The Empirical Econometrics and Quantitative Economics Letters
ISSN 2286 – 7147
Volume 4, Number 1 (March 2015): pp. 1 - 10.

Waiting time distributions for the log returns of time series of maritime transport and oil prices

Julien Guigue

Département d'Economie, Institut d' Etudes Politiques, 27, rue Saint Guillaume, 75337 Paris Cedex, France.

guiguej@free.fr

ABSTRACT

In this paper, we utilize the waiting times (WT) technique. Time intervals that define the WTs are obtained as the abscissa of the intersections of the log returns related to time series of prices with a threshold of a given amplitude. For the series of transport (provided through Baltic Dry Index (BDI)) and oil prices compiled in the period from 11 January 1985 to 30 May 2013, it is shown that, after stationnarization of these series, the profile of the distributions of the WTs are very different. The results obtained provide an insight on time relations between transport prices and oil prices. They are complementary to those displayed by such methods as cointegration and regression analyses.

1. Introduction

International maritime transport prices are a major variable of international trade, as demonstrated by Hummels [1]. Besides, oil prices are characterized by a certain volatility since the first oil crisis.

In this context, it seems interesting to study the possible time relationship between these two series, denoted by P_t^T et P_t^p, respectively.

It should be noted that a multivariable analysis can be performed in order to identify other series potentially influencing transportation prices. This study has been done [5]. It shows however that among potential influencing series, only oil prices related time series has a significant impact.

We propose therefore to limit ourselves to two series, namely transportation and oil prices. At least two types of questions must be considered.

Firstly, series with high sample size are required to obtain good statistics. Secondly, the series of data [1] being non stationary, it is necessary to transform them in order to avoid spurious regression [2]. Among the transformations widely utilized to stationnarize time series, we choose the log returns

$$r_t = \log \frac{P_{t+1}}{P_t} \qquad (1)$$

written here for one of the time series.

It can be shown that this transformation, either utilizing numerical tests, or according to the autocorrelation function (ACF) criteria, is good enough in order to make the raw series of Figure 1, quasi-stationary [3]. However, this method has a major drawback: no interesting information can be obtained from the interpretation of the ACFs.

We therefore propose another point of view which consists of using the waiting time (WT) method. To our knowledge, this is the first application of the WT to economic indicators such as series of transport and oil prices, since it was yet applied to stock indices [2], [3] only.

The interest of such a method appears at two levels. From an analytical point of view, it enables a comparison of non stationary time series through their PDFs. Secondly from a prospective standpoint, it allows a qualitative prediction: assessment of the probability for a shock to be followed by another one.

This method has an obvious interest primarily in Finance, but also in Economics, as proposed in this note.

---

[1] The data have been obtained through Bloomberg extracts.

[2] A rather complete analysis of statistical properties of several time series including transport and oil series has been performed in a recent work [5]. In this study we discuss the reliability of the BDI, due to its volatility as compared with other indices. Various programs taken from the Matlab library have been utilized to characterize basic properties of the series such as the trends, the heteroscedasticity, the cointegration relations. In this regard, we stress that transport and oil series are cointegrated variables which allows therefore to complete regression calculations for the couple but not for a direct treatment of raw data considered separately.

[3] We admit that an ACF slowly decreasing having almost all of its value (besides the intercept one) located outside the confidence limits (±0.05 in this case) characterizes a non stationary process. On the other hand, a rapidly decreasing ACF of which all the values (besides the intercept value) are located within the confidence interval (±0.05 in this case) characterizes a stationary process. This is a necessary but not sufficient condition.

The main result obtained reveals that despite a similar appearance of the series, they display significant differences in their time properties (see Figures (1)), through different PDF laws. The transport related series displays two distinct shapes, in short and long intervals respectively, whereas regarding oil prices one behavior seems sufficient to describe its properties. This can be interpreted as a higher volatility displayed by transport prices as compared with oil prices.

The paper will be organized as follows: Firstly the method will be described in section 2 together with the assumptions leading to the analytical expression of the probability density function (PDF) of the WT. Secondly some specific results will be presented in section 3. We finally conclude with some remarks of interest in the economy of the maritime transports in relation with the oil prices.

2. Method and theory of the WTs

The method of the WTs, defined on the basis of the threshold *s*, has recently proved its interest in econometrics (see for example authors cited in [2]), specially in the analysis of the statistical properties of time series.

