MODELLING ELECTRICITY PRICE SPIKES IN THE PHILIPPINES: A MARKOV REGIME SWITCHING AND TOBIT REGRESSION APPROACH

A Thesis Presented to

the Faculty of the School of Applied Economics

University of Southeastern Philippines

Obrero, Davao City

In Partial Fulfillment of the Requirements

for the degree of

MASTER OF SCIENCE IN ECONOMETRICS

GREJELL B. SEGURA

March 2016

DECLARATION

I certify that the substance of this thesis has not been submitted for any degree and is not currently being submitted for any other degree.

I certify that to the best of my knowledge, any help received in preparing this paper and all sources have been acknowledged in this thesis.

APPROVAL SHEET

The thesis attached hereto, entitled "**MODELLING ELECTRICITY PRICE SPIKES IN THE PHILIPPINES: A MARKOV REGIME SWITCHING AND TOBIT REGRESSION APPROACH**", prepared and submitted by GREJELL B. SEGURA in partial fulfillment of the requirements for the degree of Master of Science in Econometrics is hereby recommended for approval and acceptance.

AGUSTINA TAN-CRUZ, PhD
Adviser

Approved by the Thesis Advisory Committee:

RONALD D. ESTRADA
Chairman

MICHAEL L. AYALA
Member

Accepted as partial fulfillment of the requirements for the degree of Master of Science in Econometrics.

SHIRLEY S. VILLANUEVA, PhD
Vice President for Academic Affairs

BIOGRAPHICAL SKETCH

Grejell B. Segura was born on October 23, 1986 in Maragusan, Compostela Valley. He is the youngest of the four children of church ministers Ramon Francisco Segura and Elena Bacorio Segura. He was born and raised inside the premises of a Christian church where his father acted as the host pastor. The family lived a modest and humble life, just having enough to survive. He had two brothers and a sister. His sister died when she was only 16 years old due to complications of having a cerebral palsy. Grejell was 5 years old at that point hence he never had the privilege of feeling a sisterly love. His brother, the eldest among the boys, died due to a kidney failure. He was so dear to Grejell that his loss had opened his eyes to the reality of life; that everything here on earth, no matter how important they may seem to us, is just temporary.

Preschool was not required then hence his mother did not send him to any class, opting to teach him at home instead. He never was a standout genius but his wit has often been a subject of fun to his parents. They would let him recite bible phrases and sing children songs to the delight of his parent's friends. It was not until grade school when his mother brought him to school. He spent all his elementary school years at Maragusan Central Elementary School. Here he would meet friends who until now is still part of his life. He also had various awards during his stint, bagging honorable mentions for being part of the top 10 which delights his parents every end of the school year. His talent in drawing was also manifested during this period, often chosen to represent the class for poster making contests and even winning some of them. It was in 5^{th} grade when he realized he has gained love in mathematics, making it his favorite subject. He even was made a school quizzer for the said subject. He graduated with 2^{nd} honorable mention after the sixth grade. Though he may not be the best in his class, his parents always believed in his ability. They always showed support, together they dreamed of him becoming a successful professional, a Civil Engineer he would always proudly say.

High school was a little different. While being enrolled at Maragusan National High School, he never thought of acquiring awards and merits. He gave up the thought of being part of the class' top students. Adolescence has also led him to become remote to some classmates. He was then known for unconsciously having a snub look, not having the guts to talk to new people, and often satisfied observing them from a distance. This would limit his friends early on in his high school life. This however did not hinder him to gain knowledge and acquire more talents. It was in the 2^{nd} year when he started learning musical instruments and made use of them to help his father during church services. This passion has bonded him to God ever since. Because of this, his confidence was boosted on the 3^{rd} year. This time he had made personal improvements and gained a number of friends. It was the most memorable year of his high school life. He participated in most events at school, representing as the cartoonist for the Journalism Conference, then reinstated as part of the Math quiz team, and participated the search for Mr. MNHS during the Intramurals.

However, 3^{rd} year was not all nice to him. His father stopped being a host pastor during this time and the family moved to another house, away from the childhood neighbors he had known for 15 years. This was saddening and challenging to the whole family who are used to live within the church’s premises. His father resorted to farming to provide for the family. To Grejell, it was harder as he realized that finances would be tighter for the whole family and the time to go to college has gotten nearer. Though bothered by this circumstance, he never gave up his dream. He ran to become a Sangguniang Kabataan councilor shortly after that period. He served truthfully while also continuing his last year in high school. The experience of being in the public office has given him more confidence until he graduated.

No matter how hard the financial state of the family has become, his father never gave up on him. He was sent to Davao City to pursue his dream, bringing his father’s promise of providing him all he would need. At the University of Southeastern Philippines (USEP), he took the exam for the civil engineering course. The result however was not favorable to him as his remarks were only good enough for the waiting list. Not having the patience to wait, he decided to enroll the BS in Mathematics while initially thinking he would later shift to the engineering course after first year. He unwaveringly continued his study despite the challenges he met during his college days. It was during this time when his dear brother died of his illness. The family’s finances were literally drained. Though this brought him emotionally down, it only made him stronger. The thought of being his parent’s last and only hope has fueled his will to finish the course. The dream of being an engineer was gone and slowly forgotten. His only wish was to finish any course and become a professional to delight his parents.

True indeed, with God’s help, he survived college and finished BS in Mathematics in April 2007. This was a memorable moment to the whole family, making his mother cry a river during his graduation march. The success though was only just a start of his dream. When most of his classmates were applying for jobs right after college, he opted to go back to his homeland and decided to have a break. There he helped his parents on their banana farm and learned how hard it is to earn money through farming. For 10 months he lived in the province thinking about the next step he would take, careerwise. Then he decided to pursue a master's degree.

Though away from the city, he started applying for a job early in 2008 through his friend's referrals. It was also at the start of that year that he took a job from the local government where he was assigned as an office messenger. It was a job he considered temporary as he desires to move back to Davao City whenever he gets hired by a company from Davao. Fortunately, in March 2008 he was hired by JACM Software Services, a BPO company. He immediately decided to relocate to Davao City having the job that he considered would help him get through his goal. Upon reading articles explaining the 2007 financial crisis, he decided to pursue a degree related to economics. Hence, he decided to enroll MS in Econometrics in USEP's School of Applied Economics (SAEc) also considering his bachelor's degree in Mathematics.

There he had the opportunity to learn from a number of good professors, Dr. Tina Tan-Cruz, Dr. Purisima Bayacag, Dr. Ed Cruz, and the late Dr. Ed Prantilla. It was a rewarding experience for him, gaining friends in the process while also learning in the field he is interested into. However, due to some unavoidable circumstance, he stopped in the second semester. The work schedule has hindered him to pursue the course as he has to work on a graveyard shift. Unable to balance the time he has between work and school, he decided to stop until a more favorable schedule came.

After a year and a half, he decided to pursue the degree. His former classmates have already graduated with the Graduate Diploma and seeing them inspired him to continue his study. Without hesitation, and with a more favorable work schedule, he decided to pursue his study. He never resorted to stopping even when his work was hindering him to attend school at times. He eventually graduated with a Graduate Diploma in April 2013, almost five years after he first enrolled. This gave him more determination and decided to pursue the much coveted degree, the MS in Econometrics. He always bears with him his will to give his parents the happiness they deserve.

Grejell's life is truly a testament of an ordinary person setting and achieving his goal for the delight of his loved ones. A number of unavoidable circumstance may try to stop his desires but the will to pursue and the determination to finish what he has started has always been there. Even getting stronger with every struggles, he always credits his success to his parents and acknowledges God for it. With every success, he is now directed to a more satisfying life. He now plans to put up his own family and has always been eager to teach his kids what he has learned in this life.

Grejell B. Segura

This study is dedicated to my sources of inspiration:

Papa Ramon, Mama Elena, Manong Garry,

Kuya Glenn, Ditse, Keith, Czent Lee, Jazz,

Rhyne, Gad and Alen

ACKNOWLEDGMENTS

The success of this study would not be possible without the help of various individuals who believed in my capacity and undoubtedly supported my endeavor.

All glory and honor I give to God alone who, when the picture of success seemed to be fading and the situation got rougher in every step, has been there always to remind me that He is true to His promises. Relying on my own understanding and ways alone would not lead me to this success. I have trusted and leaned on Him all through my life and He never failed me. Thank you for the strength, wisdom and for providing me all the things that I need. Above all, thank you for orchestrating every detail of my life. You are indeed my all living God and savior.

To my loving parents, Papa Ramon and Mama Elena, thank you for nurturing me and making me the person that I am today. This success is the fruit of all my aspiration to please and honor both of you. My love to you has driven me to pursue this personal achievement as compensation to your never ending support. To my brother Garry; my niece, Keith; my nephews, Czent Lee, Jazz, Rhyne, and Gad; my in-laws, Diding and Maricel; to my deceased brother, Glenn; and sister, Ramilyn, thank you for inspiring me in every way.

To Dr. Agustina tan-Cruz, thank you for the patience, motivation, support, and for imparting me a reliable and invaluable knowledge. Your guidance during the course of writing this research is highly regarded. I will forever look up to you as my mentor and a great influence.

To Sir Michael Ayala and Sir Ronald Estrada, thank you for the advices. Your generosity and unselfishness have greatly helped this study. The ideas would not be in fruition without your inputs in every phase of this paper.

To the Joyful Assembly of God Church headed by Ptr. George Geronimo and his pastoral staff, Pas Nina, Ate Jen Lu and Pas Von. To my second family, the Joyful Assembly of God Music Team, thank you for the prayers and encouragement. Being part of the music ministry always gives me the happiness in my heart. Thank you for inspiring me to pursue my dreams and acting as my own panel of advisors.

Finally, thank you to my soon to be wife, Alejandra. I would not have decided to finish this study without your encouragement. Thinking about our future family is the best motivation that pushed and carried me this past year. This is for you.

-GREJ-

ABSTRACT

SEGURA, GREJELL, B. University of Southeastern Philippines, Davao City. April 2016. MODELLING ELECTRICITY PRICE SPIKES IN THE PHILIPPINES: A MARKOV REGIME SWITCHING AND TOBIT REGRESSION APPROACH.

Adviser: Agustina Tan-Cruz, Ph.D.

The main objective of this study is to model the electricity price spikes of the Philippines using a two-state Markov Regime Switching method. It utilizes the Expectation-Maximization algorithm, a more efficient method to compute the parameters compared to the Maximum Likelihood method. In addition, it investigates the effect of the natural gas price and the Indonesian coal price, two of the largest contributors to power production in the Philippines, to the spikes of the electricity price by using the Tobit regression analysis.

Results showed that the Markov Regime Switching of order 3 or MRS(3) model is capable of capturing the spikes of electricity price. The transition probabilities indicated that the normal state is more dominant than the spikes state. It also indicates that the average duration that a spike occurs is 1.9 months. Meanwhile, the average duration for the normal state is 5.9 months.

The tobit regression analysis revealed that the natural gas price and the Indonesian coal price have no significant effect to the electricity price spikes. This finding showed that their effects do not extend to the occurrence of such phenomenon and that there are other factors affecting it, but not included in the study.

Keywords: electricity price spikes, EM algorithm, Indonesian coal price, Markov Regime Switching, natural gas price, Tobit Regression,

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Electricity Market in the Philippines

The Philippine government has long been making efforts to make electricity available for all. The commodity was first used in the country in 1890 and since then it has asserted its significance. The country's economy is undeniably dependent on the availability of the said commodity. In 1970, the country's Gross Domestic Product declined due to the oil crisis that affected energy production in the country (Patalinghug, 2003). Since then, a lot of laws and mandates were made to guarantee the availability of electricity. Since market liberalization was successful in other countries, the country aimed to replicate the success. The aim was not only to distribute the commodity swiftly with lower price due to competition but to have an efficient, transparent and reliable market for electricity (www.wesm.ph).

The electric spot market of the Philippines was materialized due to the electric industry reforms mandated by the Electric Power Industry Reform Act of 2001 (EPIRA) in June 2001. The law mandated the Department of Energy to establish the Wholesale Electricity Spot Market (WESM) and come up with the rules for the operation of WESM. Together with the participation of the electric industry institutions, the WESM Rules were established and promulgated in June 2002. Following this, the Philippine Electricity Market Corporation (PEMC), a non-profit and non-stock corporation, was incorporated in November 2003. It was established to serve as the Autonomous Group

Market Operator (AGMO) assigned to prepare for the initial operations of WESM. A number of trial operations were conducted and soon after WESM started operating in the Luzon grid on June 26, 2006. The Visayas grid was added into WESM four years later, commencing its operation on December 26, 2010.

With the supply shortage experienced in Mindanao, the DOE has directed the PEMC to establish an interim electricity market which is specially created and designed for Mindanao. The PEMC was appointed by DOE as the interim market's operator and was tasked for the implementation of the Interim Mindanao Electricity Market. The operation started in December 2013 and will be integrated to WESM soon after when the DOE declares.

The power generators that participate in WESM can generate around 10000MWh at any given moment. The plants generate using various resources such as natural gas, coal, hydro, geothermal, crude, diesel, biomass, wind and solar. Coal and natural gas are accounted for almost three-fourths of the total generation mix. Coal has been observed to produce more than half of the power generation at times. Geothermal plants account for around 14% in power contribution while hydro-electric power plants contribute between 3%-10% of the total generation mix. Since the launch of wind power plants, there have been an observed increase in their production. As of 2014, production contributes around 0.65% of the total power production in the market. The distribution of the power generation can be seen in Figure 1 (www.wesm.ph).

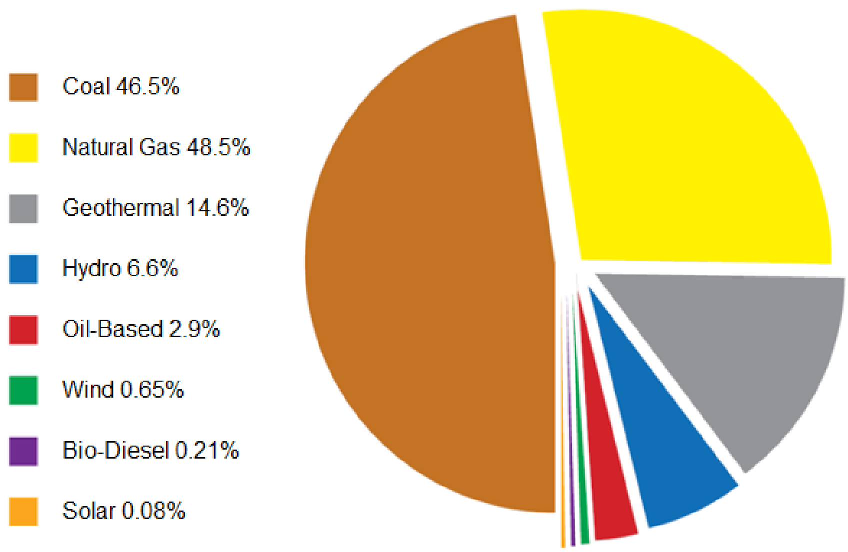

Figure 1. The distribution of Philippine power generation by Fuel Type in 2014.
Source: WESM.

1.2 Liberalization of Electric Markets

Until the early 1980s, many countries have suffered the monopolistic structure in electricity markets. Government authorities were traditionally in sole power to decide the production and distribution of electricity. It was beneficial to their citizens at first but widespread use of electricity has led to more serious issues. The problem abounds when governments are incapable of addressing the shortage and surplus of the supply which in turn wastes the public funds. In order to improve the efficiency in investment, technology, and possibly reduce the prices, electric markets were liberalized (Serati, Manera and Plotegher, 2008). Chile was first to realize the liberalization in the 1980s of which many countries have followed suit including the Philippines. Separate generation and distribution were introduced on this type of markets and power was traded based on the underlying cost (Weron, 2006). It has allowed the participation of private sectors which brought in diverse ways to produce electricity and have guaranteed the consumers the availability of supply. Generation through oil, coal, and water resources were vastly improved and exposed along with new technologies such as solar, geothermal and wind sources. The power trade is made through contracts or derivatives and transactions are sometimes settled through over-the-counter market (Noren, 2013). Markets of this types allow the supply to decrease or increase to meet the demand (Serati, *et al*., 2008). On the other hand, the government's role was reduced to regulations and policy making to secure the efficiency of the markets.

