Examining the theoretical-empirical inconsistency on stationarity of shipping freight rate

Ying KOU[1], Meifeng LUO[2], Yifei ZHAO[3]

[1]Sino-US Global Logistics Institute, Shanghai Jiao Tong University, Shanghai, P.R. China; (Address) Rm 303, Zhongyuan Building, Shanghai Jiao Tong University, 1954 Huashan Road, Shanghai, P.R. China; Email: maggiekou@sjtu.edu.cn

[2]Dept. of Logistics and Maritime Studies, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong S.A.R. (Address): Rm 639 Li Ka Shing Tower, Hung Hom, Kowloow, Hong Kong S.A.R. Email: Meifeng.luo@polyu.edu.hk

[3]Sino-US Global Logistics Institute, Shanghai Jiao Tong University, Shanghai, P.R. China; (Address) Rm 304, Zhongyuan Building, Shanghai Jiao Tong University, 1954 Huashan Road, Shanghai, P.R. China; Email: yfzhao@sjtu.edu.cn

Abstract: As a market price, shipping freight rates should be stationary in theory, but most empirical tests found them non-stationary. To examine the causes of this theoretical-empirical inconsistency, we investigate the sensitivities of the stationarity of shipping freight rates from two perspectives: Sample length and sample window. Longer samples are found not sufficient to make them stationary. Instead, sample windows separated by the structure breaks are tested stationary. Moreover, freight rates are found to have entered into a new phase since the 2008 financial crisis. This study contributes to the literature on the stationarity of shipping freight rates by providing an explanation for the theoretical-empirical inconsistency.

Keywords: Shipping freight rates; Stationarity; Economic cycle; Structural change

Funding: This work was supported by the National Natural Science Foundation of China under Grant [71603164].

1. Introduction

Stationarity, one of the stochastic properties of shipping freight rate, has been extensively studied in the literature. In theory, in a perfectly competitive market, the equilibrium market price resulting from the interaction between market supply and demand should be stationary (Rust, 1985). It may move up and down irregularly in the short-term, and in the long-term it should fluctuate along a constant mean, which is viewed as a mean-reverting process (Tvedt, 2003). Therefore, the mean and variance should be stable over time.

However, there is very little empirical evidence on the stationarity of bulk shipping freight rates that supports the above view. On the contrary, many empirical tests using different sample frequencies, test methods, ship types and sampling windows, as listed in the literature review, suggest that the shipping freight rates are not stationary. Rather, it follows a first-difference stationary process (Kavussanos, 1996; Berg-Andreassen, 1997; Alizadeh and Nomikos, 2007; Xu, Yip and Liu, 2011a; Yin, Luo and Fan, 2017).

Clearly, an inconsistency exists between the theoretical view and empirical results on the stationarity of shipping freight rates (hereinafter referred to as the theoretical-empirical inconsistency). To authors' knowledge, such inconsistency has not attracted much attention. Indeed, the reasons behind it are still unclear, as the empirical results on the stationarity of freight rates only represent the average/overall property in their respective sample periods. The conclusion that shipping freight rates are non-stationary is not totally convincing either, because the stability of the non-stationary result has not been tested. This is especially true now after the Baltic Dry Index (BDI), the drybulk shipping market indicator, dropped below its long-term mean after the outbreak of the 2008 financial crisis. Will the previous empirical result still hold if new data is included in the test?

We examine the sensitivity of stationarity to the sample duration from two aspects– sample length and sample window based on assumption that shipping freight rate reflects the short-run and long-run interaction between the unique J-shaped supply and inelastic demand in shipping market (Stopford, 2009, p.g., 165). In the long-run where the shipping supply is flexible, freight rate fluctuates along the flat part of the J-curve. Therefore, it should be stationary. In the short-run, shipping supply is fixed, or the increase in supply is not sufficient to offset the exogenous demand increase, demand intercepts with the supply curve at the vertical section, freight rates will increase and may appear to be trend stationary or non-stationary. If overcapacity exists, the freight rate can appear to be stationary because it also fluctuates along the flat part of the J-curve. To check whether this is the case, we first examine whether simply increasing the sample length is sufficient for a time series to be stationary. Second, we test the stationarity for the samples separated by Break-Point (BP)—the time when the nature of the shipping cycle changed. If the result is stationary within one window, but not with multiple windows, it can explain non-stationary result in most of the existing literature. The change of shipping cycles is the main reason that results in these non-stationary conclusions in the past empirical test.

Several interesting results are found in this study. First, sampling length does affect the stationarity test result but it is not a sufficient condition. Second, the stationarity of shipping freight rates is sensitive to the sample window. Importantly, freight rates are found to be stationary in the sample windows separated by BPs. Third, since the 2008 financial crisis, freight rates have been stationary and the stationarity has been stable, from both short-term and long-term aspects. All the above results are in line with shipping economic theories. The movement of freight rate is theoretical-empirical consistent if taking its long- and short-run properties into account.

This paper contributes to the theory in shipping economics and shipping market dynamics by pointing out the changing nature of stationarity in shipping freight rates. Theoretically, it fills in a gap in the literature on the stationarity of freight rates, by providing an explanation for the theoretical-empirical inconsistency. In practice, stationarity is a fundamental property of a time series data. Failure to recognize its stationarity may affect the results of any subsequent analysis, such as co-integration, Granger causality, or a Vector autoregressive (VAR) model. In particular, understanding the dynamic pattern of freight rates can help the shipping industry better anticipate future market changes and thus reduce business risks. However, to make an accurate prediction, merely increasing the sample length may not be enough, as the impact of structure changes must also be taken into account. In addition, this study provides a clear delineation of sample windows with different property of stationarity, which will facilitate future related studies.

The remainder of this article is as follows. Section 2 reviews previous studies; Section 3 discusses the methodology used in this paper; Section 4 presents the empirical analysis; and Section 5 is the Conclusion.

2. Literature Review

There is a large body of literature discussing the nature of shipping freight rates, either from the nature of a single time-series (Veenstra and Philip, 1997; Kavussanos and Alizadeh, 2001; Kavussanos and Alizadeh, 2003; Erol, 2017) or from its relationship with others (Beenstock, 1985; Tsolakis, Cridland and Haralambides, 2003; Alizadeh and Nomikos, 2007; Xu, Yip and Liu, 2011a; Kou and Luo, 2015; Adland, Bjerknes and Herje, 2017).

This study mainly belongs to the first group, addressing a single time-series data of shipping freight rate, and focusing on its stationarity. The modelling of shipping freight rates using a mean-reverting process first appeared in the work by Bjerksund and Ekern (1995) in modeling the cash flows in shipping. After that, Tvedt (1997) applied Geometric Mean Reversion in modelling the freight rate process, and Kou and Luo (2015) incorporated structural changes into the ocean freight rate model using the Ornstein-Uhlenbeck process.

Empirical works examining the nature of shipping freight rate stationarity are quite numerous, since this is a fundamental test of a time series. Table 1 summarizes most of the previous traditional stationarity tests on the freight rates in dry bulk shipping. Surprisingly, with such a big difference in sample frequency and window, freight rate indicator, test method, and ship size, the test results are very consistent that the shipping freight rates are non-stationary, which contradicts maritime economic theory. Only two exceptions exist. Tsolakis, Cridland and Haralambides (2003) found that the annual time charter rates are stationary; Tvedt (2003) found that the weekly Panamax spot rate is stationary after transforming the unit from US$ into Japanese yen.

