Econometric Model Design, Approach and Methodology Report
– A Review of the Current Methodology

By Feng Qu & Zhou Shihao

Contents

1 Executive Summary

The Market Surveillance and Compliance Panel (MSCP) Annual Report, prepared and submitted by the MSCP to the Energy Market Company Pte Ltd (EMC), comprises an analysis of outliers of the Uniform Singapore Energy Price (USEP) in the National Electricity Market of Singapore (NEMS) which are identified by an econometric model. The model was applied with two specifications to identify the USEP outliers – the first specification (without stepwise regression) in the MSCP Annual Report 2007 (also further detailed in How Market Fundamental Factors Affect Energy Prices in the NEMS – An Econometric Model) and the second specification (with stepwise regression) applied since the MSCP Annual Report 2009. The outliers were determined based on the upper bound of the predicted USEP computed by the model. To provide such critical prediction of the USEP, the previous two specifications were constructed based on economic theories, such as demand and supply, as well as production cost and energy price. Nevertheless, the explanatory power of the explanatory variables in the previous econometric models could be significantly improved by accommodating more empirical features.

This report proposes a revised econometric model by revising and enhancing the previous model in three aspects:

a. several important and new features are incorporated into the revised model;
b. a set of robustness checks are added; and
c. variable and model selection procedures are employed.

First, the previous models did not capture some important features of time series data, e.g. the autocorrelation of the USEP. In addition, the impact of macroeconomic policies and the fluctuations in aggregate demand and supply on the USEP had not been fully captured in the previous models. With the new features listed below, the revised econometric model could better describe how the USEP was determined, thus providing a more accurate prediction of the USEP and the identification of outliers.

Second, different from the previous econometric model, in which the outliers were determined by a single regression using one sample period, the new econometric model included a set of robustness checks. The idea was that the USEP outliers detected by a single regression using one sample could be sensitive to the choice of variables, samples or model specifications. Robustness checks with different subsamples and model specifications allowed us to examine the reliability of the USEP outliers detected. It turned out that the results obtained from the robustness checks were similar to those of the revised econometric model, indicating that the revised model was robust to different model specifications and subsamples. Thus, we could rely on the prediction based on the revised model.

Third, different from the "stepwise regression" approach applied in the MSCP Annual Report 2009, we introduced clear-cut variable and model selection procedures. With these procedures, the data could tell us which variables should be included in the regressions and which model should be assigned with more weights. When more data becomes available in the future, the model specifications could be updated based on the suggested selection procedure.

Specifically, we incorporated the following new features into the revised model:
a. creating a dynamic model with a lagged dependent variable;
b. controlling for annual macroeconomic impacts on the USEP;
c. addressing seasonality;
d. introducing a non-linear effect;
e. introducing a time trend;
f. introducing heterogeneous effects by using interaction terms; and
g. refining the one-month lag of the fuel oil price.

The analysis was based on the sample period from 2003 to 2019 and the explanatory variables were as follows:

a. Uniform Singapore Energy Price (USEP);
b. combined cycle gas turbine (CCGT) supply;
c. steam turbine (ST) supply;
d. supply cushion;
e. offers below SGD$100/MWh;
f. forecasted demand;
g. reserve cushion;
h. one-month lagged oil price;
i. square of one-month lagged oil price;
j. amount of forced outages interacted with forecasted demand;
k. amount of CCGT planned outages interacted with forecasted demand;
l. amount of ST planned outages interacted with forecasted demand;
m. time trend;
n. year dummies;
o. month dummies; and
p. day dummies.

Particularly, to identify outliers based on the regression results, we first constructed the upper bound and lower bound, which would be three Standard Deviations (SD) above and below the predicted USEP for most of the cases in this report, and two SD above and below the predicted USEP for the weekly average regression, respectively. Subsequently, we compared the actual USEP in the database against the upper bound. If the actual USEP is above the upper bound, then it is deemed as an outlier in the market. Please refer to Figures A1 to A6 in the Appendix for the visualisation of outliers.

In addition, a set of robustness checks included:

a. new model specifications to predict USEP: Hybrid model (Ordinary Least Squares and Autoregressive Integrated Moving Average, OLS+ARIMA);
b. regression with subsample of data from 2010 to 2019; and
c. regression with weekly average of the observations.

In order to provide clear criterions for model/variable selection process, we introduced several measures listed below:

a. R-squared (R^2);
b. Mean Absolute Percentage Error (MAPE);
c. Mean Absolute Error (MAE);
d. Root Mean Square Error (RMSE);
e. Akaike Information Criterion (AIC); and
f. Bayesian Information Criterion (BIC).

Based on these measures, our revised model performed consistently better than the previous model applied to identify outliers of the USEP, suggesting that the prediction of the USEP and outliers based on the revised model were more reliable. In addition, we also examined the robustness of our revised model by:

a. providing the results based on subsample of data from 2010 to 2019;
b. switching to other model specifications such as Hybrid Model (OLS+ARIMA);
c. comparing our choice of explanatory variables with the Least Absolute Shrinkage and Selection Operator (LASSO) method; and
d. smoothing out the randomness of daily observations by taking the weekly average of the observations.

It turned out that the results, based on our revised model, were robust to the choices of sample and different model specifications. According to the revised model, we had lesser outliers compared to the previous models, suggesting that the NEMS was more efficient than expected. Overall, there were several advantages of the revised model:

a. stronger explanatory power of the independent variables;
b. more flexibility in the long run, regardless of the macroeconomic state of the year;
c. clearer criterions for variable selections when more data becomes available in the future;
d. more up-to-date methodology such as machine learning algorithm to verify the findings; and
e. more refined code which could accommodate more complicated data in the future.

1.1 Data Description

Table 1 presents the definition of variables used in the regression model. For more information about the summary statistics of variables, please refer to Table A1.

Table 1: Variable Description

Notation in the Regression	Variable Description
$usep_t$	USEP is the dependent variable in the regression, where there is one record of USEP daily. The USEP less than \$50/MWh and more than \$4,000/MWh were excluded in the analysis to eliminate any abnormal values of the USEP that might affect our model estimation. Overall, the results were consistent with or without those abnormal USEP.
$ccgtsupply_t$	CCGT supply is the explanatory variable in the regression which refers to the energy offers available for dispatch from CCGT units. It is collected daily. Based on the demand and supply theory, the relationship between CCGT supply and USEP is negative.
$stsupply_t$	ST supply is the explanatory variable in the regression which refers to the energy offers available for dispatch from ST units. It is collected daily. Similarly, it has negative relationship with USEP.
$supplycushion_t$	Supply cushion is the explanatory variable in the regression which refers to the ratio between the total supply and the demand gap and supply, while the total supply is the sum of CCGT supply and ST supply. It is collected daily. An increase in supply cushion leads to a decrease in USEP.
$offers_t$	Offers is the explanatory variable in the regression which refers to the total offers that are at \$100/MWh or less. It is collected daily. The relationship between offers and USEP is expected to be negative.
$demand_t$	Demand is the explanatory variable which refers to the forecasted daily demand of electricity computed by the NEMS. Based on the demand and supply theory, an increase in demand causes a decrease in USEP.
$reservecushion_t$	Reserve cushion is the explanatory variable that refers to the spare capacity available in the reserves after dispatch. It is collected daily. The relationship between reserve cushion and USEP is negative.

Notation in the Regression	Variable Description
lagoilprice	One-month lagged oil price is the explanatory variable that refers to the production cost of electricity. As reported, fuel cost account for a significant proportion of the running costs of electricity generation. Unlike other explanatory variables, it is collected monthly. Based on the economic theory, higher oil price causes higher production costs, which in turn leads to higher USEP.
$\boldsymbol{forceoutage_t}$	Forced outage is the explanatory variable that refers to the unanticipated daily outage volume. Again, if there is an outage, the supply of energy becomes lower, which then leads to higher USEP. In addition, there may exists interaction effect between forced outages and forecasted demand.
$\boldsymbol{ccgtoutage_t}$	CCGT planned outage is the explanatory variable that refers to the anticipated daily CCGT outage volume. Similar to the forced outages, if the CCGT energy supply becomes lower due to outages, it leads to higher USEP. In addition, there may exists interaction effect between planned outages and forecasted demand.
$\boldsymbol{stoutage_t}$	ST planned outages is the explanatory variable that refers to the anticipated daily ST outages. Like the forced outages, if the ST energy supply becomes lower due to outages, it leads to higher USEP. In addition, there may exists interaction effect between planned outages and forecasted demand.
$\boldsymbol{year_k}$	Year dummies are the additional explanatory variables that account for year fixed effects such as changes in macroeconomic factors. There are total k number of dummies depending on the sample period. The base year is the first year of the sample period.
$\boldsymbol{month_m}$	Month dummies are the additional explanatory variables that addresses seasonality problems in the data. We expect higher demand for energy during months with higher average temperature, which then leads to higher USEP. The base month is January of each year.
$\boldsymbol{weekday_p}$	Weekday dummies are the additional explanatory variables that again addresses seasonality problems in the data. Like the month dummies, we expect significant positive relationship between days with higher consumption of energy and USEP. The base day is every Sunday of the week.