We start with the sequences of P_t^T et P_t^p, made of about 7200 data compiled almost daily during the period [1985−2013], and reported in Figure 1. Their ACFs which can easily be derived would clearly show that the series are non stationary. Therefore, we must deal with the sequences of log returns r_t^T, r_t^p and then seek the PDF and the cumulative distribution function (CDF) 4 of the WTs, which can be obtained given the threshold s.

We denote by v(θ,s) the PDF of the WTs where θ = ti+1−ti is measured for a given s. Each point ti is obtained as the abscissa of the intersection of rt with the line of constant value s. Before going further, we must first verify that the sequence of {θ} is random. 5

We assume that the best approximation of the conditional probability v(θ|s) can be written as

$$v(\theta|s) = \mu_1 \theta^{\beta} + \mu_2 e^{-\mu 3} \theta^{q}, \qquad (2)$$

where all the matching coefficients depend of course on *s* and are non integer. We emphasize that the equation (2) is different to the expression written under the form of a product as it has been proposed by M.S. Santhanan and Holger Kantz [4]. Note also that in the cases studied in this paper, it was not possible to show that *q* takes a value different to 1 in the equation (2).

The expression of the PDFs as the sum of two functions in the equation (2), instead of a product, helps to clearly distinguish between the domain of small WTs ($\theta << 2$) from the domain of high values of WTs (θ close to 5).

---

[4] We recall that the CDF (increasing and positive function) can be calculated from the corresponding PDF (normalized function). Thus for a given PDF $v_T(\theta)$

$$C_T(x) = \int_1^x v_T(\theta) d\theta + \text{cst}$$

.

[5] To do so, the sequence is shortened by setting a high level for *s*, in order to have a few number of points N. An usual routine indicates for the sequence of $\{\theta\}$ derived from the log returns of the transport indices, with s=0.035, N=152, nruns=$_{63}$ n1=$_{68}$, and n2=$_{72}$, that h=0, pValue=0.2060, $|z|$ = 1.237. Thus the hypothesis assuming that the sequence is random cannot be rejected since the pValue is higher than the confidence interval is $_{\alpha = 0.05}$ and $|z| < 1.96$.

3. Main Results

In this section, the main results of the WTs of the series associated with the log returns of the transport indices and the oil prices are reported.

3.1 Transport indices

We denote by $\Gamma_T(x)$ the CDF estimated from the data and we wish to compare it with two CDFs derived from two other PDFs.

a) model power (-4) and exponential,

$$v_T(\theta) = \frac{1}{M_v}\left(\frac{a_t}{\theta^4} + b_t \mathrm{e}^{-c_t\theta}\right), \qquad (3)$$

where $a_t = 0.12$, $b_t = 0.016$, $c_t = 0.07$ and M_v is the constant of normalization of v_T in the domain of variation of $\{\theta\}$.

b) generalized extreme value,

$$g_T(\theta) = \frac{1}{M_g}\left(\frac{1}{\sigma}\exp\left(-\left(1+k\Theta\right)^{-\gamma}\right)\left(1+k\Theta\right)^{-1-\gamma}\right), \qquad (4)$$

where $\gamma = \frac{1}{k},\ \Theta = \frac{\theta-\mu}{\sigma},\ \mu < \theta,\ k \neq 0,\ \ 1+k\Theta > 0$. The matching coefficients are here $k = 2.8$, $\sigma = 0.015$, $\mu = 1$ and M_g is the constant of normalization of g_T in the domain of variation of $\{\theta\}$.

A statistical comparison between these models is given in the Table 1, i.e. the CDFs derived from Γ_T and v_T on one hand, and Γ_T and g_T on the other hand. The test validates the CDF derived from the equation (3) since the pVal is much more higher than $\alpha = 0.05$. Moreover the highest difference between the two CDFs, ks, which is about 0.13 is reasonably weak.

However, given the values of the present case, the well known law in economy (4) is satisfactory even with a pVal slightly higher than α. In fact, the PDFs are close as shown in Figure 2. Therefore, both models are acceptable.

Now, a similar analysis can be done for the oil prices.