Power liberalization has proven to be effective to most of the countries. The participation of many sectors introduced the competition and was beneficial mostly to the consumers. Many countries transformed their markets to duplicate the successes of the pioneers. After the Chilean reform, the British followed and reorganized their power sector in 1990. In 1992, the Nordic market Nord Pool opened, which consists of Norway, Sweden, Finland and Denmark. It became the first multinational power exchange (Haldrup and Nielsen, 2004). Australia opened the Australian National electricity Market in 1998 while New Zealand reformed and launched their market in 1996. A number of states in the USA have also opened their respective markets beginning late 1990s (Weron, 2006). The Philippines established its market in June, 2006.

1.3 Electricity Price on the Spot Markets

Since the liberalization of electric markets, there has been better monitoring on the prices. The modernization of the markets together with the technological advances of the operation has provided a better view of its nature. Prices are now generally observed to have possessed shared characteristics. Common to markets worldwide are the observed high volatility and mean reversion. This means that prices tend to exhibit large jumps but always pull back to its long term mean. Much like other common goods, electricity is widely driven by the cost of production. Furthermore, it is observed to be stable in the long run but is usually disturbed by factors driving the demand and supply (Blochlinger, 2008). This supply disturbance for instance is caused by the weather conditions in the short run or the amount of rainfall in the medium run.

Prices also exhibit seasonality which shows different characteristics annually, weekly or even on a daily level. The sudden change on the prices is a consequence of the fluctuations on demand and shortage of supply during this seasons. For instance, spikes usually arise during the peak hours on daily or weekly level. On the other hand, extreme load fluctuations, generation outages and transmission failures cause sudden price hikes in any given time. Prices increase multiple times at this occasion but are normally short-lived after the system is fixed (Hardle and Truck, 2010).

Philippine electric price, much like in the other countries, also possesses the same common characteristic. It has shown seasonality as price usually goes up during peak hours on a daily scale or during summer time on a monthly scale. The country regularly experiences a wet and a dry season for the whole year. These seasons usually affect the market price drivers. The availability of supply during wet season for example is caused by the higher output of the hydroelectric plants (PEMC, 2014). Extreme weather conditions also greatly affect the price in the market. Since the country is prone to typhoons, it is usually common to observe a destroyed electric facility during such season. The inability of the companies to deliver electricity due to a destroyed transmission line affects the market price (www.wesm.ph). A concrete example is the impact of typhoon Glenda in July 2014 when around 86% of the Meralco customers had no power due to destroyed facilities. The event caused the price to rise in August 2014 to recuperate the damage (www.inquirer.net).

Figure 2 shows the electric price on the WESM from June 2006 to December 2013. The figure indicates the hourly price in blue while the orange line represents the monthly average price and presents how volatile the price is during the said period. The first year of WESM witnessed an average price of PhP 5,036 per MWh. In general, the WESM price history has noted high market prices during April, October and May. This is when the average monthly average supply margin is at the lowest (www.wesm.ph). Bidding behavior of the traders was also observed to be a significant driver of the prices in the market. This was witnessed starting 2010 towards 2013 in Figure 2 when the change in the offer behavior raised the level of prices in the market (PEMC, 2014). Another observable characteristic of the price is its possibility to attain a negative value. This odd phenomenon is observed on the hourly level and is common also to other markets worldwide. During this particular moment, power plant operators especially those using the renewable resources such as wind, opts to pay the distributors for economic purposes. When demand is so low for the particular hour, power plant operators are forced to sell the produced electricity as shutting down the plant and starting it again later may cost them more (www.epexspot.com). This was observed in the Philippines starting 2007 with more extreme negative values observed in 2008 and 2009.

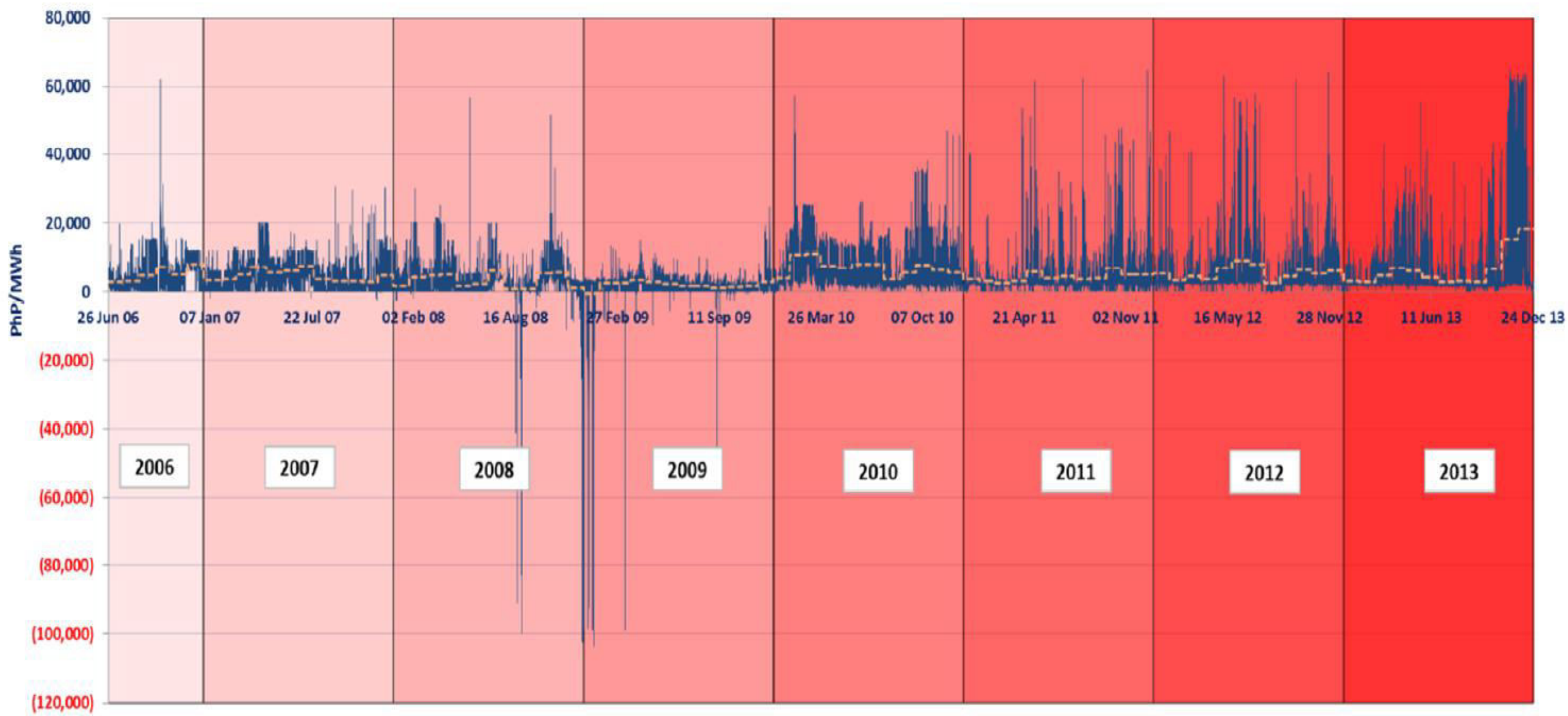

Figure 2. Hourly and monthly average price of load weighted average price (LWAP) in the WESM from June 2006 to December 2013
Source: PEMC

1.4 Rationale of the Study

Electricity prices have been largely watched not only by the government but also by the electric industry participants. These include the private companies involved in the generation and distribution of the said commodity. With the occurrence of spikes, the participants are more exposed to risks. Power plant operators are keen on price spikes as their values are more dependent on it, as spikes facilitate the recovery of high marginal costs (Higgs and Worthington, 2010). Profitability analysis and power planning is also dependent to vast knowledge in the electric prices while forecasting it in the short run is important for pricing the derivative contract (Serati, *et al*., 2009). The government on the other hand is concerned on the public interest. It is the one which determines when subsidies and other consumer policies are required. The importance of knowing the dynamics of electricity prices then is vital. Due to this, a number of studies have been conducted worldwide to rationalize the problem underlying it. In the Philippines, the level of importance of understanding the price dynamics is the same as those abroad. Consumers are very concerned about the price as evidenced by the formation of advocate groups and all sorts of information drive given by these non-government organizations. With so few literature which are available to refer to, there is a need to expose the subject to broaden the information. The lack of enough literature is evident in this research. There have been few authors who studied the electricity rates but did not have enough exposure to the concern of the public sector.

The Philippine electricity rates have been ranked among the highest in the survey of 44 countries in 2012 (www.interaksyon.com). With this problem and with the supply shortage, it is necessary to study the implications of the dynamics of the electricity prices to help both the regulators and the participants on the mitigation process. Furthermore, input prices or fuel costs have long been determined to be factors in the movement of electricity prices (www.eia.gov). The need to provide this information is important as more than 70% of the power generated in the country uses coal and natural gas as fuel. Unfortunately, the literature lacks this information.

1.5 Objectives of the Study

The main objective of this study is to investigate the occurrence of electricity price spikes in the Wholesale Electricity Spot Market (WESM) in the Philippines using the Luzon grid data. Specifically, the study attempted to:

1. model the electricity price spikes using the Markov Regime Switching method;
2. determine the periods or months when spikes on electricity price occur in the Philippines;
3. determine whether or not prices of Indonesian coal and natural gas affect the spikes of electricity in the Philippines.

1.6 Scope and Limitations of the Study

The study focuses on the electricity spot price of the Philippines particularly in the Luzon grid. Data were taken from the database of WESM on a monthly basis, from January 2009 to December 2015. The author used the Luzon data since it was the first market launched by WESM in 2006 and the Visayas market was only added in 2010. Furthermore, Luzon has the greatest consumption with more than 4000GWh per month compared to the Visayas region which only consumes 400GWh in the current market (www.wesm.ph). Mindanao meanwhile does not have a permanent electricity market as of the moment since it is only using the Interim Mindanao Electricity Market implemented by DOE on 2013.

1.7 Organization of the Study

This thesis is organized as follows: Chapter 2 presents the related literature of this study regarding the application of Markov Regime Switching (MRS) in electricity price. It also reviews the use of MRS analysis to other economic fields. Furthermore, it presents the use of other statistical methods to investigate the electricity price. The methodology is presented in Chapter 3. This includes the discussion of this study's theoretical and conceptual framework. Another section on this chapter is the econometric model which includes a thorough presentation of the Markov Regime Switching and the estimation process. Data and variables are also discussed on this chapter while the last part is a presentation of the Tobit regression. Chapter 4 shows the results and the discussion. Lastly, Chapter 5 is the presentation of the summary and conclusion together with the suggested areas for further research.

CHAPTER 2

REVIEW OF RELATED LITERATURE

2.1 Studies Using Markov Regime Switch Modelling

Despite the short history of most markets worldwide, there have been a number of studies related to modelling of the electricity price. Researchers and institutions have dedicated their time and resources on probing this matter as electricity affects all areas in the modern world. Various authors used the Markov Regime Switching approach which is by far the most popular in this field due to its ability to capture volatility, a common feature of electricity price.

In 2013, Noren studied the German and Nordic markets using the non-linear Markov-switching models. The focus of the study was to compare the two estimation methods which are popularly used to estimate the model: the maximum likelihood and the Expectation-Maximization (EM) algorithms. Both were found to be equally accurate in their estimates but the EM Algorithm was found to have a higher rate of convergence than the maximum likelihood. The author calibrated both the 2-state and 3-state models to show which model best fits the data. In both calibrations, EM algorithm provides the approximation in fewer iterations. In addition, the study found that the length of volatility of German and Nordic markets were significantly different since the probability of remaining on a drop regime is smaller for the Nordic on all occasions. However, daily spikes and hourly mean-reverting prices persist almost at the same length or period.

Wlodarczyk and Zawada (2008) estimated the daily price of the Polish electricity using a specified Markov Switching-GARCH model, taking into consideration the autoregressive dependence to both the conditional mean and variance. Variances were found to be significantly different between particular states or regimes. The model indicated that the variance for the high volatility state is 5 times higher than the low volatility state. The findings have also showed that low volatility level is expected to last for 12 days while high volatility level lasts for 5 days on the average.

Power spikes nature while using the Regime-Switching approach was the center of interest of De Jong (2006). The data was based on six European markets and two US electricity markets. These are from Scandanavia, Germany, Netherlands, France, Austria, Spain, the Pennsylvania-New Jersey-Maryland (PJM), and the New England markets. It covers around 36% of total generating capacity in Europe and the largest market in the US. The findings showed that Regime-Switch models are more capable of capturing the market dynamics of electricity prices than the GARCH (1,1) or Poisson Jump model. There are significant differences between markets having different hydro-power supply schemes. Hydro-power has a dampening effect on spikes which is caused by its ability to indirectly store the electricity by storing water resources.

Using the data from the Nordic electricity spot market, Haldrup and Nielsen (2004) developed a model that can generate a long memory in each state. As it was observed that the Nordic markets are characterized by a high degree of long memory in the spot prices, then it was assumed that an extreme form of fractional cointegration exists between markets. A regime switching long memory model or fractional integration was applied to come up with the empirical results. The study found that the regime switching and the long memory significantly co-exist in the market. The price behaviors for different markets are different which depends on the presence of the bottlenecks in the energy transmission.

Kanamura and Ohashi (2004) analyzed the transition probabilities of regime switching in electricity prices which showed its dependence to the demand. The idea is contrary to some results that the transition probabilities are constant among regimes. Data used were daily spot prices taken from the Pennsylvania-New Jersey-Maryland Interconnection (PJM) electricity market in Pennsylvania, USA. Results indicated that the increase in demand from 700,000MWh to 800,000MWh increases the price by $6/MWh while an increase of demand from 1,000,000MWh to 1,100,000MWh increases the price dramatically by $153/MWh. The large difference therefore indicates a price spike. These findings solidify the notion that the transition probabilities depend on the demand level for the electricity and are therefore ever changing. The study asserted the general consensus that electricity prices are prone to spikes during summer and winter when the demand for electricity is high.

Higgs and Worthington (2007) studied the Australian National Electricity Market (ANEM) which comprises the interconnected market of New South Wales, Queensland, South Australia and Victoria, using daily spot prices. Price spikes were found to be short lived as the mean reversion coefficients in the spike regimes are much higher than the normal regime. This implies a much faster return of the price to the equilibrium. The volatilities were also found to exceed 7% in spike periods but less than 0.5% in normal periods. In addition, the regime-switching model outperformed the basic stochastic and mean – reverting models given the lower value of the log-likelihood. The probability that a spike occurs is 5.16% in New South Wales and 9.44% in Victoria.

Bessec, Fouquau and Meritet (2014) predicted the electricity spot prices using time series models with a double temporal segmentation. The study was conducted in the French Wholesale Market by using the hourly electricity spot price. The aim was to compare several time series models and pick the most accurate model. The accuracy of the models were measured and ranked through the criterion root mean square error (RMSE), mean absolute error (MAE) and mean absolute percent error (MAPE). The three-state Markov-Switching model was found to be the best model to capture the sudden and fast-reverting spikes characterized by the market and hence yields more accurate forecasts. The model generated the best value for the RMSE, MAE and MAPE.

A different version of Markov Regime Switching (MRS) model, the MRS-Vector Autoregression (MRS-VAR) model was utilized by Haldrup, Nielsen and Nielsen (2009). The model allows fractional integration in every observed state. Using the Nord pool data, results showed that even though prices are identical when in non-congested period, the prices are not, implying they are fractionally cointegrated in the congested period. Price convergence therefore is a characteristic that happens after the regime switching instead of a conventional error correction mechanism.

Weron and Janczura (2010) proposed a new method that significantly reduces the computation process that was first introduced on the independent regimes of an MRS model. The problem of complexity on the estimation of the traditional MRS model was exposed. The increasing state process of the Expectation step in EM algorithm is said to be very complex with other authors suggesting it to be limited to the last 10 observations. Instead of storing large amount of filtered and smoothed probabilities in every possible state process paths, the author suggested to require conditional probability for only one time step. The model is a 3 regime MRS which was tested by a simulation study by using Germany's European EnergyExchange (EEX) market. The outcome showed that sample means were close to its true parameters and standard deviations decrease with the increasing sample size.