Table A1: Summary Statistics

No.	Variables	Unit	Mean	Median	Standard Deviation	Minimum	Maximum	Form in Regression
1	USEP	$/MWh	130.476	116.240	60.949	36.130	997.160	LOG
2	CCGT Supply	MW	5,685.436	5,248.150	1,722.768	1,772.920	8,601.850	LOG
3	ST Supply	MW	819.461	810.520	780.352	0.010	3,179.890	LOG
4	Supply Cushion	MW	0.227	0.225	0.041	0.032	0.447	LOG
5	Offers	MW	4,667.438	4,297.030	1,018.950	2,715.560	6,991.130	LOG
6	Demand	MW	5,006.387	5,050.530	718.327	2,782.460	6,456.010	LOG
7	Reserve Cushion	%	46.590	47.790	9.612	5.780	71.090	LOG
8	Lagged Oil Price	US$/barrel	64.400	62.180	26.065	24.110	117.830	LOG
9	Forced Outages	MW	31.248	0.000	73.086	0.000	679.020	LOG
10	CCGT Planned Outages	MW	467.269	365.000	405.243	0.000	2,476.370	LOG
11	ST Planned Outages	MW	476.576	462.000	395.043	0.000	2,070.000	LOG

2 Methodology

2.1 Previous Models

The MSCP applied static models to identify outliers of the daily USEP in the past years. The models were developed based on economic theory such as demand and supply, as well as production cost and energy price. The outlier was determined as an actual USEP higher than the upper bound, which was computed by three SD above the predicted USEP. Based on the previous models, we observed a decreasing number of outliers over the past two decades, suggesting that the energy market had become more efficient. The previous models were developed based on standard economic theory and expressed as follows:

Existing OLS1 (MSCP Annual Report 2007)[1]

$$\log(usep_t) = \alpha + \beta_3 \log(ccgtsupply_t) + \beta_4 \log(stsupply_t) + \beta_5 \log(supplycushion_t) + \beta_6 \log(offers_t) + \beta_7 \log(demand_t) + \beta_8 \log(reservecushion_t) + \beta_9 \log(oilprice_{t-30}) + \beta_{10}(\log(ccgtoutage_t) + \beta_{11}(\log(forceoutage_t) + \varepsilon_t .$$

The model considered production cost by including 30-day lagged of fuel oil prices. Moreover, it also accounted for forecasted energy demand by including $demand_t$.

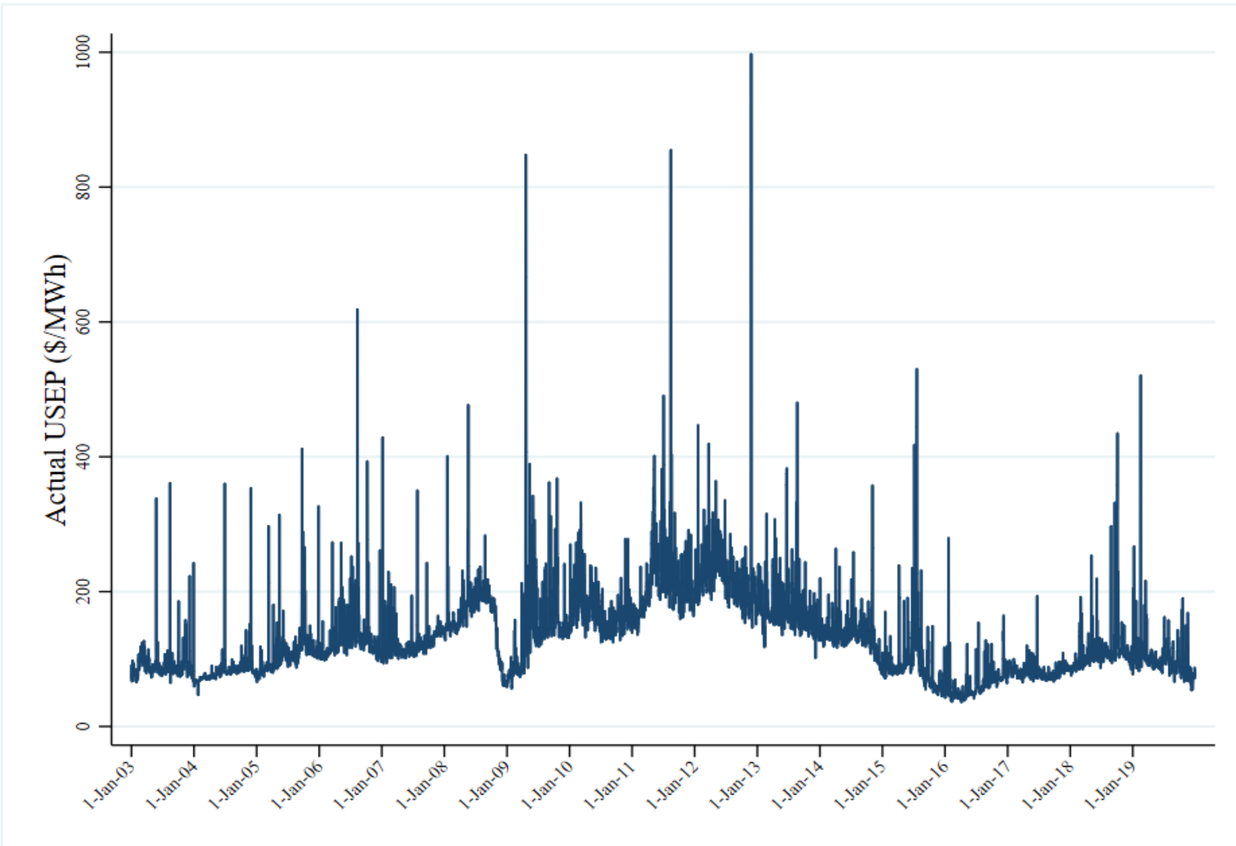
On the other hand, there were several limitations of this previous model. Firstly, the model was static which did not account for any dynamic effect of the dependent variable. In fact, time series data with high frequency usually exhibited strong serial correlation between the periods. For instance, the USEP of period *t* heavily depended on the USEP in period *t-1*. Therefore, the predicted values of the USEP based on a static model might be misleading without the dynamic term.

Secondly, the USEP could be affected by many macroeconomic factors such as the annual Gross Domestic Product (GDP) of Singapore and Consumer Purchasing Index, other than demand and supply. Especially in 2020, when the macroeconomic factors fluctuated heavily due to the outbreak of Coronavirus Disease 2019 (COVID-19). For instance, the Ministry of Trade and Industry announced that Singapore's GDP growth was expected to be from "-6.5 to -6.0 percent" in 2020. Hence, the performance of a model without taking macroeconomic factors into account might not be consistent throughout the years. Therefore, we included dummies for all the years as extra explanatory variables in the regression, to control for year-specific effects which might potentially affect both explanatory and dependent variables simultaneously.

[1] There was another model specification based on the benchmark model in 2007, where there were only three explanatory variables, namely, CCGT electricity supply, supply cushion, and lagged fuel oil price. We denoted this model specification as Existing OLS2 in this report for comparison.

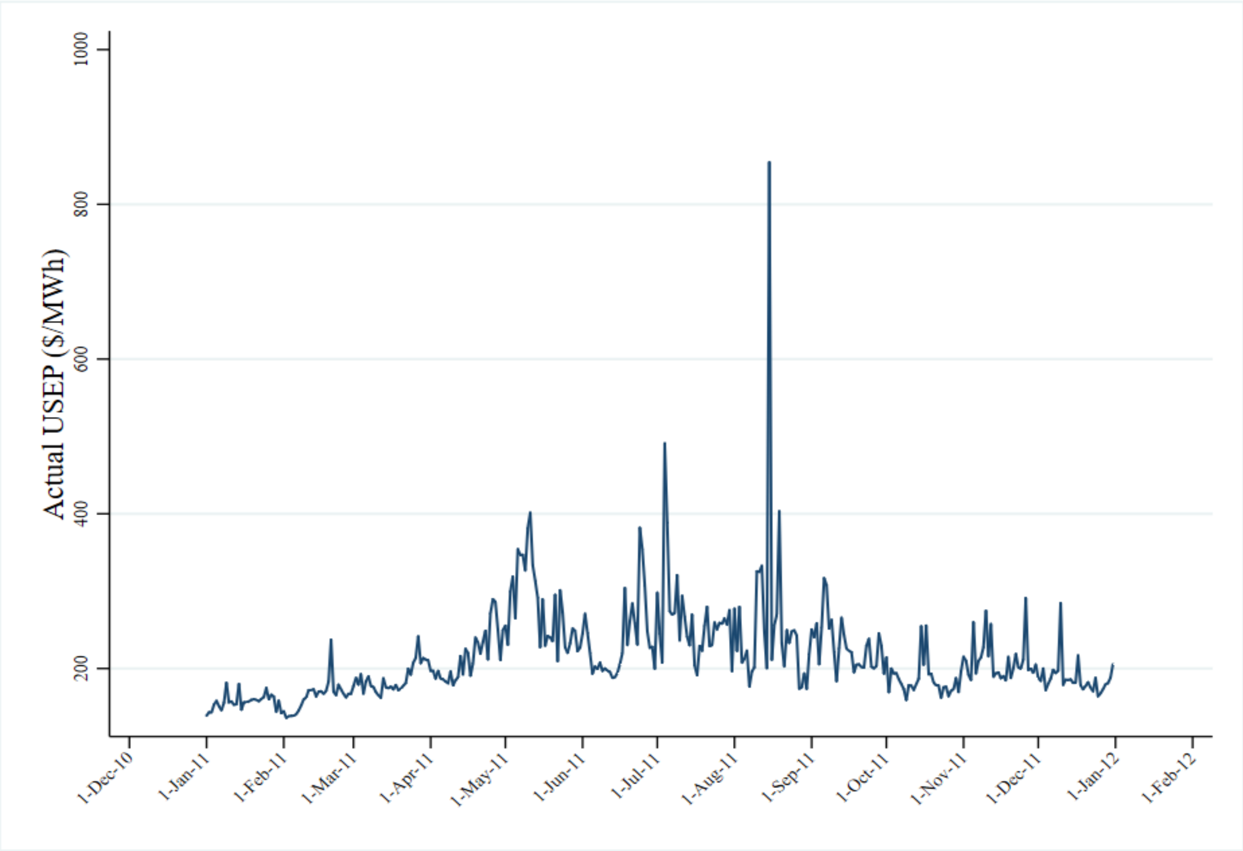
Thirdly, based on Figure 1, we observed that there were some patterns of the daily USEP during certain periods. For instance, there was an increasing trend from 2003 to 2008 and a decreasing trend from 2012 to 2016. Without accounting for this factor, the results based on certain subsamples would be affected by these time-specific trends. Hence, we included an additional time trend explanatory variable to address the trends in different periods.

