National
Electricity
Market of
Singapore

Occasional Paper by
Market Surveillance and Compliance Panel

HOW MARKET FUNDAMENTAL FACTORS AFFECT ENERGY PRICES IN THE NEMS – AN ECONOMETRIC MODEL

16 July 2007

Contents

EXECUTIVE SUMMARY **3**
INTRODUCTION **5**
CURRENT METHODOLOGY **5**
COMPARISON OF CURRENT METHODOLOGY WITH ECONOMETRIC MODELLING **6**
ECONOMETRIC MODELLING FOR THE NEMS **7**
ESTABLISHING THE QUANTITATIVE RELATIONSHIP BETWEEN MARKET FUNDAMENTAL FACTORS AND ENERGY PRICES **7**
Model Specifications 7
Modelling 7
Sample period 8
Variables 8
Functional form of the econometric model 10
Model Equation 10
Econometric Model Results 10
Explanatory Variables 11
Dummy Variables 12
Robustness of the Model 12
Conclusion on the Econometric Model 13
IDENTIFICATION OF PRICE OUTLIERS **13**
Model Methodology 13
Frequency of Outliers and Possible Reasons 16
Examples of Price Outliers 16
19 March 2006 16
12 August 2006 16
21 December 2006 17
CONCLUSION **17**
APPENDIX A **19**
Regression Analysis 19
How to Run a Regression? 19
APPENDIX B **21**
Correlation between Variables 21
Effects of Correlation between Variables 21
APPENDIX C **23**
Elasticity 23
APPENDIX D **25**
Robustness of the Model 25
APPENDIX E **27**
Heteroskedasticity 27
REFERENCES **28**

EXECUTIVE SUMMARY

Since 2003, the MSCP has used a static model to identify outlier prices. Spot prices are compared to a benchmark price and prices higher than the benchmark are identified as outliers. This model relies on the assumption that high prices provide the first indication of inefficient market outcomes.

In a competitive market, efficient prices should be the outcome of market fundamental factors which reflect normal demand and supply conditions. To understand the relationship between such factors and the energy prices in the NEMS, the MSCP has therefore developed an econometric model which takes into account such factors. Once developed, the model may be used as a benchmark to assist in the identification of outliers.

The model will serve two objectives:

1. Quantitative relationship - The model will enable the MSCP to understand how different market fundamental factors have quantitatively affected energy prices in the past.

2. With an understanding of the historical relationship between market fundamental factors and energy prices, the MSCP will be able to use the econometric model to estimate energy prices in the context of varying demand and supply conditions. To provide a margin for error, the estimated prices can be expanded to include an upper and lower price band. Prices falling outside of these bands can be identified as outliers that warrant further attention.

The econometric model is estimated based on daily average data from January 2003 to December 2006. The dependent variable for this model is the daily average energy price. The explanatory variables included in the econometric model are:

- Combined cycle gas turbine (CCGT) supply
- Steam turbine (ST) supply
- Supply cushion
- Percentage of offers at $100/MWh or less
- Demand
- Reserve cushion
- Fuel oil price
- CCGT planned outages
- Forced outages.

The MAU recognises that energy prices can differ between business and non-business days. Hence, the model has been augmented with month, day-of-week, annual dummy variables to take into account this temporal variation.

The estimation results of the econometric model with respect to the quantitative relationship between energy prices and the explanatory variables identified are set out below[1]:

Variable	Coefficient
CCGT supply	-0.016
ST supply	-0.003
Supply cushion	-0.994
Offers	-1.165
Demand	0.101
Reserve cushion	-0.303
Lag of fuel oil price	0.415
CCGT planned outages dummy A	-0.015
CCGT planned outages dummy B	0.103
Forced outages dummy	0.046

Tests have also been done to ensure the robustness and statistical soundness of the model. The explanatory variables included in the model are able to explain 78 percent of the variation in energy prices.

Based on the econometric model and the price bands constructed, outliers are identified for closer study.

The results from the econometric model suggest that simple market fundamental factors included in the model are able to explain a significant proportion of the variation in daily average energy prices between 2003 and 2006.

Divergence of the actual energy price beyond a reasonable price range may trigger closer scrutiny of market behaviour.

Recognising that the model described in this paper is a useful starting point, outliers which can be explained by normal demand and supply conditions will also prompt further work in refining the model to sharpen its analytical capabilities. Over time, the model can continue to be improved by including more market fundamental factors that have an effect on prices. With data collected over a longer period of time, the accuracy of the estimated results may improve. It will also enable the model to be evaluated.

[1] The results can be interpreted in the following manner: A 10 percent increase in CCGT supply resulted in a 0.16 percent decrease in the energy price etc.

INTRODUCTION

1. Under section 4.1.11 of Chapter 3 of the Singapore Electricity Market Rules, the Market Surveillance and Compliance Panel (MSCP) monitors conduct and activities in the National Electricity Market of Singapore (NEMS) to assess whether the underlying structure of the wholesale electricity markets is consistent with the efficient and fair operation of a competitive market.

2. This role requires focus on price outcomes and systematic analysis of prices. In the shorter run, an efficient price signal enables power plants to be able to respond to changes in demand. Prices during peak hours should also signal the profitability of peaking units which encourages more generation in order to meet demand. Therefore, in the long run, market participants can plan their generation investments based on their expectations of future energy prices.

3. While a good price signal is the desired outcome of a liberalised market, it is not a guaranteed one. Obstacles such as the market structure and firms' behaviour may prevent energy prices from converging towards efficient prices. The MSCP recognises that the occurrence of price spikes and volatility are common features in a liberalised electricity market. This is due to inflexible demand and supply conditions in the short term. However, the occurrence of persistent high prices and an increasing number of price spike incidents will require closer attention.

