DEMAND FOR ELECTRICITY: A CASE IN SOUTH KOREA

BY

TINGWEN LIU

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Economics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2013

Urbana, Illinois

Doctoral Committee:

Professor In-Koo Cho, Chair
Professor Anil Bera
Professor George Deltas
Associate Professor Seung-Hyun Hong

Abstract

This dissertation studies wholesale and sector-wise electricity demand in South Korea. Electricity demand analysis provides useful insights for market performance evaluation, load prediction, market restructuring, tariff schedule design, etc. In recent years, there has been a heated debate in Korea on how to restructure the electricity market, since low reserve margins that have been in operation (6.7% on average in 2010 for instance) have been threatening the stability and integrity of the electricity system. This dissertation thus attempts to address three important questions about Korean electricity demand-side market restructuring: (1) What are the estimates of the price elasticity of electricity demand in the wholesale and retail markets, including the residential, industrial, and commercial sectors? (2) How do inter-temporal price changes affect electricity consumption, and what are the estimates of the inter-temporal electricity cross-price elasticities in the wholesale market? (3) Except for the electricity price, what other factors affect electricity consumption in the wholesale and retail markets, including the residential, industrial, and commercial sectors?

In Chapter 2, I review current studies on electricity demand estimations, with the emphasis on price elasticity after the year 2000. Twenty papers (selected on the basis of the author's judgment) are summarized and evaluated, along with six papers that are discussed in relatively more detail. I also present evaluations and critiques of these works.

In Chapter 3, I briefly introduce the Korean electricity market and how it functions. In Chapter 4, I investigate the underlying features of the data in each market and sector and present these features both graphically and statistically.

In Chapter 5, I study the wholesale electricity market. Under the Real Time Pricing (RT-

P) structure, I discuss the model specification with respect to hourly consumption data with a consideration of aggregate utilization behaviors to control the complicated cyclical consumption patterns. Identification is established when the exclusion condition is not satisfied in the demand and supply system. The estimated real-time aggregate price elasticity, based on the whole sample, is -0.0034, the corresponding long-run price elasticity is -0.0640, and the estimated cross-price elasticities within the previous twenty-two hours are all negative, suggesting complementarity price effects. Price elasticities are also affected by the size of responsive customers. The effects of interruptible service operated by Korea Electric Power Corporation (KEPCO) and large buyers in the wholesale market with on-site generators on the demand curve are not detected based on a smooth transition model. Price elasticities with regard to each hour within a day are also estimated. Temperature and different types of the day also affect aggregate electricity consumption.

In Chapter 6, I study the retail electricity market, with a focus on the residential, industrial, and commercial sectors. Section 6.1 studies the residential sector. A basic regression model is built based on Ito (2012)'s finding that, contrary to the implications of conventional economic theory, households respond to the average electricity price rather than the marginal price when the tariff structure is increasing stepwise. I show that, on average, households respond to the previous month's average electricity price based on encompassing tests, which might be explained by the cognitive cost of a household obtaining the price information for the current monthly bill, as Ito (2012) implied. A structural time series model (STSM) with four different specifications is also applied to take account of the Underlying Energy Demand Trend (UEDT). The estimated aggregate price and income elasticities are around -0.2923 and 1.0388. Even though natural gas is a theoretical substitute for electricity, statistically, it does not affect electricity consumption. Other factors, such as temperature and holidays, have significant effects on electricity consumption. Moreover, the UEDT shows a steady decreasing usage trend, indicating, in the residential sector, that improved energy efficiency is the driving force of the UEDT.

Section 6.2 studies the industrial and commercial sectors. A simple theoretical analysis is first provided to model electricity demand for each pricing interval under the Time of Use (TOU) tariff structure. An absence of daily/monthly sector consumption data in different pricing intervals prohibited me from applying the theoretical model in practice. Instead, I take advantage of monthly aggregate data and model demand as monthly aggregate consumption against the monthly average price. This modeling compromise would introduce some bias into the price coefficients, for instance, by masking own- and cross-price effects in different pricing intervals. Except for the basic log-log specification, a seemingly unrelated regressions (SUR) model and an STSM, used to take account of the UEDT, are also applied. I find that firms in the industrial sector are responsive to electricity price variations, with the estimated price elasticity being around -0.19, but that firms in the commercial sector are not. Income elasticities in the commercial and industrial sectors are 1.7326 and 1.4585, respectively. Natural gas substitution elasticity is significant in the industrial sector with the basic and SUR models but this result is not robust to the STSM specification. Substitution effects are all insignificant in the commercial sector. Moreover, both sectors show an increasing UEDT trend. Further, once the UEDT is controlled, the estimated income elasticity becomes smaller (1.2483 in the commercial sector), indicating that part of the UEDT effects are confounded in the income coefficient when the UEDT is not specifically controlled. Other factors, such as temperature and holidays, have significant effects on electricity consumption.

