DID THE PSYCHOLOGICAL BARRIER OF US$100 CRUDE OIL PRICE HAVE AN IMPACT ON PRICE CLUSTERING BEHAVIOUR IN THE OIL FUTURES MARKET?

Murphy Lee (University of Technology, Sydney); Chee Jin Yap (Swinburne University, Melbourne); Paresh Kumar Narayan (Deakin University, Melbourne, email: narayan@deakin.edu.au)

ABSTRACT

In this paper, we test the price clustering theory in the crude oil futures market. Our novelty is that we consider the psychological barrier effect on price clustering and its determinants in the crude oil futures market. We find that prices tend to cluster on zero and five three weeks before the crude oil price reached the $100 per barrel mark and three weeks after the price rise. However, in the post- $100 price period, price clustering declines. We examine the determinants of price clustering and find that while size, volume, tick volume, and volatility have similar effects in terms of sign and significance in both periods, the spread variable is only significant (and positive) in the post-$100 rise period. The general decline in price clustering and the significance of the spread variable suggests that once a price psychological effect elapses, traders in the crude oil futures market behave differently than when the price approaches a high price level.

Keywords: Crude Oil; Futures Market; Psychological Barrier; Price Clustering.

JEL codes: C22, D81
1. INTRODUCTION

Pricing clustering at certain numbers has come to be known as the price clustering theory (Harris 1991). A wide range of empirical studies shows that stock prices cluster around certain round numbers; see for instance, Sonnemans (2006), Ascioglu et al. (2007), Cai et al. (2007), Brown and Mitchell (2008), and Ikenberry and Weston (2008).

The emergence of price clustering theory has provided a strong suggestion that prices are not uniformly distributed and do not follow a random walk, thus contradicting traditional finance theory. Therefore, this phenomenon has become a major topic of interest. A number of theories have been proposed to explain the reasons for price clustering and we discuss these briefly in the next section.

Odd pricing is widely acceptable in the marketing literature. The theory describes that sellers price their goods just below some round number (Sonnemans 2006). This can be casually observed where prices of certain goods are set at $9.99 instead of $10.00. This is so as consumers consider round number prices significantly higher than odd prices. Brenner and Brenner (1982) documented that humans view the first digits in numbers as of significant importance as compared to the last digits. Therefore, consumers are more likely to purchase goods with a price of $9.99 than similar goods that are priced at $10.00.

While our focus in this paper is on price clustering, our research question is novel, in that we ask: did price clustering behaviour change in the oil futures market when oil prices reached $100 per barrel? A second related contribution of this study is that we examine the determinants of price clustering in the market in a period before the crude oil price reached the $100 mark (pre-$100) and in a period after the oil price reached the $100 mark (post-$100). The idea behind this is to examine whether the psychological barrier of
reaching the $100 price level had any impact on market fundamentals affecting price clustering. These are fresh insights from the price clustering hypothesis tested on the oil futures market, for never in the history of the global economy has the oil price reached the $100 mark. However, the probability of the price reaching the $100 mark in the future is high given the scarcity of the oil resource. Hence, our study provides fresh evidence of investor behaviour when prices are about to reach, and when they actually reach, the $100 mark.

We organize the rest of the paper as follows. In the next section, we discuss the conceptual framework that motivates price clustering behavior. In Section 3, we discuss the estimation approach. In section 4, we discuss the data, while section 5 contains the results.

2. PRICE CLUSTERING HYPOTHESES

The main hypotheses proposed to explain the price clustering phenomenon in the literature are the preference hypothesis, the price resolution hypothesis, the negotiation hypothesis, the attraction hypothesis, the collusion hypothesis, and the cultural preference hypothesis. Mitchell (2001) stated that it is possible that a combination of hypotheses may explain price clustering instead of a single hypothesis.

Niederhoffer (1965) first proposed the preference hypothesis, which perceives people as having a preference for round numbers. This may result because people are more accustomed to round numbers. This is consistent with evidence from the psychology literature, which has found that people tend to round their answers in order to simplify
calculations and the degree of rounding is positively correlated with the degree of difficulty (Loomes 1998).