CURRENT METHODOLOGY

4. Since 2003, the MSCP has used a static model to identify high prices. The model relies on the assumption that high prices are possible signals of inefficient market outcomes.

5. High prices are established by comparing spot prices to a benchmark price.

6. After identifying the number of spot prices that are considered high prices, further studies are conducted to analyse the demand and supply conditions for the period under review. This enables the MSCP to understand whether the occurrence of high prices is the result of variation in market fundamental factors or other undesirable market behaviour.

7. However, prices which are high are not necessarily inefficient. Low prices may also point to inefficiency. In a competitive market, efficient prices should be the outcome of market fundamental factors which reflect normal demand and supply conditions. To understand the relationship between such factors and the energy prices in the NEMS, the MSCP has therefore developed an econometric model which takes into account such factors. Once developed, the model may be used as a benchmark to assist in the identification of outliers.

COMPARISON OF CURRENT METHODOLOGY WITH ECONOMETRIC MODELLING

8. To facilitate understanding as to how the use of an econometric model differs from the current static benchmark price methodology, Table 1 contains a comparison between the two.

Table 1: Comparison between Current Methodology and Econometric Model

	Current Methodology	Econometric Model
Methodology	A static benchmark price is established.	An econometric model is established based on historical demand and supply conditions and price behaviour to estimate how changes in such market fundamental factors affect most prices (i.e. excluding extreme ones).
Identification of Outliers	Prices higher than the benchmark price are identified as outlier prices.	With the relationship between supply/demand conditions and price behaviour established in the econometric model, estimated prices under dynamic demand and supply conditions may be predicted. Pre-defined upper and lower band from the estimated prices are then set. Prices falling outside of these bands are identified as outlier prices.
Rationale	High prices provide the first indication of an inefficient market.	An efficient market provides price signals that appropriately reflect demand and supply conditions.
Advantages	1. Simple to implement	1. Both high and low prices are considered. 2. Takes into account dynamic demand and supply conditions.
Disadvantages	1. High prices do not necessarily indicate an inefficient market. 2. Low prices are not identified using this methodology	1. Results depend on the robustness of the econometric model.

Table 1: Comparison between Current Methodology and Econometric Model		
	Current Methodology	**Econometric Model**
	although they can be indicative of inefficient outcomes. 3. Unable to take into account dynamic demand and supply conditions.	

ECONOMETRIC MODELLING FOR THE NEMS

9. The econometric model for the NEMS has been developed based on economic theories and power system principles, using historical NEMS data. The model will serve two objectives:

    a. Quantitative relationship – The model will enable the MSCP to understand how different market fundamental factors have quantitatively affected energy prices in the past;

    b. Identification of outliers – With an understanding of the historical relationship between market fundamental factors and energy prices, the MSCP will be able to use the econometric model to estimate energy prices in the context of varying demand and supply conditions. To provide a margin for error, the estimated prices can be expanded to include an upper and lower price band. Prices falling outside of these bands can be identified as outliers that warrant further attention.

10. Each of these objectives is examined more closely below.

ESTABLISHING THE QUANTITATIVE RELATIONSHIP BETWEEN MARKET FUNDAMENTAL FACTORS AND ENERGY PRICES

11. The purpose of this section is to establish how different market fundamental factors quantitatively affect energy prices.

Model Specifications

Modelling

12. Ordinary least squares (OLS) regression analysis[2] is used to model the relationship between market fundamental factors and energy prices.

[2] Please refer to Appendix A for more information on regression analysis.

Sample period

13. For the purpose of developing the econometric model, we have used daily average data from January 2003 to December 2006.

Variables

14. In selecting market fundamental factors as variables for developing the model, we have taken into account the availability of data, their relevance and the correlation[3] between variables.

15. The interpretation of each variable included in the model is based on the assumption that other variables remain constant.

16. It is also recognised that energy prices can vary between business and non-business days. Therefore, the econometric model accounts for this regular temporal variation through the inclusion of month, day-of-week, and annual dummy variables[4].

17. However, fluctuations in energy prices may not be confined to the above-mentioned factors. The model developed as explained in this paper is intended as a starting point. Over time, the model can continue to be improved by including more market fundamental factors that can have an effect on prices.

18. We now describe the variables used in the model and the methodology behind the construction of these variables. We also outline our initial expectations as to how a change in an explanatory variable may affect the dependent variable. Results obtained were subsequently checked to ascertain if they were in line with these expectations. If not, further investigations were carried out to understand the reasons and modifications made to the model where necessary.

    a. Uniform Singapore electricity price (USEP) – The dependent variable refers to the daily average USEP. USEP less than SG $50/MWh and more than SGD $4000/MWh are excluded from the model. This ensures that the econometric model is not estimated based on data from a period with abnormally high prices that skew estimations.

    b. Combined cycle gas turbine (CCGT) supply – This explanatory variable refers to the energy offers available for dispatch from CCGT units. Based on historical data, an increase in CCGT supply is expected to lower USEP.

    c. Steam turbine (ST) supply – This explanatory variable refers to the energy offers available for dispatch from ST units. Similar to CCGT supply, an increase in ST supply is expected to lower USEP.

---

[3] Please refer to Appendix B for an explanation on the effects of correlation on the econometric model and steps taken to ensure there is no correlation between variables.