Keywords: Electricity market, Demand, Price elasticity

To my parents, for their love and support.

Acknowledgments

This dissertation would not have been possible without the guidance of my committee members, help from my friends, and support from my family.

I would like to express my gratitude to my adviser, Professor In-Koo Cho, for his guidance, patience, encouragement, and providing me with the financial support and an excellent atmosphere for doing research. I would like to thank Professor George Deltas for his insight, thoughtful criticism, and time and attention during busy semesters. I would also like to thank Professor Anil Bera and Professor Seung-Hyun Hong, for their feedback and help. I am also thankful to Carol Banks, and other members of the department staff for their patience through this process.

Table of Contents

Chapter 1

Introduction

"Teach a parrot the terms supply and demand and you've got an economist." – Thomas Carlyle

Since 2006, the reserve margin in the Korean electricity market has been decreasing to a one-digit level; it was 6.7% in 2010 for instance. This has been threatening the stability and integrity of the electricity supply and demand system.[1] Korea, like many other countries, applies an inflexible electricity tariff design in the retail market; that is, the tariff does not adjust with hourly supply and demand and the price paid by retail customers thus does not reflect the true marginal cost of production. Economic theory indicates that this kind of arrangement will lead to market distortion and inefficiency. As Borenstein (2002) points out, the lack of demand response is the fundamental problem with the electricity market after the devastating 2000-01 California electricity crisis.[2] To mitigate the problem of a reduced

[1] A low reserve margin indicates possible electricity shortages, which might result in regular rolling blackouts. Blackouts have negative impacts on production and business operations, safety, and comfort at home and work, and other social activities. The reason for the low reserve margin in Korea is due to weather conditions, population increase, and economic activities. The optimal reserve margin generally depends on various factors: input prices, generation capacity, customer price response, load shapes, and system load factors. According to the Energy Vortex website, "regulatory bodies usually require producers and transmission facilities to maintain a constant reserve margin of 10–20% of normal capacity." A North American electric reliability corporation assigned a 15% reserve margin for its thermal system.

[2] After June 2000, California's wholesale electricity price skyrocketed, the wholesale price increased 800% from April 2000 to December 2000, and the price went back to normal in September 2001. During the crisis, the state suffered from large-scale rolling blackouts and one of the state's largest utilities, Pacific Gas & Electric, declared bankruptcy. The crisis cost $40 to $45 billion. The consensus view about the primary cause of this crisis is that it was the significant market power exerted by power plants. According to Wikipedia, installed generating capacity in California during the crisis was 45GW, while demand was only 28 GW at the time of the blackouts. Energy companies, mainly Enron, created this artificial discrepancy between demand and supply. Borenstein, Bushnell, and Wolak (2002) find that 21% of the price increase that occurred in the crisis was due to production costs, 20% due to competitive rents, and 59% due to market power. Other relevant work includes Ioskow (2001), Joskow and Kahn (2002), Wolk, Nordhaus, and Shapiro (2000), and Friedman (2009).

reserve margin, policy makers have begun to emphasise not only supply-side methods by increasing generation capacity, developing alternative energy sources, and other innovations, but also demand-side restructuring, by encouraging demand response and conservation through both pecuniary and non-pecuniary incentives,[3] to maintain system reliability. One important method of demand-side restructuring is to reform the tariff structure, such as by adopting real-time pricing (RTP) or critical-time pricing (CTP). This ideological shift generated heated debate on the demand-side electricity market restructuring, especially in relation to tariff structure reform, in Korea. Policy makers are concerned with how price changes impact on aggregate and individual customers. However, despite its importance, there is limited research available on electricity demand in Korea to enable an evaluation of the performance of different proposals, creating uncertainty about the reform and provoking controversy. This dissertation, therefore, attempts to shed some light on three important questions about Korean electricity demand-side restructuring: (1) What are the estimates of the price elasticity of electricity demand in the wholesale and retail markets, including the residential, industrial, and commercial sectors? This question identifies whether customers will respond to price changes and feeds directly into evaluating the conservation, welfare, and revenue impacts under alternative tariff structures. (2) How do inter-temporal price changes affect consumption patterns, and what are the estimates of the inter-temporal electricity cross-price elasticities in the wholesale market? This question concerns the shift of electricity consumption, which has important implications for the impacts of different tariff structures. For instance, the conservation goal of price reform will be difficult to realize if the aggregate cross-price effect causes reduced peak-time consumption shift to other hours, leaving total consumption, in the extreme case, unchanged. (3) Except for the electricity price, what other factors affect electricity consumption in the wholesale and retail markets,