Table 1: Summary of past studies on the traditional stationary tests of shipping freight rates

Literature	Sample Frequency	Representative Variables	Methodology	Ship type	Sample window	Result
Kavussanos (1996)	M	BFI,TC	ADF	Cap,Pan,Haz	1973.01-1992.12	N

Berg-Andreassen (1997)	Q	SPR,TC	ADF	Cap,Pan,Haz	1980-1989	N
Glen and Rogers (1997)	W	Index	ADF,PP	Atlantic-Cap	1989.10-1996.06	N
				Pacific-Cap	1991.09-1996.06	N
Veenstra and Philip (1997)	M	SPR	DF	Cap,Pan	1983.09-1993.08	N
Veenstra (1999)	M	SPR,TC	ADF	Cap,Pan,Haz	1980.10-1993.10	N
Kavussanos and Alizadeh (2002)	M	TCE,TC	ADF,PP	Cap,Pan,Hax	1980.01-1997.08	N
Tsolakis, Cridland and Haralambides (2003)	A	TC	ADF	Cap,Pan,Hax	1960-2001	Y
Tvedt (2003)	W	TCE	ADF	Cap,Pan,Haz	1988-1999	N
	W	SPR	ADF	Cap,Haz	1984-1999	N
				Pan	1984-1999	Y
	D	BFI	ADF		1995.01-1999.10	N
Kavussanos and Visvikis (2004)	D	SPR,FFA	ADF,PP,KPSS	Atlantic-Pan	1997.01-2000.07	N
				Pacific-Pan	1997.01-2001.04	N
Koekebakker, Adland and Sødal (2006)	W	TCE	ADF	Cap,Pan,Hax	1990.01.05-2005.05.20	N
Alizadeh and Nomikos (2007)	M	TC1	PP,KPSS	Cap	1979.04-2004.09	N
				Pan,Haz	1976.01-2004.09	
van Dellen (2011)	W	SPR	ADF,PP,KPSS	Cap,Pan	1989.01-2009.06	N
Xu, Yip and Liu (2011a)	M	BDI,TC	IPS,ADF,PP	Cap	1999.03-2009.05	N
				Pan	1998.05-2009.05	N
				Hax	2000.09-2009.05	N
Xu, Yip and Marlow (2011b)	M	SPR	ADF	Cap,Pan	1973.01-2010.10	N
		TC1		Cap	1977.01-2010.10	N
				Pan	1976.01-2010.10	N
Yin, Luo and Fan (2017)	M	BPI,BCI	ADF,PP	Cap,Pan	2007.01-2013.11	N

Notes: (1) A, Q, M, D represents annual, quarterly, monthly and daily data frequencies, respectively; (2) SPR, TC, FFA and BFI stands for spot rate, time charter rate, forward freight agreement and Baltic freight index, respectively; (3) ADF, PP, KPSS and IPS means Dickey and Fuller *(1981)*, Phillips and Perron *(1988)*, Kwiatkowski et al. *(1992)* and Im, Pesaran and Shin *(2003)*, respectively; (4) Cap, Pan, Haz and Hax means Capesize, Panamax, Handysize and Handymax ship carriers, respectively; (5) Y indicates the examination result is stationary, while N indicates the result is non-stationary.

Apart from traditional stationary test on the stationarity of shipping freight rates, there are also studies with advanced test method. Koekebakker, Adland and Sødal (2006) pointed out that traditional test method is based on linear form regression while freight rates may follow a non-linear form stationary process. They applied non-linear statistical model in the test, but still get the non-stationary result (Table 1). van Dellen (2011) argued that the order of integration of freight rates may lie in the middle of 0 and 1, which is fractional integrated. He examined the freight rates in both dry bulk and tanker shipping sector and found that the integration orders of freight rates in drybulk sector are higher than that in the tanker sector. Fractional integration has also been applied to other transportation time-series such as traffic volume (Karlaftis and Vlahogianni, 2009). However, their result is also non-stationary.

Existing studies on the separation of different shipping cycles are also limited. Quantitatively, only Kou and Luo (2015) identified the structural changes based on a modified mean-reverting process. They pointed out that a traditional mean-reverting assumption may not be appropriate for the freight rate process. Alternatively, a shifting mean with structural changes should be taken into account, due to the existence of the long term imbalance between demand and supply, and the technology progress in shipping. However, their main concern at that time was not the stationarity of shipping freight rates. To the authors' knowledge, up to now no studies have calculated the shipping cycles based on the stationarity of freight rate.

To sum up, this theoretical-empirical inconsistency and its possible impact on the statistical analysis of the shipping market motivate us to explore further the sensitivity of such a statistical test when changes in sample size and duration occur.

3. Methods

The most common method for stationarity test of a time series includes the Augmented Dickey Fuller (ADF) test (Dickey and Fuller, 1981), the Phillips-Perron (PP) test (Phillips and Perron, 1988), and the KPSS test (Kwiatkowski, Phillips, Schmidt and Shin, 1992). As shown in Table 1, most existing studies on the stationarity analysis of shipping freight rates used ADF test. Although some also used PP or KPSS, but the result are the same. To enable the comparison with existing studies, ADF test is adopted. The PP and

KPSS test results are also provided. In addition to the test method, the linear statistical model that allows slope and intercept changes is adopted. It is recognized that there are other high-power regression models, such as the non-linear model by Koekebakker, Adland and Sødal (2006). For the purpose of this study, if a general model can prove stationary, using high-power model is not necessary.

3.1. Sensitivity of stationarity to the sampling length

Traditional stationarity test does not take into account the structure changes in the time series data. Use y_t to denote the freight rate time series, the statistical equation for ADF test can be written as follows:

$$\Delta y_t = u + \beta t + \alpha y_{t-1} + \sum_{j=1}^{k} c_j \Delta y_{t-j} + e_t \quad (1)$$

where u is the intercept indicating a drift, t is the time index, β is the coefficient representing the time trend, α is the coefficient presenting process root, k is the lag length and e_t is an independent identically distributed residual term. The main concern of this testing is whether the coefficient α equals to zero. Thus, the null hypothesis of $\alpha=0$ is tested against the one-side alternative hypothesis $\alpha<0$ (stationary process). The conventional t-ratio for α, $t_\alpha{}^{ADF} = \hat{\alpha}/(se(\hat{\alpha}))$, is used to evaluate the hypothesis, where $\hat{\alpha}$ is the estimate of α, and $se(\hat{\alpha})$ is the coefficient standard error.

Separate ADF tests are carried out based on whether or not $\beta=0$. If β is significant, it indicates that the freight rate fluctuates around the trend line $u+\beta t$, so we then check whether α is significant. If β is not significant, then we set $\beta=0$, and run the ADF test again. The optimal lag k is selected by the Schwarz Information Criterion (SIC) (Schwartz, 1978). A specified upper bound k_{max}=4 is chosen to keep it consistent with some of the past studies (Kavussanos and Alizadeh, 2002; Kavussanos and Visvikis, 2004).

Sensitivity analysis is conducted by examining whether the ${t_\alpha}^{ADF}$ is sensitive to the sampling length. A series of ADF tests is conducted by iteratively increasing the sample size. For a data with sample size N, the test starts with the first i observations ($i<N$). Then, every subsequent test increases the size of the test sample by one. Since economic theory suggests that, in the long term, freight rate should move around its mean level, β=0 is appropriate for the sensitivity analysis of the sampling length.