[4] Please refer to Appendix A for a more detailed explanation on dummy variables.

d. Supply cushion – The supply cushion is the ratio between the supply and the demand gap and supply. This explanatory variable refers to the spare capacity available after energy has been dispatched. Based on historical data, an increase in supply cushion is expected to lower USEP.

e. Offers – Offers refer to the percentage of total offers that are at SGD $100/MWh or less. Based on historical data, an increase in the percentage of offers at SGD $100/MWh or less is expected to lower USEP.

f. Demand – This variable refers to the daily average demand in the NEMS. Based on historical data, an increase in demand is expected to lead to an increase in USEP.

g. Reserve cushion – The reserve cushion refers to the spare capacity available in reserves after dispatch. Similar to the supply cushion, a decrease in USEP is expected to follow an increase in the reserve cushion.

h. One month lag of fuel oil price – Fuel costs account for a significant proportion of the running costs of the thermal - fired generators that make up 97 percent of Singapore's generation capacity. This means that the majority of the production costs for Singapore's generation assets are either directly or indirectly (through the pegging of natural gas prices to an oil benchmark) determined by international oil prices. Hence, changes in fuel input costs, such as the price of fuel oil or natural gas, have a significant influence on USEP. For the NEMS, the most relevant oil benchmark is the HSFO 180 cst (High Sulphur Fuel Oil).

   Prior to the inclusion of fuel oil prices in the model, the correlation between USEP and fuel oil price up to a lag of six months was tested. Fuel oil price with a lag of one month was found to have the highest correlation with USEP and was therefore used for the model.

i. CCGT planned outages dummy variable – CCGT generators are base load units in the NEMS. They supplied 78 percent of the energy in the NEMS in 2006. For inclusion into the econometric model, CCGT planned outages are categorised into two groups – Group A and Group B.

   Group A refers to generation companies which have ST generators in their portfolio. For generation companies that are categorised under Group A, there is a tendency for them to offer into the market their ST units when their CCGT units are on planned maintenance. The effect of CCGT planned outages from Group A on USEP can therefore affect USEP positively or negatively and is inconclusive. This means that the coefficient of this variable can take on a positive or negative value.

   Group B refers to generation companies which only have CCGT units to offer into the market. An increase in outages of CCGT units from Group B is expected to result in an increase in USEP.

j. Forced outages dummy variable – Forced outages are unanticipated and are expected to have a negative effect on energy prices. Based on historical data, an increase in forced outages is expected to lead to an increase in USEP.

As forced outages rarely occur[5], adding variables that capture these factors to the model do little to explain variation in energy prices over a long period of time. However, this variable is important in understanding the behaviour of energy prices at a specific point in time. The supply dips when a generation plant previously supplying electricity to the grid trips. This is likely to lead to an increase in prices in the short term. Due to competitive pressure and contractual obligations, supply conditions can improve quickly as generation companies respond to the higher price. The result is a temporary price spike in the energy market followed by a prompt decline in prices despite the tripped plant staying on the sideline

Functional form of the econometric model

19. For the purpose of this study, the econometric model relies on the logarithm of prices. This is because supply curves in most wholesale electricity markets tend to be non-linear. This means that an increase in demand tends to result in much larger increases in prices when supply is tight. Similarly, the supply curve for the NEMS is non-linear. The log-log model allows us to capture some of the non-linearities.

20. The use of a log-log model also allows us to interpret the coefficients as elasticities[6].

Model Equation

21. The specification for the regression[7] is as follows:

$$USEP = \alpha + \beta_1\ CCGT\ supply + \beta_2\ ST\ supply + \beta_3\ Supply\ cushion + \beta_4\ Offers + \beta_5\ Demand + \beta_6\ Reserve\ cushion + \beta_7\ Lagged\ oil\ price + \beta_8\ CCGT\ planned\ outages\ (Group\ A) + \beta_9\ CCGT\ planned\ outages\ (Group\ B) + \beta_{10}\ Forced\ outages$$

Econometric Model Results

22. Table 2 shows the regression results for the econometric model.

---

[5] Viewing the frequency of occurrence in the context of a total of 48 periods a day and 365/366 days a year.
[6] Please refer to Appendix C for more detailed information on the concept of elasticity.
[7] In our preliminary research, outages from ST units were considered in the model. However, as they are not base load units in the NEMS, the impact of ST planned outages on energy prices were inconclusive. Hence, we have chosen not to include ST planned outages in the model.

Table 2: Estimation results of the econometric model		
Variable	**Coefficient**	**P-value**
Constant	11.891	0.000
LOG (CCGT supply)	-0.016	0.811
LOG (ST supply)	-0.003	0.932
LOG (Supply cushion)	-0.994	0.000
LOG (Offers)	-1.165	0.000
LOG (Demand)	0.101	0.571
LOG (Reserve cushion)	-0.303	0.000
LOG (Lag of fuel oil price)	0.415	0.117
CCGT planned outages dummy A	-0.015	0.262
CCGT planned outages dummy B	0.103	0.000
Forced outages dummy	0.046	0.000
Model Diagnostics		
R-squared	0.79	
Adjusted R-squared	0.78	
LM test	Present	
White heteroskedasticity test	Present	
Number of observations	1461	

Explanatory Variables

23. The model results indicate the following:

a. Higher impact from base load offers – With CCGT units supplying more than 50 percent of energy in the NEMS from 2003, it is not surprising that the regression results show that offers from CCGT units had a bigger influence on energy prices compared to ST offers. The coefficient estimates suggest that an increase in offer availability from CCGT units by 10 percent led to a decrease in energy prices by approximately 0.16 percent. This is higher in comparison to the 0.03 percent decrease in energy prices for ST units.

b. Supply cushion – The estimation results show that an increase in the supply cushion by 10 percent led to a decrease in energy prices by approximately 9.9 percent.

c. Offers – An increase in offers at SGD $100/MWh or less by 10 percent led to a decline in energy prices by approximately 11.7 percent.

d. Demand condition – The model results show that an increase in average demand by 10 percent resulted in an increase in prices by 1.0 percent.

e. Reserve cushion – An increase in the reserve cushion by 10 percent led to a decrease in prices by approximately 3.0 percent.

f. Importance of fuel cost – The results confirm that fuel cost is one of the most important explanatory variables that are responsible for the variation in energy prices over time. A 10 percent surge in the fuel price led to a 4.2 percent increase in energy price.