The price resolution hypothesis was proposed by Ball et al. (1985). It argued that prices cluster at particular focal points. The idea is that such clustering reduces search costs in the presence of market friction and uncertainty. Grossman et al. (1997) argued that valuation of asset is made more difficult when quotes and trades are infrequent or the market is illiquid. The uncertainty forces traders to round quotations and, therefore, results in greater price clustering. Similarly, Gwilym and Alibo (2003) pointed out that clustering increases with price volatility as traders reduce their exposure by transacting quickly, therefore relying on a less precise valuation process.

Harris (1991) proposed the negotiation hypothesis, which perceives clustering as a convenient practice to reduce negotiation costs. Prices will converge quicker as offers and counter-offers are restricted to a reduced set of prices, therefore the tendency for prices to cluster is significantly higher. The level of clustering is expected to be greater in high-priced stocks as the cost that traders expect from any rounding errors decreases with price.

The attraction hypothesis explains that investors might have a basic attraction to some numbers (Goodhart and Curcio 1991). Aitken et al. (1996) found people are more attracted to numbers zero and five. These findings are justified from the psychology literature, where Shepard et al. (1975) argued that some numbers are easier to process than others as less time and energy are spent to process these numbers because people tend to remember even numbers much faster than odd numbers. Ascioglu et al. (2007) suggested that people are attracted to numbers that end with zero and five because they are easier to recall.
The collusion hypothesis states that tacit collusion among dealers is highly possible when there is a defect in the market structure (Christie and Schultz 1994). Christie and Schultz (1994) explained that there is a possibility that dealers may collude implicitly in the form of using round numbers as ‘focal points’ to coordinate prices. Dealers are able to take advantage of this situation if only a small price set is being traded on a wider bid/ask spread, thus earning more commissions, causing prices to cluster.

The most recent hypothesis relating to price clustering is the cultural influence hypothesis proposed by Brown et al. (2002), which essentially perceives some numbers as having cultural significance. Therefore, it is expected that prices have a tendency to avoid or cluster at those significant numbers. Brown et al. (2002), Cai et al. (2007), and Brown and Mitchell (2008) have found that prices tend to cluster at the number ‘eight’ in the Hong Kong and Chinese stock markets, thus signifying that cultural influences do play an important role in price clustering behaviour.

3. APPROACH AND MODEL

In this section, we discuss our estimation approach and the estimable model relating to the determinants of price clustering.

3.1. Approach

We use three different methods to count the frequency of the transacted price over the entire sample period based on the last digit. These tests are the standard Chi-square goodness-of-fit statistic, the Hirshmann-Herfindal Index (HHI), and the Standardised Range (SR) test. In this sub-section, we briefly discuss each of these.

---

1 In the Chinese culture, the number ‘8’ is considered to be a lucky number because it sounds like ‘prosperity’ in Cantonese, whereas the number ‘4’ is considered to be an unlucky number as it sounds like ‘die’ in Cantonese.
3.1.1. Chi-square goodness-of-fit statistic

Niederhoffer (1965), Chueh (2000), Chung and Chiang (2006), and Ikenberry and Weston (2007) have used the standard chi-square goodness-of-fit test statistic to determine if the frequency distribution of the last digit of the traded price followed a uniform distribution. The chi-square goodness-of-fit statistic is calculated as the summation of all squared deviations between the observed level of price clustering and the expected level of price clustering under uniform distribution:

\[ W = \sum_{i=1}^{k} \frac{(O_i - A_i)^2}{A_i} \]

Where \( W \) is chi-square distributed with \((k-1)\) degrees of freedom, \( O_i \) is the observed frequency from the last digit of the transacted prices, and \( A_i \) is the expected quotation frequency under the uniform distribution. If \( W \) is equal to zero, this signifies that there is no price clustering. A large value for \( W \) represents a significant deviation from uniform distribution, thus this is evidence of price clustering.