Figure 1: Time Series Plot of Daily USEP From 2003 to 2019



Besides, in time series data with high frequency, there might exist seasonality. Based on economic theory, the prices of goods are higher when the demand increases. Thus, since the average temperature of Singapore is generally higher between June and August, compared to between January and March, the USEP may be higher from June to August. This hypothesis is supported by Figure 2, which shows the time series plot of daily USEP in 2011.

Figure 2: Time Series Plot of Daily USEP in 2011



We observed the USEP during first quarter of the year were generally lower compared to the USEP in the second and third quarters. Similarly, as the consumption of electricity on weekends was higher compared to that on weekdays, we would expect higher USEP on Saturdays and Sundays. Therefore, we included both weekly and monthly dummies in the revised model to address such seasonality effects. This allowed us to control for season-specific factors in the regression model and facilitate comparisons between different time periods. By implementing seasonal adjustment, we could then examine the effect of each month, instead of the average annual impact, on the USEP. Hence, the revised model would have a stronger explanatory power of dependent variable, R^2 and a higher prediction accuracy.

Next, the previous model included one-month lagged fuel oil price as the additional explanatory variable to account for the production cost of electricity. However, it might have limited explanatory power of the dependent variable as the relationship between the lagged fuel oil price and the USEP might not be linear. When the fuel oil price increased by a certain amount, it would lead to a higher production cost of electricity and hence, the USEP increases. Nevertheless, if the fuel oil price increased exponentially to an extremely high level, the electricity consumers might reduce their consumption of electricity, while the electricity suppliers would incur a higher production cost. Either way, the quadratic form of the lagged fuel oil price might have a significant impact on the USEP. To account for this, it was necessary to add the higher order term of lagged fuel oil price. In the revised model, we took the square of the one-month lag of fuel oil price to examine the higher order effect of fuel oil price.

Besides, there might exist some heterogenous effects which could also potentially affect the USEP. By standard economic theory, lower supply leads to higher prices, given demand being constant. Conversely, higher demand also causes higher prices, given supply being constant. It would be interesting to examine the USEP when both supply and demand are high at the same time. By multiplying the outages and forecasted demand, we constructed new interaction terms to explain the dependent variable. More specifically, we let both planned and forced outages interact with the forecasted demand as there might exist heterogeneity within the interaction terms. For instance, the amount of unavailable electricity supply due to planned outage might have a greater impact on the daily USEP because the amount is usually significant, while a forced outage may have less impact on USEP since the unavailable energy supply is small[2].

Lastly, we refined the algorithm in two ways. First, we reassured and incorporated a command that detect and drop any USEP less than $50/MWh and more than $4,000/MWh. As mentioned in the previous report, this ensured that the econometric model was not estimated based on data from a period with abnormally high prices that skewed estimations. Also, we developed an additional algorithm to compute one-month lagged fuel oil price for the model. The previous approach computed 30-day lagged fuel oil prices, which was not always the case for every month. This generally would not affect the results significantly when the fuel oil prices are collected on a monthly basis and do not fluctuate by much. Nevertheless, our approach would be more rigorous and flexible, as the revised algorithm would compute an exact one-month lag of fuel oil price instead of 30 days and hence, it could accommodate more granular data in the future (e.g. daily fuel oil prices instead of monthly fuel oil prices).

2.2 Revised Model

To address all the potential problems that mentioned earlier, our revised OLS model is expressed as follows:

<u>Revised Model (Revised OLS)</u>

$$\log(usep_t) = \alpha + \beta_1 \mathbf{\log(\boldsymbol{usep_{t-1}})} + \beta_2 trend + \beta_3 \log(ccgtsupply_t) + \beta_4 \log(stsupply_t) + \beta_5 \log(supplycushion_t) + \beta_6 \log(offers_t) + \beta_7 \log(demand_t) + \beta_8 \log(reservecushion_t) + \beta_9 \log(lagoilprice) + \beta_{10} \mathbf{\log(lagoilprice)^2} + \boldsymbol{\beta_{11}(\log(forceoutage_t) * \log(demand_t))} + \boldsymbol{\beta_{12}(\log(ccgtoutage_t) * \log(demand_t))} + \boldsymbol{\beta_{13}(\log(stoutage_t) * \log(demand_t))} + \gamma_k \sum_{k=1}^{k} year_k + \delta_m \sum_{m=1}^{12} month_m + \pi_p \sum_{p=1}^{7} weekday_p + \varepsilon_t .$$

---

[2] By adding the three variables of outages in the regression, we found that the coefficient of forced outages was statistically insignificant, indicating that it had no power in explaining the daily change of USEP. The *t*-statistics of the other two coefficients were greater compared to that of forced outages.

3 Econometric Model Results

We first run the revised regression model with full sample (from 2003 to 2019) and the results are shown in column (1) of Table 2.

Table 2: Regression Results of Full Sample From 2003 to 2019

Dependent Variable	Logarithm of Actual USEP Below $100/MWh		
	(1)	(2)	(3)
Specification	**OLS**	**ARIMA**	**LASSO**
Lagged USEP	0.388***	0.989***	0.390***
	(15.635)	(498.267)	(15.801)
Lagged Moving Average		-0.666***	
		(-113.680)	
Trend	-0.000		
	(-1.212)		
CCGT Supply	-0.234***		-0.267***
	(-6.082)		(-7.991)
ST Supply	0.014***		
	(3.952)		
Supply Cushion	-1.595***		-1.534***
	(-13.010)		(-12.372)
Offers	-0.568***		-0.560***
	(-12.517)		(-12.609)
Demand	0.819***		0.880***
	(10.600)		(11.465)
Reserve Cushion	-0.248***		-0.243***
	(-13.304)		(-13.194)
Lagged Oil Price	-0.109		
	(-0.524)		
Squared Lagged Oil Price	0.054**		0.041***
	(2.221)		(11.537)
ST Planned Outages			0.000
			(0.382)
Forced Outages*Demand	0.001***		0.001***
	(5.459)		(5.374)
CCGT Planned Outages*Demand	-0.000**		
	(-2.099)		
ST Planned Outages*Demand	0.000		
	(0.608)		
2004	-0.153**		-0.244***

Dependent Variable	Logarithm of Actual USEP Below $100/MWh		
	(1)	(2)	(3)
Specification	**OLS**	**ARIMA**	**LASSO**
	(-2.024)		(-13.381)
2005	0.008		-0.185***
	(0.051)		(-8.821)
2006	0.071		-0.216***
	(0.310)		(-9.114)
2007	0.148		-0.233***
	(0.487)		(-9.268)
2008	0.275		-0.199***
	(0.722)		(-7.163)
2009	0.424		-0.142***
	(0.929)		(-5.204)
2010	0.529		-0.132***
	(0.995)		(-4.691)
2011	0.662		-0.089***
	(1.089)		(-2.770)
2012	0.738		-0.111***
	(1.080)		(-3.083)
2013	0.813		-0.152***
	(1.070)		(-3.915)
2014	0.972		-0.140***
	(1.162)		(-3.271)
2015	1.136		-0.067
	(1.246)		(-1.608)
2016	1.171		-0.117***
	(1.185)		(-2.818)
2017	1.270		-0.122***
	(1.195)		(-2.838)
2018	1.279		-0.207***
	(1.123)		(-4.555)
2019	1.389		-0.189***
	(1.143)		(-4.042)
February	0.014		0.005
	(1.270)		(0.517)
March	0.005		-0.010
	(0.353)		(-1.048)
April	0.033		0.008
	(1.570)		(0.701)
May	0.040		0.007
	(1.489)		(0.686)
June	0.051		0.011

Dependent Variable	Logarithm of Actual USEP Below $100/MWh		
	(1)	(2)	(3)
Specification	**OLS**	**ARIMA**	**LASSO**
	(1.591)		(1.014)
July	0.062		0.014
	(1.580)		(1.220)
August	0.057		0.001
	(1.249)		(0.104)
September	0.056		-0.009
	(1.103)		(-0.847)
October	0.064		-0.010
	(1.100)		(-0.905)
November	0.067		-0.014
	(1.056)		(-1.249)
December	0.056		-0.032***
	(0.799)		(-3.120)
Monday	0.011		0.012
	(1.015)		(1.134)
Tuesday	-0.012		-0.011
	(-1.138)		(-1.052)
Wednesday	-0.024**		-0.022**
	(-2.454)		(-2.328)
Thursday	-0.024**		-0.023**
	(-2.567)		(-2.419)
Friday	-0.024***		-0.023**
	(-2.618)		(-2.496)
Saturday	0.012		0.013
	(1.524)		(1.639)
Constant	7.579**	4.781***	3.157***
	(2.323)	(56.953)	(6.297)
Number of observations	6,032	6,095	6,032
R^2	0.864		0.863