3.2 Oil prices

Let us denote by $W_p(\theta)$ the PDF estimated from the data and $\Gamma_p(\theta)$ its corresponding CDF. We want to compare it with two models.

a) model power (-2) and exponential,

$$v_p(\theta) = \frac{1}{N_v}\left(\frac{a_p}{\theta^2} + b_p \mathrm{e}^{-c_p\theta}\right), \qquad (5)$$

where $a_p = 0.295$, $b_p = 0.39$, and $c_p = 0.57$ and where N_v is the constant of normalization of v_p in the domain of variation of $\{\theta\}$.

b) generalized Pareto,

$$g_p(\theta) = \frac{1}{N_g}\left(\frac{1}{\sigma}\left(1 + k\Theta\right)^{-1-\gamma}\right) \quad (k \neq 0, 1 + k\Theta > 0) \quad . \tag{6}$$

where $\gamma = 1/k$; $\Theta = \frac{\theta-\mu}{\sigma}, \mu < \theta$. Here $k = -0.175$, $\sigma = 3$, $\mu = 0$ and N_g is the constant of normalization of g_p in the domain of variation of $\{\theta\}$.

As previously noted, the comparison of the CDFs derived from Γ_p and v_p on the one hand, and Γ_p and g_p on the other hand, is made. Here again, in view of the Table 1, the test validates the hypothesis that the CDF obtained from the equation (5) yielding a pVal significantly higher than α, and a value of ks of about 0.10, cannot be rejected. On the other hand, the well known Pareto law (6) is satisfactory although the value of pVal is lesser and 0.26, the difference between the CDFs is relatively high. As before, both modelizations can be acceptable.

The PDFs and CDFs are finally reported in Figures 2 and 3 together with the theoretical expressions as summarized in section 2 and drawn in continuous lines. The normalization of the PDFs are made numerically only on time intervals in which the distributions are finite. It is clearly seen that the theoretical expressions are in excellent agreement with the data.

The method of the WTs being validated for one value of *s*, we want to study the behavior of $v(\theta,s)$ for several values of $s \geq 0$. This behavior is measured by the first two moments of θ.

3.3 The first two moments of θ

The method of the WTs being validated for one value of *s*, we want to study the behavior of $v(\theta,s)$ for several values of $s \geq 0$. This behavior is measured by the first two moments of θ.

a) Average values,

The curve displaying the first moments $<\theta>T$ and $<\theta>p$ derived from the distributions (3) and (4) for the log returns of transport and oil respectively, is plotted in figure 4. It can be seen that the linear dependence between the average values of θ is given by

$$<\theta>T = 2.5 + 1.8 <\theta>p \tag{7}$$

for $<\theta>p \geq 10$ corresponding to approximately 10 days.

b) Variances,

The second moments are related to the standard deviations $\sigma(s)$, again derived from (3) and (4). It can also be shown in figure 5 that there exists a domain of *s*, $0 \leq s \leq 0.04$, in which the variances are proportional

$$\sigma_T = -5 + 5.5\, \sigma_p. \tag{8}$$

for $1 < \sigma_p \leq 100$. This linear relation, verified in a rather large range of *s*, is important for economic reasons but for statistical reasons as well since no such relation does exist for the average values of θ. Now, when *s* increases, the size of the intervals decreases significantly, e.g. for *s* close to 0.1, n_θ is about 15, whereas n_θ is about 1000 for $s = 0.008$. However, despite the low value of n_θ, the accuracy of the numerical calculations of the first two moments is still acceptable.

4. Specific comments and concluding remarks

The first result stresses the differences of time properties related to the transport and oil prices. We emphasize that the power and exponential functions (3) correctly describe the distribution corresponding to the transport indices for the short intervals of time and long time intervals respectively. Concerning oil prices, the exponential function of (5) is an excellent approximation of the distribution.

These differences prove that despite a quite close aspect of the time series as shown in Figure 1, the time properties of the series of transport and oil prices are different.

In addition the results obtained provide information on the volatility of transportation related series and more precisely on the BDI. Firstly we note that the range of time intervals is wider for the transport indices, as compared with oil prices, which is a sign of higher volatility. Secondly, the exponential law describing the PDFs in the short time intervals, which we do not note on oil prices, testifies in the same extent of a higher volatility of the BDI as compared with oil prices

We have shown elsewhere that both series are dependent [5], notably from the cointegration and multivariable standpoints. However, the results described above reveal the dependence is not so strong.

Lastly, this method allows us to make some qualitative predictions: two close shocks on prices occur more frequently for the transport indices than for the oil prices.