3.2. Stationarity sensitivity to the sample window

To examine whether the freight rates in different sample windows are stationary, we first identify the Break Points (BPs) by taking into account the drift and trend changes together with stationarity. The pioneering works in this direction are Perron (1989), Zivot and Andrews (1992) (referred to as ZA model here-after), Lumsdaine and Papell (1997) and Lee and Strazicich (2003). The BP in Perron (1989) is predetermined by the author, while that in other works are identified by the data-dependent method. Among them, ZA model only considered one BP, while the other two works considered two BPs. However, none of the above methods can be applied directly in our case. As the BPs of freight rate process indicates the change of shipping cycles, both their number and timing are unknown. Therefore, a methodology suitable for finding BPs for shipping freight rate needs to be designed.

Structure changes considered in ZA model include the changes in the intercept and/or the trend, which are considered as the fundamental changes in shipping cycles. Thus, ZA models are adopted as the basic models in identifying BPs. These models are based on Equation (1) with $\beta \neq 0$:

Model A $$\Delta y_t = \mu^A + \beta^A t + \theta^A DU_t + \alpha^A y_{t-1} + \sum_{j=1}^{k} c_j^A \Delta y_{t-j} + e_t \quad (2A)$$

Model B $$\Delta y_t = \mu^B + \beta^B t + \gamma^B DT_t + \alpha^B y_{t-1} + \sum_{j=1}^{k} c_j^B \Delta y_{t-j} + e_t \quad (2B)$$

Model C $$\Delta y_t = \mu^C + \beta^C t + \theta^C DU_t + \gamma^C DT_t + \alpha^C y_{t-1} + \sum_{j=1}^{k} c_j^C \Delta y_{t-j} + e_t \quad (2C)$$

where $DU_t = \begin{cases} 0, t \leq T_b \\ 1, t > T_b \end{cases}$ and $DT_t = \begin{cases} 0, t \leq T_b \\ t - T_b, t > T_b \end{cases}$, T_b is the time of BP. The null hypothesis of Model A-C is that y_t is a non-stationary process, and the alternative hypothesis that y_t is a stationary process with a one-time structural change. Models B and C belong to the trend-stationary case, while Model A is a trend-stationary ($\beta \neq 0$) or a non-trend stationary ($\beta = 0$) case.

Based on the Model A-C, the steps to find BPs are as follows:

① The terminal condition for the procedure to stop searching a BP in the sample is specified. If the sample length T>48, the process continues; Otherwise, it ends.

② All the models A-C are applied to test all possible break times between [24, T-24].

③ If it is not possible to find any break point to improve the sample's stationarity, or the structure change at the break point is not significant, the procedure ends.

Otherwise, collecting all the $t_\alpha^{A^*}$, $t_\alpha^{B^*}$ and $t_\alpha^{C^*}$. Since they are all negative, the model with smallest t_α is chosen, *i.e.* $t_\alpha^*=\min(t_\alpha^{A^*}, t_\alpha^{B^*}, t_\alpha^{C^*})$, and go to next step.

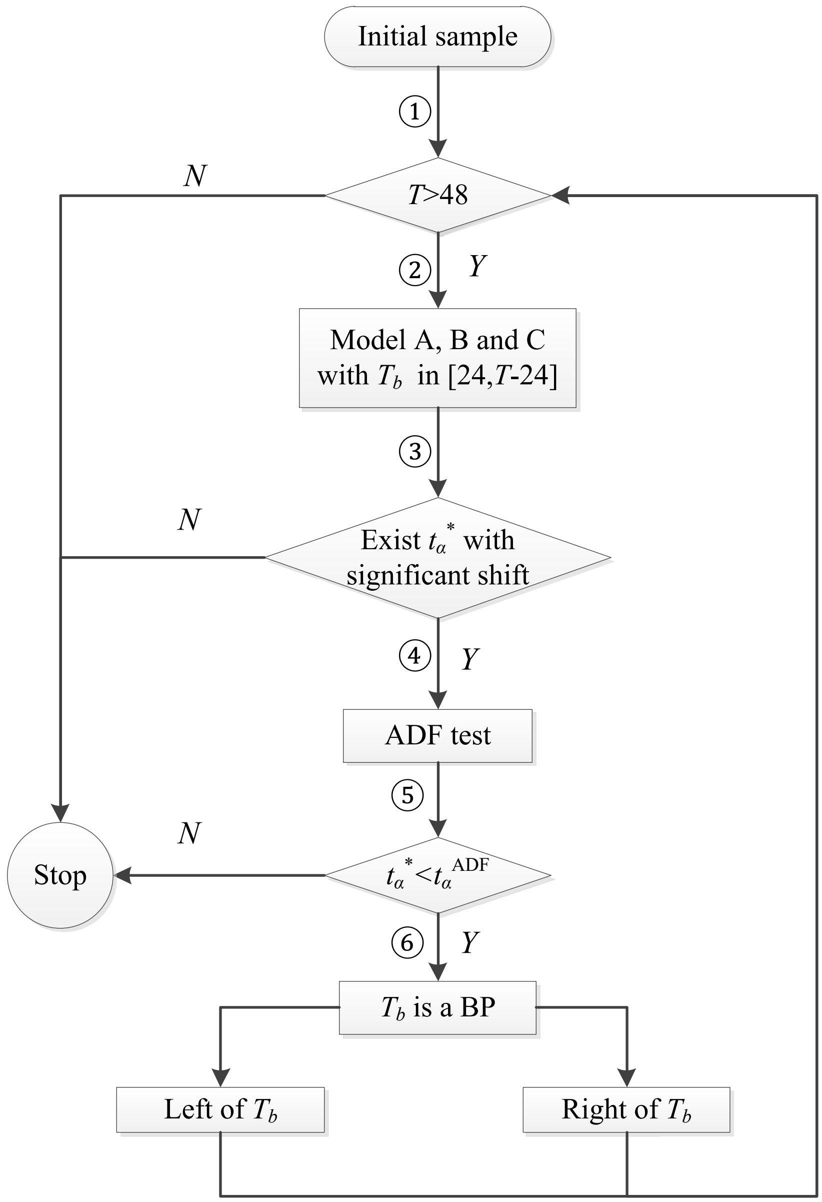
④ An ordinary ADF test will be applied to get the *t*-value (t_α^{ADF}) for the stationarity test.

⑤ If $t_\alpha^* \geq t_\alpha^{ADF}$, since they are both negative, it indicates that the structure change at T_b cannot improve the stationarity. Then it is not a BP, the procedure ends. Otherwise,

⑥ the T_b is considered as a BP. It splits the sample into two parts. For each part, repeat the process from step 1 to step 6.

From the above steps, a series of BPs in the shipping freight rate can be identified. The flow chart of the procedure is shown in Figure 1.

Figure 1: Flow diagram for identifying the break points

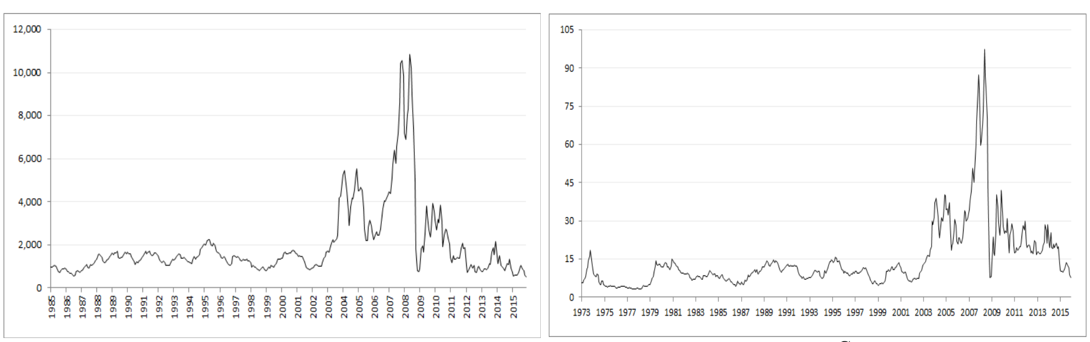


There are three differences comparing our method with ZA's: (I) The selection of the model among A, B or C is pre-determined in ZA model, while it is data-dependent in our study. (II) The BP in ZA model maximizes the chance for the series to be stationary, ignores whether structural change at BP is significant. Whereas the BP in this study is the point that improves the stationarity of the series with significant structural change. (III) ZA model can only find one BP at a time, while our method can obtain multiple BPs.