3.1.2. Hirshmann-Herfindal Index

The HHI is a measure of price concentration. It has been widely used in the literature; see, for instance, Grossman et al. (1997), Chung and Chiang (2006), and Ikenberry and Weston (2007) to study price clustering. The HHI sums the squared values of the market shares of all market participants, and is calculated as:

\[ HHI = \sum_{i=1}^{t} (f_i)^2 \]

Where \( f_i \) is the frequency of trades, measured in percentage that occurs at fraction \( i \) and which are the possible ticks. Under the null hypothesis of no price clustering, the HHI
equals to $1/k$, which is 0.1 as there are 10 observations in total. Therefore, if the HHI is equal to 0.1, this signifies that there is no price clustering.

3.3.3. Standardised Range

An alternative measure of price concentration is the SR, which was proposed by Grossman et al. (1997). The SR is computed through taking the difference between the highest and lowest quotation frequencies and dividing it by the expected quotation frequency under the null distribution. The SR is computed as:

$$SR = \frac{HF - LF}{A_i}$$

Where $HF$ is the highest recorded quotation frequency, $LF$ is the lowest recorded quotation frequency, and $A$ is the expected quotation frequency under the uniform distribution. This measure allows us to rank the degree of price clustering between different markets or trading periods. The market with a higher estimated value of the SR would have a higher degree of price clustering.

3.2. Model of the determinants of price clustering

We use a multivariate model to explore the interaction between price clustering in the crude oil futures market and its determinants, namely the bid-ask spread, trade size, volume, tick volume, and volatility.$^2$ Our proposed regression model for price clustering in the crude oil futures has the following form:

$$PC_t = \beta_0 + \beta_1 SPR_t + \beta_2 SIZE_t + \beta_3 VOL_t + \beta_4 TIC_t + \beta_5 SIG_t + \epsilon_t$$

$^2$ Open interest is excluded from the analysis as it was found to have a high correlation with volume; this approach is similar to Schwartz et al. (2004).
Here, \( t \) represents time, \( PC \) is the percentage of transacted prices that occur at zeros and fives for each half-hourly interval, \( SPR \) is the bid-ask spread, \( SIZE \) is the trade size, \( VOL \) is the half-hourly volume, \( TIC \) is the tick volume, \( SIG \) is the intraday half-hourly volatility, and \( \varepsilon \) is the residual term.

The bid-ask spread widens when the market becomes more illiquid, leading to greater uncertainty of the true price. As a result, investors will transact at a certain price, leading to price clustering. The realised bid-ask spread is calculated as the average absolute transacted price change in the opposite direction in an interval (see Wang and Michalski, 1994). It is estimated by computing returns from transacted prices; then sets of price changes that exhibit price continuity are dropped. Following this, absolute values of the price changes that are reversed are taken and averages of these absolute values are computed.\(^3\)

The trade size variable is obtained by averaging trading size of all transactions during the half-hour period. As Easley and O’Hara (1987) pointed out, informed investors sometimes transact in large orders. Informed investors will then quote a more clustered price in large orders to hide their knowledge (Aitken 1996). This suggests that clustering should share a positive relationship with trade size. This intuition is refuted by the negotiation hypothesis, which perceives that investors trading on smaller orders will limit the numbers of bids and asks. Thus, trades will be transacted at a more clustered price. By trading with a smaller price set, traders are able to lower their costs of negotiation, whereas investors who transact on larger orders will find it more beneficial to negotiate on a wider range of prices. Butler and Loomes (1988) explained that small and uninformed investors trade on clustered prices in order to hide in the crowd as well as being uncertain about the true

---

value of the asset. It follows that if the negotiation hypothesis is valid, clustering should be inversely correlated with trade size.

Volume is usually used as a proxy for liquidity. The more frequently an asset is traded, the more precisely its true value is known. Thus, clustering of prices should be less when volume increases. We compute volume by inversing the square root of the average number of trades. Harris (1991) explained that due to information-theoretic considerations, each transaction conveys information about the underlying value of the asset coupled with some noise, thus these standard errors are proportional to the inverse square root of the number of transactions.

A variant of the volume variable, the tick volume, also captures trading activities (Huimin and Shumei 2006). This independent variable measures the total number of ticks and is formulated by summing all price changes at half-hourly periods. The expected relationship between tick volume and price clustering is negative.