Possible reasons to explain the different behaviors relate to: firstly the BDI is a spot index provided on a daily basis having experienced changes in its compilation methodology since its creation in early 1980's, which induce a high volatility, whereas oil price related index (EUBRCRDT) is an average value which smoothes price fluctuations.

From a methodological standpoint, we can conclude by noticing that the method based on the waiting time intervals is an alternative method to the autocorrelation function, which is inefficient for the time series stationnarized either by diffserence or by the log returns. It is also a complementary method to cointegration approach, which in our case had revealed that both series were cointegrated of order 1 [5].

Acknowledgments

The author is grateful to Professor J. Le Cacheux for his interest in this work. The author would also like to thank Professors G. Bresson and A. Charpentier for their advice and suggestions regarding the statistical tests utilized in this paper.

References

[1] Hummels D., Transportation Costs and International Trade in the Second Era of Globalization, Journal of Economic Perspectives, 21, No. 3, pp. 131-154, 2007

[2] Lee J.W., Lee K.E. and Rikvold P.A., Waiting-Time Distribution for Korean Stock-Market Index KOSPI, Journal of the Korean Physical Society, Vol. 48, pp. S123-S126, 2006

[3] Sabatelli L., et al., The Waiting-Time Distribution in financial markets, Euro Physics Journal B, Vol. 12, pp. 273-275, 2002

[4] Santhanam M.S. and Holger Kantz, Return interval distribution of extreme events and long term memory, Physical Review E, Vol. 78, No. 5, p. 78.05113, 2008

[5] Guigue J., Les coûts de transport maritime international de marchandises: évolutions et perspectives, Thèse de Doctorat, 2015

Transport Indices			
	h	pVal	ks
CDF of (3)	0	0.9360	0.1333
CDF of (4)	0	0.0591	0.40
Oil Prices			
	h	pVal	ks
CDF of (5)	0	0.9998	0.1063
CDF of (6)	0	0.4622	0.2632

Table 1 : Table of the results of the tests for the comparison of the CDFs.