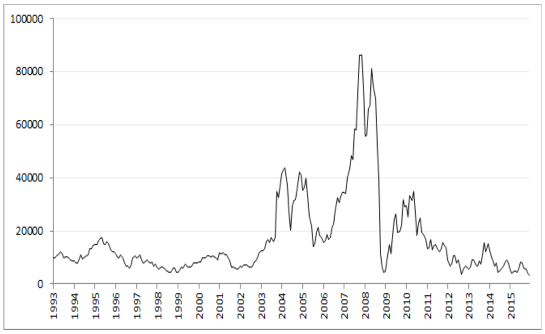
4. Empirical results

Many previous studies have used various market indicators, such as the BDI, SPR, TCE, and TC, with specific ship size categories, including Capesize, Panamax and Handysize, and different sampling frequencies such as daily, weekly, monthly and annual data. To be comparable with past results (Alizadeh and Nomikos, 2007; Xu, Yip and Marlow, 2011b), monthly SPR for Capesize carriers (SPR^{cap}) with 176,000 DWT and monthly 1-year TC for Handysize carries ($TC1^{Haz}$) with 30,000 DWT are collected from Clarkson Shipping Intelligence Network. To cover all the main drybulk ship types and examine different freight rate indicators, we also collect daily and monthly BDI, monthly TCE for Panamax carriers (TCE^{Pan}) with 72,000 DWT capacity, and weekly 1-year TC for Handysize carries. Figure 2 plots monthly freight rate indicators. Their movements are quite similar.

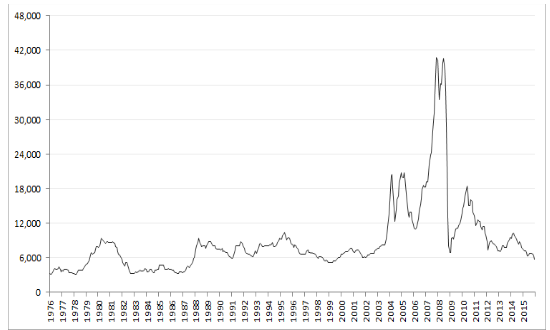
Figure 2: Monthly freight rate indicators



(1) BDI (2) SPR^{Cap}



(3) TCE^{Pan}



(4) $TC1^{Haz}$

4.1. Traditional unit root test results

Table 2 provides a description of all the data used in this study and their results from the traditional ADF, PP and KPSS tests. The unit root nulls are rejected for all the indicators and the βs are equal to zero, indicating that they are all stationary with a constant mean in the long term. This clearly supports the economic theory, but contradicts previous empirical results. Since these tests include the new market conditions, it seems that the current reduction in market freight just reverts from the previous exceptionally high freight rate to its long term mean, which makes the test results stationary. To check its stability, a sensitivity analysis is next carried out.

Table 2: Freight rate indicators and their traditional ADF test results

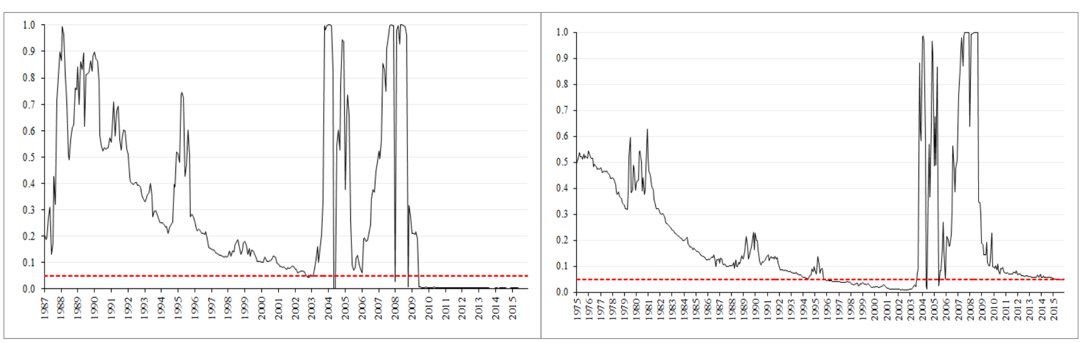
						t_α			
Variables	Frequency	Sample duration	Obs.	Trends	k	ADF	PP	KPSS	Results
BDI	D	1985.01.04-2015.12.24	7779	β=0	12	-3.56***	-3.83***	0.30	Y
	M	1985.01-2015.12	372	β=0	1	-3.81***	-3.83***	37.5	Y
SPR^{Cap}	M	1973.01-2015.12	516	β=0	2	-3.78***	-2.95**	266.8	Y+
TCE^{Pan}	M	1993.01-2015.12	276	β=0	1	-3.01**	-3.04**	27.8	Y
$TC1^{Haz}$	M	1976.01-2015.12	480	β=0	1	-3.78***	-3.78***	124.5	Y+
	W	1989.01.06-2015.12.25	1408	β=0	8	-3.49***	-2.72*	1257.1	Y

Notes: (1) D, M and W mean daily, monthly and weekly data, respectively; (2) The lag length k is selected by SIC; (3) ** and *** stands for rejecting the null hypothesis at 5% and 1% significant levels, respectively; (4) Y signifies stationary, and + signifies non-stationary in both Alizadeh and Nomikos (2007) and Xu, Yip and Marlow (2011b).

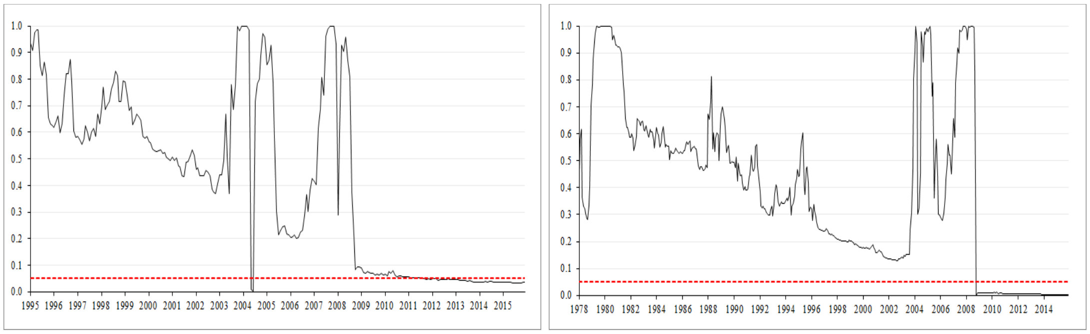
4.2. Sensitivity to sample length

This section describes the sensitivity of the stationarity test to sampling length, using monthly BDI, SPR^{Cap}, TCE^{Pan} and $TC1^{Haz}$. The original ADF test discussed in Section 3.1 is applied iteratively, starting with the initial sample size 24 observations (two years), and expanding by one month in every subsequent test. Figure 3 displays all the *p*-value of α (p_α) from every ADF test. The horizontal axis indicates the ending time of each test sample. From the figure, it is clear that up to Sep. 2002, the BDI tends to be non-stationary, which is consistent with the past studies in Table 1. However, this property is not stable when the sample ends after Oct. 2002 and before 2008. If it ends at Nov. 2002, Feb. 2003, May 2004, Jun. 2004, and Jan. 2008, the test results are stationary. Interestingly, after Sep. 2008, and especially after May 2009, BDI becomes a highly stable stationary process. This explains why, if we use recent data to test the stationarity of freight rates (as in section 4.1), the result is contradictory to the findings in the existing literature.