The volatility variable is measured using Parkinson’s (1980) extreme value estimator. Alizadeh et al. (2002) found this measure of volatility to be a highly efficient proxy for volatility and is also robust to microstructure noise. Volatility is computed as follows:

\[ \text{SIG}_t = 0.361 \left[ \log \left( \frac{H_t}{L_t} \right) \right]^2 \]

Where, \( H \) and \( L \) represent the highest and lowest prices during the time period \( t \). In periods of high volatility, the true value of an asset is highly uncertain to investors, thus leading them to trade at round numbered prices.
3.3. Testing for Endogeneity in Explanatory Variables

Gwilym et al. (1998) and Chung and Chiang (2006) found that the level of price clustering could also influence the bid-ask spread, suggesting the possibility of an endogenous relationship. Following these studies, we used the Hausman (1978) specification test to test for endogeneity. The residual term, \( \mu \), from the bid-ask spread regressed against all other explanatory variables is obtained, then the Hausman (1978) specification test is estimated as follows:

\[
SPR_t = \alpha_0 + \alpha_1 SIZE_t + \alpha_2 VOL_t + \alpha_3 TIC_t + \alpha_4 SIG_t + \mu_t
\]

The original price clustering determinants model is then augmented with the residual term as follows:

\[
PC_t = \beta_0 + \beta_1 SPR_t + \beta_2 SIZE_t + \beta_3 VOL_t + \beta_4 TIC_t \\
+ \beta_5 SIG_t + \beta_6 \mu_t + \epsilon_t
\]

Our finding differs from Gwilym et al. (1998) and Chung and Chiang (2006) in that we do not find evidence of causation running from price clustering to the bid-ask spread.\(^4\)

3.4. Generalised Least Squares Estimator

Consistent with this literature, we use a generalised least squares (GLS) approach, which essentially transforms the regression into one with homoskedastic errors. Three steps are carried out for this transformation, starting with estimating the variance function, followed by calculating the transformed variables, and then using the least squares to estimate the

\(^4\) For the purpose of brevity, we have omitted the test results here. However, results are available upon request from the authors.
transformed model. These steps are widely known, so we do not report them here to conserve space.

4. DATA

This paper uses the crude oil futures prices data obtained from the Securities Industry Research Centre of Asia-Pacific (SIRCA). The analysis of price clustering is based upon tick-by-tick transaction price from the electronically traded contracts traded in NYMEX. Following the work of Chueh (2000) and Chung and Chiang (2006), we picked the nearest to maturity futures contracts as they are the most actively traded. The first nearby futures contract is perhaps a better proxy for spot price as is not influenced by considerations of physical commodities such as shipping (Geman and Kharoubi 2008).

To avoid the shrinking of trading volume and extreme price volatility as the maturing contract approaches settlement date, we have excluded the last trading month from the analysis.

The sample period covers three weeks period prior to, and following, the date when the crude oil futures prices peaked at $100 for the first time, which was on the 27th February 2008. Thus, the sample period ranges from 6th February 2008 to 19th March 2008. This particular sample period is chosen because we examine two periods: one, three weeks prior to the price hitting the US$100 mark and two, the three week period when price was over US$100. In addition to obtaining tick-by-tick transaction prices, the dataset also includes bid-ask prices, quotes and traded volume. All these data are also sourced from the SIRCA.

The electronic-trading virtually transacts for 24 hours a day. We consider data from 0900 to 1430 GMT for the period 6th February 2008 to the 19th March 2008. This ensures that we have a total of 36008 observations for the electronic-trading of crude oil futures.
For the purpose of the regression analysis, the data sample is divided into half-hourly intervals, leading to 322 intervals in total. We excluded 19 intervals because no trading occurred during these times. The failure to exclude outliers will lead to inaccurate coefficient estimates and the inclusion of such intervals will cause the explanatory variables to have extreme values that do not represent the true activities of the market. Therefore, we excluded the half-hourly intervals that record infrequent trading. This ensures a consisting of 34,151 observation, equivalent to 164 half-hourly intervals.

5. RESULTS

5.1. Descriptive statistics

Some commonly used descriptive statistics are reported in Table 1. These relate to the full sample period, the pre-$100 per barrel price period, and the post-$100 per barrel price period. We notice that in the post-$100 price period, the standard deviation is lowest compared with the pre-$100 price period and the full sample. This seems to suggest that market uncertainty subsided when prices eventually reached the $100 per barrel mark.