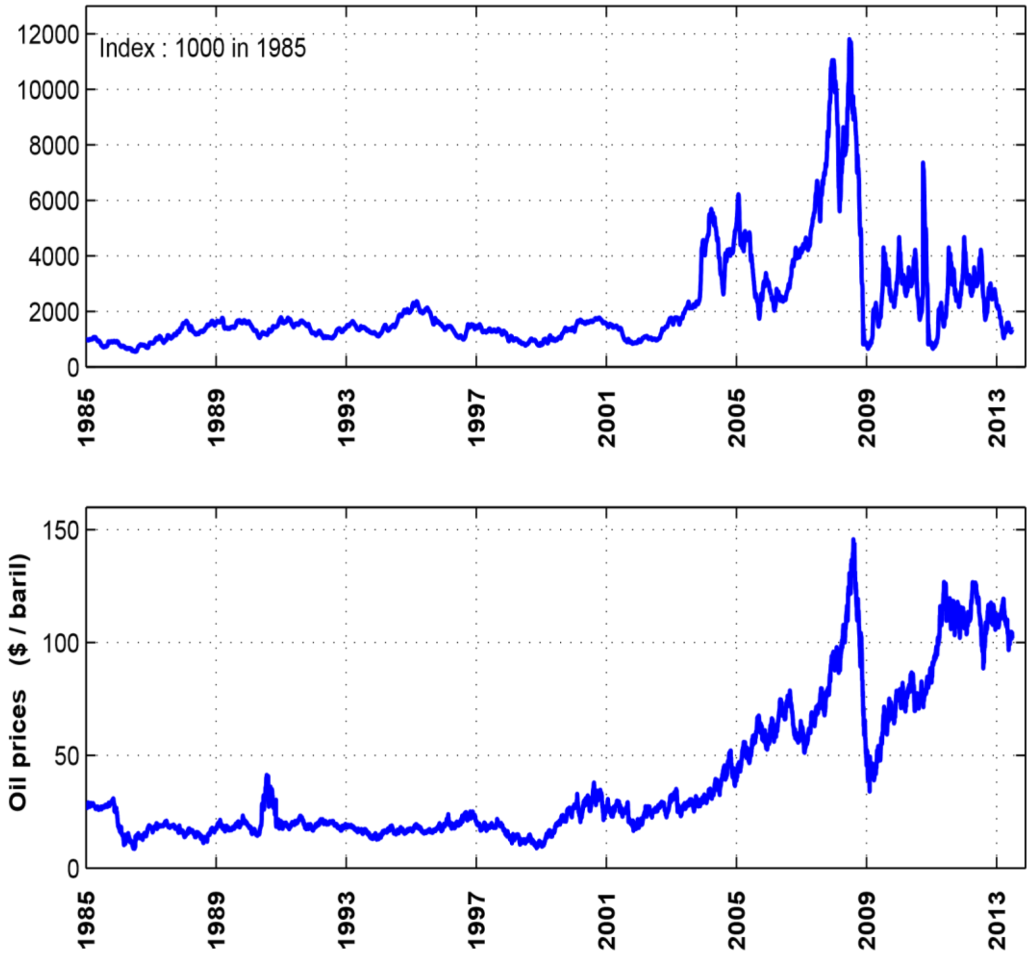
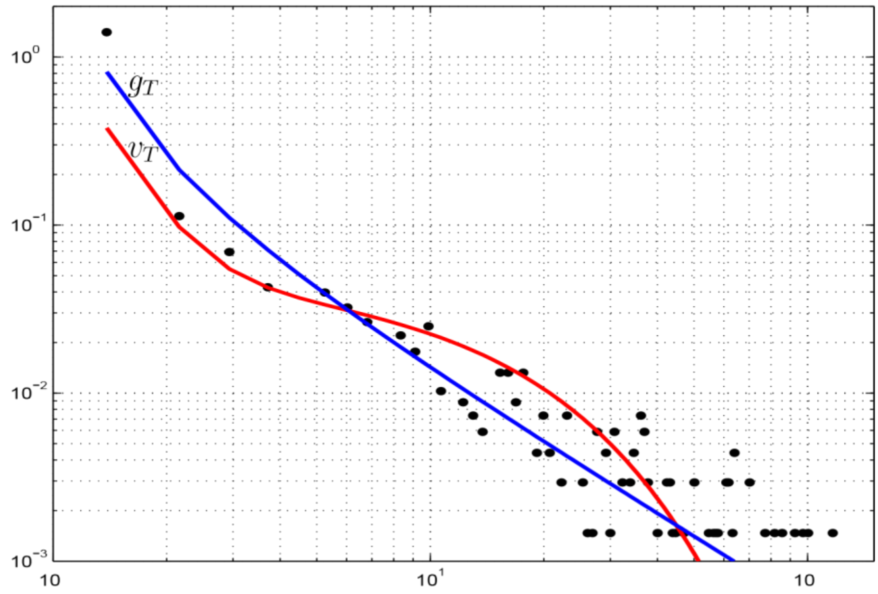
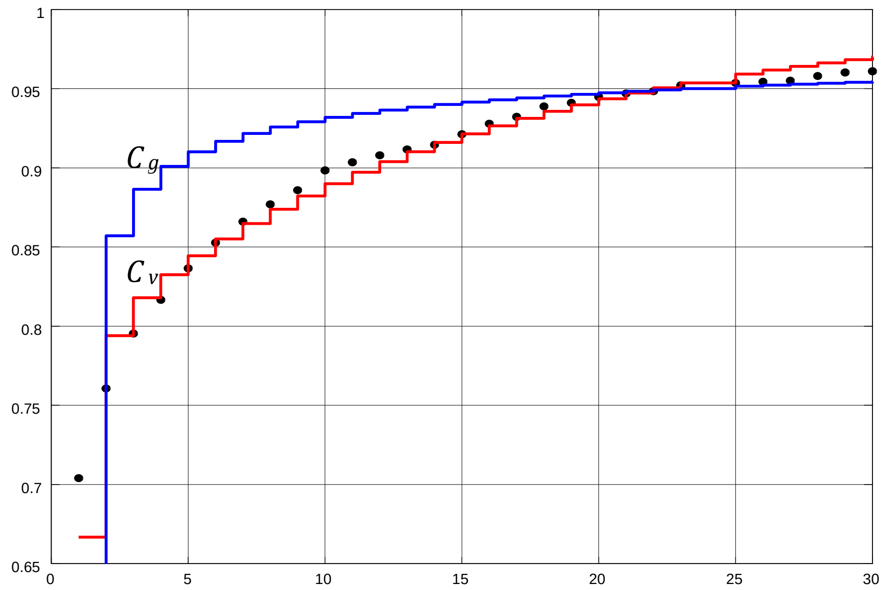


Figure 1: Sequence of time series of the transport and oil prices for the period 11 January 1985 - 30 May 2013.