Dummy Variables

24. In the case of dummy variables, it is positive or negative nature of the relationship that is of significant.

a. Higher impact from base load offers – With CCGT units supplying more than 50 percent of energy in the NEMS

b. CCGT planned outages – The model confirms that outages of CCGT units from Group B (i.e. generation companies which only have CCGT units to offer into the market) led to an increase in energy prices as the effect of lost output could not be replaced with any alternative sources. Outages of CCGT units from Group A (i.e. generation companies which have ST generators in their portfolio) led to a decrease in energy prices.

c. Marginal price effect from forced outages – The model confirms that forced outages had a negative impact on energy prices.

Robustness of the Model

25. The model diagnostics and p values in Table 2 indicate the tests that have been done with regard to ensuring the robustness and statistical soundness of the model. Further details are available in Appendix D.

Conclusion on the Econometric Model

26. The current econometric model seeks to provide a quantitative explanation of the factors behind the variation in energy prices. Although the model only successfully explains close to 80 percent of the variations in energy prices, it affirms that several indices in the catalogue of monitoring indices[8] adopted by the MSCP in 2003 are major contributors to variation in energy prices over time. The MSCP plans to continue to research and refine the model's analytical capability.

IDENTIFICATION OF PRICE OUTLIERS

Model Methodology

27. The purpose of this section is to explain the methodology used to identify outlier prices.

28. It is assumed that an econometric model with good explanatory capability will be able to predict energy prices based on observed market fundamental factors reflecting demand and supply.

29. On this basis, energy prices under the prevailing market conditions for each day are estimated using the econometric model developed. The prevailing market conditions are specified using the market fundamental factors identified as the relevant explanatory variables. Actual energy prices are then compared with estimated prices.

30. Energy prices are estimated using the econometric model equation below:

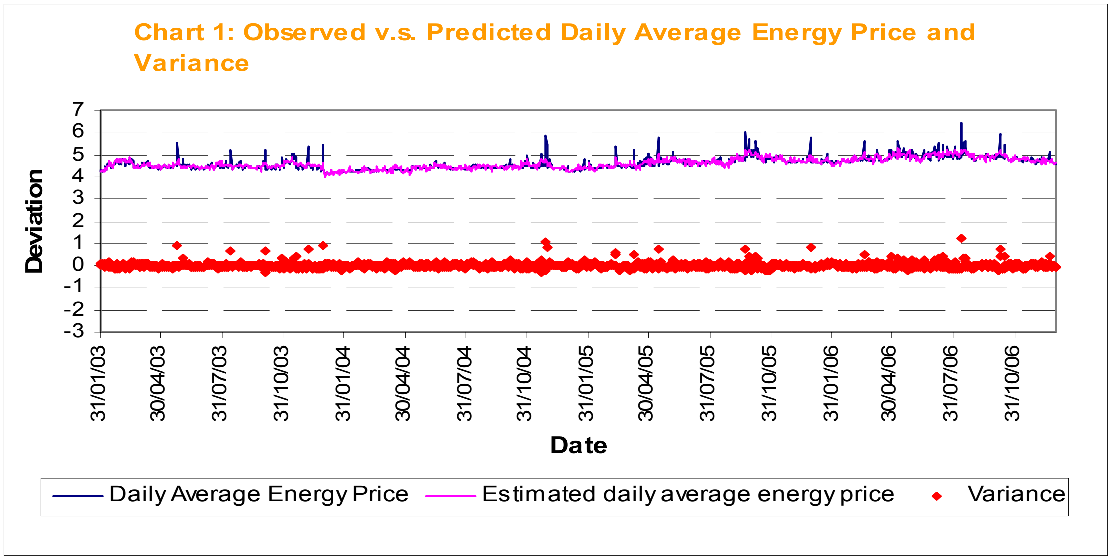
    Log USEP = 11.891 - 0.016 Log CCGT supply - 0.003 Log ST supply - 0.994 Log Supply cushion – 1.165 Log Offers + 0.100 Log Demand - 0.303 Log Reserve cushion + 0.415 Log Oil price - 0.015 Log CCGT planned outages (Group A) +0.103 Log CCGT planned outages (Group B) +0.046 Log Forced outages.

31. The variances between actual and estimated energy prices for 31 January 2003[9] to 31 December 2006 are shown in Chart 1. The variance between actual and estimated energy prices is calculated using the following equation:

    Variance = Actual Price – Estimated Price

[8] Please refer to the catalogue of indices available at http://www.emcsg.com/f6234,93137/93137_Catalogue_of_Monitoring_Indices_29_July_04.pdf

[9] Due to the use of data of fuel oil price that lags by a month, the estimated model uses observations between 31/01/2003 and 31/12/2006.



32. It is recognised that departures of actual energy prices from estimated energy prices be due to the exclusion of market fundamental factors that can be important in explaining the specific incident. Hence, a reasonable price range based on three standard deviations[10] is constructed in order to identify possible price outliers. Chart 2 shows the actual energy prices, the upper and lower bands and the outliers that fall outside the reasonable price range.

[10] Yoo and Meroney (2005) also used the three standard deviation criterion in their econometric model to identify abnormal prices.