Notes:
1. *Lagged USEP is the dynamic term.*
2. *Robust t-statistics are stated in parentheses.*
3. **** p<0.01, ** p<0.05, * p<0.1.*

Overall, the signs of coefficients associated with explanatory variables were consistent with our expectations. Particularly, we observed significant negative coefficients associated with supply side variables such as CCGT supply, supply cushion and reserve cushion. In contrast, the coefficient of demand was significantly positive, which was consistent with standard demand and supply theory. Also, the coefficient of 0.388 of the lagged USEP suggested that the lagged term had significant dominant power in determining the USEP in the current period. This again revealed the strong serial correlation of the USEP between the periods. Besides, the interaction term of forced outage and demand had a significant positive coefficient, implying that there was heterogenous effect. More specifically, a forced outage leads to a higher USEP. This was supported by the coefficient of 0.001 for *forcedoutages*demand* in column (1) of Table 2. On the other hand, the rest of the two interaction terms exhibited either negative and/or insignificant relationships with the USEP. One possible explanation would be that the planned outages were anticipated by both electricity suppliers and consumers and they might adjust their supply or consumption patterns before the outage. Thus, there was no prominent impact on the USEP. Lastly, using January as the base month, we observed an increasing trend of coefficients from February to December suggesting higher energy price in these months compared to January. Similarly, we observed that the Wednesday dummy exhibited a significant negative impact on the USEP compared to Sunday (base day), even after accounting for higher electricity demand during weekdays by including the forecasted demand. This implies that, given other things being constant, the USEP during weekdays are relatively higher compared to those of weekends. In addition, a coefficient of -0.153 showed that the overall macroeconomic conditions in 2004 causes a drop in the USEP compared to the base year, which was 2003. Similarly, coefficients of 1.279 and 1.389 in year 2018 and 2019 suggest that there was an increase in the USEP by 11% from 2018 to 2019 due to year-specific factors.

3.1 Comparison With Previous Model

R^2 increased from 0.75 to 0.86 in the revised econometric model. This was because more variation in the USEP could be explained by the revised econometric model which accommodated more model features. Therefore, we observed significant coefficients of the dynamic term, interaction terms and dummies.

4 Robustness of the Revised Model

We examine the robustness of our revised OLS by using alternative methods commonly applied in the literature. In particular, we proposed two alternative methods: the Hybrid model with 50% OLS and 50% ARIMA[3], and LASSO. The results are presented in columns (2) and (3) of Table 2.

4.1 Hybrid Model

Instead of relying on the prediction of a single model, a more common way in the literatures of energy price is to construct a Hybrid model, which is a combination of two individual models with weights. Hybrid approaches provide prediction with higher accuracy. Bissing et al. (2019) proposed a hybrid model that combined results from linear regression model with ARIMA and Holt-Winters models to forecast the hourly spot price of electricity in the Iberian market. First, we started with the ARIMA (p,d,q) model, where p was the lag of Autoregressive term, d was the differencing order and q was the lag of Moving Average. ARIMA model is a commonly used

[3] We imposed a weight of 50% for the USEP values predicted by the OLS and 50% for those predicted by the ARIMA model and took the weighted average as the predicted USEP of the Hybrid model.

forecasting method which accounts for trend and seasonality. Since most of our variables were stationary, we constructed the ARIMA with $p=1$, $d=0$, and $q=1$. The proposed ARIMA (1,0,1) model is expressed as follows:

$$\log(usep_t) = \alpha + \beta_1 \log(usep_{t-1}) + \theta_1 \varepsilon_{t-1} + \epsilon_t$$

Then, we combined the results obtained by the ARIMA (1,0,1)[4] model with the revised OLS model.

4.2 Subsample: From 2010 to 2019

As shown in Figure 1, we observed fluctuations of the USEP over the years. We addressed such issue with the addition of time trend and year dummies. The results with full samples from 2003 to 2019 were found to be robust to different model specifications. In this section, we examined whether our revised model was sensitive to the choices of sample. Hence, we run the revised model with subsample from 2010 to 2019 instead of the full data from 2003 to 2019. The regression results of revised model are shown in column (1) of Table 3.

Table 3: Regression Results of Subsample From 2010 to 2019

Dependent Variable	Logarithm of Actual USEP Below $100/MWh		
	(1)	(2)	(3)
Specification	**OLS**	**ARIMA**	**LASSO**
Lagged USEP	0.367***	0.990***	0.368***
	(12.752)	(421.415)	(12.714)
Lagged Moving Average		-0.643***	
		(-76.831)	
Trend	-0.001**		
	(-2.101)		
CCGT Supply	-0.441***		-0.437***
	(-5.918)		(-5.843)
ST Supply	0.006		0.006
	(1.435)		(1.433)
Supply Cushion	-1.475***		-1.476***
	(-6.992)		(-7.032)
Offers	-0.763***		-0.770***
	(-11.413)		(-11.576)
Demand	1.344***		1.338***
	(10.870)		(10.868)
Reserve Cushion	-0.270***		-0.271***
	(-7.859)		(-7.883)
Lagged Oil Price	1.008**		0.979**

[4] The regression results based on the ARIMA (1,0,1) model with full sample and subsample are presented in column (2) of Table 2 and 3 respectively. Also, please refer to Figure A2 in the Appendix for the outliers obtained by Hybrid model from 2003 to 2019.

Dependent Variable	Logarithm of Actual USEP Below $100/MWh		
	(1)	(2)	(3)
Specification	**OLS**	**ARIMA**	**LASSO**
	(2.488)		(2.447)
Squared Lagged Oil Price	-0.072		-0.069
	(-1.477)		(-1.427)
CCGT Planned Outages			-0.002**
			(-2.141)
Forced Outages*Demand	0.001***		0.001***
	(4.252)		(4.173)
CCGT Planned Outages*Demand	-0.000**		
	(-2.132)		
ST Planned Outages*Demand	0.000		
	(0.510)		
2011	0.233**		0.036**
	(2.446)		(2.320)
2012	0.425**		0.031
	(2.263)		(1.427)
2013	0.609**		0.020
	(2.156)		(0.827)
2014	0.863**		0.077**
	(2.295)		(2.349)
2015	1.196**		0.215***
	(2.551)		(5.980)
2016	1.368**		0.190***
	(2.428)		(5.139)
2017	1.524**		0.152***
	(2.330)		(4.308)
2018	1.597**		0.028
	(2.136)		(0.755)
2019	1.822**		0.058
	(2.167)		(1.518)
February	0.025		0.009
	(1.572)		(0.643)
March	0.018		-0.013
	(0.877)		(-0.978)
April	0.042		-0.005
	(1.569)		(-0.366)
May	0.061*		-0.003
	(1.780)		(-0.187)
June	0.078*		-0.001
	(1.948)		(-0.084)
July	0.111**		0.015

Dependent Variable	Logarithm of Actual USEP Below $100/MWh		
	(1)	(2)	(3)
Specification	**OLS**	**ARIMA**	**LASSO**
	(2.242)		(0.962)
August	0.114**		0.001
	(2.033)		(0.047)
September	0.124*		-0.005
	(1.957)		(-0.357)
October	0.131*		-0.015
	(1.822)		(-1.025)
November	0.148*		-0.014
	(1.879)		(-0.950)
December	0.137		-0.042***
	(1.568)		(-3.303)
Monday	0.012		0.013
	(0.891)		(0.955)
Tuesday	-0.027**		-0.026**
	(-2.162)		(-2.083)
Wednesday	-0.028**		-0.027**
	(-2.428)		(-2.346)
Thursday	-0.029**		-0.028**
	(-2.573)		(-2.485)
Friday	-0.026**		-0.025**
	(-2.349)		(-2.270)
Saturday	0.006		0.006
	(0.627)		(0.656)
Constant	10.010**	4.817***	0.322
	(2.157)	(37.685)	(0.306)
Number of observations	3,477	3,539	3,477
R^2	0.898		0.898

Notes:
1. *Lagged USEP is the dynamic term.*
2. *Robust t-statistics are stated in parentheses.*
3. **** p<0.01, ** p<0.05, * p<0.1.*

4.3 Comparison With Full Sample (From 2003 to 2019)

Overall, we obtained similar coefficients compared to the full sample results. Furthermore, the R^2 increased from 0.864 to 0.898. This implied better fitness of model using the subsamples. One possible reason is that the results of 2003 to 2019 are heavily affected by the observations during past years (2003-2009). However, more recent data (2010-2019) may better explain the current energy market.

4.4 Results Based on Weekly Average Data

As we observed in Figures 1 and 2, there existed significant volatility in daily USEP. This was partly due to the mechanism of the market clearing process. Generators could submit up to 10 price-quantity pairs of energy offers (electricity supply). Thus, there was some randomness in the offers that met the total forecasted demand. Since outliers were defined as abnormal high prices due to market inefficiency instead of randomness, prediction which relied on the daily USEP might be misleading. To avoid such randomness, we took the weekly average of all the variables to smooth out the volatility of the daily data from 2003 to 2019. The results are presented in Table 4.