Three weeks prior to the price reaching the $100 per barrel, market volatility, as measured by the standard deviation, was around 76 per cent higher than what it was three weeks after the price had risen above the $100 per barrel mark. This indicates that the crude oil futures market was different in the two price periods. We explore whether this difference existed in terms of the degree of price clustering and the determinants of price clustering.
5.2. Evidence on price clustering

We begin the discussion of the results with an analysis of the frequency distribution of price clustering. In Figure 1, we plot the frequency distribution of the last digit of closing prices. Strong presence of price clustering is noticed. Evidence suggests that prices tend to cluster on 0 and 5.

Table 2 presents a more formal test of price clustering in the form of statistical measures. Table 2 reports the frequency distribution of the last digit, the HHI and the SR for the crude oil futures contracts. The results reveal that the last digit 0 occurs most frequently – 15.79 per cent of the time. The last digit 5 is the next most common, recording a frequency of 12.71 per cent. All other possible occurrence of last digits recorded a frequency of no more than 10.5 per cent.

5.2. Determinants of price clustering

The determinants of price clustering for the full sample period are reported in Table 3. We find that except for volatility, all variables turn out to be statistically significant and have the expected signs. Size, volume, and tick volume are statistically significant at the 1 per cent level, while spread is statistically significant at the 5 per cent level.

Consistent with Harris (1991), Aitken et al. (1996), and Chung and Chiang (2006), the results show that clustering increases with the bid-ask spread. This suggests that investors do trade on familiar or round numbers when the crude oil futures market has a larger bid-ask spread. Thus, they face a higher level of uncertainty in determining the true price of the asset, which leads them to trade at a more clustered price.
The trade size variable is found to have a positive and significant effect on clustering. This suggests that clustering is likely to be driven by large traders as they are more likely to have the capacity to transact in larger trade sizes. Additionally, this finding also rejects Harris’s (1991) negotiation hypothesis. Our study supports the notion by Aitken et al. (1996) that these informed investors quote a more clustered price in large orders to hide their knowledge of the market.

The results also show that a drop in both volume and tick volume will cause clustering to increase. This result is generally consistent with the findings of Harris (1991), Aitken et al. (1996), Gwilym et al. (1998), and Ikenberry and Weston (2007).

The impact of volatility on price clustering is found to be positive, a result consistent with Harris (1991), and Aitken et al. (1996), and Chung and Chiang (2006). However, this result is statistically insignificant. This means that volatility does not affect the level of price clustering in the crude oil futures market, a result which is highly inconsistent with most studies on price clustering in other markets. Two other volatility measures, namely the ranged-based procedure of Alizadeh et al. (2002) and the Garman and Klass (1980) method were also computed and used in this analysis. It was concluded that all volatility measures yielded similar results; these results are available from the authors upon request.

5.3. Price clustering when prices are trading around $100

Table 4 presents the results of the frequency distribution of price clustering when prices are trading before and after $100. The results show that the last digit of zero occurs most frequently followed by the last digit as five in both periods. This attraction on digits zero and five is consistent with our results for the full sample period. Three weeks prior to the oil price reaching the $100 per barrel mark, around 19 per cent of prices clustered on the last digit of zero. However, three weeks after the crude oil price rose to $100 per barrel
mark, while prices still clustered strongly on zero, it declined to around 15 per cent. Similarly, we observe that the second most common last digit on which prices clustered was five: before the price hit the $100, around 16 per cent of the prices clustered at five but after it reached $100 only 12 per cent of the prices clustered on five.

This finding is corroborated by the standardised range statistic, which suggests that price clustering on last digits of zero and five were greater in the pre-$100 price period than in the post-$100 price period. There are two significances of these findings. First, it signifies that price clustering persisted in the crude oil futures market despite the crude oil prices hitting the physiological barrier of the $100 per barrel mark. Second, it suggests that there is greater price uncertainty among investors in the pre-$100 price period. This uncertainty is not surprising because of the expectation that prices will continue to increase and will eventually reach the $100 per barrel mark for the first time in history.