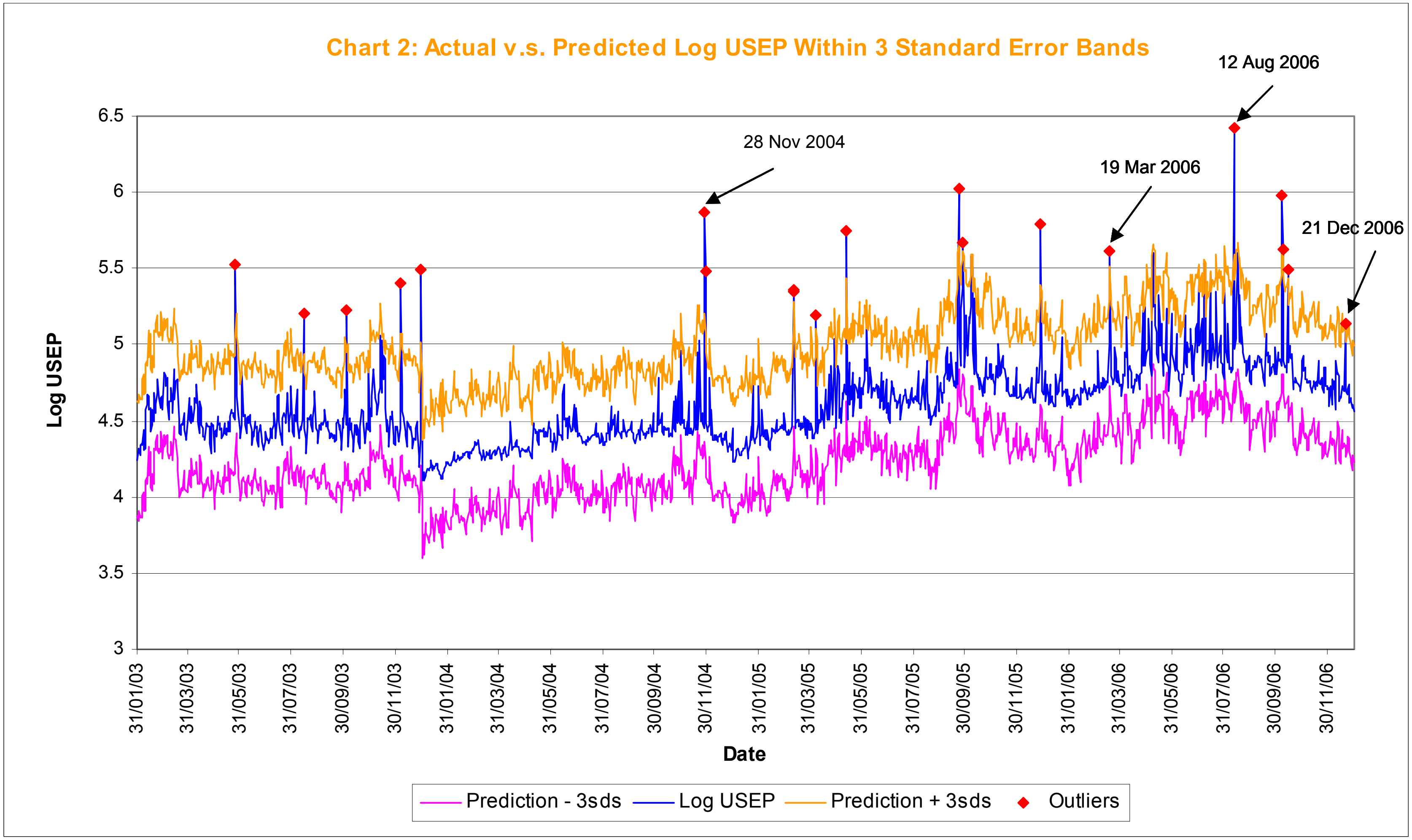


Chart 2: Actual v.s. Predicted Log USEP Within 3 Standard Error Bands
12 Aug 2006
28 Nov 2004
19 Mar 2006
21 Dec 2006
Log USEP
6.5
6
5.5
5
4.5
4
3.5
3
31/01/03
31/03/03
31/05/03
31/07/03
30/09/03
30/11/03
31/01/04
31/03/04
31/05/04
31/07/04
30/09/04
30/11/04
31/01/05
31/03/05
31/05/05
31/07/05
30/09/05
30/11/05
31/01/06
31/03/06
31/05/06
31/07/06
30/09/06
30/11/06
Date
Prediction - 3sds
Log USEP
Prediction + 3sds
Outliers

Frequency of Outliers and Possible Reasons

33. According to the bands established, the number of outliers identified as a percentage of the total number of observations in each year since market start falls within a marginal one to two percent range. This information is summarised in Table 3.

Table 3: Percentage of outliers

Year	Outliers (%)
2003	1.4
2004	0.5
2005	1.9
2006	1.6

34. The outliers identified by the econometric model are price outcomes that warrant further attention.

Examples of Price Outliers

35. A sample of our analysis of recent instances of outliers is provided below.

36. There are a number of instances in which the actual energy price falls outside the established bands. This section provides a sample of our analysis of recent instances.

a. *19 March 2006*

The demand for 19 March 2006 was one of the highest Sunday averages since the start of the market. The supply cushion variable included in the model captured the low spare capacity available. USEP hovered around $130/MWh for most of the day, with USEP increasing to above $300/MWh for a few periods.

However, when three gas turbines (GT) units were offered into the market, USEP spiked to $1,023/MWh.

b. *12 August 2006*

This price outlier detected by the model is mainly due to six consecutive periods of high prices that are above $3000/MWh. A CCGT unit had tripped in period 15. However, offers from this unit only dropped to zero from period 21 onwards, resulting in a dip in the supply cushion from 17.6 to 12.2 percent. USEP peaked for six consecutive periods at above $3000/MWh as the supply cushion hovered between 10 and 12 percent.