Figure 3: p_α from ADF test with increasing sample length



(1) BDI (2) SPR^{Cap}



(3) TCE^{Pan} (4) $TC1^{Haz}$

For SPR^{Cap}, Alizadeh and Nomikos (2007) found that it is non-stationary in the period of 1973.01–2010.10. Our sensitivity analysis shows that if the sample size ends between Nov. 1995 and Jul. 2003, the series is actually stationary. In addition, after Jun. 2010, it tends to be stationary by 10% significant level, and even becomes stationary at 5% level after 2015.

Results from TCE^{Pan} and $TC1^{Haz}$ are quite different before and after the financial crisis. TCE^{Pan} gradually becomes stationary after 2008.10, first at 10% significant level, then at 5% level. $TC1^{Haz}$ is a non-stationary process prior to 2008.10, and after that it becomes strongly stationary.

Generally, prior to 2003 a longer sample length does show a trend of decreasing *p*-value. However, the sudden increase in the freight rate after 2003 brings back the *p*-value to a very high level, making it appear non-stationary.

4.3. Sensitivity to sample window

In this section, we examine the stationarity's sensitivity to the change of the sample window. We will apply the newly designed method in Section 3.2 to identify the BPs in the whole sample, and test the stationarity in different sample windows separated by the BPs.

Table 3 provides the steps and results of finding BPs for BDI, SPR^{Cap}, TCE^{Pan} and $TC1^{Haz}$. Taking BDI as an example, in step 1, result from Model C is selected. $t_{\alpha}^{*} < t_{\alpha}^{ADF}$, and both intercept and slope dummies (t_{θ} and t_{γ}) are significant, so that 2006.05 is the first BP. This separates the whole sample into two sub-samples: pre-2006.05 and post-2006.05 (not included). Then, in step 2, BPs are searched in these two sub-samples, and another two BPs are found, at 2003.09 and 2008.05. In step 3, only two sub-samples need to be further examined, since the other two contain less than 49 observations. In the sub-sample 2008.06–2015.12, due to $t_{\alpha}^{*} > t_{\alpha}^{ADF}$, 2010.05 will not be viewed as a BP. Finally, after a total of 7 steps, 7 BPs are established.

The results in Table 3 show that for all the indicators, Sep. (Oct.) 2003, Apr. (May) 2006 and Jul. (Aug.) 2008 are established as BPs. At the same time, these three points are the most significant change points for all the freight indicators. The BPs before 2003 are not consistent. Among them, the most common BPs are 1995 and 2001.

Table 3: Steps in identifying BPs for monthly BDI, SPR^{Cap}, TCE^{Pan} and $TC1^{Haz}$

			Traditional test		New method					
Step	**Sample**	**Obs.**	***k***	**t_{α}^{ADF}**	**Model**	***k***	**BP**	**t_{α}^{*}**	**t_{θ}**	**t_{γ}**
					BDI					
1	1985.01-2015.12	372	1	-3.814^{***}	C	4	2006.05	-4.44^{**}	2.87^{***}	-3.69^{***}
2	1985.01-2006.05	257	2	-2.268	C	2	2003.09	-6.70^{***}	6.80^{***}	-5.77^{***}
	2006.06-2015.12	115	1^{T}	-3.633^{**}	C	1	2008.05	-7.40^{***}	-6.85^{***}	-4.80^{***}
3	1985.01-2003.09	225	1	-1.894	C	1	2001.06	-3.35	-2.67^{***}	3.92^{***}
	2008.06-2015.12	91	1^{T}	-7.476^{***}	C	1	$2010.05^{\times}$	-7.40^{***}	-3.51^{***}	-2.77^{***}
4	1985.01-2001.06	198	1	-2.684^{*}	A	1	1995.08	-4.00^{**}	-3.28^{***}	-
5	1985.01-1995.08	128	1^{T}	-2.515	C	1	1991.11	-3.84^{**}	-3.18^{***}	2.19^{**}
	1995.09-2001.06	70	1	-2.967^{**}	+					
6	1985.01-1991.11	83	1^{T}	-2.814	A	1	1987.07	-3.90^{**}	2.67^{***}	-
7	1987.08-1991.11	52	2	-3.513^{**}	+					
					SPR^{Cap}					
1	1973.01–2015.12	516	4	-2.869^{**}	C		2006.05	-5.26^{**}	3.98^{***}	-4.21^{***}
2	1973.01–2006.05	401	2	-2.273	C		2003.09	-5.59^{**}	5.58^{***}	-4.31^{***}
	2006.06–2015.12	115	1^{T}	-3.805^{**}	C		2008.08	-8.52^{***}	-7.30^{***}	-5.9^{***}
3	1973.01–2003.09	369	1	-2.625^{*}	C		2001.06	-3.82^{**}	-2.46^{**}	3.69^{***}
	2008.09–2015.12	88	1^{T}	-8.562^{***}	+		-	-	-	-
4	1973.01–2001.06	342	1	-3.348^{**}	C		1978.03	-4.43^{**}	2.82^{***}	2.13^{**}

5	1973.01–1978.03	63	2^{T}	-3.399*	B	1974.12	-6.68***	-	5.77***
	1978.04–2001.06	279	1	-3.556***	+	-	-	-	-
					TCE^{Pan}				
1	1993.01–2015.12	276	1	-3.011**	C	2006.05	-4.51**	2.56**	-3.52***
2	1993.01–2006.05	161	1	-2.207	C	2003.09	-6.30***	6.05***	-5.20***
	2006.06–2015.12	115	1^{T}	-3.344*	C	2008.07	-6.24***	-5.51***	-3.73***
3	1993.01–2003.09	129	0	-0.928	C	2001.06	-2.66	-2.05**	3.56***
	2008.08–2015.12	89	2^{T}	-7.275***	C	2010.09	-7.60***	-5.06***	-5.25***
4	1993.01–2001.06	102	0	-1.683	A	1995.09	-3.05	-3.30***	-
	2010.10–2015.12	63	0^{T}	-3.973**	+	-	-	-	-
5	1995.10–2001.06	69	0	-3.038**	+	-	-	-	-
					$TC1^{Haz}$				
1	1976.01–2015.12	480	1	-3.782***	C	2006.04	-5.01**	3.57***	-4.07***
2	1976.01–2006.04	364	3	-1.958	C	2003.10	-4.21**	5.65***	-4.75***
	2006.05–2015.12	116	1^{T}	-3.387*	C	2008.08	-7.44***	-6.49***	-5.06***
3	1976.01–2003.10	332	1	-2.367	C	1980.04	-2.59	-3.24***	-2.21**
	2008.09–2015.12	88	1^{T}	-7.875***	C	2010.06$^{\times}$	-2.94	-4.57***	-2.57***
4	1976.01–1980.04	52	0	2.634	B	1978.04	-4.00**	-	-4.48***
	1980.05–2003.10	282	1^{T}	-2.648	A	1995.09	-3.39	-3.05***	-
5	1980.05–1995.09	185	1^{T}	-2.833	+	-	-	-	-
	1995.10–2003.10	97	1^{T}	-1.240	C	2001.07	-1.84	-2.30	3.51***
6	1995.10–2001.07	70	1	-3.301**	+	-	-	-	-

Notes: (1) ^{T} means the ADF test includes trend term β; (2) *, ** and *** stands for rejecting the null hypothesis at 10%, 5% and 1% significant levels, respectively; (3) $^{\times}$ means the time point is not to be viewed as a BP because of $t_{\alpha}^{ADF}>t_{\alpha}^{*}$; (4) + indicates that no BP can be found in the sample; (5) The 1% and 5% critical values for t_{α} in Model A-C refer to Zivot and Andrews (1992); (6) Sub-samples with less than 48 observations are not listed.