Capturing a fat tailed distribution was the main focus of Murara (2010). They utilized the Regime Switching model while using the NordPool (Sweden) data. The study considered three criteria for modelling: the price difference, capacity/flow difference and spikes in Finland electric prices. They also determined the regular and non-regular regime as states as well. The empirical results show that price separation and capacity-flow difference is correlated at 0.10 when comparing the Finland and Sweden prices. Furthermore, it is more probable for a price to remain on regular regime to not change than change states, or move from non-regular regime to the other.

Markov Regime Switching method is popular not only on the electricity price modelling but to other fields as well. Its ability to predict and forecast the spikes or crisis periods has been widely used. Some authors conducted their studies on identifying crisis situations in countries. Cruz and Mapa (2013) have modeled an Early Warning System for the Philippines by using the Markov-Switching and Logistic Regression models. The system used a 2 regime MS-AR(2) model which shows that the estimated average inflation rate for high inflation exceeds the threshold set by the Bangko Sentral ng Pilipinas (BSP). The study suggested that it is more likely for the country to be on the low inflation stage than in a high inflation stage. The duration of low inflation is also estimated to be twice as long as that of high inflation.

MRS has also been used in the labor market. Krolzig, Marcellino and Mizon (2000) utilized the 3-regime cointegrated vector autoregressive MRS to study the labor market in the United Kingdom. The regime is represented by the recession, growth and high growth and the result provided a good characterization of the sample data. Kano and Ohta (2002) also studied the empirical matching function of the Japanese labor market by using a 2-state Markov Regime Switching model. The result showed that the matching function is regularly changing between the increasing and the constant returns period. In addition, the durations of the regimes were found to be very short.

Alizadeh, Nomikos and Pouliasis (2008) used the Markov Regime Switching (MRS) approach to determine the time-varying minimum variance hedge ratio in the energy markets. West Texas Intermediate (WTI) crude oil, unleaded gasoline, and heating oil were the variables used in the study. Hedge ratio is the ratio of future contracts to buy or sell for each unit of underlying asset on which the hedger takes risk. It is used to manage potential adverse effects of price changes in the market. The authors used two models using univariate MRS and bivariate MRS – VAR with GARCH error structure. These models were compared to the existing results given by previous authors. The sample tests showed that MRS model outperforms the other methods in reducing the portfolio risk in the petroleum products and crude oil markets.

2.2 Electricity Price Using Other Statistical Methods

Studies on electricity prices using various methods are popular in the literature. Hardle and Truck (2010) used the dynamic semiparametric factor method (DSFM), which focused on the specification of the behavior of the hourly prices of the European Energy Exchange (EEX). Findings showed that a small number of dynamic factors can explain around 80% of variation in the prices. However, the explanatory power decreases for periods with higher number of spikes. A deseasonalized data was also calibrated and was found to have just a slight improvement in the performance. Overall, the DSFM approach has well represented the structure of the hourly prices.

Eriksrud (2014) compared the Nordic pool electricity price by using a seasonal ARIMA (SARIMAX) model to panel, neural network and a random forest models. It was found that SARIMAX is better than all other three models. It is the only model that beat the naive predictions in winter and summer season prices. Neural Network was good in predicting price in summer but is not in the winter period. Random forest on the other hand overfits the predictions for both season. The result is an approval to previous literatures that time series models are better than machine learning methods when it comes to electricity price modelling.

Christensen, Hurn, and Lindsay (2011) focused on the prediction of the electricity prices, using the half-hourly spot price in four Australian markets, covering the period from March 2001 to June 2007. The nonlinear variant of the Autoregressive Conditional Hazard (ARCH) was employed to model the prices. It was found that durations between price spikes were nonlinearly dependent on previous expected and observed durations.

Additionally, probit model was applied to study the nature of spikes which was identified as $100 to $300 for normal and $300 to 10,000 as severe. Abnormal loads and temperature extremes were found to have significant impact to the spikes while the maximum temperature was culminating a negative effect.

Table 1 summarizes all the literature discussed in this chapter. It includes the author, methodology used, the location of the study and summary of the result.

Table 1. Summary of Related Literature

Author	Methodology	Location	Result
Noren (2013)	Nonlinear Markovswitching models	German and Nordic markets	EM Algorithm have a higher rate of convergence than the MLE
Wlodarczyk and Zawada (2008)	Markov SwitchingGARCH model	Polish market	Low volatility level lasts for 12 days while high volatility level lasts for 5 days on the average
De Jong (2006)	Markov RegimeSwitching	Scandanavia, Germany, Netherlands, France, Austria, Spain, the PennsylvaniaNew JerseyMaryland (PJM), and the New England markets	MRS models are more capable of capturing the market dynamics of electricity prices than the GARCH (1,1) or Possoin Jump model
Haldrup and Nielsen (2004)	Regime switching long memory model	Nordic market	Price behaviors for different markets are different which depends on the presence of the bottlenecks in the energy transmission

Kanamura and Ohashi (2004)	Markov RegimeSwitching	PennsylvaniaNew JerseyMaryland Interconnection (PJM) electricity market	Transition probabilities depend on the demand level for the electricity and are therefore ever changing
Higgs and Worthington (2007)	Markov RegimeSwitching	Australian National Electricity Market (ANEM)	Probability that a spike occurs is 5.16% in New South Wales and 9.44% in Victoria
Bessec, Fouquau and Meritet (2014)	3state MarkovSwitching model	French Wholesale Market	3 state MRS model generated the best value for the RMSE, MAE and MAPE
Haldrup, Nielsen and Nielsen (2009)	MRSVector Autoregression (MRSVAR) model	Nordic market	Price convergence is a characteristic that happens after the regime switching instead of a conventional error correction mechanism
Weron and Janczura (2010)	3state MarkovSwitching model	Germany's European EnergyExchange (EEX) market	Sample means were close to its true parameters and standard deviations decrease with the increasing sample size
Murara (2010)	Markov RegimeSwitching	Nordic market	Price separation and capacityflow difference is correlated at 0.10 when comparing the Finland and Sweden prices
Cruz and Mapa (2013)	MarkovSwitching and Logistic Regression models	Philippines	More likely for the country to be on the low inflation stage than in a high inflation stage
Krolzig, Marcellino and Mizon (2000)	3regime cointegrated MRSVAR	United Kingdom	The regime is represented by the recession, growth and high growth and the result provided a good characterization of the sample data

Kano and Ohta (2002)	2state Markove Regime Switching model	Japan	Matching function is regularly changing between the increasing and the constant returns period
Alizadeh, Nomikos and Pouliasis (2008)	Markov Regime Switching	USA	Sample tests showed that MRS model outperforms the other methods in reducing the portfolio risk in the petroleum products and crude oil markets
Hardle and Truck (2010)	Dynamic semiparametric factor method (DSFM)	Europe	Small number of dynamic factors can explain around 80% of variation in the prices
Eriksrud (2014)	Seasonal ARIMA (SARIMAX) model	Nordic market	Time series models are better than machine learning methods when it comes to electricity price modelling
Christensen, Hurn, and Lindsay (2011)	Autoregressive Conditional Hazard (ARCH) model	Australia	Durations between price spikes were nonlinearly dependent on previous expected and observed durations

CHAPTER 3

METHODOLOGY

This chapter discusses the research methodology and the econometric model using the Markov Regime Switch – Tobit Regression approach. R Statistical Software, particularly the MSWm package, was used for the Markov Regime Switching estimation. MSWm is the package especially made to estimate the MRS by using the EM Algorithm, which in this study is the estimation process of choice of the author. Another package called “tseries” and “stats” were used for the Augmented Dickey-Fuller and the Phillips-Perron stationarity tests respectively. Tobit regression estimation on the other hand was done through Shazam Version 11.

3.1 Theoretical Framework

Spikes in electricity price are common to spot markets. Its liquidity is sometimes attributed to the frequency and volume of the trades happening in the modern and complex world of trading. Electricity prices are determined by the equilibrium of supply and demand like any other commodity. However, it is much different from the common goods because storing it is very uneconomical. This characteristic limits the market players’ strategies in trading (Noren, 2013).

Common to electric markets are the existence of the three different types of seasonality. This seasonality adds distortion to the prices and sometimes triggers abrupt changes in the prices. The first is the greater demand in summer due to air conditioning. Second is the significant change in the use of electricity between weekends and weekdays. Finally, the daily variation of consumption which is evident when day and night usage are compared.

Electric energy consumption is largely considered inelastic implying sudden changes in prices can occur anytime even if supply changes slightly. Supply is usually capped for all markets due to the limited availability of generators while exporting and importing are constrained to the capacity of the grids. This means that an incorporation of markets together with power outages during certain periods will cause an increase in demand. Prices may cause sudden spike at this stage (Noren, 2013).

Figure 3 shows the dynamics of the electricity price based on the movement of supply and demand. It represents two different scenarios of having a higher market price due to the shifts. Demand curve is usually steeper in an electricity market as consumption is considered inelastic. The first figure shows the positive shift of demand from D to D'. This change pushes the market price upward from P_1 to P_2, resulting to a higher quantity of production from Q_1 to Q_2. The second figure on the other hand shows the negative shift of supply from S to S' and a steeper demand curve D. This happens when there is a lower quantity of production noted by the change of Q_1 to Q_2. This negative change pushes the market price upward from P_1 to P_2 even when supply changes slightly. This depicts the scenario of having sudden decrease of power generation due to weather conditions or transmission failures.

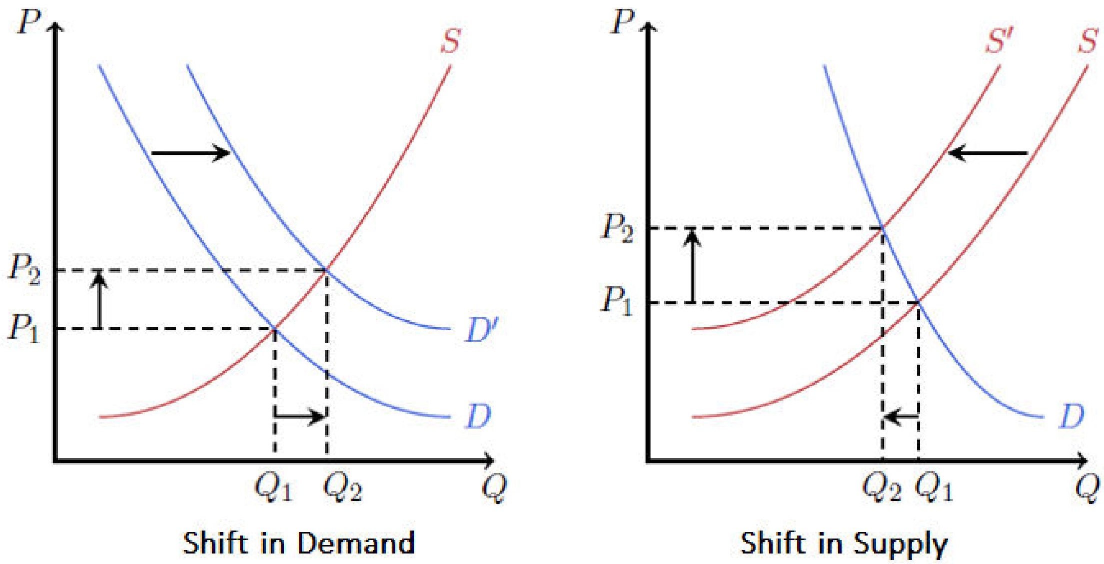

Figure 3. Electricity price based on the movement of supply and demand.
Source: Noren, 2013.

3.2 Conceptual Framework

This study implements the Markov Regime Switch – Tobit Regression approach inspired by the Markov Regime Switch – Logistic Regression approach of Cruz and Mapa (2010). The estimation process for the MRS is much different from that of Cruz and Mapa as this study utilizes the EM Algorithm, perceived to be a more efficient process compared to the traditional Maximum Likelihood. In addition, using Tobit regression is also justified given the nature of the dependent variable given by the result in the MRS analysis. The spikes are identified through the decision rule imposed by Hamilton (1989). This decision rule is used as the threshold for the method. The dependent variable therefore is left censored at 0.5 and the influence of the natural gas and Indonesian coal price to the price spike are investigated through this process.

The Maximum Likelihood Estimation (MLE) introduced by Hamilton (1989) is limited to relatively small systems. MLE computes the parameter estimates one by one given the loglikelihood function. The loglikelihood function becomes more complex when more unknown parameters are added to the systems. Thus the calculation requires more time and additional computer programming (Hamilton, 1990). In addition, the estimation suffers difficulties when finding global maxima particularly in small sample sizes (Noren, 2013). The EM Algorithm on the other hand is not dependent on the size of the system. The process only requires the trivial calculation from the smoothed inferences about the unobserved regime. LogLikelihood functions were computed given the most recent parameters in every iteration. Convergence is known to be slow but the estimates are approximately obtained through relatively fewer iteration (Noren, 2013). Additionally, in this type of estimation, every iteration increases the value of the likelihood function which guarantees a better Maximum Likelihood estimate.

Using the natural gas and Indonesian coal price is based on previous assumptions. The price was identified to be driven by a number of contributing factors. According to the Energy Information Agency (EIA), input prices of power production were found to be factors of the movement of the electricity prices (www.eia.gov). Its effect to the price spikes however are yet to be determined. Natural gas and coal comprises around 70% of the total power produced throughout the country. This study attempts to determine their effects on the Philippine setting and investigate using the Tobit regression.

Figure 4 shows how the estimation is conducted for this study in order to achieve the objective. There are two phases of the estimation process: first is the Markov

Regime Switching analysis to determine the spikes of the electricity price. From there, the independent variables are incorporated to the system by using the Tobit Regression for the second phase. The results then yield an output that would enable the prediction of the future spikes.

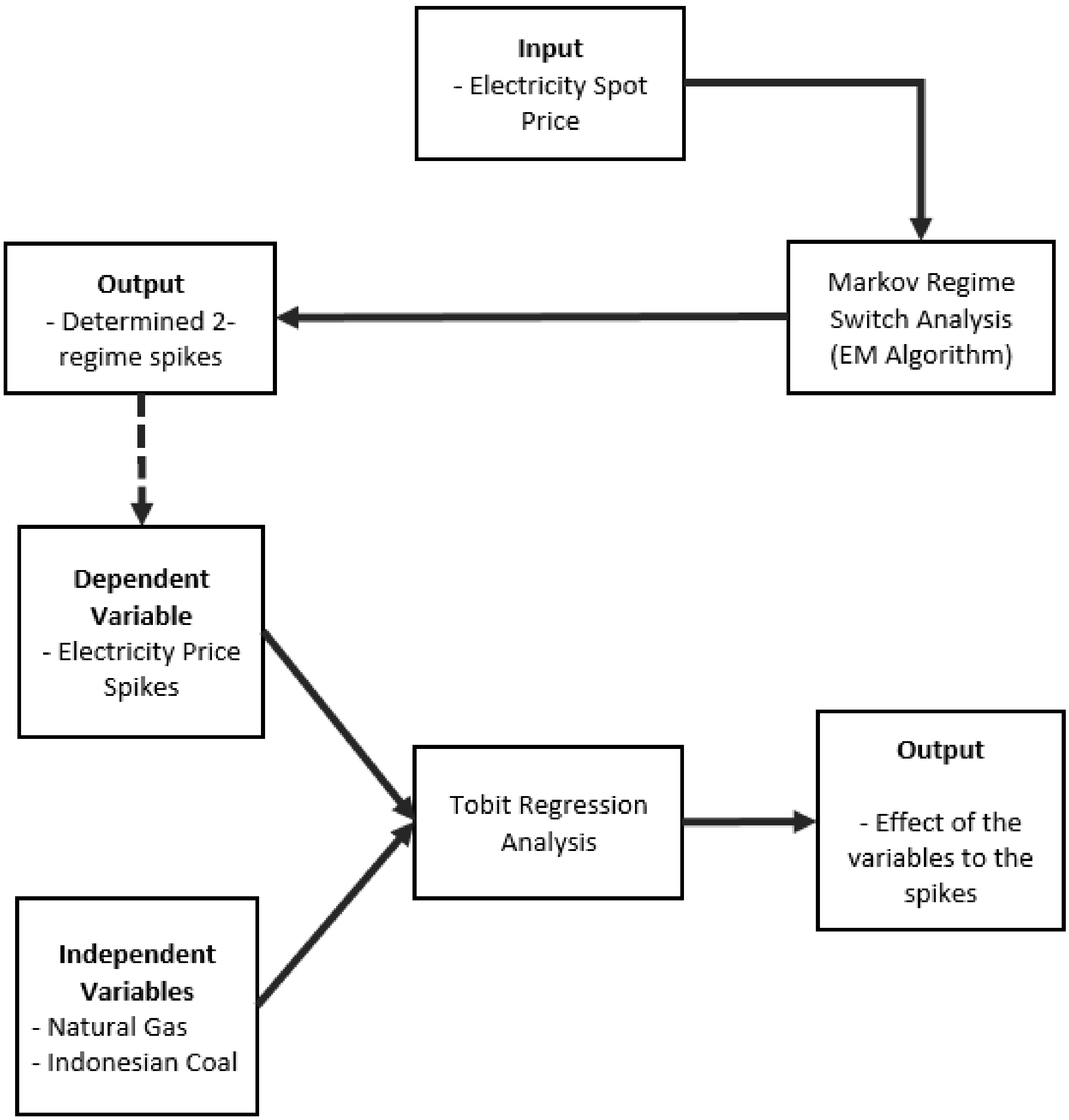

Figure 4. Diagram showing the estimation process.