During these high price periods, contingency reserve and regulation also hit their price caps of $3250/MWh and $2750/MWh respectively. The supply cushion further tightened when a GT unit tripped in period 25, as 100MW of capacity was lost. The effect of the loss in supply due to the trip in GT unit was minimal as demand started to decline from period 25. Energy prices normalised in period 32 when the tripped CCGT unit resynchronised to the grid.

c. *21 December 2006*

This price outlier was detected by the model mainly due to the price spikes between periods 34 to 36. There was an unplanned disruption to the piped gas supply from Malaysia, which led to the tripping of two CCGT units. In period 34, USEP hit the price cap of $4500/MWh and load shedding occurred in some areas. The sudden dip in supply caused the supply cushion to decline to 9.8 percent.

Due to declining demand when the gas interruption occurred, three periods of high prices above $1000/MWh were recorded. Supply on average recovered to the level prior to the gas interruption in period 40 as the tripped units returned to normal operation.

Therefore, in this instance, the price spikes had been due to a shortfall in the system.

37. The MSCP also recognises that outliers identified may be a result of market fundamental factors that affect USEP but have not been captured by the model at present or which are difficult to be adequately captured in the model. Factors falling under the latter category include the following:

    a. As a result of co-optimisation[11] by the market clearing engine (MCE), some of the cheaper energy offers may not be taken into consideration in arriving at the final price outcome.

    b. Transmission line constraints which may affect price outcomes.

CONCLUSION

38. The results from the econometric model explained in this study suggest that simple market fundamental factors included in the model are able to explain a significant proportion of the variation in daily average energy prices between 2003 and 2006.

[11] The MCE clears the energy, reserve and regulation markets simultaneously and the merit order dispatch in one of these product markets may result in out-of-merit order dispatch in another product market. The MCE endeavours to find the best solution that achieves a globally optimal result by maximising the objective function.

39. The MSCP intends to use this econometric model to analyse the influence of market fundamental factors on energy prices in the NEMS. This model will help the MSCP to understand the dynamics behind changes in energy prices.

40. The MSCP also intends to use this model as a screening tool for potential inefficient price outcomes in the future. Divergence of the actual energy price beyond a reasonable price range may trigger closer scrutiny.

41. Recognising that the model described in this paper is a useful starting point, outliers which can be explained by normal demand and supply conditions will prompt further work in refining the model to sharpen its analytical capabilities. Over time, the model can continue to be improved by including more market fundamental factors that can have an effect on prices. A longer period of time will allow more data to be collected which may bring about greater accuracy in the estimated results. It will also enable the model to be evaluated.

Professor Lim Chin
for Chair, Market Surveillance and Compliance Panel

APPENDIX A

Regression Analysis[12]

1. Regressions are used to quantify the relationship between one variable and the other variables that are thought to explain it. Regressions can also identify how close and well determined the relationship is.

How to Run a Regression?

2. Before a regression is run, a theoretical model can help explain how and why the **dependent variable** is determined by one or more **explanatory variables**.

3. For example, assuming that an individual's wealth depends on his or her level of education is an example of a simple model with one explanatory variable. A corresponding equation would look like:

$$Y = a + \beta X + e$$

On the left-hand side is Y, the dependent variable, wealth. On the right-hand side are *a*, the constant (which may be the person's inheritance), and *β*, the coefficient (or slope) multiplied by X our explanatory variable, education. In algebra, the regression says that "wealth depend only on education and in a linear way". In this instance, the other explanatory factors, if there are any, are omitted. The error term (e) accounts for factors that influence earnings but are not included in the econometric model.

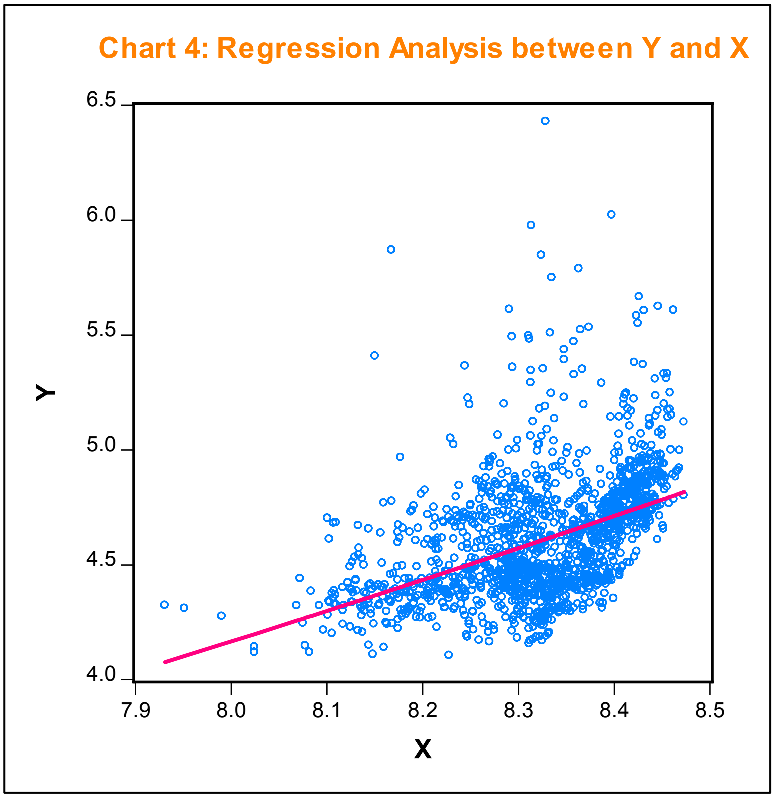
4. However, wealth may be affected by a variety of other factors other than education. The econometric model that is estimated will be more complicated. A multiple-variable regression would be estimated. The model would resemble:

$$Y = a + \beta_1 X_1 + \beta_2 X_2 + \ldots$$

Several *X* variables help explain *Y* (wealth) – like ability, intelligence, age, education, marital status, and parental education. The *β* coefficients measure the impact of these variables on earnings, assuming the other variables are constant.

5. Graphically as shown in Chart 4, the regression analysis attempts to find the best fit line that relates the dependent variable Y to the independent variable X.