After identified the BPs, the conventional ADF and PP tests are conducted for BDI, SPR^{Cap}, TCE^{Pan} and $TC1^{Haz}$ in each sample window. For the non-stationary sub-period, the test by ZA model with one structural change point (SCP) is applied. The results are summarized in Table 4. It is quite clear that all the indicators in different sub-periods is either stationary or stationary with one structure break (except for SPR^{Cap} in the period of 1975.01-1978.03), which indicates that the selection of sample window is the key factor in the theoretical-empirical inconsistency. Within one window, freight rates are stationary. If multiple windows are included, the test result may be non-stationary. This result is in line with the theory explained in the introduction. In addition, the period following the 2008 financial crisis shows a strong stationary trend, which is also in line with our sensitivity analysis, implying that the freight rates entered into a new phase at that time.

Table 4: Traditional unit root test and Zivot and Andrew's test for monthly BDI, SPR^{Cap}, TCE^{Pan} and $TC1^{Haz}$ in each sample window

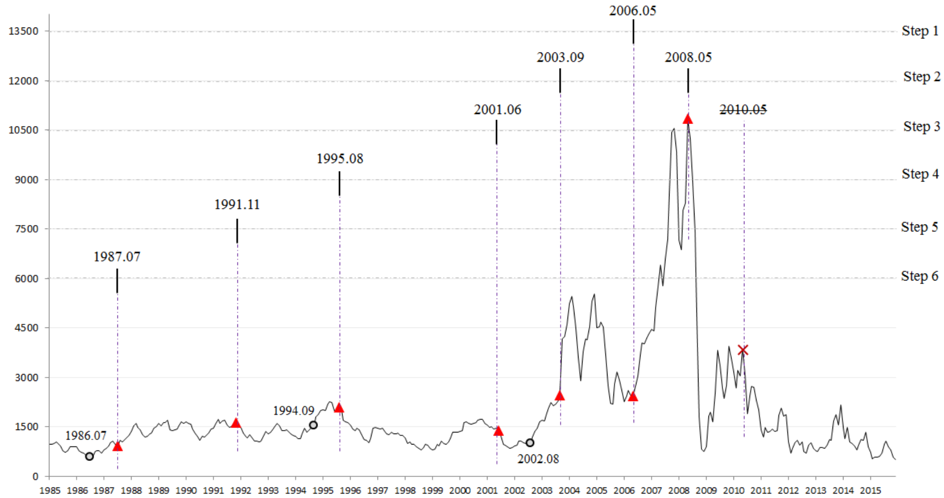
			Traditional unit root			Zivot and Andrews's test					
No.	Sample	Obs.	k	t_α^{ADF}	t_α^{PP}	Model	SCP	t_α^{*}	λ	5% C.V.	Result
					BDI						
1	1985.01–1987.07	31	1	-2.32	-2.39	B	1986.07	-4.91^{***}	0.6	-3.95	Y
2	1987.08–1991.11	52	2	-3.51^{**}	-2.67^{*}						Y
3	1991.12–1995.08	45	0	-2.32	-0.33	A	1994.09	-4.30^{**}	0.8	-3.75	Y
4	1995.09–2001.06	70	1	-2.97^{**}	-2.77^{*}						Y
5	2001.07–2003.09	27	0^{T}	-3.11	-3.11	C	2002.08	-4.52^{**}	0.5	-4.24	Y
6	2003.10–2006.05	32	1^{T}	-4.14^{**}	-3.58^{**}						Y
7	2006.06–2008.05	24	2^{T}	-4.06^{**}	-24.8^{***}						Y
8	2008.06–2015.12	91	1^{T}	-7.48^{***}	-5.18^{***}						Y
					SPR^{Cap}						
1	1973.01–1974.12	24	0	-1.02	-1.75	B	1973.09	-5.50^{***}	0.4	-3.94	Y
2	1975.01–1978.03	39	0^{T}	-2.75	-2.75	C	1976.05	-3.81	0.4	-4.22	N
3	1978.04–2001.06	279	1	-3.56^{***}	-3.53^{***}						Y
4	2001.07–2003.09	27	0	-2.84	1.86	C	2002.08	-5.58^{***}	0.5	-4.24	Y
5	2003.10–2006.05	32	1^{T}	-3.90^{**}	-3.61^{**}						Y
6	2006.06–2008.08	27	2^{T}	-3.73^{**}	-25.8^{***}						Y
7	2008.09–2015.12	88	1^{T}	-8.56^{***}	-7.10^{***}						Y
					TCE^{Pan}						
1	1993.01–1995.09	33	0	-0.76	-0.76	A	1994.09	-3.43^{*}	0.6	-3.76	Y
2	1995.10–2001.06	69	0	-3.04^{**}	-3.04^{**}						Y
3	2001.07–2003.09	27	0^{T}	-2.97	-2.97	B	2002.07	-4.15^{*}	0.5	-3.96	Y
4	2003.10–2006.05	32	0^{T}	-3.59^{**}	-3.58^{**}						Y
5	2006.06–2008.07	26	1^{T}	-2.74	-3.07	C	2007.08	-5.09^{***}	0.6	-4.24	Y
6	2008.08–2010.09	26	0^{T}	-4.41^{***}	-4.41^{***}						Y
7	2010.10–2015.12	63	0^{T}	-3.97^{**}	-3.97^{**}						Y
					$TC1^{Haz}$						
1	1976.01–1978.04	28	0	-1.82	-1.82	B	1976.05	-4.23^{**}	0.2	-3.80	Y
2	1978.05–1980.04	24	1^{T}	-3.05	-3.15	C	1979.04	-6.46^{***}	0.5	-4.24	Y
3	1980.05–1995.09	185	1^{T}	-2.83	-2.64	A	1981.05	-3.83^{**}	0.1	-3.68	Y
4	1995.10–2001.07	70	1	-3.30^{**}	-3.11^{**}						Y
5	2001.08–2003.10	27	1^{T}	-1.05	-1.99	C	2003.06	-4.37^{**}	0.9	-3.80	Y
6	2003.11–2006.04	30	1	-2.81^{*}	-2.78^{*}						Y
7	2006.05–2008.08	28	1	-1.48	-1.36	C	2007.09	-5.04^{***}	0.6	-4.24	Y
8	2008.09–2015.12	88	1^{T}	-7.88^{***}	-6.53^{***}						Y

Notes: (1) ^{T} means includes trend term β; (2) *, ** and *** stands for rejecting the null hypothesis at 10%, 5% and 1% significant levels, respectively; (3) $\lambda = T_b/T$; (4) C.V. are the critical values based on λ in Zivot and Andrews (1992).