[3]Public awareness affects electricity usage. As Joskow (2001) points out, the decline in electricity demand from February to May of 2001 in California was partly "due to the effects of all the publicity about electricity on consumer behavior and the formal energy conservation programs implemented by the state "since the retail price at that period did not rise significantly. The most effective approach to reduce system stress (rationing supply shortages) is by price signals.

including the residential, industrial, and commercial sectors? This question helps policy makers better understand and model electricity demand, and might also provide other available policy choices to reduce overall electricity load. To answer these three questions, I study the wholesale and sector-wise electricity demand in Korea based on hourly and monthly aggregate data to understand the demand side of the Korean electricity market. An absence of micro-level data is the limitation of my data-set, and prohibited me from investigating customer-level demand as in Taylor et.al. (2006), Allcott (2011), Patric and Wolak(2001), Resis and Wright (2004), and providing individual-level results as a counterpart to aggregate-level results. In general, both individual and aggregate level results are useful, since directly specifying aggregate elasticities from individuals, or vice versa, requires clear insight into individual behaviors and their interactions, which are hard to fathom. Therefore, my results should be interpreted carefully as aggregate effects, rather than effects for individual households.[4] Alternatively, this dissertation can also be viewed as an investigation that takes place when a researcher cannot access micro-level data, and reveals what is the best that he or she can interpret about the market.[5]

This dissertation speaks not only to policy issues surrounding the demand side of the electricity market, but also more broadly to the electricity demand literature on model specification, identification, and theoretical research on market evaluation,[6] as summarized in the following:

- Empirically, I show that it is deficient to model demand separately for each hour of the day without taking account of consumption influenced by the previous hour's

[4]Taylor (1975 p.104) points out that "the coefficients and elasticities must be interpreted for what they are, as representing effects for the aggregates involved, rather than effects for individual consumers."

[5]Griliches (1985 p.198) mentions possible responses to the concern about data qualities, as "we need to learn how to live with data limitations and adjust for their foibles. That is all there is – it is the only game in town and we have to make the best of it." Agrist and Pischke (2010 p.23) emphasize the importance of accumulating evidence, as "The process of accumulating empirical evidence is rarely sexy in the unfolding, but accumulation is the necessary road along which results become more general".

[6]Market evaluation results might be sensitive to the level of the assumed price elasticities, as the results in Green and Newbery (1992) for instance, and therefore a demand study could provide guidance for the choice of price elasticity.

consumption when the tariff structure is RTP.

- I establish the identification condition for the demand equation when no suitable exclusion condition is available in the demand and supply system.

- I investigate model specification with incorruptible services by taking advantage of the smooth transition model in the wholesale market.

- Empirically, I show that, in aggregate, households in residential sectors respond to the previous month's average electricity price.

- Empirically, I illustrate the Underlying Energy Demand Trend (UEDT) in different sectors in the retail market.

The study of electricity demand can be dated back to the 1960s, but the estimates are still vague and there has been no consensus about the appropriate estimation methods until recently, except for the agreement that the short-run demand response with respect to price is less elastic than the long-run and that both short- and long-run price elasticities are smaller than the other general products in absolute value. Furthermore, most research deals with quarterly or yearly data, and fewer study daily or hourly data.[7] Two aspects require special attention when specifying an electricity demand model. (1) Demand for electricity is time-oriented since demand for electricity is a derived demand — electricity is not used for its own sake. Consumers gain satisfaction through the services provided by electrical appliances. Electricity usage, therefore, depends on consumers' utilization behaviors and electrical appliance purchase decisions. (2) The costly storage property implies that fluctuations in electricity consumption are difficult to smooth through the use of inventories, like other commodities. From a time series point of view, consumption fluctuations comprise a variety of cyclical patterns: an hourly cycle within a day depending on different day types, a weekly cycle, and a seasonal cycle. Generally speaking, electricity consumption time series

[7] For example, Patric and Woolak (2001), Taylor et.al. (2006), Lijesen (2007), and Allcott (2011) adopt hourly data in their analysis.

data contain high frequencies, multiple seasonality, and calendar effects, although differences exist for data with more refined micro-level information or in different sectors.