3.3 Econometric Model

The econometric model in this study explains the dynamics of electricity price spike. The study used the Markov Regime Switch model to model and determine the period of spikes. Tobit Regression meanwhile was used to determine the influence of the input prices to the existence of the spikes. Electricity prices are observed to be volatile, with strong mean reversion and frequently exhibits spikes, thus using the Markov Regime Switch is a viable choice. The model captures the spikes in the price levels which allows the prediction and forecast to be significantly improved. It has been introduced by various studies using different types of approach. The 2-regime MSAR model is used for this study which is basically the simplest form of the Markov Regime Switch model using the monthly spot price of the electricity as an exogenous variable. The next section gives a thorough discussion of the Markov Regime Switching models.

3.3.1 Markov Regime Switching Model

One of the most popular approaches in modeling nonlinearity in a time series is taking an assumption of multiple behavior or structural breaks. That is, modeling separately the different periods of a time series as normal and spike periods (Mahieu and Huisman, 2001). Such method is also widely known as Markov Regime Switching and was popularized by Hamilton (1989). The original version introduced by Hamilton focuses on the mean behavior of the variables. However, due to its success and popularity in various studies, many have considered exploring the method and incorporating the switching dynamics to the variance models, i.e, Haldrup, *et al.* (2009), and Hamilton and Susmel (1994). The advantage of using Markov Switch Model in an economic variable is its ability to capture different patterns over time. Financial variables in particular, may exhibit patterns that a conditional mean model is unable to represent as variables of this kind display multiple structures or behavior. Instead of using one model, it is more feasible to introduce several models to capture these patterns (Kuan, 2002). This is also the reason why using MRS to model electricity prices is popular.

3.3.2 Simple and Generalized Markov Switch Model

This section focuses on the explanation of Markov Switching on the AR model as it is the model suitable for the data of this research. To better understand the Markov Regime Switch model, it is preferable to start presenting it through a simple form. It is important to remember that this is only for the simplification of the presentation and the model could be any type of a time series model for this study. Thus, consider a stationary AR(1) model. Assuming it has two structural breaks or regimes, a twostate Markov Switching Model should be represented as (Kuan, 2002):

$$Y_t = \begin{cases} \alpha_0 + \beta Y_{t-1} + \varepsilon_t, & S_t = 0 \\ \alpha_1 + \beta Y_{t-1} + \varepsilon_t, & S_t = 1 \end{cases}$$

where S_t is the state or period of the variable at time *t* and assumes the values 1 or 0, it is governed by a Markov process and changes based on its underlying transition probabilities. The parameter $|\beta| < 1$ and the error term ε_t is normally distributed with mean 0 and variance of σ_ε^2, while the coefficient α_1 is the parameter of the time series with switching effect. The equation can also be viewed as $Y_t = \alpha_0 + \alpha_1 S_t + \beta Y_{t-1} + \varepsilon_t$ for $S_t = 0, 1$. The same idea also applies to more complex model of AR or ARCH to which more parameters define the states. To generalize, take for example an AR(k) model (Chen, 2013):

$$Y_t = \alpha + \sum_{i=1}^{k} \beta_i Y_{t-i} + \varepsilon_t.$$

A generalized two-state model for Markov Regime Switch should look like:

$$Y_t = \begin{cases} \alpha_0 + \sum_{i=1}^{k} \beta_{0i} Y_{t-i} + \varepsilon_{0t} , & S_t = 0 \\ \alpha_1 + \sum_{i=1}^{k} \beta_{1i} Y_{t-i} + \varepsilon_{1t} , & S_t = 1 \end{cases}, \qquad (1)$$

where we let the vectors $\emptyset_{\mathbf{0}} = \{ \alpha_0, \beta_{01}, \ldots, \beta_{0k}, \sigma_0^2, \varepsilon_{0t}\}$ and $\emptyset_{\mathbf{1}} = \{ \alpha_1, \beta_{11}, \ldots, \beta_{1k}, \sigma_1^2, \varepsilon_{1t}\}$ are the parameters for the models at $S_t = 0$ and $S_t = 1$ respectively. It should also be noted that the models are stationary at any given state.

3.3.3 Likelihood Function

The likelihood function plays a vital role in the statistical inference for estimating the set of parameters of the MRS models. It summarizes the evidence about the unknown parameters of the given data. In most cases, because of the computational expediency, many have chosen to work with the log likelihood functions (mathworld.wolfram.com). As for this study, the log likelihood function plays a significant role in estimating the model through the EM algorithm.

For previous observations of Y_t at time t, let $F_t = \{ Y_t, Y_{t-1}, Y_{t-2}, Y_{t-3} \ldots\}$ and the state variable S_t. Under the assumption of normality, the mean for an AR(k) model is given by $\mu = \beta_{S_t,0} + \sum_{i=1}^{k} \beta_{S_t,i} Y_{t-i}$. Hence the probability distribution function of Y_t is (Chen, 2013):

$$f(Y_t|S_t) = \frac{1}{\sqrt{2\pi\sigma_{S_t}^2}} \exp\left\{-\frac{(Y_t - \beta_{S_t,0} - \sum_{i=1}^{k} \beta_{S_t,i}\, Y_{t-i})^2}{2\sigma_{S_t}^2}\right\}$$

and the likelihood function should be presented as,

$$L(f(Y_t|S_t)) = \sum \left(\frac{1}{\sqrt{2\pi\sigma_{S_t}^2}} exp\left\{-\frac{(Y_t - \beta_{S_t,0} - \sum_{i=1}^{k} \beta_{S_t,i} Y_{t-i})^2}{2\sigma_{S_t}^2}\right\} \right) . P(S_t = i|F_{t-1}) \,. \qquad (2)$$

3.3.4 Markov Chain Process and the Transition Probability

The change of states in Markov Regime Switching model is a random process that follows an *n* - state Markov chain. Thus, it is characterized as "memoryless" where the preceding change of state is not dependent on the previous states of the model but on the current state instead. To simplify, the state S_{t+1} is not dependent on S_{t-1} but on S_t alone. Consider a 2-state *i* and *j* markov process illustrated in the Figure 5:

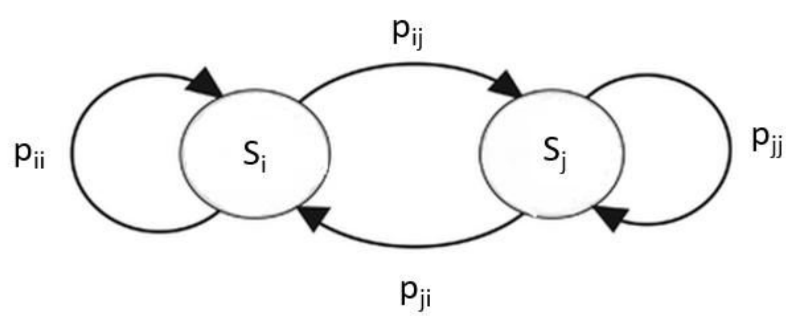

Figure 5. A 2-state Markov Process S_i and S_j with the underlying probabilities.

An important property of a Markov Chain is its distribution. If a Markov Chain is at time *t*, the marginal distribution $P(S_t = j)$ for $j = 0,1$ is given by the equation (Vigoda, 2003),

$$P(S_t = j) = \sum_{i=0}^{1} p_{ij} \,.\, P(S_{t-1} = i). \tag{3}$$

It follows that if an initial distribution is given by $P(S_1 = j)$, then the *t*-step distribution of a Markov Chain is given by,

$$P(S_t = j) = \sum_{i=0}^{1} p_{ij}^{(t)} . P(S_1 = i).$$

The transition probability p_{ij} is the probability that S_t will move from state *i* to *j* at time t_1 to time *t*. This is denoted as:

$$p_{ij} = p_{i|j} = P(S_t = i | S_{t-1} = j), \qquad i, j = 0,1$$

In matrix form, the probabilities for the 2-state process are presented as:

$$\boldsymbol{P} = \begin{bmatrix} P(S_t = 0|S_{t-1} = 0) & P(S_t = 1|S_{t-1} = 0) \\ P(S_t = 0|S_{t-1} = 1) & P(S_t = 1|S_{t-1} = 1) \end{bmatrix} = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix},$$

where the transition probabilities in every row are restricted to equate to 1, i.e., $p_{i0} + p_{i1} = 1$ and $p_{j0} + p_{j1} = 1$.

Another interesting property of the transition probability is its ability to identify the expected duration of the states. It can be calculated given the formula (Hamilton, 1989):

$$\frac{1}{1 - p_{ii}} = \text{the time duration that a state stays at i.}$$

For this study the duration is expressed in months. This means that smaller values for $1 - p_{ii}$ tends to stay longer in the state *i* while larger values will otherwise implicate a shorter duration for the said state.

3.3.5 Filtered and Smoothed Probabilities for S_t

The values of Y_t is always presumed to be observed. On the other hand, the actual value of S_t is unobservable as it is latently impacting the model. The problem is solvable through a probabilistic inference given the observed values of Y_t and the transition probabilities (Kole, 2010). This is achieved through the use of the Bayes' Rule:

$$P(A|B) = \frac{P(B|A).P(A)}{P(B)}$$

Let $F_t = \{ Y_t, Y_{t-1}, Y_{t-2}, Y_{t-3}, \dots, Y_1\}$ be the set of previous observations of Y, and $\emptyset_i$ as the vector of parameters for the Markov Regime Switch model. In application of Bayes' Rule, the inference of the regime at time *t*=1 given the probability distribution of Y_1 is,

$$\begin{aligned} P(S_1 = 0|Y_1 = y_1) &= \frac{P(Y_1 = y_1|S_1 = 0).P(S_1 = 0)}{P(Y_1 = y_1)} \\ &= \frac{P(Y_1 = y_1|S_1 = 0).P(S_1 = 0)}{P(Y_1 = y_1|S_1 = 0).P(S_1 = 0) + P(Y_1 = y_1|S_1 = 1).P(S_1 = 1)} \\ &= \frac{f(y_1;\mu_0;\emptyset_0).P(S_1=0)}{f(y_1;\mu_0;\emptyset_0).P(S_1=0)+f(y_1;\mu_1;\emptyset_1).P(S_1=1)} \end{aligned} \qquad (4)$$

The third equality here presents a probability $P(S_1 = 0)$. This is the probability for the initial regime S_1 to be equal to 0. Similarly, $P(S_1 = 1)$ is the probability for the initial regime S_1 to be equal to 1. Naturally, these probabilities are unobserved just like all that is in S_t. Thus this probability needs to be estimated just like all parameters. Let this new parameter be noted as follows:

$$P(S_1 = 0) = \tau$$

$$P(S_1 = 1) = 1 - \tau$$

Replacing this to the equation, the *inference probability* becomes

$$P(S_1 = 0|Y_1 = y_1) = \frac{f(y_1; \mu_0; \sigma_0^2).\tau}{f(y_1; \mu_0; \sigma_0^2).\tau + f(y_1; \mu_1; \sigma_1^2).(1 - \tau)}$$

After this inference at time *t* = 1 has been computed, this result is then used to determine the distribution for time *t* = 2. Using equation (3) or the marginal distribution of Markov Chain at time *t* = 2 then,

$$\begin{aligned} P(S_2 = 0|Y_1 = y_1) = {} & P(S_2 = 0|S_1 = 0, Y_1 = y_1).P(S_1 = 0|Y_1 = y_1) \\ & + P(S_2 = 0|S_1 = 1, Y_1 = y_1).P(S_1 = 1|Y_1 = y_1) \\ & = p_{00}\, P(S_1 = 0|Y_1 = y_1) + p_{01}\, P(S_1 = 1|Y_1 = y_1) \qquad (5) \end{aligned}$$

These calculations can be repeated to time *t* = 3 up to any given time of *t* by substituting equation (5) to (4). This recursive calculation gives the values for what is called the *filtered probabilities* which are used to forecast the probabilities at time *t +1*. To summarize, a *filtered probability* is the probability that $S_t = i$ given the observations including that of time *t*,

$$P(S_t = i|F_t; \emptyset_{\boldsymbol{i}}).$$

In addition to this, the probabilities can also be calculated conditional on all samples of Y_T. This probability, based on the filtered probability is called the *smoothed probability*.

$$P(S_t = i|F_{\boldsymbol{T}}; \emptyset_{\boldsymbol{i}}).$$

The starting value for the smoothed probability is taken from the final iterated value in the filtered probability (Kim, 1994). This starting value would then generate a new value until convergence.

3.3.6 Estimation

Let all the parameters be collected into a vector $\boldsymbol{\Omega}$. That is in equation:

$$\boldsymbol{\Omega} = \{\emptyset_{\boldsymbol{0}}, \emptyset_{\boldsymbol{1}}, \boldsymbol{P}, \tau\}$$

where $\emptyset_{\boldsymbol{0}}$ and $\emptyset_{\boldsymbol{1}}$ are parameter vectors of the time series of the 2 different regimes, ***P*** as the vector of the transition probabilities and τ as the probability for the initial regime. The estimation of $\boldsymbol{\Omega}$ is not straightforward since the regimes are not directly observable (Weron and Janczura, 2010). If the states were observable then the problem would be easier to estimate by using the regression models. Considering these issues, Hamilton pointed out 3 problems concerning the estimation. First is the problem of inference: this questions the logical conclusion of the researcher to the values of the regime at time *t* given the observations of *Y* up to the time *t* and beyond. Second, the problem of forecasting: given the prior observations of Y_t , what should be the best forecasting value of Y_{t+j}? And lastly, the problem of estimation: what are the values of $\boldsymbol{\Omega}$ given the observed values of Y_t (Hamilton,1990).

Hamilton (1989) answered these problems and introduced a feasible way for estimation using the Maximum Likelihood. However, while the method gives an analytical solution of the model through recursive process, it is relatively difficult given the needed calculation time for each parameter especially during larger systems of models. The method introduced by Hamilton on his first paper (1989) becomes more difficult to compute as the system becomes larger. Due to this, Hamilton (1990) pointed that the method is limited only to relatively small systems. He then suggested an alternative estimation using the ExpectationMaximization algorithm in his paper (1990) pioneered by Dempster, Laird and Rubin (1977). The estimation utilizes the inferences calculated from the smoothed probability to be discussed on the next section. It requires relatively lesser time to calculate as the method is independent of the size of the systems of models (Hamilton, 1990). In this paper, the author has chosen to favor the EM algorithm as the method of estimation based on the reason imposed by Hamilton (1990).

3.3.7 The EM Algorithm

The use of EM Algorithm in Markov Switch models was introduced by Hamilton (1990) and the idea was generally taken from Dempster, *et al* (1977). Later on, it was widely used by researchers using Switching models. It is composed of the Expectation and Maximization stages, thus the name EM Algorithm. Hamilton (1990), utilizes the smoothed probabilities of S_t to provide the maximum likelihood estimates of the parameters. Taking the smoothed probability is accepting the assumption that the unobserved state is s_t on time *t* based on the entire sample **T**.

$$P(s_t = i | F_{\boldsymbol{T}}; \boldsymbol{\Omega}).$$

The probabilities were used to reweigh the observed data of Y_t. After reweighing the observations, regression is performed to generate values for the parameter $\boldsymbol{\Omega}$. The new parameter is then used to obtain the values of the smoothed probability $P(s_{t+1} = i|F_{\boldsymbol{T}}; \boldsymbol{\Omega})$. The process is repeated until a convergence of the values of $\boldsymbol{\Omega}$ is obtained. Hamilton (1990) has proven that upon iteration of this process, the values of the likelihood function increases which means it improves on every yield of the parameters. Thus, the converged values of $\boldsymbol{\Omega}$ is a maximum likelihood estimate.