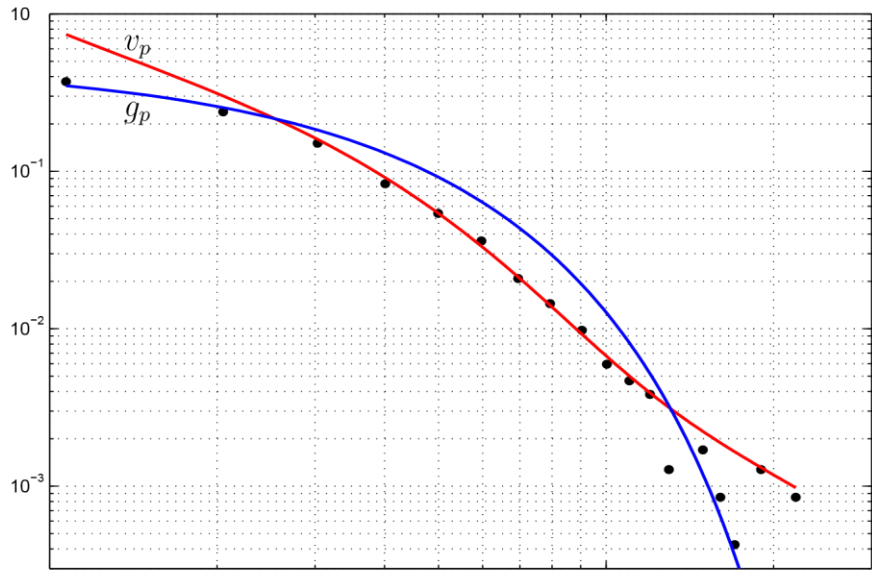


(a)

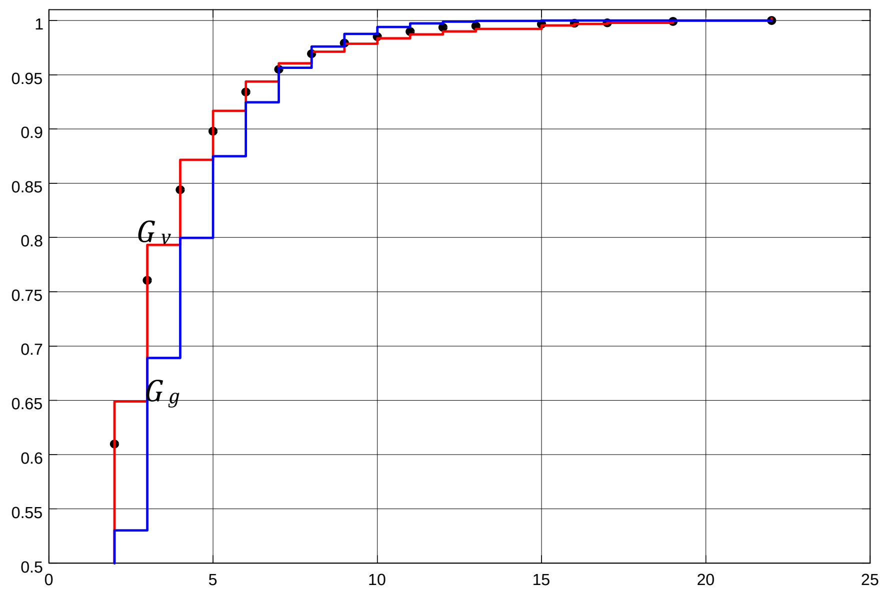


(b)

Figure 2: The transport indices. (*a*): PDFs of the waiting times $\{\theta\}$ in log-log coordinates. The PDF of data (points) and the PDFs given by equations (3) and (4), drawn in lines and quoted v_T and g_T respectively. (*b*): Cumulative Distribution Functions (CDF) in linear coordinates. CDF of data (points) and CDFs derived from the equations (3) and (4), quoted C_v and C_g respectively.



(a)



(b)

Figure 3: The oil prices. (*a*): PDFs of the waiting times $\{\theta\}$ in log-log coordinates. The PDF of data (points) and the PDFs given by equations (5) and (6), drawn in lines and quoted v_p and g_p respectively. (*b*): The Cumulative Distribution Functions (CDF) in linear coordinates. The CDF of data (points) and the CDFs derived from the equations (5) and (6), quoted G_v and G_g respectively.