Table 4: Regression Results of Weekly Sample From 2003 to 2019

Dependent Variable	**Logarithm of Actual USEP Below $100/MWh**
	(1)
Specification	**OLS**
Lagged USEP	0.496***
	(12.127)
Trend	-0.000
	(-0.266)
CCGT Supply	-0.315***
	(-4.444)
ST Supply	-0.005
	(-0.805)
Supply Cushion	-1.264***
	(-4.684)
Offers	-0.628***
	(-7.710)
Demand	1.157***
	(7.406)
Reserve Cushion	-0.063**
	(-2.430)
Lagged Oil Price	-0.748***
	(-2.863)
Squared Lagged Oil Price	0.118***
	(3.830)
Forced Outages*Demand	0.000
	(0.433)
CCGT Planned Outages*Demand	0.000
	(0.623)
ST Planned Outages*Demand	-0.000
	(-0.805)
Constant	2.169*
	(1.743)
Number of observations	877

Dependent Variable	Logarithm of Actual USEP Below $100/MWh
	(1)
Specification	**OLS**
R^2	0.897

Notes:
1. *Lagged USEP is the dynamic term.*
2. *Robust t-statistics are stated in parentheses.*
3. **** p<0.01, ** p<0.05, * p<0.1.*
4. *All the variables in the regression are converted to weekly averages, instead of daily observations.*

4.5 Comparison With Daily Observations

In Table 4, we observe similar coefficients for all the explanatory variables compared to the daily observations, implying that our revised model is robust to different frequencies of observations.

5 Variable and Model Selection

We introduced several measures for variable selection. More specifically, we applied two information criteria: Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC); and three measures for prediction accuracy: Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), to help us select the explanatory variables for our OLS model. *Lower* values of AIC and BIC indicate better fitness of the model and *lower* values of the three prediction accuracy measures also imply more powerful prediction. For the computation of all the five measures, please refer to the Appendix. In addition, the commonly used R^2 was also included in our variable selection process, where higher R^2 values suggest higher explanatory power of explanatory variables.

5.1 Model Selection: Full Sample Results

In this section, we compared the revised model with the previous ones based on the six measures of variable and model selection, using the full sample. The results are shown in panel (A) of Table 5.

Table 5: Performance of Models

		R^2	MAPE	MAE	RMSE	AIC	BIC
(A)	Revised OLS	0.86	1.82	8.96	0.15	-5,942	-5,627
	Existing OLS1	0.80	2.35	11.60	0.18	-3,570	-3,503
	Existing OLS2	0.75	2.72	13.30	0.20	-2,268	-2,241
	Hybrid		1.73	8.55	0.15		
	LASSO	0.86	1.83	9.00	0.15	-5,930	-5,642
(B)	Revised OLS	0.90	1.82	8.97	0.14	-3,763	-3517
	Existing OLS1	0.86	2.21	10.88	0.17	-2,654	-2,592
	Existing OLS2	0.81	2.57	12.56	0.19	-1,717	-1,692

Notes:
1. *Panel (A) denotes full sample from 2003 to 2019.*
2. *Panel (B) denotes subsample from 2010 to 2019.*

As we observed above, our model (revised OLS) had a higher R^2, lower values for AIC and BIC, as well as lower values for prediction accuracy measures, compared to the existing OLS models, suggesting the better fitness and forecasting performance of our model.

5.2 Model Selection: Subsample Results

Similarly, we also compared the variable and model selection criteria of the three models using subsample from 2010 to 2019. The results are presented in panel (B) of Table 5. Again, all the measures suggested better fitness and forecasting performance of our model (revised OLS) compared to the previous ones. Interestingly, we observed that the R^2 based on the subsample was slightly larger compared to that of the full sample, suggesting better explanatory power of the subsample. One possible reason would be that the full sample from 2003 to 2019 included too many observations from the past, which could no longer explain the current market very well, compared to the subsample. Since the subsample still covered daily observations of nearly 10 years, the sample size was sufficient to generate valid and reliable results if our focus was on the current electricity market.

5.3 Hybrid Model

In addition, we also examined the difference between the Hybrid model and our revised OLS. Since the Hybrid model is a combination of two models, measures such as the R^2, AIC, and BIC are not presented here. Also, as shown in panel (A) of Table 5, we observed similar measures for the Hybrid model. In particular, the forecasting performance was slightly better compared to the revised OLS.

5.4 Selecting Variables: LASSO

Next, we applied the LASSO method. The LASSO is a popular machine learning method which helps the user select the most suitable explanatory variables for the regression.

Therefore, the LASSO-suggested OLS model with the full sample from 2003 to 2019 is presented as follows[5]:

$$\log(usep_t) = \alpha + \beta_1 \log(usep_{t-1}) + \beta_2 \log(ccgtsupply_t) + \beta_3 \log(supplycushion_t) + \beta_4 \log(offers_t) + \beta_5 \log(demand_t) + \beta_6 \log(reservecushion_t) + \beta_7 \log(lagoilprice^2) + +\beta_8(\log(stoutage_t) + \beta_9(\log(forceoutage_t) * \log(demand_t)) + \gamma_k \sum_{k=1}^{k} year_k + \delta_m \sum_{m=1}^{12} month_m + \pi_p \sum_{p=1}^{7} weekday_p \, \varepsilon_t \, .$$

We listed the six measures for variable and model selection in panel (A) of Table 5. We compared the model suggested by LASSO with our revised model based on the criteria suggested by us. Overall, we observed similar results for all the six measures of model fitness. Moreover, our revised OLS seemed to have even slightly better performance, indicating that our results were robust to different model specifications and hence, we could rely on the conclusion obtained from the revised OLS model.

6 Identification of Outliers

The definition of outliers in our revised methodology mainly remained the same as that since the MSCP Annual Report 2007 – any actual energy price which lies above the predicted USEP with additional 3 SD is considered as an outlier. Moreover, we also imposed a more stringent constraint on weekly average regression, where the upper bound was defined as 2 SD above the predicted USEP.

6.1 Full Sample: From 2003 to 2019

Outliers identified by revised model based on full sample are presented in panel (A), column (1) of Table 6[6].

Table 6: Outliers Based on Daily Sample

	(1)	(2)	(3)
Specification	**OLS**	**Hybrid**	**LASSO**

[5] The regression results based on LASSO with full sample and subsample are presented in column (3) of Table 2 and 3, respectively.

[6] Please refer to Figure A1 in the Appendix for the outliers obtained by OLS from 2003 to 2019.

	Outlier Date	Difference	Outlier Date	Difference	Outlier Date	Difference
(A)	26-May-03	0.1126299	26-May-03	0.1706276	26-May-03	0.1155314
	14-Aug-03	0.2173858	14-Aug-03	0.2785182	14-Aug-03	0.2224932
	29-Jun-04	0.333075	30-Dec-03	0.0121093	29-Jun-04	0.3328147
	28-Nov-04	0.1794577	29-Jun-04	0.3259635	28-Nov-04	0.1778307
	12-Aug-06	0.3166251	28-Nov-04	0.2272849	12-Aug-06	0.316752
	6-Jan-07	0.2347178	12-Mar-05	0.0694818	6-Jan-07	0.2360306
	21-Apr-09	0.2796235	12-Aug-06	0.4054794	21-Apr-09	0.2753592
	15-Aug-11	0.1714087	6-Jan-07	0.3027496	15-Aug-11	0.1786952
	26-Nov-12	0.3752928	20-Apr-09	0.0358791	26-Nov-12	0.3609414
	6-Jul-15	0.1452098	21-Apr-09	0.3174586	6-Jul-15	0.144259
	20-Jul-15	0.0716343	15-Aug-11	0.1903973	20-Jul-15	0.0657816
	22-Jan-16	0.2855797	26-Nov-12	0.5105553	22-Jan-16	0.2793746
	16-Feb-19	0.1873651	6-Jul-15	0.1680222	16-Feb-19	0.1932535
			20-Jul-15	0.0288577		
			22-Jan-16	0.3433232		
			16-Feb-19	0.3740387		
(B)	15-Aug-11	0.0264816	15-Aug-11	0.0538282	15-Aug-11	0.0275631
	26-Nov-12	0.2356396	26-Nov-12	0.3753591	26-Nov-12	0.2304339
	22-Jun-16	0.1785808	6-Jul-15	0.0163937	22-Jan-16	0.1766915
	16-Feb-19	0.0249434	22-Jan-16	0.2305002	16-Feb-19	0.02526
			16-Feb-19	0.2272258		

Notes:
1. *Panel (A) denotes full sample from 2003 to 2019.*
2. *Panel (B) denotes subsample from 2010 to 2019.*
3. *Columns (1), (2), and (3) denote the OLS, Hybrid, and LASSO specification respectively.*

In comparison, as shown in the table below, the previous model detected two outliers in 2019, while our method detected only one outlier. Looking at the true values of the USEP on those two dates, it is obvious that the USEP on 16 February 2019 was much higher than that on 8 January 2019. Hence, the probability of the USEP on 16 February 2019 being an outlier was also higher. This implied that the electricity market became more efficient over time as shown by our model, while the previous model might be misleading as some outliers detected were not significant.

Outlier Date	USEP (Previous Model)	USEP (Revised Model)
8-Jan-19	$266.91/MWh	-
16-Feb-19	$520.58/MWh	$520.58/MWh

6.2 Hybrid Model

Also, the outliers detected by the Hybrid model with full sample are listed in panel (A), column (2) of Table 6[7]. Overall, we obtained similar results compared to the revised OLS model. Particularly, unlike the previous model 1 which identified two outliers in 2019, the Hybrid model excluded 18 January 2019 as the outlier. This is again consistent with findings of our revised model.