[12] This section has been adapted from Ramcharan (2006).



6. A **dummy variable** is one that takes the value 0 or 1 to indicate the absence or presence of an effect that can change the outcome. Dummy variables are used when we want the independent explanatory variable to take on two or more distinct categories reflecting different characteristics. Using the same example, assume that an individual's wealth can be explained by two factors – age and education. An individual's wealth can then be expressed in the following equation:

$$Y = a + \beta_1 X_1 + \beta_2 X_2 + e$$

β_1 measures an individual's wealth associated with his age. β_2 represents an individual's level of education. This is a dummy variable that takes the value of 1 if the individual has a university degree and 0 if the individual does not have a university degree.

$$X_2 = \begin{cases} 1 & \text{If the individual has a university degree} \\ 0 & \text{If the individual does not have a university degree} \end{cases}$$

Therefore, an individual's expected wealth is dependent on whether the individual has a university degree. This can be seen by taking expected values on both sides of the equation for $X_2 = 1$ and $X_2 = 0$.

$$E(Y) = \begin{cases} \beta_1 & \text{If } X_2 = 0 \\ \beta_{1+}\beta_2 & \text{If } X_2 = 1 \end{cases}$$

APPENDIX B

Correlation between Variables

1. One of the assumptions of the econometric model we estimated in this study is that there is no correlation between the explanatory variables included. The interpretation of the coefficients in the econometric model is dependent on this assumption. Correlation between explanatory variables can result in inaccurate standard errors which affects the accuracy of results.

Effects of Correlation between Variables

2. For example, we consider the regression model:

$$Y = a + \beta_1 X_1 + \beta_2 X_2$$

3. β_1 (the coefficient of X_1) is interpreted to measure the change in Y that is due to a change in X_1. The interpretation of β_1 is also dependent on the assumption that β_2 remains constant.

4. If β_1 and β_2 are correlated, a given change in X_1 is expected to lead to a predictable similar change in X_2. The interpretations of the coefficients in such a model will be difficult as it is unclear how X_1 or X_2 affects Y.

5. The existence of correlation between explanatory variables implies that there will be very little data in the sample to give one confidence of an accurate interpretation.

6. Two variables are said to have the same information if their correlation in absolute value is greater than or equal to 0.95 (Wessel et al., 1998).

7. Table 4 provides a summary of the correlation between the explanatory variables included in the econometric model. As seen, the correlation between different variables all has absolute values of less than 0.95. Therefore, none of the variables included in the econometric model are considered highly correlated.

Table 4: Correlation	LOG (CCGT Supply)	LOG (Supply cushion)	LOG (Offers)	LOG (Demand)	LOG (Reserve cushion)	LOG (Lag of fuel oil price)
LOG (CCGT supply)	1.00	-0.22	-0.32	0.49	-0.38	0.77
LOG (Supply cushion)	-0.22	1.00	0.22	-0.65	-0.03	-0.36
LOG (Offers)	-0.32	0.22	1.00	-0.21	-0.12	-0.63
LOG (Demand)	0.49	-0.65	-0.21	1.00	-0.15	0.50
LOG (Reserve cushion)	-0.38	-0.03	-0.12	-0.15	1.00	-0.28
LOG (Lag of fuel oil price)	0.77	-0.36	-0.63	0.50	-0.28	1.00

APPENDIX C

Elasticity

1. Elasticity measures the effect on the dependent variable of a 1 percent change in an independent variable.

2. Mathematically, elasticity can be expressed in the following equation:

$$\text{Elasticity} = \frac{Y_2 - Y_1}{Y_1} \bigg/ \frac{X_2 - X_1}{X_1}$$

$$= \frac{\partial Y}{Y} \bigg/ \frac{\partial X}{X}$$

$$= \frac{\partial Y}{\partial X} * \frac{X}{Y} = \mathbf{A}$$

3. A log-log model means that the dependent variable and all explanatory variables included in the model are logged.

4. The economic interpretation of the coefficients of the variables in a log-log model is different from that of a linear model. The coefficients of the variables in a log-log model can be interpreted as its elasticities. The next example explains why this is the case.

5. Assume we have a simple log-log model expressed in the following equation:

$$\log Y = \log C + \alpha \log X$$

Differentiating both sides with respect to X, we get:

$$\frac{\partial Y}{\partial X}\left(\frac{1}{Y}\right) = \alpha * \frac{1}{X}$$

$$\frac{\partial Y}{\partial X} = \frac{Y * \alpha}{X}$$

$$\frac{\partial Y}{\partial X} = \left(\frac{Y}{X}\right) * \alpha$$

$$\alpha = \frac{\partial Y}{\partial X} * \frac{X}{Y} = \mathbf{A}$$

As seen, the coefficient of the variable X gives us the elasticity of X with respect to Y. Therefore, the coefficients of variables included in a log-log model can be interpreted as elasticities.

APPENDIX D

Robustness of the Model

1. The model diagnostics are checked to ensure that the model is statistically sound and robust.

2. A **p-value** describes the exact significance level associated with a particular econometric result. A low p-value indicates that the explanatory variable included in the model plays a significant role in explaining variations in the dependent variable.

3. The **R-squared** statistic measures the success of the regression in predicting the values of the dependent variable within the sample. The R-squared statistic can be interpreted as the fraction of the variation of the dependent variable that can be explained by the independent variables. The R-squared statistic will be equivalent to one if the regression fits perfectly and vice versa. The R-squared statistic tends to over estimate the strength of the association especially if the model has more than one independent variable.

4. A problem with using the R-squared as a measure of goodness of fit is that the R-squared will never decrease when more independent variables are added. Hence, the **adjusted R-squared** is a more appropriate measure. It penalises the R-squared for the addition of independent variables that do not contribute to the explanatory power of the model. Therefore, the adjusted R-squared is never larger than the R-squared and can decrease as more independent explanatory variables are added.