Chen (2013) has summarized the E and M steps for Markov Switch model to the following details:

1. Pick an initial value for $\boldsymbol{\Omega} = \boldsymbol{\Omega_0}$
2. The Expectation stage:

   Calculate the values of the following probabilities:

   a. $P(s_t = 1|F_{\boldsymbol{T}}; \boldsymbol{\Omega_0})$ for all values of *t*.
   b. $P(s_t = 0|F_{\boldsymbol{T}}; \boldsymbol{\Omega_0})$ for all values of *t*.
   c. $P(s_t = 1, s_{t-1} = 1|F_{\boldsymbol{T}}; \boldsymbol{\Omega_0})$ for all values of *t*.
   d. $P(s_t = 0, s_{t-1} = 1|F_{\boldsymbol{T}}; \boldsymbol{\Omega_0})$ for all values of *t*.
   e. $P(s_t = 1, s_{t-1} = 0|F_{\boldsymbol{T}}; \boldsymbol{\Omega_0})$ for all values of *t*.
   f. $P(s_t = 0, s_{t-1} = 0|F_{\boldsymbol{T}}; \boldsymbol{\Omega_0})$ for all values of *t*.
3. The Maximization stage:

   Set the next value of $\boldsymbol{\Omega_1} = \arg\max_{\boldsymbol{\Omega}} E[\log f(F_T|S_T, \boldsymbol{\Omega_0})]$
4. Repeat steps 2 and 3 until convergence of $\boldsymbol{\Omega}$.

Randomly selecting the starting values for $\boldsymbol{\Omega_0}$ is easy. For as long as the values picked are feasible or positive and probability values are from 0 to 1. However it should be noted that the distribution parameters for the 2 regimes is significantly different from each other (Kole, 2010). As remarked by Kole (2010), variances should be set such that regime 1 is three to four times larger than regime 0.

Iteration will also stop at some point when convergence is obtained. In this case, a criterion is imposed based on a certain value to when iteration should stop. To determine this, take note that the likelihood function is improved for every yield of new values. Setting the difference of the present and preceding likelihood functions in its minimum value is a welcome criterion for the convergence. Again, Kole (2010) remarked that this value should be set at below 10^{-8}. At this point, it is assumed that the converged values are obtained.

3.4 Data and Variables

This study is based on a time series data taken from the Wholesale Electricity Spot Market (WESM). It is a monthly Load Weighted Average Price (LWAP) for the Luzon grid and covers the period of January 2009 – December 2015. It was downloaded from the WESM website by collating the monthly summary reports they have provided. The explanatory variables on the other hand were taken from the data provider *www.Indexmundi.com* and *www.minerba.esdm.go.id* where the primary sources are the World Bank and the International Monetary Fund. These variables are of time series forms namely natural gas price and Indonesian coal price.

3.5 Dependent Variable

LWAP

This refers to the monthly Load Weighted Average Price of electricity generated in the Luzon grid. The variable's unit is in peso as it is the medium used by the participants of the market. This is the exogenous variable studied in the first phase of this study. As its name suggests, the variable is a weighted monthly average of the spot price. Power plants generate different amount of loads and sells their production on WESM. Therefore, prices differ with every plant and weighing the price through the load is necessary (www.wesm.ph).

3.6 Inputs

Fuel cost is a contributing factor to the movement of the electricity prices (www.eia.gov). Fuel such as natural gas and coal are mostly used for power generation especially in the case of Philippines. These commodities are traded worldwide and its prices are observed to be volatile. These input variables are used as the explanatory variable on the second part of the modelling which is the Tobit regression. These are the variables that are hypothesized to affect the occurrence of high price periods of the *LWAP:*

1. ***Natural Gas Price***

    This variable refers to the spot price of the natural gas being traded at the Henry Hub Terminal sourced from the International Monetary Fund. Its unit is in US Dollars per Million Metric British Thermal Unit.

2. ***Indonesian Coal Price***

This refers to the price of reference for the Indonesian coal index. It is one of the most common coal which is being traded worldwide and composes around 90% of the Philippine's total coal imports (www.manilatimes.net). Its unit is in US dollar per metric tons.

3.7 Tobit Regression

The term Tobit was derived from the name of James Tobin, the author who proposed such statistical model. It is used for estimating the relationship of a non-negative dependent variable and a vector of independent variables. However, unlike other methods, tobit is only used when there is censoring from below or above the dependent variable (www. ats.ucla.edu). The model assumes that a latent variable exists which linearly depends on the values of the independent variable. This study entails to quantify the influence of natural gas and Indonesian coal to the occurrence of electricity price spikes and using the tobit regression is logical. The estimation of this model is the second part of the study.

The spikes are predetermined through the result given by the smoothed probabilities in the MRS analysis. Based on the decision rule imposed by Hamilton (1989), all probabilities above 0.5 or $P(S_t = 1) > 0.5$ are considered spikes. This decision rule has been used by various studies not only on the subject of electricity but to all similar studies using MRS method, e.g. Weron and Janzcura (2010). The decision rule therefore can be used as a threshold for the tobit regression analysis. The model is specified by left-censoring at 0.5.

Let *Y* be the dependent variable that is left-censored at 0.5, there exist a latent variable *Y** such that *Y* is equal to *Y** whenever the latent variable is above 0.5 and 0 otherwise. This relationship is shown as an equation below:

$$Y^* = \beta X'_i + \epsilon_i$$

$$Y = \begin{cases} 0 & if\ Y^* \leq 0.5 \\ Y^* & if\ Y^* > 0.5 \end{cases}$$

where X_i is a vector of independent variables, in this case the natural gas and Indonesian coal price. The term β is a vector of the parameters and ∈ is the error term which is assumed to be randomly distributed.

Estimation of the parameters is done by using the maximum likelihood method. Thus, it is based on the loglikelihood function by assuming that the error is normally distributed. The loglikelihood function is first specified by an indicator function given as follows,

$$I = \begin{cases} 1 & if\ \ Y = 0 \\ 0 & if\ \ Y > 0 \end{cases}$$

With this indicator function, the loglikelihood function is determined, assuming the mean is 0 and the variance is σ^2:

$$\log L = \sum \left[I\ log\varphi\left(\frac{0.5 - \beta X'}{\sigma}\right) + \ (1 - I)\left(log\omega\left(\frac{Y - \beta X'}{\sigma}\right) - log\sigma\right)\right]$$

where φ and ω are noted as the probability density function and the cumulative density distribution function, respectively (Henningsen, 2015).

The coefficients β are interpreted in the same manner with the ordinary least square regression coefficients. However, the effect is linear only to the uncensored latent variable and not directly to the observed dependent variable (Long, 1997).

CHAPTER 4

RESULTS AND DISCUSSION

This chapter presents the trend of electricity price in the Philippines. The stationarity test, AR specification, Markov Regime Switching analysis and Tobit regression was also discussed in this chapter. The results were obtained using the R statistical software and Shazam.

4.1 Price Trend

Figure 6 shows the monthly Load Weighted Average Price (LWAP) of the electricity from January 2009 to December 2015. It is visually observable that the price exhibited spikes at times. The price started low from February 2009 up to December 2009 but started to indicate high variation after that period up to December 2015. As recorded by WESM, the demand for energy throughout the year 2010 increased by 13.77% from 2009. The supply meanwhile was insufficient through the first quarter of 2010 due to the El Nino phenomenon. This results to an upward trend during that period. The following year, 2011, the highest price was observed in October which immediately declined until the 1st quarter of 2012. The year 2012 on the other hand saw the highest price during the 2nd through the 3rd quarter. In 2013 meanwhile, the extreme values were observed and the highest price of the period was recorded during the last quarter of the year. During the same period, the Malampaya power plant was shut down and a series of outages from major coal plants occurred. In addition, Yolanda, the strongest typhoon ever recorded, hit the country in November 2013 causing the outages on various locations (PEMC, 2014). Though the price varies throughout the year, price patterns especially during summer and wet seasons are not clearly observable on some periods (www.wesm.ph).

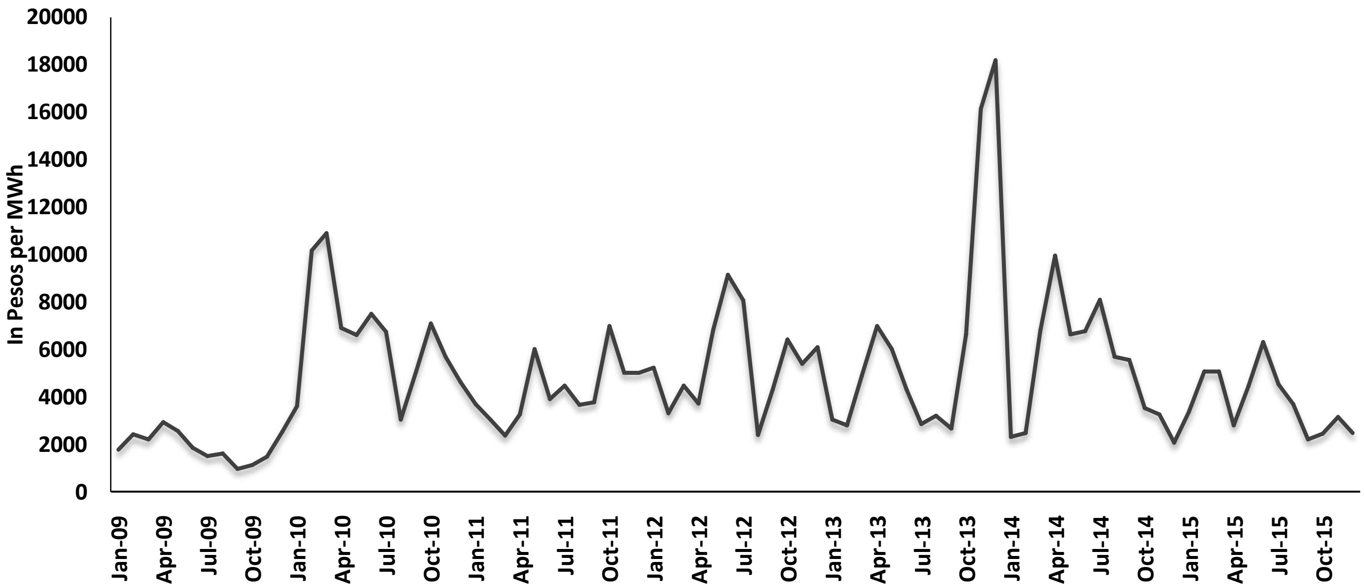

Figure 6. Monthly LWAP: January 2009 – December 2015.
Source: WESM.

The Natural Gas price exhibited fluctuations throughout the period January 2009 to December 2015. Noticeable in Figure 7 is the steep increase of price in November 2009 to January 2010 where it reached 5.84 US dollar per unit, after which the price fell until it reached the lowest point in April 2012. The price reached 1.95 US dollar per unit at that point and was the lowest observed at that time. From there, the price again climbed and reached its highest point in February 2014. This was the highest price recorded throughout the whole period at 5.97 US dollar per unit. After peaking, the price was observed to be on steady decline again. This continued until the last observation which also exhibited the lowest price of the whole period of observation at 1.92 US dollars per unit.

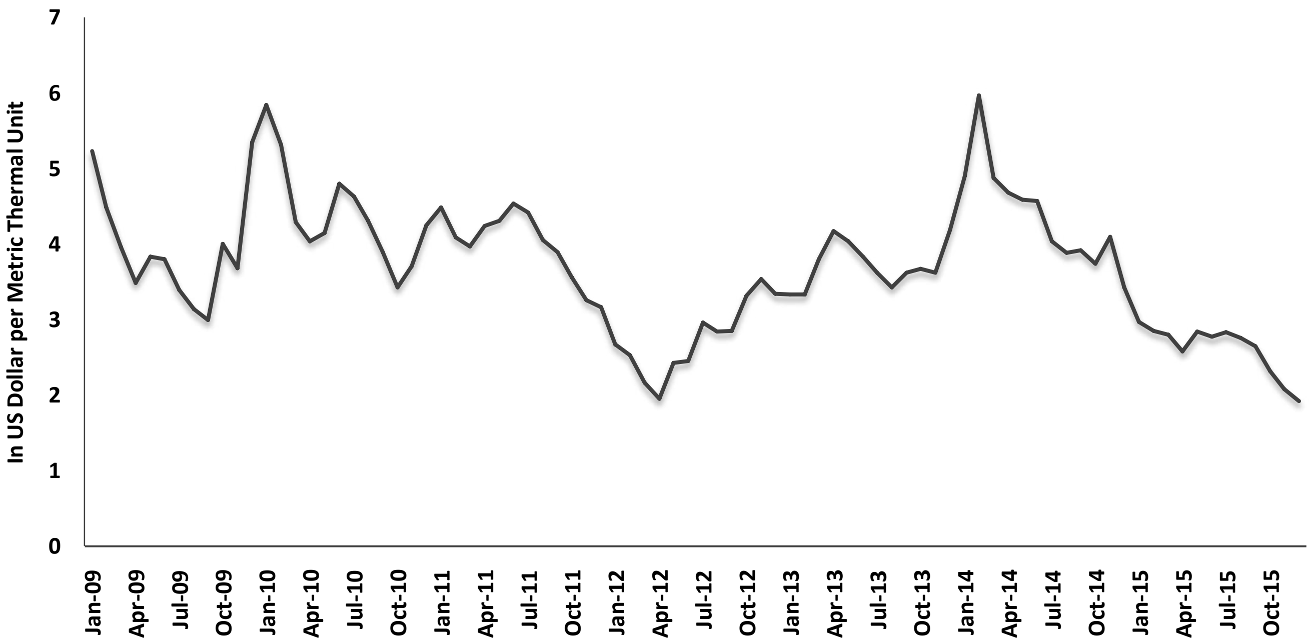

Figure 7. Natural Gas Price: January 2009 – October 2015.
Source: www.indexmundi.com.

The Indonesian Coal price does not exhibit much volatility as observed in Figure 8. The price is lower in 2009 and increased through the early months of 2011, where it peaked in February. It also exhibited the highest jump of price from January 2011 to February 2011. The price increased an average of 1.9 dollars per unit from 2009 up to the highest point in February 2011 resulting for the trend to go upwards. After peaking at 127.05 dollars, the price gradually declined up to the last observed period of December 2015. During this period, the price decreased on an average of 0.96 dollars per unit.

Figure 8. Indonesian Coal Price: January 2009 – October 2015.
Source: www.minerba.esdm.go.id.

4.2 Stationarity Test

Before starting the estimation procedure, a stationarity test was first conducted to all the variables. Common to all time series modelling, stationarity tests are applied to guarantee a more reliable prediction which is also appropriate to MRS. Mean, variance and autocorrelation should be constant overtime to assure the consistency of the model. The author implemented two different stationarity tests to come up with a reliable result. The first method used was the Augmented Dickey-Fuller test, then the Phillips-Perron test. Both tests hypothesize that there is a unit root on some level of confidence and therefore nonstationary. A package from R called “tseries” has the function called “adf.test” was used to test the existence of unit root. The function is an application of the

Augmented DickeyFuller test as the name “adf.test” suggests. Another package “stats” was used for the Phillips-Perron test.

The LWAP was first tested and the result is shown in Table 2. It indicates that for the Augmented Dickey-Fuller test, the statistic is determined to be -3.23 with the p-value of 0.089. This implies that the null hypothesis is rejected and therefore LWAP is stationary at 90% confidence. On the other hand, the statistic for the Phillips-Perron test is determined to be -4.75 with p-value of 0.01. This indicates that the null hypothesis is rejected and therefore LWAP is stationary with 99% confidence. This also implies that the raw data do not need to be differenced as they are already stationary in level form. Furthermore, this is a fair indication that the model to be calibrated will be dependable and allows the study to proceed to the calibration of the MRS model.