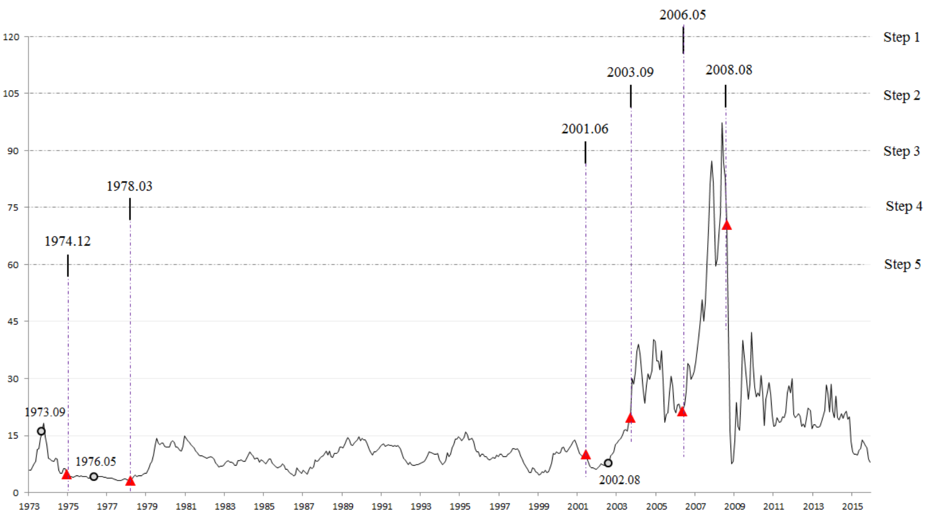
Next, we plot all the above steps in finding BPs and SCPs in Figure 4. The triangles indicate the BPs, and the circles are SCP identified in the stationary test using ZA model.

The BPs in the four indicators after 1995.08 are quite consistent except SPR^{Cap}. In general, at least five major periods can be established, pre-2001, 2001–2003, 2003–2006, 2006–2008 and post-2008. The BPs in 2001, 2003 and 2006 correspond to the interplay of the fast growths in China and a series of economic events, and the BP in 2008 corresponds to the worldwide financial crisis. The *1st period* is prior to Jun. (Jul.) 2001, which corresponds to the period where supply and demand are roughly balanced. Freight rate in this period shows no great jump. The *2nd period* starts from Jun. (Jul.) 2001 to Sep. (Oct.) 2003, which is the first period of the freight rate jump. Although Chinese economic reform started to show some positive impacts, the negative impacts from 9-11 in 2001, and the corporate fraud scandals (Enron) have resulted in the SCP in 2001.08. The *3rd period* is from Sep. (Oct.) 2003 to Apr. (May) 2006, which can be viewed as the second fast booming period in the shipping industry where demand is intercept with the J-shaped supply at the vertical section. The *4th period* is from May (Jun.) 2006 to the global financial crisis in 2008, which is historically the greatest booming period where the demand grows much faster than the capacity. In the 2nd, 3rd and 4th periods, demand-supply gap widens over time. Thus, the freight rate shows several large jumps. The *last period* is the post-financial crisis. Overcapacity is the well-known property during this period, and the freight rate drops to its historically low level.

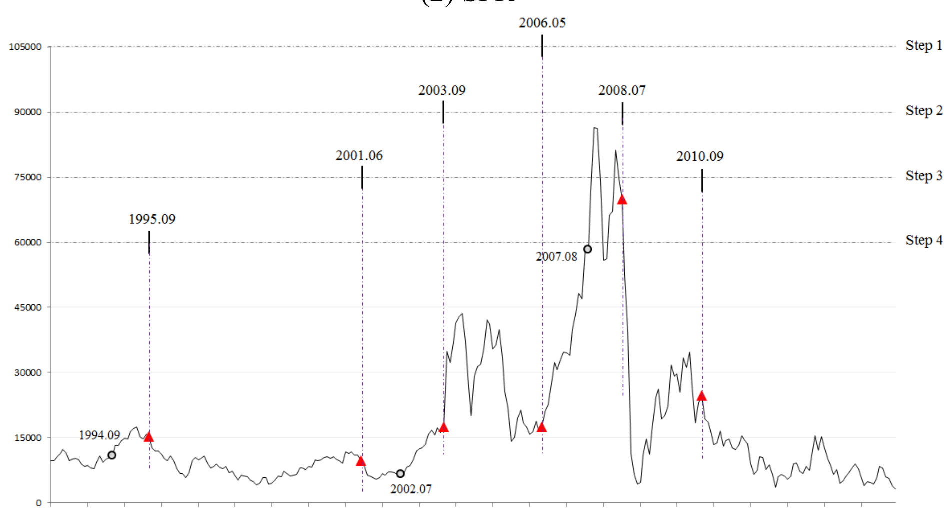
Figure 4: Steps, BPs and SCPs of monthly BDI, SPR^{Cap}, TCE^{Pan} and $TC1^{Haz}$



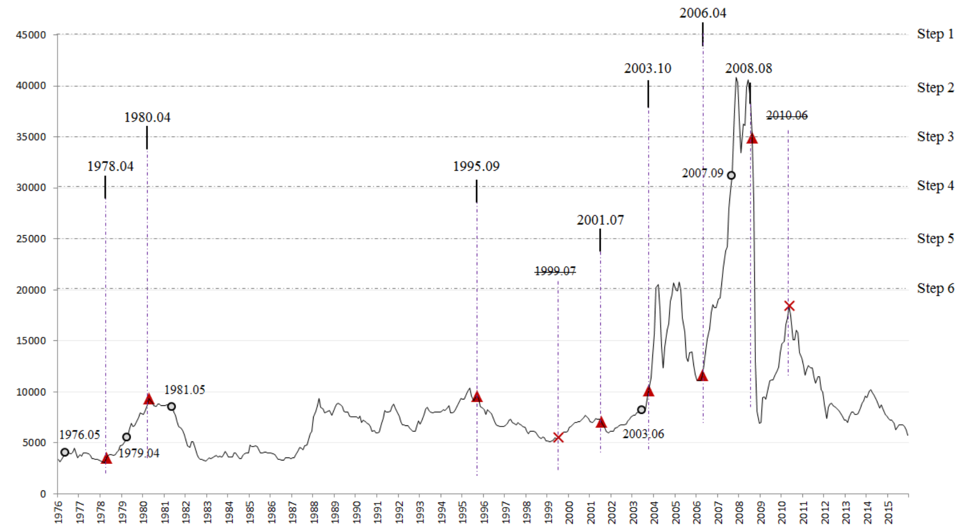
(1) BDI



(2) SPR^{Cap}



(3) TCE^{Pan}



(4) $TC1^{Haz}$

4.4. Results discussions and implications

The major findings and their implications are summarized in this section.

First, for the samples that include the most recent shipping freight rates, even the traditional stationarity test without considering the mean and trend shift can find them stationary. This finding contradicts with most of previous empirical results. It reminds the researchers doing co-integration or causality analysis that the stationarity of shipping freight rates is not invariant with time.

Second, the stationarity of all the freight indicators is found to be sensitive to the change in sample length, mainly because the recent shipping market recession has reverted the freight rate to its long-term mean. However, merely increasing the sample length would not make the freight rates stationary. This result indicates that the stationarity of shipping freight rate can change overtime. Non-stationary results are common if the demand-supply gap is large, i.e., the demand constantly crossing the vertical section of the supply function at the end point of the sample.

Third, the selection of sample window is the key factor resulting in the theoretical-empirical inconsistency. The stationarity of freight rate is highly sensitive to the location of sample windows. When the sample covers a shipping cycle with similar nature, freight rates are found to be stationary. Understanding the changes in the property of freight rates can help the industry participants to better anticipate the future freight rate and reduce financial risks.

Fourth, all the indicators exhibit a remarkable change following the financial crisis in 2008. Before that, most of them are non-stationary; after that, these indicators show strong stationary trends. It suggests that, in order to make accurate predictions, future studies on freight rate time-series should separate samples into pre-2008 and post-2008.