6.3 Subsample: From 2010 to 2019

Outliers detected by the revised model based on subsample are listed in panel (B), column (1) of Table 6[8]. As shown, 6 July 2015 and 20 July 2015 were sensitive to the choice of subsample from 2010 to 2019. We observed the USEP on those two dates were $417.9/MWh and $444.53/MWh respectively. However, the average USEP in July of 2015 was relatively higher compared to other months in 2015. Therefore, the differences between the upper bound (3 SD above the predicted USEP) and the actual USEP on both days were significantly smaller. In conclusion, since the subsample from 2010 to 2019 could explain the current market better, the two dates were less likely to be outliers compared to other dates.

Outlier Dates	Difference
15-Aug-11	0.1714087
26-Nov-12	0.3752928
6-Jul-15	0.1452098
20-Jul-15	0.0716343
22-Jan-16	0.2855797
16-Feb-19	0.1873651

6.4 Weekly Average

Similarly, we also detected outliers from 2003 to 2019 using weekly average observations. For the identification of outliers, instead of 3 SD, we used 2 SD above the USEP to determine outliers. This implied that the upper bound and lower bound from the actual USEP became narrower, hence, we should observe more outliers compared to the previous results in section 6.1. The outliers are presented in Table 7.

Specification	Date	Difference
Weekly	16-Apr-09	0.3379416
	2-Jul-15	0.1412997
Monthly	1-Jul-15	0.0133953

[7] We also present the results of Hybrid model based on subsample in panel (B), column (2) of Table 6.

[8] Please refer to Figure A3 and A4 for the outliers obtained by OLS and Hybrid model from 2010 to 2019 respectively.

Notes:
1. *The results are based on weekly and monthly average of full sample from 2003 to 2019.*
2. *The upper bound is constructed by 2 SD above the predicted USEP.*
3. *There is no outlier if the upper bound is defined as 3 SD above the predicted USEP.*

We observed significantly fewer outliers after smoothing out the randomness from the daily USEP. Interestingly, we obtained no outliers if we switched the definition of upper bound from 2 SD above the predicted USEP to 3 SD above the predicted USEP. Therefore, these results indicated that the electricity market was efficient when we smoothed out all the daily fluctuations. In addition, we also presented the outliers detected based on monthly average observation. Again, we observed fewer outliers after further smoothing out the randomness. Also, July 2015 was still categorised as the outlier, which was consistent with our weekly findings.

6.5 LASSO

We presented outliers based on full sample and subsample in panels (A) and (B), column (3) of Table 6[9]. Overall, we obtained the same outlier dates as detected by our revised OLS, suggesting that our model was robust to different model specifications.

6.6 Interpretation of Outliers

Overall, the outliers detected were associated with high actual USEP. However, we could not simply interpret the outliers from the perspective of supply and demand, and regressors used in general, since their impact on the USEP had been reflected in the predicted USEP.

For instance, 26 May 2003, detected as an outlier with the actual USEP of $338.63/MWh. The corresponding demand was 4,266.43 MW, while the supply generated by CCGT was 3,126.96 MW. The supply was significantly lower than the market demand. Based on the market clearing mechanism, the USEP became exceptionally higher and was eventually detected as an outlier. However, such shortage of electricity supply could be due to the randomness from generators' offers. Since outliers are defined as the abnormally high prices due to market inefficiency instead of randomness, prediction relying on daily USEP may be misleading. Therefore, the results based on weekly average observations show that the market was actually efficient as fewer outliers were detected, after smoothing out the daily randomness from generators' offers.

[9] Please refer to Figure A5 and A6 in the Appendix for the outliers obtained by LASSO from full sample and subsample respectively.

7 Conclusion

In conclusion, there were several advantages of the revised model[10].

Table A2: Model Comparison

Revised Model				Previous Model			
Model Specification	**Factors Considered**	**Robust-ness**	**Statistical Inference**	**Model Specification**	**Factors Considered**	**Robust -ness**	**Statistical Inference**
Dynamic Model	Demand	Hybrid model	*t*-statistics/ *p*-values	**Static Model**	Demand		*t*-statistics/ *p*-values
	Supply	LASSO	AIC		Supply		
	Cost	Sub-samples	BIC		Cost		
	Seasonal effect	Weekly average	MAPE		Seasonal effect		
	Heterogeneous effect	Monthly average	MAE				
	Non-linear effect		RMSE				
	Dynamic effect						

Firstly, the revised model had a stronger explanatory power of the USEP as it included more explanatory variables such as dynamic term, time trend and interaction terms, which were shown to be statistically significant. More specifically, the interaction terms accounted for time invariant factors from heterogenous effect, while the dynamic term and time trend captured time variant factors. These new features ensured the revised model would retain a strong predictive power of outliers when more observations are available in the future. The results were shown to be robust with different model specifications and subsamples. Particularly, we included more up-to-date machine learning algorithm (such as the LASSO) to examine the robustness of our model. Next, as we constructed year dummies to address any change in macroeconomic state through the years, the revised model would more flexible in the long run. Especially when there were significant fluctuations of economic factors in 2020 due to the COVID-19, we could still rely on the prediction of our revised model. Thirdly, we introduced two information criteria and three measures for prediction accuracy, on top of the existing standard R^2, which made a clear-cut variable and model selection process. The policymaker should consider the overall performance based on the six measures to decide on the explanatory variables and model specifications, instead of relying on any individual measure. In addition to the model revision, we also refined the programming codes to accommodate more granular data. This created greater possibility to improve the prediction accuracy when more data is available in the future. Finally, the model would need to be revised again when the structure of the NEMS or relevant market rules and regulations change drastically. For instance, new regulations such as extra emission costs for non-eco-friendly electricity generation could be imposed to promote a green electricity market in Singapore. This may change the behaviour of electricity suppliers and well as electricity consumers in the context of private provision of public goods (Kotchen, 2006). Then, extra explanatory variables related to electricity generation from renewable sources would need to be considered in the new model.

[10] For further comparison between the revised model and existing model, please refer to Table A2.

8 References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, *19*(6), 716-723.

Bissing, D., Klein, M. T., Chinnathambi, R. A., Selvaraj, D. F., and Ranganathan, P. (2019). A hybrid regression model for day-ahead energy price forecasting. *IEEE Access*, *7*, 36833-36842.

de Marcos, R. A., Bello, A., and Reneses, J. (2019). Electricity price forecasting in the short term hybridising fundamental and econometric modelling. *Electric Power Systems Research*, *167*, 240-251.

Dickey, D. A., and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, *74*(366a), 427-431.

Granger, C. W. J., and Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics*, *2*(2), 111-120.

Kotchen, M. J. (2006). Green markets and private provision of public goods. *Journal of Political Economy*, *114*(4), 816-834.

Market Surveillance & Compliance Panel Annual Report 2007.

Market Surveillance & Compliance Panel Annual Report 2009.

Occasional Paper by Market Surveillance and Compliance Panel, How Market Fundamental Factors Affect Energy Prices in the NEMS – An Econometric Model, 2007

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

Said, S. E., and Dickey, D. A. (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, *71*(3), 599-607.

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

9 Appendix

9.1 Stationary Tests

There are two common stationarity tests, namely, Augmented Dickey–Fuller (ADF) (Said and Dickey, 1984) test and Phillips–Perron (PP) test (Phillips and Perron, 1988).

Suppose we have a model:
$\Delta y_t = \alpha + \rho y_{t-1} + \gamma_1 \Delta y_{t-1} + \varepsilon_t,$

where α is a constant, ρ and γ_1 are the coefficients, and ε_t exhibits Gaussian distribution. The ADF test is to test the null hypothesis of $\hat{\rho} = 1$. If the null hypothesis is not rejected, it indicates that y_t is non-stationary.

Similarly, the PP test builds on the Dickey-Fuller (DF) test (Dickey and Fuller, 1979), where:
$\Delta y_t = \alpha + \rho y_{t-1} + \varepsilon_t$

Unlike the ADF test, there is no Δy_{t-1} term in the PP test. Again, the null hypothesis if PP test is to test whether $\hat{\rho} = 1$. The difference between the two tests is that PP test is robust to heteroskedasticity and autocorrelation of ε_t, while the ADF test is not.

Non-stationary data may lead to spurious regression (Granger and Newbold, 1974), which then in turn causes misleading statistical inference. Also, the ARIMA (*p*,*d*,*q*) model with *d=0* also requires stationarity condition. Therefore, it is critical to check stationarity of data before running the regression.

We run both stationarity tests for the main variables in our model. It turned out that most of the variables were stationary and the only exception was the one-month lagged oil price. Since the lagged oil price was first difference stationary (i.e. $fdlagoilprice_t = lagoilprice_t - lagoilprice_{t-1}$ is stationary), we run the regression model with $fdlagoilprice_t$ and $fdlagoilprice_t^2$ instead of $lagoilprice_t$ and $lagoilprice_t^2$ to check the impact of non-stationarity on the empirical results. Overall, the regression results were similar and outlier dates remained the same for the full sample regression. Hence, we still used $lagoilprice_t$ instead of $fdlagoilprice_t$ in the main regression.

Variable	ADF Test	PP Test
$\boldsymbol{usep_t}$	✓	✓
$\boldsymbol{ccgtsupply_t}$	✓	✓
$\boldsymbol{stsupply_t}$	✓	✓
$\boldsymbol{supplycushion_t}$	✓	✓
$\boldsymbol{offers_t}$	✓	✓
$\boldsymbol{demand_t}$	✓	✓
$\boldsymbol{reservationcushion_t}$	✓	✓
$\boldsymbol{lagoilprice_t}$	×	×

9.2 Forecasting Performance

To evaluate the forecasting performance of our regression model, there are three commonly used measures in literature (de Marcos et al., 2019). They are the Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). These error measures are computed as follows:

$MAPE = \frac{100}{N}\sum_{i=1}^{N} |\frac{\widehat{Y}_i - Y_i}{Y_i}|$;
$MAE = \frac{1}{N}\sum_{i=1}^{N} |\widehat{Y}_i - Y_i|$; and
$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{Y}_i - Y_i\right)^2}$,

where $\widehat{Y}_i$ is the forecast value of dependent variable, Y_i is the actual value, and N is the sample size. In our context, $\widehat{Y}_i$ denotes the predicted value of USEP and Y_i represents the actual USEP. In general, all three indices measure how well the predictions match the observed data. The three indices will be small if the predicted values are very close to the true observations and will be large if for some of the observations, the predicted values deviate from the true observations substantially.