Table 2. Stationarity Test for LWAP

Augmented Dickey-Fuller Test

data:	LWAP
DickeyFuller (statistic)	-3.23
pvalue	0.09
alternative hypothesis:	stationary

Phillips-Perron Test

data:	LWAP
DickeyFuller (statistic)	-4.75
Pvalue	0.01
alternative hypothesis:	stationary

As for the natural gas and the Indonesian coal, both were found to be non-stationary at the raw level as shown on Table 3. Both of the tests show large p-values and are incapable of rejecting the hypothesis that a unit root exists. The variables are then differenced to investigate further. The results show that the variables are stationary at the first difference. All p-values given by the ADF and Phillips-Perron tests are significant with 99% confidence.

Table 3. Stationarity Test for Natural Gas and Indonesian Coal Price: First Differenced.

Augmented Dickey-Fuller Test

data:	Nat Gas	Indo Coal	Nat Gas (1)	Indo Coal (1)
DickeyFuller (statistic)	-2.09	-2.06	-4.20	-4.21
pvalue	0.54	0.55	0.01	0.01
alternative hypothesis:	stationary			

Phillips-Perron Test

data:	Nat Gas	Indo Coal	Nat Gas (1)	Indo Coal (1)
DickeyFuller (statistic)	-2.76	-1.30	-7.68	-6.36
Pvalue	0.26	0.86	0.01	0.01
alternative hypothesis:	stationary			

4.3 Lag Order Specification

After the stationarity of the series was confirmed, the estimation proceeds to finding the order of the MRS model. To have a view of the structure of the MRS model, it is necessary to determine the number of the exogenous variable’s lags that affect the system. The author chose to compare up to 6 lags of the AR to be able to capture the structure of the desired model. First, a correlogram of the partial autocorrelation function was investigated. As shown on Figure 9, the PACF suggests that the lag order of the series is 3. This is because the partial autocorrelation for the first 3 lags are statistically significant at 95% level and is above the threshold whereas the partial autocorrelations for the following lags are statistically insignificant.

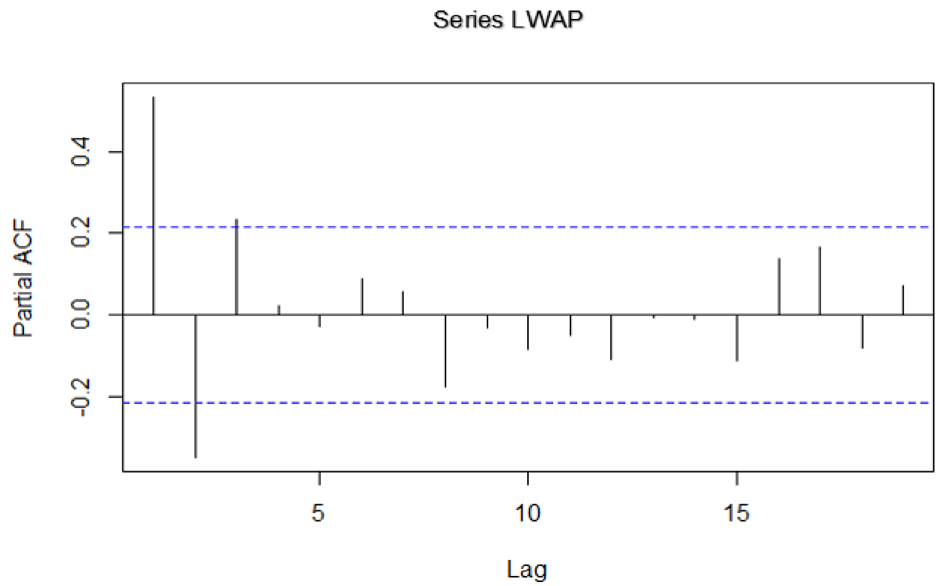

Figure 9. Partial Autocorrelation Correlogram.

The correlogram alone would not guarantee the true lag order of the series. It is therefore necessary to statistically compare each lags by a proven criterion. The Akaike Information Criterion (AIC) is the best criterion to estimate an autoregressive lag length (Liew, 2004). The lesser the value of AIC, the better is the fit of the model. Here in this study, AIC was used to select the best AR model for up to 6 lags. Table 4 gives the comparison of all 6 lags and consistent with the information given by the PACF, the model with the least AIC is the AR(3). It is therefore the best fit model among the 6 lags compared.

Table 4. Akaike Information Criterion Comparison.

Order	df	AIC	LogLikelihood
1	3	1537.4	765.68
2	4	1529.1	760.53
3	5	1526.4*	758.19
4	6	1528.3	758.16
5	7	1530.3	758.13
6	8	1531.5	757.74

4.4 Regime Identification

The order has been identified and the structure is now found to be up to 3 lags. The next step is then the identification of the regimes or spikes. This was attained by modeling through the MRS analysis. As discussed, EM Algorithm was used for this study to estimate the parameters of the MRS model. To do this the author used the MSwM package of R Statistical Software. The package is especially designed for the said algorithm method. In order to attain the best result, the author set the maximum iteration to 500. However, the parameters already converged on the 58th iteration as the difference for the likelihood functions were below the tolerance level 10^8 at that point.

The estimated parameters for MRS with order 3 were shown in Table 5. Regime 1 represents the model for the normal state while Regime 2 is considered the spike state or the period when the spikes were observed. All the parameters for both models were found to be highly significant. The R^2 for Regime 1 is at 0.65 which indicates that 65% of the variations on the normal state were explained by the model. Meanwhile, the R^2 for Regime 2 indicates that it explains 80% of the variations on the spike state. The intercepts (α) were both observed to be significantly different between both models with

782.73 for Regime 1 and 9794.70 for Regime 2. Regime 2's intercept then is more than 10 times larger than the intercept of Regime 1. This alone indicates that Regime 2 is indeed the spike state between the 2 models.

The coefficients for $LWAP_{t-1}$ were all positive for both regimes. This indicates a positive effect to the current observed price. To be specific, when all factors are to be held constant, Regime 1 has $\beta_1 = 0.70$ which implies that every peso increase of $LWAP_{t-1}$ increased the current price by 0.70 pesos given that it is in normal state. As for Regime 2, $\beta_1 = -1.03$ which also indicates that the price has been increased by 1.03 pesos for every peso increase of $LWAP_{t-1}$ when it is in a spike state. In addition, both models have shown that the coefficient of the $LWAP_{t-2}$, β_2 has a negative effect to the current price. For Regime 1, $\beta_2 = -0.19$ implies that if the given month is at normal state, every peso increase of its $LWAP_{t-2}$ has reduced the current price by 0.19 pesos. It follows that for Regime 2, $\beta_2 = -1.26$ implies that every peso increase of $LWAP_{t-2}$ reduces the price by 1.26 pesos if the month is at the spike state.

Lastly, $LWAP_{t-3}$ was also found to have a positive effect to the current price. At $\beta_3 = 0.26$, when the regime is on normal state, this means that every peso of $LWAP_{t-3}$ has increased the current price by 0.26 peso. On the other hand, at $\beta_3 = 0.67$, when the regime is on the spike state, the coefficient implies that every peso of $LWAP_{t-3}$ has decreased the current LWAP by 0.67 pesos.

Table 5. Markov Regime Switching Model Coefficients.

	Intercept (α)	β1	β2	β3	Standard Error
Regime 1	805.65**	0.70**	-0.19**	0.26**	1063.98
Regime 2	9807.66**	-1.03**	-1.26**	0.67**	1909.71

Significance Level: 0.001 '***' 0.01 '**' 0.05 '*'

Multiple Rsquared:

Regime 1	0.65
Regime 2	0.80

This result can be summarized into an equation in the form of equation (1) from Chapter 3. If $S_t = 0$ represent the regime 1 and $S_t = 1$ represent regime 2, then the MRS(3) model for LWAP is presented as follows,

$$LWAP_t = \begin{cases} 805.65 + 0.70LWAP_{t-1} - 0.19LWAP_{t-2} + 0.26LWAP_{t-3} + \varepsilon_{0t}\,, & S_t = 0 \\ 9807.66 - 1.03LWAP_{t-1} - 1.26LWAP_{t-2} + 0.67LWAP_{t-3} + \varepsilon_{1t}\,, & S_t = 1 \end{cases}$$

Another important result of the estimation is the transition probabilities of the regimes. Shown in Table 6 are the probabilities to which the states may change from one to another. The results indicate that the probability that a price spike may occur from the normal state or as the table configures, from regime 1 to regime 2 is $p_{12} = 0.17$. Meanwhile, the probability that the price will shift from spiked state to normal or regime 2 to regime 1 is $p_{21} = 0.53$. Additionally, the probabilities that the regimes will not change were $p_{11} = 0.83$ and $p_{22} = 0.47$ for regimes 1 and 2 respectively.

Table 6. Transition Probabilities

	Regime 1	Regime 2
Regime 1	0.83	0.17
Regime 2	0.53	0.47

The expected duration or months that the regimes will occur were also calculated given the probabilities. It was found that the average month that the price spike occurs is 1.9 months. This implies that whenever the price will shift to the spike state, it will stay there for about 1.9 months. On the other hand, the average duration that the normal state occurs is 5.9 months. Accordingly, this also implies that whenever the price will shift to the normal state, it will stay there for about 5.9 months. The periods indicate that the normal state occurs around 3 times as long as the spiked state.

Figure 10 shows the charted values of both the smoothed and filtered probabilities. Red lines indicate the smoothed probabilities while the black lines represent the filtered probabilities. As discussed, the smoothed probabilities are based on the filtered probabilities. It can be noticed that there were no spikes observed after the 61st month based on the chart from regime 2. The dominance of the normal state compared to the spike state is clearly observable on the graphs as well.

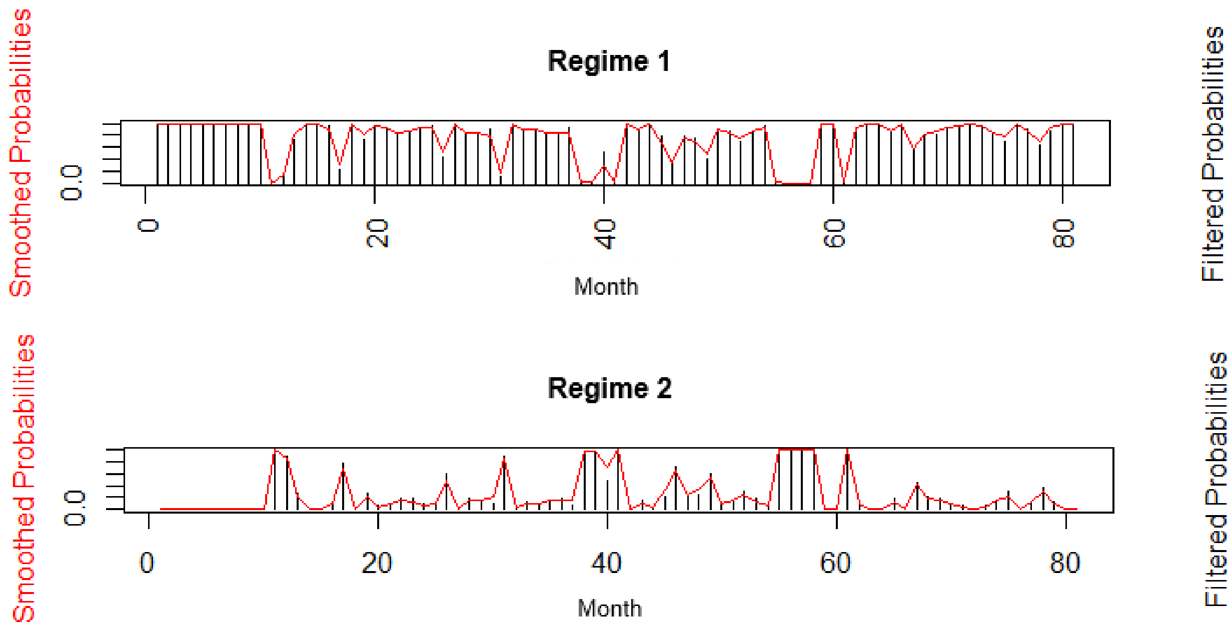

Figure 10. Smoothed and Filtered Probabilities.

The smoothed probabilities were used to classify the months into the normal state and the spike state. Table 7 shows the classification. The classification indicates that there are about 5 times that the spike occurred for just a month. The longest happened twice for about 4 months. First was during May 2012 – Aug 2012 and again in Oct 2013 – Jan 2013. To further investigate this period, the monthly report of WESM was referenced and discussed.

In Feb 2010 – March 2010, the WESM reported that the demand for energy during this 2 months increased and peaked a new high. A lot of factors have contributed to this per the institution's report. One of the major factors was the lesser contribution of natural gas due to the outage and limited supply by Malampaya. Generation from coal and geothermal also increased during this period (WESM Summary Report, March 2010).

For the period May 2012 – Aug 2012, the high price was attributed to unavailability of the power plants on most occasions. The demand increased therefore during this period which alerted the system operators. In fact, there were yellow alerts declared. Malampaya natural gas was also limited during this stretches which significantly reduced the capacity of the natural gas plants. Coal and natural gas was accounted for the 41.3% and 32.9% of the total generation during this period (WESM Summary Report, June - August 2012).

Lastly, for Oct 2013 – Jan 2014, the tighter supply has brought the higher price in the market. This was a result of the unavailability of major generating units. The period was maligned by the effect of the strongest typhoon Yolanda. A market suspension was implemented after the typhoon. In addition to this, the Malampaya onshore natural gas complex has undergone a scheduled maintenance that started on November 11, 2013 to December 10, 2013 (WESM Summary Report, October – January 2014).

Table 7. Date Classification of the Regimes.

Regime 1	Regime 2
Jan 2009 Jan 2010	Feb 2010 March 2010
Apr 2010 Jul 2010	10Aug
Sep 2010 Sep 2011	11Oct
Nov 2011 Apr 2012	May 2012 Aug 2012
Sep 2012 Dec 2012	13Jan
Feb 2013 Mar 2013	13Apr
May 2013 Sep 2013	Oct 2013 Jan 2014
Feb 2014 Mar 2014	14Apr
May 2014 Dec 2015	

4.5 Effect of Natural Gas and Indonesian Coal Prices to Electricity Price Spikes

The second stage is the investigation of the effect of the natural gas and Indonesian coal prices to the electricity price spikes. Tobit regression was used to quantify this effect. This method requires that the dependent variable has been censored either from the left or right. For this study, the censoring was based on the value imposed by the decision rule in the MRS analysis. The author used the decision rule implemented by Hamilton (1989) which is also followed by some authors in the study of electricity prices, e.g. Weron and Janzcura (2010). The notion was to categorize any probabilities higher than 0.5 as spikes. Hence, this value was used to censor the dependent variable from the left. The MRS procedure reduces the original data points to 82. Thus, for the tobit regression estimation, the total number of the uncensored observation is 15 out of the 82 observed events or around 18.5% of the data. To capture the relationship of the variables, the estimation first used the raw data of the natural gas and Indonesian coal prices as the independent variables. This is to show if there is some relationship between the price spikes and the input prices at this form. Note that the independent variables were all non-stationary at this level as shown in section 4.2.

Table 8 shows the result of the tobit regression for the raw form of the input prices. It is noticeable that the relationship is statistically non-existent at all of the independent variables. Neither of the prices of natural gas and Indonesian coal exhibited a significant coefficient as all t-ratios are too small.

Table 8. Summary of Results for Tobit Regression – Data in Level

Variable	Normalized Coefficient	Standard Error	t-ratio	Regression Coefficient
natural gas	0.01	0.18	0.54	0.13
Indo coal	0.001	0.01	0.81	9.37
Intercept	-1.84	0.98	-1.86*	-2.47

The investigation however should not stop at this stage. As the independent variables were found to be non-stationary, it is therefore necessary to examine the effect further through the stationary form of the data. The data is stationary at its first difference as also shown on section 4.2. With this form, the study proceeded using the tobit regression again. Table 9 shows the result for this investigation but it propagates the same outcome to the first tobit regression. The coefficients were all found to be insignificant for both the natural gas and the Indonesian coal prices. The relationship therefore is non-existent and the movement of the input prices does not affect the spikes in the electricity price.

