2012; 97(9): 84-90.
- [20] Zhang Y.J., Wang Z.Y.. Investigating the price discovery and risk transfer functions in the crude oil

-
- and gasoline futures markets: Some empirical evidence. *Applied Energy*, 2013, 104(4):220-228.
- [21] Oladosu G. Identifying the oil price–macroeconomy relationship: An empirical mode decomposition analysis of US data. *Energy Policy* 2009; 37(12):5417-5426.
- [22] Tiago CJ, Tovar J. Macroeconomic effects of oil price shocks in Brazil and in the United States. *Applied Energy*, 2013, 104(4):475-486.
- [23] Huang Y, Guo F. The role of oil price shocks on China’s real exchange rate. *China Econ Rev* 2007; 18(4):403–416.
- [24] Cong RG, Wei YM, Jiao JL, Fan Y. Relationships between oil price shocks and stock market: an empirical analysis from China. *Energy Policy*. 2008; 36(9):3544-3553
- [25] Keyi Ju, Bin Su, Peng Zhou, Dequn Zhou, Yuqiang Zhang. Oil price crisis response: Capability assessment and key indicator identification. *Energy*, 2015, 32:1353-1360.
- [26] Keyi Ju, Bin Su, Dequn Zhou, Yuqiang Zhang. An incentive-oriented early warning system for predicting the co-movements between oil price shocks and macroeconomy. *Applied Energy*, 2016, 163:452-463.
- [27] Lee BR, Lee K, Ratti RA. Monetary policy, oil price shocks, and the Japanese economy. *Japan and the World Economy* 2001; 13: 321-349.
- [28] Rumi, M., Sanjay, P., Lurion D.M.. Oil price volatility and stock price fluctuations in an emerging market: evidence from South Korea. *Energy Economics*, 2011, 33(5): 975-986
- [29] Dogrul HG, Soytas U. Relationship between oil prices, interest rate, and unemployment: evidence from an emerging market. *Energy Economics* 2010; 32: 1523-1528.
- [30] Sajjadur, R., Apostolos S. Oil price uncertainty and the Canadian economy: evidence from a VARMA, GARCH-in-Mean, asymmetric BEKK model. *Energy Economics*, 2012, 34(2): 603-610.
- [31] Ramaprasad B, Malliaris AG. Oil prices and the impact of the financial crisis of 2007–2009. *Energy Economics* 2011; 33(6):1049–1054
- [32] Gisser M, Goodwin TH. Crude oil and the macroeconomy: test of some popular notions: note. *J Money Credit Bank* 1986; 18(1):95–103
- [33] Lee K, Ni S, Ratti RR. Oil shocks and the macroeconomy: the role of price variability. *Energy J* 1995; 16(4):39–56
- [34] Hamilton JD. This is what happened to the oil price macroeconomy relationship. *J Monet Econ* 1996; 38(2):215–220
- [35] Jimenez-Rodriguez R, Sánchez M. Oil price shocks and real GDP growth: empirical evidence for some OECD countries. *Appl Econ* 2004; 37(2): 201–228
- [36] Gronwald M. Large oil shocks and the US economy: infrequent incidents with large effects. *Energy J* 2008, 29(1):151–172.
- [37] Huang B.N.. Factors affecting an economy’s tolerance and delay of response to the impact of a positive oil price shock. *The Energy Journal*, 2008, 29:1-34.
- [38] Hamilton JD. Nonlinearities and the Macroeconomic Effects of Oil Prices. *Macroeconomic Dynamics*, 2011, 15(S3):1006-1026
- [39] Ji Q., Fan Y. Dynamic integration of world oil prices: A reinvestigation of globalization vs. regionalization. *Applied Energy*, 2015, 155:171-180.