This evidence generally supports price barrier theories: that there are significant changes in investors’ trading behaviour when prices are trading around a significant price level or a psychological barrier price level.

5.4. **Determinants of price clustering when prices are trading around $100**

The results on the determinants of share price clustering based on the generalised least squares approach, as explained earlier, are reported in Table 5. Essentially, given the objective of this paper, we compare the determinants of price clustering in the two periods: the pre-$100 per barrel price period and the post-$100 per barrel price period. Column 2 reports the results for the pre-$100 price period while column 3 reports the results for the post-$100 price period.
The results are as follows. In the pre-$100 price period, three of the five variables, namely size, trading volume, and tick volume, have the expected signs and are statistically significant. Size has a statistically significant positive effect on price clustering while volume and tick volume have statistically significant negative effects on price clustering on the crude oil futures market.

By comparison, in the post-$100 price period, more variables are statistically significant: four out of the five variables are statistically significant. Spread which was statistically insignificant in the pre-$100 price period, becomes significant in the post-$100 price period, and as expected has a positive effect on price clustering.

In sum, we notice four features of the results on the determinants of price clustering in the oil futures market when we compare the pre-$100 price period with the post-$100 price period. First, the spread variable is only statistically significant in the post-$100 price period. Second, volatility while having a negative effect on price clustering is statistically insignificant in both periods. Third, size, volume, and tick volume have the expected signs and are statistically significant in both periods.

Fourth, we notice that the magnitude of the impacts of some variables is greater in the pre-$100 price period. For example, the coefficients on trade size and volume are significantly higher in the pre-$100 price period.

Doucouliagos (2004) argues that investors are likely to become more anxious and uncertain when prices reach a certain significant (or psychological barrier price) price level. Therefore, it can be concluded that when crude oil futures prices are approaching a significant price level of US$100 per barrel, there is greater uncertainty among investors.
which causes them to trade at a more clustered price which is why a greater level of price clustering and a larger impact of some of the determinants of price clustering are observed in the pre-$100 price period. Doucouliagos (2004) further explained that generally once prices crash through these price levels, there is a strong indication that the directional move is powerful and investors will expect prices to continue in its current direction. The results also support this because a lower level of clustering and the determinants having a weaker impact on price clustering are observed as investors are more certain of prices in the post-$100 price period.

6. CONCLUDING REMARKS

The goal of this paper was to apply the price clustering theory in the crude oil futures market. The novelty of this paper was that we considered the psychological barrier effect on price clustering and its determinants in the crude oil futures market. Essentially, given that the crude oil price hit the $100 per barrel mark for the first time in history, we considered the price of $100 per barrel as a psychological price barrier. This led us to divide the futures contract price into two sub-samples: a pre-$100 price period and a post-$100 price period, each consisting of a three week window. We then examined evidence for price clustering in these two periods and considered the determinants of price clustering in the two periods. We unraveled a number of interesting findings, which provide insights on the behavior of traders prior to the price hitting a psychological barrier of $100 and immediately after the price rise.

Our main findings can be summarized as follows. First, while prices tend to cluster on digits zero and five in both the pre-$100 price and post-$100 price periods, clustering declines in the post-$100 price period. This implies less panic and subdued volatility of the market in the post-$100 price period. Second, we find that in both periods, size has a
statistically significant positive effect and volume and tick volume have statistically significant negative effects on price clustering. However, the magnitudes of the impacts are relatively small in the post-$100 price period. This signifies the subdued nature of the market after the crude oil price had hit the $100 per barrel mark. Third, importantly the spread variable is only significant in the post-price rise period. This implies that when prices had risen above the US$100 mark, the bid-ask spread had increased. Fourth, volatility is statistically insignificant in both periods. Taken together, our results provides evidence that investors tend to behave differently once crude oil prices reach a perceived psychological barrier.
REFERENCES