In comparison, MAPE is the percentage version of MAE, while RMSE takes square of the prediction errors before they are averaged. This implies that the weight of large errors in RMSE is larger compared to MAE. Hence, it is more sensitive to extreme values of Y_i. Since the USEP less than \$50/MWh and more than \$4,000/MWh were excluded in our analysis, the conclusion drawn from RMSE is reliable.

Overall, the forecasting performance should not be evaluated based on any one of the three measures alone. For instance, if we want to penalise larger prediction errors more, then RMSE is more appropriate than MAE or MAPE. Therefore, all the measures need to be taken into consideration when we evaluate the prediction accuracy.

9.3 Information Criteria

Two information criteria are often used for model selection, namely, Akaike's information criterion (AIC) (Akaike, 1974) and Schwarz's Bayesian information criterion (BIC) (Schwarz, 1978). They are computed as follows:

$AIC = -2\ln L + 2k$; and

$BIC = -2\ln L + k\ln N$,

where $\ln L$ is the maximized log-likelihood of the model, k is the number of parameters estimated, and N is the sample size. They are typically used for selecting the best regressors in the regression models, especially when we do not have access to out-of-sample data. Since the main objective of our model is to predict future USEP based on historical data, it is essential to include the two information criteria in the model selection procedure. Intuitively, BIC penalises complex model and large sample size more heavily compared to AIC. Smaller values of AIC and BIC indicate better fitness of the model. Larger value of $\ln L$ indicates greater goodness of fit of the model, while smaller k suggests the model is more parsimonious. It is noteworthy that one can only compare either AIC or BIC of different models based on the same data set. Also, we cannot solely rely on the two criteria when selecting the model as it is possible that both models have poor fitness of data. Hence, we should take all the measures (including the three prediction accuracy measures and R^2) into account when we decide the model specifications.