-
- [40] Bal DP, Rath BN. Nonlinear causality between crude oil price and exchange rate: A comparative study of China and India. *Energy Economics*, 2015, 51:149-156.
 - [41] Zhang X, Yu LA, Wang SY, Lai KK. Estimating the impact of extreme events on crude oil price: an EMD-based event analysis method. *Energy Econ* 2009; 31(5):768–778
 - [42] Yu LA, Wang SY, Lai KK. Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. *Energy Econ* 2008; 30(5):2623–2635.
 - [43] Jammazi R, Aloui C. Crude oil price forecasting: Experimental evidence from wavelet decomposition and neural network modeling. *Energy Economics*, 2012, 34(3): 828-841.
 - [44] Yu L, Zhao Y, Tang L. A compressed sensing based AI learning paradigm for crude oil price forecasting. *Energy Econ* 2014; 46:236–245
 - [45] Ju K.Y., Zhou D.Q. Zhou P., Wu Junmin.. Macroeconomic effects of oil price shocks in China: an empirical study based on Hilbert–Huang transform and event study. *Applied Energy*, 2014, 136: 1053-1066.
 - [46] Zhang YJ, Zhang L. Interpreting the crude oil price movements: evidence from the Markov regime switching model. *Applied Energy*, 2015, 143:96-109.
 - [47] Zhang YJ, Wang J. Exploring the WTI crude oil price bubble process using the Markov regime switching model. *Physica A: Statistical Mechanics and its Applications*, 2015, 421:377-387
 - [48] Zhang YJ, Yao T. Interpreting the movement of oil prices: driven by fundamentals or bubbles?. *Economic Modelling*, 2016, 55:226-240.
 - [49] Huang HY, Lin JX, Chen CC. Fan MH. Review of outlier detection. *Application Research of Computers*, 2006, 8:8-13.
 - [50] Barnett V, Lewis T. *Outliers in Statistical Data*. New York: John Wiley & Sons. 1994
 - [51] Knorr E. M, Ng R T. Algorithms for mining distance-based outliers in large datasets. New York: Proc. Of Int. Conf. Very Large Data-Bases (VLDB'98), 1998, 392~403
 - [52] Spiros P, Hiroyuki K. LOCI: fast outlier detection using the local correlation integral. The 19th International Conference on Data Engineering. 2003: 315.
 - [53] Jagadish HV, Koudas N, Muthukrishnan S. Mining deviants in a time series databases. Edinburgh: Proc. of Int. Conf. Very Large Databases (VLDB'99), 1999.102-113.
 - [54] Zheng JG, Jiao LC. Using deviation detection for data mining. *Computer Engineering*. 2001, 27(8):33-35.
 - [55] Hübner O. Estimation of standard errors and treatment effects in empirical economics-methods and applications. *Journal of Labour Market Researches*. 2014, 47:43-62
 - [56] Kershaw D, Gao QG, Wang H. Anomaly-based network intrusion detection using outlier subspace analysis: A case study. *Advances in Artificial Intelligence*, 2011, 6657:234-239.
 - [57] Todorov V, Templ M, Filzmoser P. Detection of multivariate outliers in business survey data with incomplete information. *Advances in Data Analysis and Classification (ADAC)*. 2011, 5(1): 37-56.
 - [58] Rousseeuw PJ. Least median of squares regression. *Journal of American Statistical Association*. 1984, 79(338):871-880.
 - [59] Filzmoser P. A multivariate outlier detection method. *Proceedings of the Seventh International Conference on Computer Data Analysis and Modeling*, volume 1, pp.18-22, Belarusian State

University, Minsk, 2004

- [60] Hubert M, Rousseeuw PJ, Van Aelst S, High breakdown robust multivariate methods. *Statistical Science*, 2008, 23:92–119.
- [61] Rousseeuw PJ, Hubert M. Robust statistics for outlier detection, *Data Mining and Knowledge Discovery*, 2011, 1(1): 73-79.
- [62] Filzmoser P, Todorov V. Robust tools for the imperfect world. *Information Sciences*, 2013, 245:4-20.
- [63] Gschwandtner M, Filzmoser P. Outlier detection in high dimension using regularization. *Advances in Intelligent Systems and Computing*, 2013, 190:237-244.
- [64] Donoho DL, Gasko M. Influential observations in data analysis. In *Proceedings of the Business and Economics Statistics Section, American Statistical Association*. Washington: ASA, 1982.
- [65] Hardin J, Rocke DM. The distribution of robust distances. *Journal of Computational and Graphical Statistics*, 2005, 14(4):928-946.
- [66] Lopuhaä H, Rousseeuw P. Breakdown of points of affine equivariant estimators of multivariate location and covariance matrices. *The Annals of Statistics*, 1991, 19:229-248.
- [67] Schölkopf B. Estimating the support of a high-dimensional distribution. *Neural Computation*, 2001, 13(7):1443-1471.
- [68] Mork KA. Oil and the macroeconomy when prices go up and down: an extension of Hamilton's results. *Journal of Political Economy*, 1989, 97 (3):740-744.
- [69] Mork K A, Mysen H T, Olsen. Macroeconomic responses to oil price increases and decreases in seven OECD countries. *The Energy Journal*, 1994, 15(4):19-35.
- [70] Hooker M A. What happened to oil-price relationship? *Journal of Monetary Economics*, 1996, 38(2): 195-213.
- [71] Hooker M A. Are oil shocks inflationary? Asymmetric and nonlinear specification versus changes in regime. *Federal Reserve Board Working paper*, 1999.