Table 1: Descriptive Statistic of the Crude Oil Futures Prices

This table presents the descriptive statistics of the crude oil futures prices ranging from the 6th February 2008 to the 19th March 2008 with trading hours ranging from 0900 to 1430 daily. All statistics are obtained from the series of the crude oil futures prices. These statistics include the number of observations (N) mean, median, standard deviation (Std. Dev.), minimum (Min.), maximum (Max.), and skewness.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>Std. Dev.</th>
<th>Min.</th>
<th>Max.</th>
<th>Skewness</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Traded Contract</strong></td>
<td>36008</td>
<td>102.75</td>
<td>104.22</td>
<td>4.270</td>
<td>86.26</td>
<td>108.65</td>
<td>-1.597</td>
</tr>
<tr>
<td><strong>Contract Trading before $100 (over a 3-week window)</strong></td>
<td>5828</td>
<td>94.86</td>
<td>97.08</td>
<td>3.744</td>
<td>86.26</td>
<td>99.17</td>
<td>-0.844</td>
</tr>
<tr>
<td><strong>Contract Trading after $100 (over a 3-week period)</strong></td>
<td>30180</td>
<td>104.28</td>
<td>104.60</td>
<td>2.167</td>
<td>98.41</td>
<td>108.65</td>
<td>-0.566</td>
</tr>
</tbody>
</table>
Table 2: Comparison of Price Clustering in Electronic-Traded Crude Oil Futures Contract

Cell frequencies are determined based on the last digit and their distribution and is measured with the Chi-square goodness-of-fit statistic. The HHI and the SR are also used to test for price concentration. Numbers in the parentheses of the row labelled HHI are the expected HHI values under the null hypothesis of no price clustering.

<table>
<thead>
<tr>
<th>Last Digit</th>
<th>Frequency</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>5685</td>
<td>15.79</td>
</tr>
<tr>
<td>1</td>
<td>3102</td>
<td>8.61</td>
</tr>
<tr>
<td>2</td>
<td>3216</td>
<td>8.93</td>
</tr>
<tr>
<td>3</td>
<td>3115</td>
<td>8.65</td>
</tr>
<tr>
<td>4</td>
<td>3261</td>
<td>9.06</td>
</tr>
<tr>
<td>5</td>
<td>4575</td>
<td>12.71</td>
</tr>
<tr>
<td>6</td>
<td>3144</td>
<td>8.73</td>
</tr>
<tr>
<td>7</td>
<td>3052</td>
<td>8.48</td>
</tr>
<tr>
<td>8</td>
<td>3642</td>
<td>10.11</td>
</tr>
<tr>
<td>9</td>
<td>3216</td>
<td>8.93</td>
</tr>
</tbody>
</table>

No. of obs. 36008  100.00

Goodness of fit ($\chi^2$) 0.052
HHI 0.105 (0.10)
Standardised range 0.731
Table 3: GLS Estimation Results on Price Clustering

The overall sample period covers 30 trading days, ranging from the 6th February 2008 to the 19th March 2008 at trading times of 0900 to 1430 daily. This has produced a total of 164 half-hourly intervals. The dependent variable is the percentage of clustering at zero and five at half-hourly intervals; SPR is the bid-ask spread obtained by using the estimator suggested by the Commodity Futures Trading Commission; SIZE is the trade size calculated through averaging the trading size of all transactions; VOL is the half-hourly volume, computed by taking the inverse square root of the average number of trades per half an hour and TIC is the total number of ticks or price changes for each half-hourly interval. SIG is the half-hourly volatility, calculated based on the Parkinson (1980) extreme value estimator. Values in parentheses are the p-values and ***,**,* denote statistical significance at the 1%, 5% and 10% levels, respectively.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-89.354***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
</tr>
<tr>
<td>SPR</td>
<td>0.641**</td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
</tr>
<tr>
<td>SIZE</td>
<td>0.944***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
</tr>
<tr>
<td>VOL</td>
<td>-3.498***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
</tr>
<tr>
<td>TIC</td>
<td>-1.628***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
</tr>
<tr>
<td>SIG</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>(0.940)</td>
</tr>
<tr>
<td>R²</td>
<td>0.266</td>
</tr>
</tbody>
</table>
Table 4: Comparison of Price Clustering when Prices are Trading Before and After US$100

Cell frequencies are determined based on the last digit and their distribution is measured with the Chi-square goodness-of-fit statistic. The HHI and the SR are also used to test for price concentration. Numbers in the parentheses of the row labelled HHI are the expected HHI values under the null hypothesis of no price clustering. The sample period when prices are trading before $100 ranges from 6th February 2008 to 26th February 2008. The sample period when prices are trading after $100 ranges from 27th February 2008 to 19th March 2008.