5. The adjusted R-squared for our econometric model is 0.78. This implies that the independent explanatory variables included in the econometric model are able to explain 78 percent of the variations in daily average energy price. Although there is no standard guideline in interpreting R-squared, various studies[13] relating to the econometric modelling of energy prices have achieved similar adjusted R-squared values.

6. The **Lagrange Multiplier (LM) test** examines the model for the existence of serial correlation between independent explanatory variables included in the econometric model. Although serial correlation was present in our initial econometric model, it has been accounted for in the estimation of the model to ensure that standard errors are accurate.

7. The **White Heteroskedasticity test** examines the model for the existence of heteroskedasticity.[14] Heteroskedasticity was found to be present in our initial econometric model and was corrected.

[13] Examples of studies that have used econometric modelling to analyse energy prices include ISO-New England (2003) and United States General Accounting Office (2006). See the references section for a more detailed list of the other studies we have consulted.
[14] Please refer to Appendix E for more information on heteroskedasticity.

8. The **Wald test** conducted allows us to test for the joint significance of the independent explanatory variables included in the econometric model. The test confirms that the explanatory variables included in the model are jointly significant in explaining variations in energy prices.

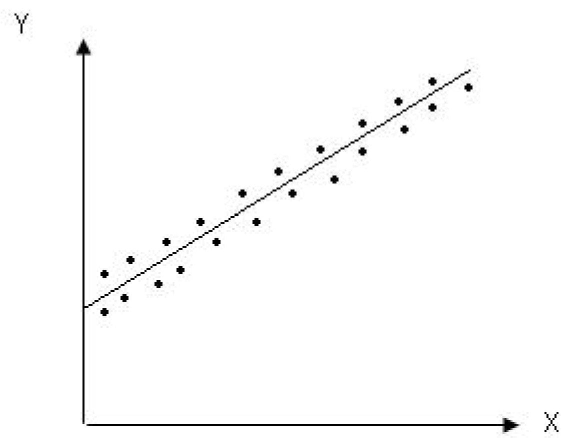
APPENDIX E

Heteroskedasticity

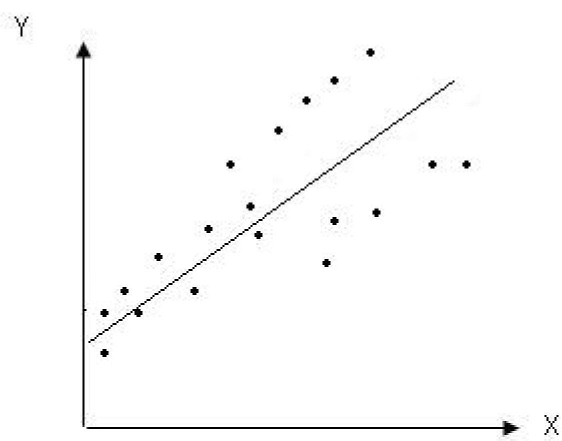
1. Assume that we have a regression model that has the following equation:

$$Y = a + \beta_1 X + e$$

2. A standard statistical assumption of our regression model is that the error term has a constant variance. This error term is referred to as being *homoskedastic.* This means that the vertical spread of the data around the predicted line will be fairly constant as X changes as seen below.



3. If the error term is heteroskedastic, it implies that the vertical spread of the data around the predicted line changes as the value of X changes. In the figure below, we see that the vertical spread of data around the predicted line increases as the value of X increases.



REFERENCES

Arciniegas Rueda, Ismael E. and Marathe Achla 2005. “Important variables in explaining real-time peak price in the independent power market of Ontario,” *Utilities Policy*, Volume 13: pages 27-39.

Barmack, M., E. Kahn, and S. Tierney, Analysis Group; and C. Goldman, Lawrence Berkeley National Laboratory, *A Regional Approach to Market Monitoring in the West,* available at
http://eetd.lbl.gov/ea/EMS/reports/61313.pdf

ISO-New England 2003. *Annual Market Reports – Technical Review, May – December 2002,* available at
http://www.iso-ne.com/markets/mkt_anlys_rpts/annl_mkt_rpts/2002/2002_Annual_market_Technical_Review_Final.pdf

Ontario Energy Board 2006. *Monitoring Report on the IESO-Administered Electricity Markets, May 2006 – October 2006*, available at
http://www.oeb.gov.on.ca/documents/msp/msp_report_final_20061222.pdf

Ontario Energy Board 2006. *Monitoring Report on the IESO-Administered Electricity Markets, November 2005 – April 2006*, available at
http://www.oeb.gov.on.ca/documents/msp/msp_report%20final_131205.pdf

Pindyck, Robert S. and Rubinfeld, Daniel L. *Econometric Models and Economic Forecasts.* McGraw-Hill, 1998.

Ramcharan, Rodney. “Back to Basics – Regressions: Why are Economists obsessed with them?”, *Finance and Development*, Volume 43, Issue 1.

United States General Accounting Office, *Restructured Electricity Markets.* GAO-02-828. Washington, D.C., available at
http://www.ksg.harvard.edu/hepg/Papers/GAO-6-02_Restructured_Elec_mkts.pdf

Wessel, Jurs, M.D., Tolan, J.W., Muskal, S.M. 1998. “Prediction of human intestinal absorption of drug compounds from molecular structure,” *Journal of Chemistry, Informatics, and Computer Science,* Volume 39: pages 726-735.

Yoo, Young and William Meroney 2005. “A regression Model of Gas/Electricity Price Relationship and Its Application for Detecting Potentially Anomalous Electric Prices.” Presented at 25th USAEE/IAEE Conference, Denver, CO, September 19, 2005.