Table 9. Summary of Results for Tobit Regression – Differenced Data

Variable	Normalized Coefficient	Standard Error	t-ratio	Regression Coefficient
natural gas(1)	0.13	0.35	0.38	0.18
Indo coal(1)	0.005	0.03	0.13	0.007
Intercept	-0.86	0.16	-5.30*	-1.17

CHAPTER 5

SUMMARY AND CONCLUSION

This paper investigated the electricity price spikes and the relationship of the input prices such the natural gas and the Indonesian coal to the occurrence of the spikes. By implementing two different stages of estimation, it has yielded two different results regarding the subject. First, the spikes were identified by using a 2-state process of the popular method Markov Regime Switching model, popularized by Hamilton (1989). It is best used for modeling nonlinearity in a time series and is applicable to electricity price due to the volatile characteristic. The model has fairly classified the spikes and the underlying coefficients were satisfyingly significant. Second, a Tobit Regression was applied to quantify the effects to the electricity price spikes of two of the largest contributing resources to Philippine power generation namely Natural Gas and the Indonesian Coal. The study used the monthly price these popular resources for the tobit regression.

The MRS was estimated using the EM Algorithm. Such method was used by numerous authors and was said to be more efficient than the traditional maximum likelihood estimation. The results show that the electricity price is very well explained by an MRS(3) model. The coefficients were highly significant up to the 3^{rd} lag of the series. Furthermore, much of the pre-identified spikes was captured by the model. The transition probabilities also indicated that the normal state is more dominant which was found to have occurred more than 3 times longer than the spike state.

The tobit regression found that the input prices has no significant effect to the price spikes. This is despite much of the power produced in the Philippines comes from this two resources. The findings proved that, in the Philippine setting, natural gas and Indonesian coal prices' effect to the electricity price does not extend to the occurrence of the spikes. There may have been other factor affecting the spikes and input prices are not one of them, at least on the scope of this study.

AREAS FOR FURTHER RESEARCH

Though this study has provided important information on the subject of electricity price, it is still recommended to conduct a research which will further enhance the literature and more importantly provide more knowledge regarding the subject matter. The author believes there are more areas that need to be explored to provide a different view of the dynamics in Philippine electricity price. The following are the suggested areas for further research:

- The study was conducted using a constrained amount of data points. It was implemented using the monthly average of the electricity price for a seven year period. Meanwhile, the trading transaction in WESM is conducted on hourly basis. Thus, having an hourly data may provide an interesting extension to this study. Furthermore, hourly data is known to have a negative value. This will be an interesting subject to extend the MRS modeling to 3-states, a model with a negative, normal, and the spike state.

- When conducting the tobit regression, there were other factors not considered for this study due to the unavailability of the data as far as the author is concerned. Typhoons and temperature were identified to be factors by some research. If availing the data is possible, then incorporating them may provide more information regarding the factors affecting the price spikes.
- Another way of modeling the MRS is by using the MRS-Vector Autoregression method. This method incorporates the input variables directly to the system. This type of modeling will give a different view in interpreting the price spike by the direct effects of the factors.

BIBLIOGRAPHY

Alizadeh, A., N. Nomikos and P. Pouliasis, 2008. *"A Markov Regime Switching Approach for Hedging Energy Commodities"*. Cass Business School. London, UK.

Bessec, M., J. Fouquau, and S. Meritet, 2014. *"Forecasting electricity spot prices using time series models with a double temporal segmentation"*. pp. 1821.

Blochlinger, L., 2008. *"Power Prices – A RegimeSwitching Spot/Forward Price Model with Kim Filter Estimation"*. University of St. Gallen, St. Gallen, Switzerland.

Chauvet, M., C. Juhn, and S. Potter, 2002. *"Markov Switching in Disaggregate Unemployment Rates"*. Empirical Econometrics 27, pp. 205232.

Chen, M.Y., 2013. *"MarkovSwitching Models"*. National Chung Hsing University. Taichung, Taiwan.

Cruz, C.J. and D. Mapa, 2013. *"An Early Warning System for Inflation in the Philippines Using MarkovSwitching and Logistic Regression Models"*. University of the Philippines Diliman. MPRA Paper No. 50078.

De Jong, C., 2006. *"The Nature of Power Spikes: A RegimeSwitch Approach"*. Studies in Nonlinear Dynamics & Econometrics. Vol.10, Issue 3. pp. 725.

Dempster, Laird and Rubin, 1977. *"Maximum Likelihood from Incomplete Data via the EM Algorithm"*. Harvard University and Educational Testing Service. Massachusetts, USA.

Engel, R., 1982. *"Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation"*. Econometrica, Vol. 50, No. 4. (Jul., 1982), pp. 9871007.

Eriksurd, A. L., 2014. *"Short Term Electricity Price Forecasting in the Nordic Region".*

Haldrup, N. and M. Nielsen, 2004. *"A Regime Switching Long Memory Model for Electricity Prices"*. University of Aarhus, Denmark. Working Paper No. 200402.

Haldrup, N., F. Nielsen, and M. Nielsen, 2009. *"A vector autoregressive model for electricity prices subject to long memory and regime switching"*. Queen's Economics Department Working Paper, No. 1211.

Hamilton, J., 1989. *"A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle"*. Econometrica. Vol. 57, No. 2, pp. 357384

Hamilton, J., 1990. *"Analysis of Time Series Subject to Changes in Regime"*. Journal of Econometrics 45 (1990), pp. 3970.

Hamilton, J. and R. Susmel, 1994. *"Autoregressive conditional heteroskedasticity and changes in regime"*. Journal of Econometrics 64 (1994), pp. 307333.

Hardle, W.K. and S. Truck, 2010. *"The dynamics of hourly electricity prices"*. Humboldt Universitat zu Berlin, Germany. SFB 649 Discussion Paper 2010013.

Henningsen, A., 2015. *"Estimating Censored Regression Models in R using the censReg Package".* University of Copenhagen, Denmark.

Higgs, H. and A. Worthington, 2010. *"Modelling spot prices in the Australian wholesale electricity Market"*. Energy Studies Review Vol. 17, No.1, 2010.

Kanamura, T. and K. Ohashi, 2004. *"On Transition Probabilities of Regime Switching in Electricity Prices".*

Kano, S. and M. Ohta, 2002. *"An Empirical Matching Function with Regime Switching: The Japanese Case"*. University of Tsukuba, Japan.

Kim, C. J., 1994. *"Dynamic linear models with Markovswitching".* Korea University. Seoul, South Korea.

Kole, E., 2010. *"Regime Switching Models: An Example for a Stock Market Index"*. Erasmus University Rotterdam. The Netherlands.

Krolzig, H.M., M. Marcellino, and G. Mizon, 2000. *"A Markov–Switching Vector Equilibrium Correction Model of the UK Labour Market"*.

Kuan, C.M., 2002. *"Lecture on the Markov Switching Model"*. Institute of Economics, Academia Sinica, Taipei 115, Taiwan.

Liew, Venus KhimSen, 2004. "Which Lag Length Selection Criteria Should We Employ?". Economics Bulletin, Vol. 3, No. 33, pp. 1−9.

Lindsay, K., A. Hurn and T. Christensen, 2011. *"Forecasting Spikes in Electricity Prices"*. NCER Working Paper Series. pp. 24.

Long, J. S. (1997). *"Regression Models for Categorical and Limited Dependent Variables".* Thousand Oaks, CA.

Mahieu, R. and R. Huisman, 2001. *"Modelling Electricity Prices: A Markov Switching Approach"*. Erasmus University. The Netherlands.

Murara, J. P., 2010. *“Regime Switching Models for Electricity Time Series that Capture Fat Tailed Distributions”*. Lappeenranta University of Technology. Finland.

Noren, V., 2013. *“Modelling Power Spikes with Inhomogeneous MarkovSwitching Models”*. Lund University. Sweden.

Patalinghug, E., 2003. *“An Analysis of the Philippine Electric Power Industry”.* University of the Philippines. Quezon City, Philippines.

Philippine Electricity Market Corporation (PEMC), 2014. *“Study of Mitigating Measures for the Philippine Electricity Market”*.

Serati, M., M. Manera, and M. Plotegher, 2008. *“Modeling Electricity Prices: From the State of the Art to a Draft of a New Proposal”*.

Tian Tse, L., 1988. *"Application of Logistic Regression and Survival Analysis to the Study of CEP, Manpower Performance and Attrition"*. National University of Singapore.

Vigoda, E., 2003. *“Markov Chains, Coupling, Stationary Distribution”*. University of Chicago. Illinois, USA.

Weron, R., 2006. *“Market price of risk implied by Asianstyle electricity options and futures”*. Energy Economics, 2008, vol. 30, issue 3. pp. 10981115.

Weron, R. and J. Janczura, 2010. *“Efficient estimation of Markov regimeswitching models: An Application to electricity wholesale market prices”*. Munich Personal RePec Paper No. 26628.

Wholesale Electricity Spot Market (WESM), 2010. *“Monthly Summary Report March 2010”.*

Wholesale Electricity Spot Market (WESM), 2012. *“Monthly Summary Report June 2012”.*

Wholesale Electricity Spot Market (WESM), 2012. *“Monthly Summary Report July 2012”.*

Wholesale Electricity Spot Market (WESM), 2012. *“Monthly Summary Report August 2012”.*

Wholesale Electricity Spot Market (WESM), 2013. *“Monthly Summary Report October 2013”.*

Wholesale Electricity Spot Market (WESM), 2013. *“Monthly Summary Report November 2013”.*

Wholesale Electricity Spot Market (WESM), 2013. *"Monthly Summary Report December 2013".*

Wholesale Electricity Spot Market (WESM), 2014. *"Monthly Summary Report January 2014".*

Wlodarczyk, A. and M. Zawada, 2008*. "Markov Switching Models for the Price of Electric Energy on the Energy Stock Market in Poland"*. Dynamic Econometric Models. Vol.8 pp. 177178

Online References:

http://www.interaksyon.com/business/40664/phelectricityratesamong10highestinsurveyo f44countries

http://www.eia.gov/energyexplained/index.cfm?page=electricity_factors_affecting_prices

http://newsinfo.inquirer.net/620354/around90ofmeralcocustomerswithoutelectricitydueto glenda

http://www.Indexmundi.com

http://www.wesm.ph

http://mathworld.wolfram.com

http://www.manilatimes.net/anybodysaycoalpowerpolicywillsignificantlyreducecostelectri city/118479/

https://www.epexspot.com/en/companyinfo/basics_of_the_power_market/negative_pric es

http://www.ats.ucla.edu/stat/stata/dae/tobit.htm

https://www.minerba.esdm.go.id/

APPENDICES

Appendix A. Data used in the study

Date	LWAP (in peso per MWh)	Natural Gas Price (in US dollar per MMBTu)	Indonesian Coal Price (in US dollar per metric tons)
Jan-09	1769	5.23	78.7
Feb-09	2413	4.49	81.35
Mar-09	2197	3.96	75.11
Apr-09	2949	3.49	63.08
May-09	2546	3.83	62.83
Jun-09	1864	3.8	63.87
Jul-09	1499	3.39	71.29
Aug-09	1623	3.14	71.47
Sep-09	965	2.99	70.44
Oct-09	1131	4	66.71
Nov-09	1480	3.68	68.99
Dec-09	2529	5.35	74.51
Jan-10	3620	5.84	77.39
Feb-10	10168	5.32	87.81
Mar-10	10898	4.29	86.64
Apr-10	6902	4.04	86.58
May-10	6597	4.15	92.07
Jun-10	7510	4.8	97.22
Jul-10	6733	4.63	96.65
Aug-10	3046	4.31	94.86
Sep-10	5058	3.89	90.05
Oct-10	7092	3.43	92.68
Nov-10	5692	3.71	95.51
Dec-10	4643	4.25	103.41
Jan-11	3680	4.49	112.4
Feb-11	3055	4.09	127.05
Mar-11	2380	3.97	122.43
Apr-11	3261	4.24	122.02
May-11	6008	4.31	117.61
Jun-11	3895	4.54	119.03
Jul-11	4480	4.42	118.24
Aug-11	3672	4.05	117.21
Sep-11	3772	3.89	116.26
Oct-11	6976	3.56	119.24

Date	LWAP (in peso per MWh)	Natural Gas Price (in US dollar per MMBTu)	Indonesian Coal Price (in US dollar per metric tons)
Nov-11	5019	3.26	116.65
Dec-11	5026	3.16	112.67
Jan-12	5218	2.67	109.29
Feb-12	3324	2.53	111.58
Mar-12	4476	2.16	112.87
Apr-12	3719	1.95	105.61
May-12	6835	2.43	102.12
Jun-12	9141	2.45	96.65
Jul-12	8068	2.96	87.56
Aug-12	2392	2.84	84.65
Sep-12	4301	2.85	86.21
Oct-12	6409	3.32	86.04
Nov-12	5394	3.54	81.44
Dec-12	6107	3.34	81.75
Jan-13	3038	3.33	87.55
Feb-13	2792	3.33	88.35
Mar-13	4890	3.8	90.09
Apr-13	6993	4.17	88.56
May-13	6023	4.04	85.33
Jun-13	4310	3.83	84.87
Jul-13	2865	3.62	81.69
Aug-13	3202	3.43	76.7
Sep-13	2678	3.62	76.89
Oct-13	6641	3.67	76.61
Nov-13	16122	3.62	78.13
Dec-13	18194	4.19	80.31
Jan-14	2309	4.9	81.9
Feb-14	2488	5.97	80.44
Mar-14	6788	4.88	77.01
Apr-14	9950	4.68	74.81
May-14	6642	4.59	73.6
Jun-14	6773	4.57	73.64
Jul-14	8104	4.04	72.45
Aug-14	5694	3.88	70.29
Sep-14	5544	3.92	69.69
Oct-14	3539	3.74	67.26
Nov-14	3253	4.1	65.7

Date	LWAP (in peso per MWh)	Natural Gas Price (in US dollar per MMBTu)	Indonesian Coal Price (in US dollar per metric tons)
Dec-14	2068	3.43	64.65
Jan-15	3367	2.97	63.84
Feb-15	5079	2.85	62.92
Mar-15	5075	2.8	67.76
Apr-15	2810	2.58	64.48
May-15	4457	2.84	61.08
Jun-15	6322	2.77	59.59
Jul-15	4518	2.83	59.16
Aug-15	3691	2.76	59.14
Sep-15	2220	2.65	58.21
Oct-15	2454	2.32	57.39
Nov-15	3158	2.08	54.43
Dec-15	2491	1.92	53.51