<table>
<thead>
<tr>
<th>Last Digit Of Prices</th>
<th>Before $100</th>
<th>Frequency</th>
<th>(%)</th>
<th>After $100</th>
<th>Frequency</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1111</td>
<td>19.06</td>
<td></td>
<td>4574</td>
<td>15.16</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>463</td>
<td>7.94</td>
<td></td>
<td>2639</td>
<td>8.74</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>524</td>
<td>8.99</td>
<td></td>
<td>2692</td>
<td>8.92</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>445</td>
<td>7.64</td>
<td></td>
<td>2670</td>
<td>8.85</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>413</td>
<td>7.09</td>
<td></td>
<td>2848</td>
<td>9.44</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>921</td>
<td>15.80</td>
<td></td>
<td>3654</td>
<td>12.11</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>458</td>
<td>7.86</td>
<td></td>
<td>2686</td>
<td>8.90</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>401</td>
<td>6.88</td>
<td></td>
<td>2651</td>
<td>8.78</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>558</td>
<td>9.57</td>
<td></td>
<td>3084</td>
<td>10.22</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>534</td>
<td>9.16</td>
<td></td>
<td>3682</td>
<td>8.89</td>
<td></td>
</tr>
<tr>
<td><strong>No. of obs.</strong></td>
<td><strong>5828</strong></td>
<td><strong>100.00</strong></td>
<td></td>
<td><strong>30180</strong></td>
<td><strong>100.00</strong></td>
<td></td>
</tr>
</tbody>
</table>

Goodness of fit ($\chi^2$)  
- Before $100$: 0.150  
- After $100$: 0.039  

HHI  
- Before $100$: 0.115 (0.10)  
- After $100$: 0.104 (0.10)  

Standardised range  
- Before $100$: 1.218  
- After $100$: 0.641
Table 5: GLS Estimation Results on Price Clustering when Prices are Trading Before and After US$100

The overall sample period covers 30 trading days, ranging from the 6th February 2008 to the 19th March 2008 at trading times of 0900 to 1430 daily. This has produced a total of 53 half-hourly intervals or 4,956 observation points when prices are trading in the pre-$100 price period and 111 half-hourly intervals or 29,195 observation points when prices are trading in the post-$100 price period. The dependent variable is the percentage of clustering at 0 and 5 at half-hourly intervals; SPR is the bid-ask spread obtained by using the estimator suggested by the Commodity Futures Trading Commission; SIZE is the trade size calculated through averaging the trading size of all transactions; VOL is the half-hourly volume, computed by taking the inverse square root of the average number of trades per half an hour; TIC is the total number of ticks or price changes for each half-hourly interval; and SIG is the half-hourly volatility, calculated based on the Parkinson (1980) extreme value estimator. Values in parentheses are the p-values and ***, **, * denote statistical significance at the 1%, 5% and 10% levels, respectively.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Before $100</th>
<th>After $100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-75.709*</td>
<td>-105.396***</td>
</tr>
<tr>
<td></td>
<td>(0.082)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>SPR</td>
<td>0.154</td>
<td>0.732**</td>
</tr>
<tr>
<td></td>
<td>(0.798)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>SIZE</td>
<td>1.628***</td>
<td>0.820**</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.013)</td>
</tr>
<tr>
<td>VOL</td>
<td>-3.231***</td>
<td>-3.901***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>TIC</td>
<td>-1.972***</td>
<td>-1.812***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>SIG</td>
<td>-0.409</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td>(0.117)</td>
<td>(0.103)</td>
</tr>
<tr>
<td>R²</td>
<td>0.62</td>
<td>0.248</td>
</tr>
</tbody>
</table>
Figure 1: Frequency Distribution of Prices

This figure plots a histogram of transacted prices based on the frequency of the last digit of the transacted price. Reported frequencies are computed over the sample period from 6th February 2008 to 19th March 2008 at trading hours of 0900 to 1430.