9.4 Figures

Figure A1: Outliers Obtained Via OLS From 2003 to 2019

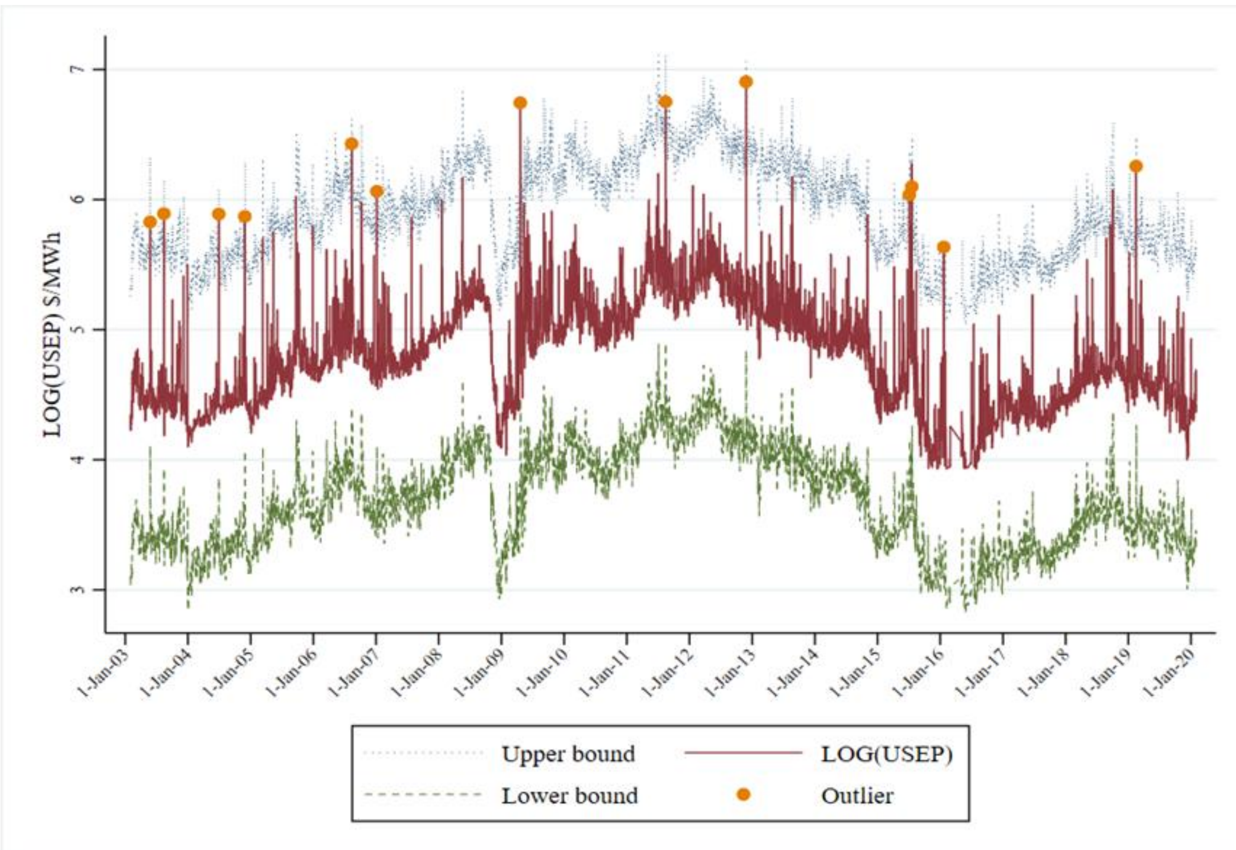


Figure A2: Outliers Obtained Via Hybrid Model From 2003 to 2019

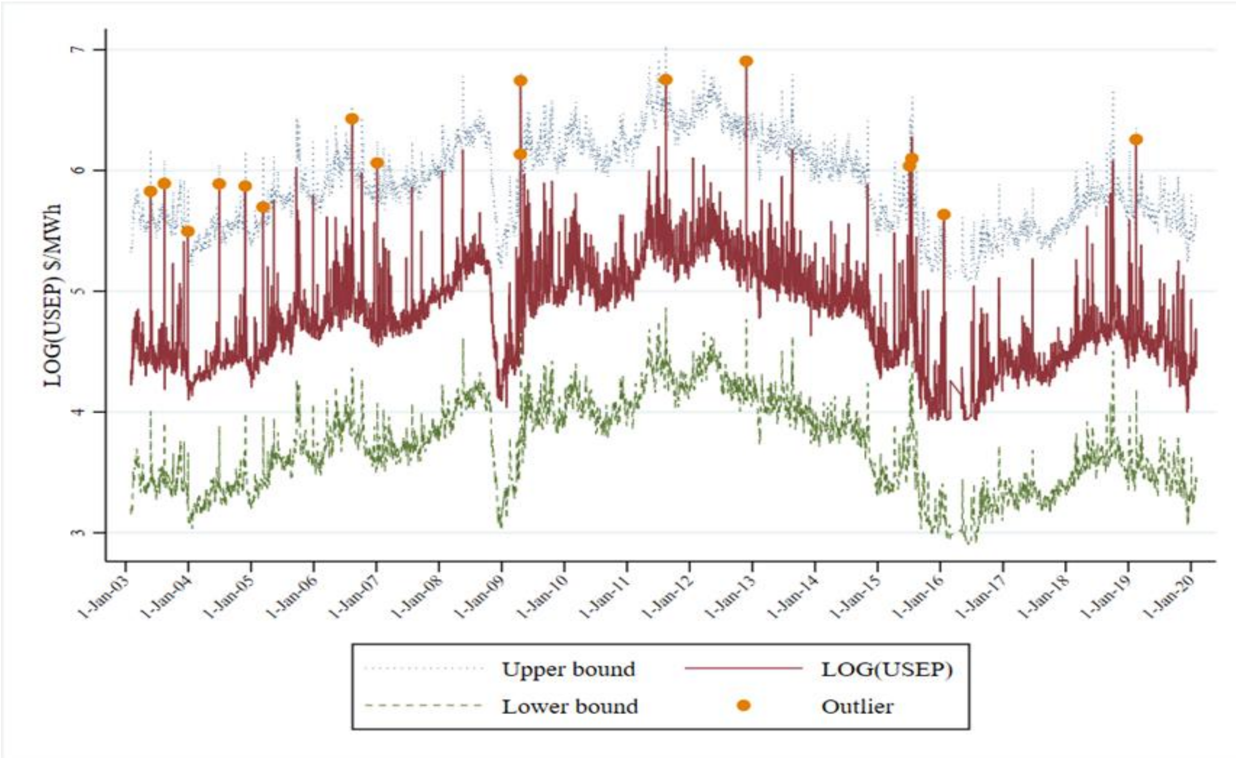


Figure A3: Outliers Obtained Via OLS From 2010 to 2019

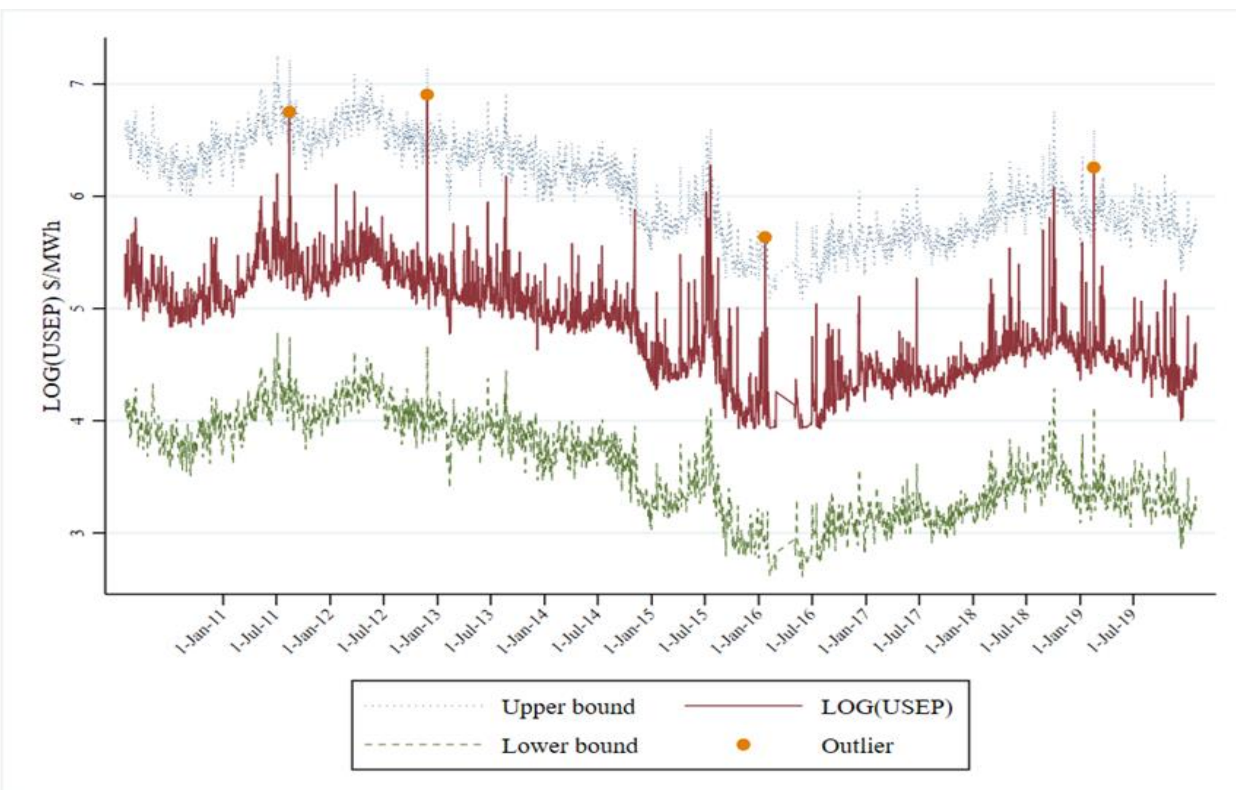


Figure A4: Outliers Obtained Via Hybrid Model From 2010 to 2019

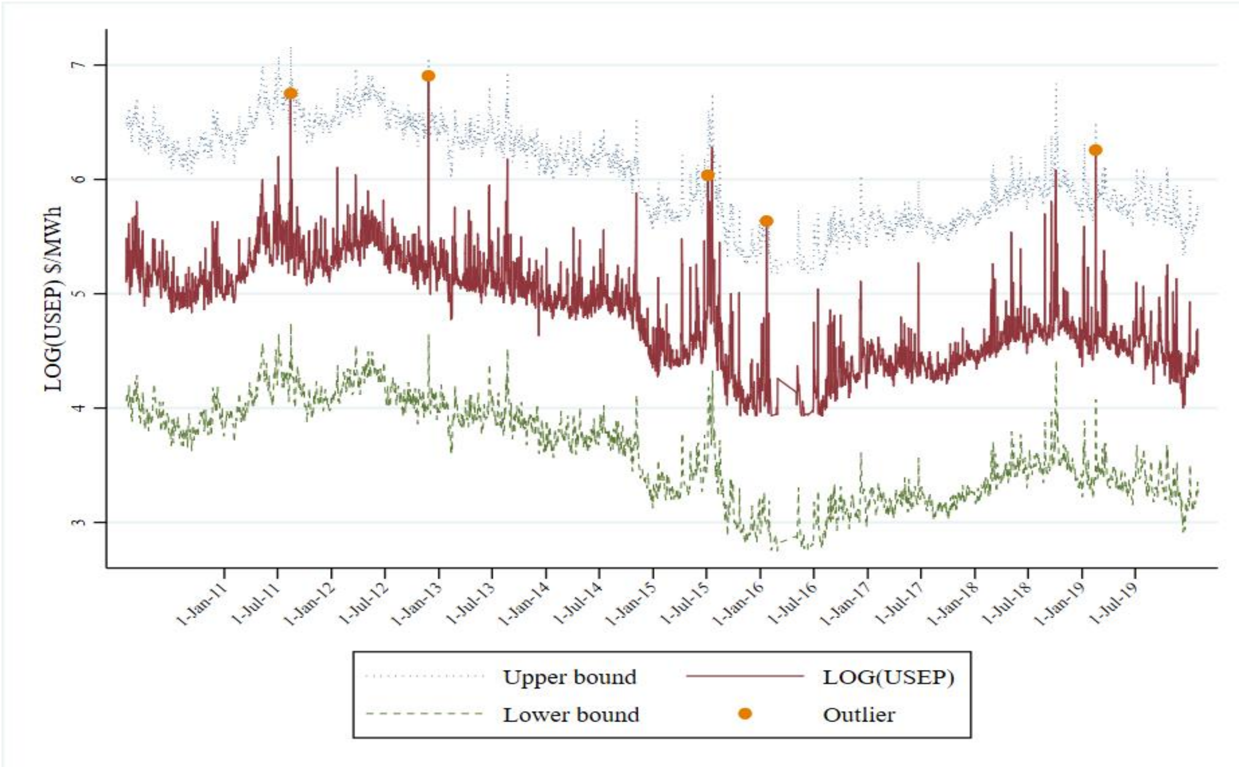


Figure A5: Outliers Obtained Via LASSO From 2003 to 2019

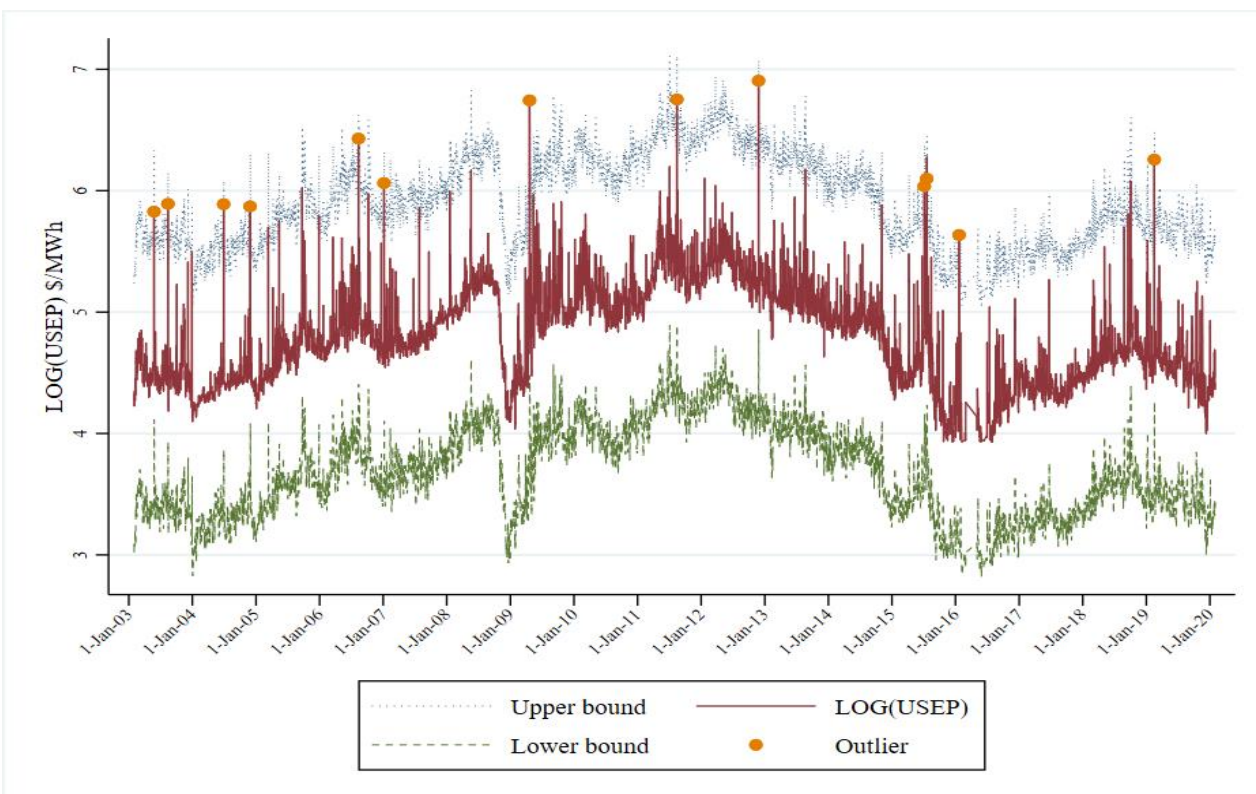


Figure A6: Outliers Obtained Via LASSO From 2010 to 2019