Appendix B. R Statistical Software – Markov Regime Switching Output

```
# load the package MSwM (Markov Switch Model)
>library(MSwM)

#load the data

>data=read.csv(file.choose(),header=F, sep="\t")
>lwap<-as.ts(data$V1)
>markovmod=lm(lwap~1,data)

#make a MRS model with k=2 states and p=3 lags – max iteration is 200

>markov=msmFit(markovmod, k=2,sw=c(T,T,T,T,T),p=3,
        control=list(trace=TRUE,maxiter=200,tol=10e-8,maxiterOuter=25,maxiterInner=50))

>summary(markov)

#to get the smoothed probabilities
>markov@Fit@smoProb

RESULT:

Initial Value: 704.6986
 Inner Iter. 1  logLikel= 704.6986
 Inner Iter. 2  logLikel= 704.6986
 Inner Iter. 3  logLikel= 704.6986
 Inner Iter. 4  logLikel= 704.6986
 Inner Iter. 5  logLikel= 704.6986
 Inner Iter. 6  logLikel= 704.6986
 Inner Iter. 7  logLikel= 704.6986
 Inner Iter. 8  logLikel= 704.6986
 Inner Iter. 9  logLikel= 704.6986
 Inner Iter. 10  logLikel= 704.6986
 Inner Iter. 11  logLikel= 704.6986
 Inner Iter. 12  logLikel= 704.6986
 Inner Iter. 13  logLikel= 704.6986
 Inner Iter. 14  logLikel= 704.6986
 Inner Iter. 15  logLikel= 704.6986
 Inner Iter. 16  logLikel= 704.6986
 Inner Iter. 17  logLikel= 704.6986
 Inner Iter. 18  logLikel= 704.6986
 Inner Iter. 19  logLikel= 704.6986
 Inner Iter. 20  logLikel= 704.6986
```

```
Inner Iter. 21  logLikel= 704.6986
Inner Iter. 22  logLikel= 704.6986
Inner Iter. 23  logLikel= 704.6986
Inner Iter. 24  logLikel= 704.6986
Inner Iter. 25  logLikel= 704.6986
Inner Iter. 26  logLikel= 704.6986
Inner Iter. 27  logLikel= 704.6986
Inner Iter. 28  logLikel= 704.6986
Inner Iter. 29  logLikel= 704.6986
Inner Iter. 30  logLikel= 704.6986
Inner Iter. 31  logLikel= 704.6986
Inner Iter. 32  logLikel= 704.6986
Inner Iter. 33  logLikel= 704.6986
Inner Iter. 34  logLikel= 704.6986
Inner Iter. 35  logLikel= 704.6986
Inner Iter. 36  logLikel= 704.6986
Inner Iter. 37  logLikel= 704.6986
Inner Iter. 38  logLikel= 704.6986
Inner Iter. 39  logLikel= 704.6986
Inner Iter. 40  logLikel= 704.6986
Inner Iter. 41  logLikel= 704.6986
Inner Iter. 42  logLikel= 704.6986
Inner Iter. 43  logLikel= 704.6986
Inner Iter. 44  logLikel= 704.6986
Inner Iter. 45  logLikel= 704.6986
Inner Iter. 46  logLikel= 704.6986
Inner Iter. 47  logLikel= 704.6986
Inner Iter. 48  logLikel= 704.6986
Inner Iter. 49  logLikel= 704.6986
Inner Iter. 50  logLikel= 704.6986
Inner Iter. 51  logLikel= 704.6986
Inner Iter. 52  logLikel= 704.6986
Inner Iter. 53  logLikel= 704.6986
Inner Iter. 54  logLikel= 704.6986
Inner Iter. 55  logLikel= 704.6986
Inner Iter. 56  logLikel= 704.6986
Inner Iter. 57  logLikel= 704.6986
Inner Iter. 58  logLikel= 704.6986
Calculating standard errors...
> summary(markov)
Markov Switching Model

Call: msmFit(object = markovmod, k = 2, sw = c(T, T, T, T, T), p = 3,
    control = list(trace = TRUE, maxiter = 200, tol = 1e-07,
        maxiterOuter = 25, maxiterInner = 50))
```

```
    AIC     BIC    logLik
  1425.397 1479.51 -704.6986

Coefficients:

Regime 1
---------
             Estimate Std. Error t value  Pr(>|t|)
(Intercept)(S) 782.7286   456.7409  1.7137   0.08658 .
lwap_1(S)       0.6988     0.0937  7.4578 8.793e-14 ***
lwap_2(S)      -0.1886     0.0768 -2.4557   0.01406 *
lwap_3(S)       0.2595     0.0578  4.4896 7.136e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1071.19
Multiple R-squared: 0.6535

Standardized Residuals:
        Min           Q1          Med           Q3          Max
-1634.1502356  -673.9720787    0.2103016   807.3739874  1869.4599773
Regime 2
---------
              Estimate Std. Error t value  Pr(>|t|)
(Intercept)(S) 9794.6928  1786.5131  5.4826 4.191e-08 ***
lwap_1(S)        1.0291     0.1962  5.2452 1.561e-07 ***
lwap_2(S)       -1.2759     0.3037 -4.2012 2.655e-05 ***
lwap_3(S)       -0.6650     0.3859 -1.7232   0.08485 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1905.952
Multiple R-squared: 0.8038

Standardized Residuals:
       Min         Q1        Med         Q3        Max
-2825.91436  -642.02079   -36.95208   318.74489  5039.46882


Transition probabilities:
         Regime 1  Regime 2
Regime 1 0.821904 0.1780964
Regime 2 0.5370636 0.4629364
```

Appendix C. R Statistical Software – Stationarity Test Output

```
>library(tseries)
>data=read.csv(file.choose(),header=F, sep="\t")
>LWAPrice<-as.ts(data$V1)
>htest.obj<-adf.test(LWAPrice)
>htest.obj


RESULT:
Augmented Dickey-Fuller Test

data:  Raw
Dickey-Fuller = -3.2312, Lag order = 4, p-value = 0.08854
alternative hypothesis: stationary
```

```
# Phillips-Perron Test
>library(stats)
>data=read.csv(file.choose(),header=F, sep="\t")
>LWAPrice<-as.ts(data$V1)
>pptest<-PP.test(LWAPrice)
>pptest

RESULT:
Phillips-Perron Unit Root Test

data:  LWAPrice
Dickey-Fuller = -4.7504, Truncation lag parameter = 3, p-value = 0.01
alternative hypothesis: stationary
```

Appendix D. Shazam Version 11 – Tobit Regression Output

```
Welcome to SHAZAM (Double Precision) v11.0 -  JUNE 201 Windows7 PAR=    78
 ...NOTE..CURRENT WORKING DIRECTORY IS: C:\Users\user\Documents\SHAZAM

 |_SET NOWIDE
 |_READ (C:\Users\user\Desktop\Raw Data - wesm indo natgas.csv) lwap natgas indo
smooth natgas1 indo1 tobit /skiplin
 | es=3
 ...NOTE..UNIT 88 IS NOW ASSIGNED TO: C:\Users\user\Desktop\Raw Data - wesm indo
natgas.
 csv
 ...NOTE..    9 VARIABLES AND       20 OBSERVATIONS STARTING AT OBS       1


 |_TOBIT tobit natgas indo/ LIST

 REQUIRED MEMORY IS PAR=      14 CURRENT PAR=      78

 TOBIT ANALYSIS, LIMIT=    0.00        25 MAX ITERATIONS
        17 LIMIT OBSERVATIONS
         3 NON-LIMIT OBSERVATIONS

 ITERATION  0  NORMALIZED COEFFICIENTS
  0.20942      0.17313E-01  -2.8523      0.98850

 ITERATION  1  NORMALIZED COEFFICIENTS
  0.40614      0.36992E-01  -5.6937      0.91138

 ITERATION  2  NORMALIZED COEFFICIENTS
  0.54275      0.52397E-01  -7.7183      0.91169

 ITERATION  3  NORMALIZED COEFFICIENTS
  0.60566      0.59811E-01  -8.6804      0.90974

 ITERATION  4  NORMALIZED COEFFICIENTS
  0.61753      0.61225E-01  -8.8636      0.90920

 ITERATION  5  NORMALIZED COEFFICIENTS
  0.61788      0.61267E-01  -8.8691      0.90918

 ITERATION  6  NORMALIZED COEFFICIENTS
  0.61788      0.61267E-01  -8.8691      0.90918

 ITERATION  7  NORMALIZED COEFFICIENTS
  0.61788      0.61267E-01  -8.8691      0.90918

  FIRST DERIVATIVES OF LOG OF LIKELIHOOD FUNCTION EVALUATED AT MAXIMUM
   -0.14876989E-13  -0.34461323E-12  -0.36637360E-14   0.17763568E-14

 NUMBER OF ITERATIONS =  7

  DEPENDENT VARIABLE = TOBIT
 VARIANCE OF THE ESTIMATE =   1.2098
 STANDARD ERROR OF THE ESTIMATE =   1.0999

                             ASYMPTOTIC
 VARIABLE  NORMALIZED     STANDARD   T-RATIO    REGRESSION ELASTICITY ELASTICITY
          COEFFICIENT      ERROR                COEFFICIENT  OF INDEX   OF E(Y)
 NATGAS    0.61788      0.58734       1.0520      0.67961      21.8506   5.7077
 INDO      0.61267E-01  0.47640E-01   1.2860      0.67387E-01  41.8495  10.9317
```

```
CONSTANT  -8.8691       5.3453      -1.6592      -9.7551
TOBIT     0.90918      0.45286       2.0076

THE PREDICTED PROBABILITY OF Y > LIMIT GIVEN AVERAGE X(I) = 0.0730
 THE OBSERVED FREQUENCY OF Y > LIMIT IS = 0.1500
 AT MEAN VALUES OF ALL X(I), E(Y) =       0.0358

                                                  DEPENDENT VARIABLE
 OB   INDEX      PROB(X)     DENSITY(X)  OBSERVED    EXPECTED   CONDITIONAL

   1  -1.8205     0.34339E-01 0.76070E-01 0.00000E+00 0.14908E-01  ------
   2  -2.8480     0.21999E-02 0.69125E-02 0.00000E+00 0.71186E-03  ------
   3  -2.6532     0.39865E-02 0.11811E-01 0.00000E+00 0.13575E-02  ------
   4  -2.6080     0.45532E-02 0.13302E-01 0.00000E+00 0.15694E-02  ------
   5  -2.4068     0.80473E-02 0.22033E-01 0.00000E+00 0.29318E-02  ------
   6  -2.5502     0.53829E-02 0.15441E-01 0.00000E+00 0.18847E-02  ------
   7  -2.7060     0.34050E-02 0.10253E-01 0.00000E+00 0.11433E-02  ------
   8  -2.3105     0.10431E-01 0.27652E-01 0.00000E+00 0.39057E-02  ------
   9  -2.3685     0.89304E-02 0.24142E-01 0.00000E+00 0.32888E-02  ------
  10 -0.99843     0.15904     0.24235     0.00000E+00 0.91912E-01  ------
  11 -0.51922     0.30180     0.34863     0.00000E+00 0.21110      ------
  12 -0.20212     0.41991     0.39088      1.0000     0.33657     0.80153
  13 -0.91022     0.18135     0.26364     0.87278     0.10841     0.59778
  14  -1.0684     0.14268     0.22545     0.00000E+00 0.80316E-01  ------
  15 -0.66404     0.25333     0.32001     0.00000E+00 0.16695      ------
  16  0.53106E-01 0.52118     0.39838     0.00000E+00 0.46862      ------
  17 -0.86857E-01 0.46539     0.39744     0.00000E+00 0.39268      ------
  18 -0.39425     0.34670     0.36911     0.68789     0.25565     0.73737
  19 -0.94845     0.17145     0.25443     0.00000E+00 0.10099      ------
  20  -1.0715     0.14196     0.22469     0.00000E+00 0.79818E-01  ------

LOG-LIKELIHOOD FUNCTION= -8.3955652
MEAN-SQUARE ERROR= 0.84460265E-01
MEAN ERROR=-0.11797309E-01
MEAN ABSOLUTE ERROR= 0.17420659
SQUARED CORRELATION BETWEEN OBSERVED AND EXPECTED VALUES=  0.12608
|_stop
```

Appendix E. Shazam Version 11 – Tobit Regression: First Differenced Output

```
Welcome to SHAZAM (Double Precision) v11.0 -  JUNE 201 Windows7 PAR=    78
 ...NOTE..CURRENT WORKING DIRECTORY IS: C:\Users\user\Documents\SHAZAM

 |_SET NOWIDE
 |_READ (C:\Users\user\Desktop\Raw Data - wesm indo natgas.csv) lwap natgas indo
smooth natgas1 indo1 tobit /skiplin
 | es=3
 ...NOTE..UNIT 88 IS NOW ASSIGNED TO: C:\Users\user\Desktop\Raw Data - wesm indo
natgas.
 csv
 ...NOTE..    9 VARIABLES AND       20 OBSERVATIONS STARTING AT OBS       1


 |_TOBIT tobit natgas1 indo1/ LIST

 REQUIRED MEMORY IS PAR=      14 CURRENT PAR=      78

 TOBIT ANALYSIS, LIMIT=    0.00        25 MAX ITERATIONS
        17 LIMIT OBSERVATIONS
         3 NON-LIMIT OBSERVATIONS

 ITERATION  0  NORMALIZED COEFFICIENTS
 -0.58346      0.44301E-01 -0.66058      1.0781

 ITERATION  1  NORMALIZED COEFFICIENTS
  -1.3565      0.76521E-01  -1.2316      1.0922

 ITERATION  2  NORMALIZED COEFFICIENTS
  -2.5000      0.91841E-01  -1.7426      1.2377

 ITERATION  3  NORMALIZED COEFFICIENTS
  -3.2297      0.94865E-01  -2.0991      1.3214

 ITERATION  4  NORMALIZED COEFFICIENTS
  -3.3464      0.95200E-01  -2.1641      1.3311

 ITERATION  5  NORMALIZED COEFFICIENTS
  -3.3485      0.95208E-01  -2.1654      1.3311

 ITERATION  6  NORMALIZED COEFFICIENTS
  -3.3485      0.95208E-01  -2.1654      1.3311

 ITERATION  7  NORMALIZED COEFFICIENTS
  -3.3485      0.95208E-01  -2.1654      1.3311

  FIRST DERIVATIVES OF LOG OF LIKELIHOOD FUNCTION EVALUATED AT MAXIMUM
    0.55511151E-15   0.26645353E-14  -0.77715612E-15   0.66613381E-15

 NUMBER OF ITERATIONS =  7

  DEPENDENT VARIABLE = TOBIT
 VARIANCE OF THE ESTIMATE =  0.56437
 STANDARD ERROR OF THE ESTIMATE =  0.75125

                              ASYMPTOTIC
 VARIABLE  NORMALIZED     STANDARD   T-RATIO    REGRESSION ELASTICITY ELASTICITY
          COEFFICIENT      ERROR                COEFFICIENT  OF INDEX   OF E(Y)
 NATGAS1   -3.3485       1.5696     -2.1333      -2.5155       1.0413   0.4659
 INDO1     0.95208E-01  0.74884E-01   1.2714      0.71524E-01   0.3165   0.1416
```

```
CONSTANT  -2.1654      0.80060      -2.7047      -1.6267
TOBIT      1.3311      0.64641       2.0592

THE PREDICTED PROBABILITY OF Y > LIMIT GIVEN AVERAGE X(I) = 0.0266
 THE OBSERVED FREQUENCY OF Y > LIMIT IS = 0.1500
 AT MEAN VALUES OF ALL X(I), E(Y) =       0.0076

                                                 DEPENDENT VARIABLE
 OB   INDEX      PROB(X)     DENSITY(X)  OBSERVED    EXPECTED   CONDITIONAL

   1 -0.98479     0.16236     0.24565     0.00000E+00 0.64425E-01  ------
   2  -1.7369     0.41198E-01 0.88263E-01 0.00000E+00 0.12549E-01  ------
   3  -3.3277     0.43787E-03 0.15716E-02 0.00000E+00 0.86042E-04  ------
   4  -1.9659     0.24654E-01 0.57766E-01 0.00000E+00 0.69852E-02  ------
   5 -0.86070E-01 0.46571     0.39747     0.00000E+00 0.26848      ------
   6  -1.3111     0.94907E-01 0.16890     0.00000E+00 0.33401E-01  ------
   7  -1.7612     0.39104E-01 0.84601E-01 0.00000E+00 0.11818E-01  ------
   8  -5.9025     0.17904E-08 0.10856E-07 0.00000E+00 0.21636E-09  ------
   9 -0.87680     0.19030     0.27163     0.00000E+00 0.78711E-01  ------
  10  -7.2318     0.23831E-12 0.17552E-11 0.00000E+00 0.23914E-13  ------
  11  -3.5319     0.20626E-03 0.77997E-03 0.00000E+00 0.38673E-04  ------
  12  0.56788     0.71494     0.33953      1.0000     0.56008     0.78340
  13   1.1722     0.87943     0.20071     0.87278     0.92519      1.0520
  14  -1.3340     0.91105E-01 0.16387     0.00000E+00 0.31805E-01  ------
  15  -2.0110     0.22161E-01 0.52810E-01 0.00000E+00 0.61925E-02  ------
  16  -3.8516     0.58679E-04 0.23966E-03 0.00000E+00 0.10260E-04  ------
  17  -1.6504     0.49429E-01 0.10220     0.00000E+00 0.15488E-01  ------
  18  -1.2643     0.10306     0.17940     0.68789     0.36883E-01 0.35787
  19  -1.2170     0.11181     0.19024     0.00000E+00 0.40700E-01  ------
  20 -0.37469     0.35395     0.37190     0.00000E+00 0.17976      ------

LOG-LIKELIHOOD FUNCTION= -6.5126534
MEAN-SQUARE ERROR= 0.36961530E-01
MEAN ERROR=-0.14403156E-01
MEAN ABSOLUTE ERROR= 0.94689000E-01
SQUARED CORRELATION BETWEEN OBSERVED AND EXPECTED VALUES=  0.61704
|_stop
```