Fifth, in general, five shipping cycles are established for all the freight rate indicators, prior to 2001, 2001-2003, 2003-2006, 2006-2008 and 2008-recent. The demand-supply gaps before 2001 are small, hence the freight rates in this period are relatively stable. After that, China-factor increased the transportation demand for bulk cargoes, and the demand-supply gaps widen over time. Now, freight rates in the post-financial crisis exhibit a clear and significant stationary property.

5. Conclusions

This paper examines the theoretical-empirical inconsistency on freight rate stationarity. We investigate two possible causes of the inconsistency: sample length and sample window. For the former, we applied the traditional stationary test on the whole sample, including the most recent freight rate data. Interestingly, freight rate turns out to be stationary, which is consistent with the economic theory. To examine whether this result is stable, a recursive ADF test is applied to check the sensitivity of stationarity to the sample length. The results show that although the stationarity of all the freight indicators is sensitive to a change in sample length, long-term data does not automatically derive a stationary result, i.e. a long-term data sample in itself is not a sufficient condition for the freight rate to be stationarity. In addition, following the financial crisis in 2008, the freight rate exhibits a clear and significant move to stationarity. This implies that the freight market entered a new phase, one that is quite different from the previous supply and demand relationship.

To investigate the possible effects caused by the sample window, we developed a new method to find the BPs in the freight rate indicators. In general, 2001.06, 2003.10, 2006.05, 2008.08 are the important times that mark structural changes in the freight market. More importantly, the empirical results demonstrate that the freight rate is stationary in each window separated by BPs, which also explains the existence of the theoretical-empirical inconsistency. The key reason is the selection of the sample window in the empirical test. In summary, freight rates should be stationary within a window having similar nature, while its stationarity is uncertain in a sample having multiple windows. It depends on whether the freight rate reverts back to its equilibrium level.

References

Adland, R., F. Bjerknes, and C. Herje. 2017. "Spatial efficiency in the bulk freight market." *Maritime Policy & Management* 44(4): 413-425.

Alizadeh, A. H., and N. K. Nomikos. 2007. "Investment time and trading strategies in the sale and purchase market for ships." *Transportation Research Part B* 41(1): 126-143.

Beenstock, M. 1985. "A theory of ship prices." *Maritime Policy & Management* 12(3): 215-225.

Berg-Andreassen, J. A. 1997. "The relationship between period and spot rates in international maritime markets." *Maritime Policy & Management* 24(4): 335-350.

Bjerksund, P., and S. Ekern. 1995. "Contingent claims evaluation for mean-reverting cash flows in shipping", in Trigeorgis, L. (ed.) *Real Options in Capital Investment, Models, Strategies, and Applications*. London：Praeger.

Dickey, D., and W. Fuller. 1981. "Likelihood ratio statistics for autoregressive time series with a unit root." *Econometrica* 49(4): 1057-1072.

Erol, S. 2017. "Calculation of the freight revenues in Turkey-focused maritime transportation." *Maritime Policy & Management*.

Glen, D.R., and P. Rogers. 1997. "Does Weight matter? A Statistical analysis of the SSY Capesize index." *Maritime Policy & Management* 24(4): 351-364.

Im, K. S., M. H. Pesaran, and Y. Shin. 2003. "Testing for unit roots in heterogeneous panels." *Journal of econometrics* 115(1): 53-74.

Karlaftis, M. G., and E. I. Vlahogianni. 2009. "Memory properties and fractional integraion in trasportation time-series." *Transportation Research Part C* 17(4): 444-453.

Kavussanos, M. G. 1996. "Comparison of volatility in the dry-cargo ship sector." *Journal of Transport Economics and Policy* 30(1): 67-82.

Kavussanos, M. G., and A. H. Alizadeh. 2001. "Seasonality patterns in dry bulk shipping spot and time charter freight rates." *Transportation Research Part E: Logistics and Transportation Review* 37(6): 443-467.

Kavussanos, M. G., and A. H. Alizadeh. 2002. "The expectations hypothesis of the term structure and risk premiums in dry bulk shipping freight markets." *Journal of Transport Economics and Policy* 36(2): 267-304.

Kavussanos, M. G., and A. H. Alizadeh. 2003. "Price discovery, causality and forecasting in the freight futures market." *Review of Derivatives Research* 6(3): 203-230.

Kavussanos, M. G., and I. D. Visvikis. 2004. "Market interactions in returns and volatilities between spot and forward shipping freight markets." *Journal of Banking & Finance* 28(8): 2015-2049.

Koekebakker, S., R. Adland, and S. Sødal. 2006. "Are spot freight rates stationary?" *Journal of Transport Economics and Policy* 40(3): 449-472.

Kou, Y., and M. Luo. 2015. "Modelling the relationship between ship price and freight rate along with structural changes." *Journal of Transport Economics and Policy* 49(2): 276-294.

Kwiatkowski, D., P. C. Phillips, P. Schmidt, and Y. Shin. 1992. "Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?" *Journal of econometrics* 54(1): 159-178.

Lee, J., and M. C. Strazicich. 2003. "Minimum Lagrange multiplier unit root test with two structural breaks." *Review of Economics and Statistics* 85(4): 1082-1089.

Lumsdaine, R. L., and D. H. Papell. 1997. "Multiple trend breaks and the unit-root hypothesis." *Review of Economics and Statistics* 79(2): 212-218.

Perron, P. 1989. "The great crash, the oil price shock, and the unit root hypothesis." *Econometrica: Journal of the Econometric Society* 57(6): 1361-1401.

Phillips, P. C., and P. Perron. 1988. "Testing for a unit root in time series regression." *Biometrika* 75(2): 335-346.

Rust, J. 1985. "Stationary equilibrium in a market for durable assets." *Econometrica: Journal of the Econometric Society* 53(4): 783-805.

Schwartz, G. 1978. "Estimating the dimension of a model." *Annals of Statistics* 6(2): 461-464.

Tsolakis, S. D., C. Cridland, and H. E. Haralambides. 2003. "Econometric modelling of second-hand ship prices." *Maritime Economics & Logistics* 5(4): 347-377.

Tvedt, J. 1997. "Valuation of VLCCs under income uncertainty." *Maritime Policy & Management* 24(2): 159-147.

Tvedt, J. 2003. "A new perspective on price dynamics of the dry bulk market." *Maritime Policy & Management* 30(3): 221-230.

van Dellen, S. 2011. "An examination into the structure of freight rates in the shipping freight markets.", Cass Business School.

Veenstra, A. W. 1999. "The term structure of ocean freight rates." *Maritime Policy & Management* 26(3): 279-293.

Veenstra, A. W., and H. F. Philip. 1997. "A co-integration approach to forecasting freight rates in the dry bulk shipping sector." *Transportation Research Part A* 31(6): 447-458.

Xu, J. J., T. L. Yip, and L. Liu. 2011a. "A directional relationship between freight and newbuilding markets: A panel analysis." *Maritime Economics & Logistics* 13(1): 44-60.

Xu, J. J., T. L. Yip, and P. B. Marlow. 2011b. "The dynamics between freight volatility and fleet size growth in dry bulk shipping markets." *Transportation Research Part E* 47(6): 983-991.

Yin, J., M. Luo, and L. Fan. 2017. "Dynamics and interactions between spot and forward freights in the dry bulk shipping market." *Maritime Policy & Management* 44(2): 271-288.

Zivot, E., and D. W.K. Andrews. 1992. "Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis." *Journal of business & economic statistics* 10(3): 251-270.

Zivot, E.., and D. W.K. Andrews. 1992. "Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis." *Journal of business & economic statistics* 10(3): 251-270.