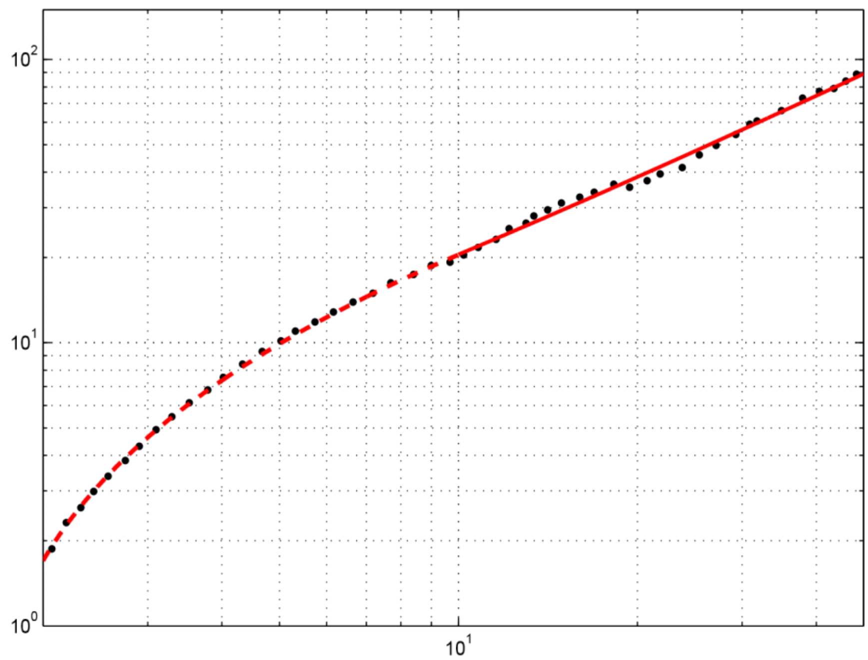


Figure 4: The variations in logarithmic scales of <θ>*T*, the average values of θ for the log returns of the transport indices, are plotted versus <θ>*p*, the average values of θ for the log returns of the oil prices (points). The best fits to the data are plotted in the interrupted and the continuous lines for short and long values of the average values of θ and given by equations (7) and (8) respectively.

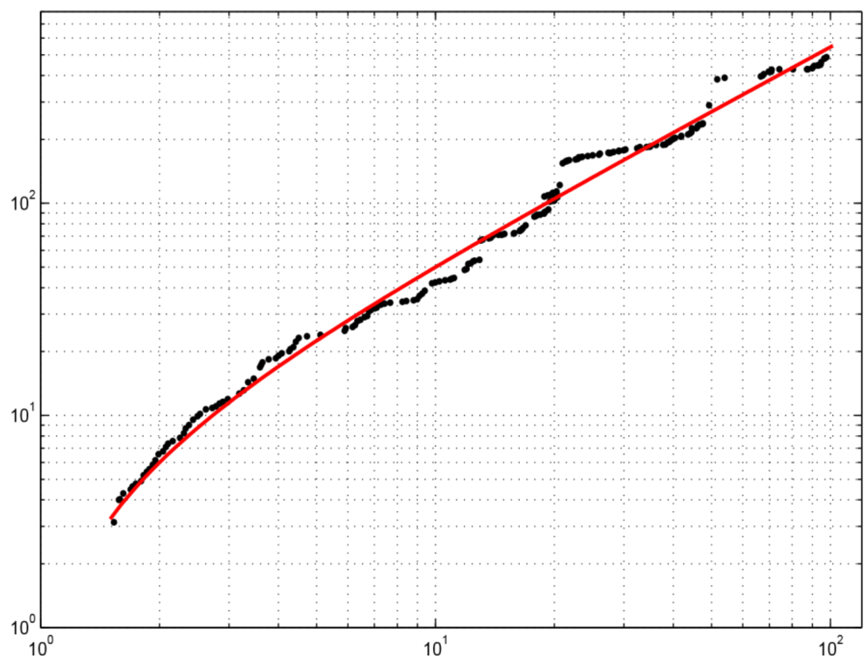


Figure 5: The variations in logarithmic scales of σ_T the standard deviations of θ, for the log returns of the transport indices are plotted versus σ_p, the standard deviations of θ for the log returns of the oil prices (points). The equation (8) is plotted in continuous line.