In this dissertation, I first study the wholesale electricity market with hourly aggregate data. The pricing structure in this market is RTP. I discuss model specification and possible proxies for aggregate utilization behaviors. Without an exclusion condition for the demand equation, the identification condition is established with the help of a non-linear supply curve. Hourly data facilitate the study of the load curve as regards inter-temporal electricity cross-price effects. I show that an analysis in which demand is studied separately for each hour of the day does not take full advantage of the information and might suffer from the omitted-variable problem due the dynamic dependency of consumption and the correlation between consumption and price under RTP. The estimated aggregate real-time (hourly) price elasticity equals -0.0034, and long-run price elasticity equals -0.0640. The highest price elasticities during the day is around noon, and the lowest in the early hours of the morning, in absolute value. Furthermore, my results show that there is no strong evidence of an aggregate consumption shift induced by lagged prices. The price elasticities are heterogeneous with respect to the size of the responsive consumers. Temperature elasticities are 0.0153 and -0.0109 in the cooling and heating degree days. I also investigate model specification by taking into account the impact of KEPCO's interruptible services and large customers' on-site generators by taking advantage of a smooth transition model. The results show that the linear demand curve does not kink inward when prices are high.

I then investigate the retail market, since the fact that end consumers do not react to the wholesale market price does not mean that they are not price sensitive. In the residential sector, the tariff has six increasing steps. My basic regression model is built on a log-log framework with the assumption that households respond to average prices. The legitimacy of this assumption is based on Ito (2012). He provides empirical results to show that, when residential households face an increasing stepwise tariff, they respond to the average monthly cycle bill rather than the marginal price, which is contrary to the conventional economic theory. The theoretical foundation of his results is based on cognitive cost: the gain associated with knowing the exact marginal price for most consumers is very small, less than the cost invested to gather the relevant information. Consistent with Ito (2012), in my sample, I conclude that, in aggregate, consumers respond to the previous monthly bill and I do this by using an encompassing test based on Davidson and MacKinnon (1993). A possible explanation is that consumers generally are able to know their exact monthly bill at the end of the billing cycle without special effort, while they still need to struggle to keep track of their usage within the billing cycle to estimate their current average monthly price. I further generalize the basic model to an STSM to control the UEDT by characterizing the original constant intercept coefficient into a stochastic process. I apply four different structural specifications and my basic results are generally robust to these specifications. The aggregate price and income elasticities are around -0.2923 and 1.0388, respectively. The natural gas substitution effect is statistically insignificant once the UEDT is controlled. Temperature elasticity is around -0.32, and important holidays have positive effects on electricity consumption. Moreover, the UEDT shows a steady decreasing usage trend, indicating, in the residential sector, that improved energy efficiency is the driving force of the UEDT.

In the industrial and commercial sectors, the tariff structure is Time Of Use (TOU). I first build a heuristic theoretical model for electricity demand in each pricing interval following the procedure of Patric and Wollack (2001). An absence of daily/monthly sector consumption data in different pricing intervals prohibited me, however, from applying the theoretical model in practice. Instead, I take advantage of monthly aggregate data and model demand as monthly aggregate consumption against the monthly average price. This modeling compromise would be expected to introduce some bias into the price coefficients, for instance, by masking own- and cross-price effects in different pricing intervals. I find that firms in the industrial sector are responsive to electricity price variations, with the estimated price elasticity being around -0.19, but that firms in the commercial sector are not. Income elasticities in the commercial and industrial sectors are 1.7326 and 1.4585, respectively. The substitution elasticity of natural gas is around 0.1345 in the industrial sector, but there is no statistically significant effect of high-sulfur fuel oil. Both natural gas and high-sulfur fuel oil have insignificant effects in the commercial sector. Furthermore, holidays have negative effects on electricity usage in the industrial sector, but have mixed effects in the commercial sector. Moreover, firms in the commercial sector respond more to temperature changes than firms in the industrial sector. The corresponding temperature elasticities are -0.0915 and -0.0375 in cooling-degree days, and -0.1112 and -0.0442 in heating-degree days in the commercial and industrial sectors, respectively. Furthermore, the SUR model and STSM, used to take account of the UEDT, are also applied. Both sectors exhibit an increasing UEDT trend. Further, once the UEDT is controlled, the estimated income elasticity becomes smaller (1.2483 in the commercial sector), and natural gas substitution elasticity becomes insignificant in the industrial sector, indicating that part of the UEDT effects are confounded in the income coefficient in the commercial sector and in the natural-gas coefficient in the industrial sector when the UEDT is not specifically controlled in sectoral demand models.

Chapter 2

Literature Review

The study of demand for electricity can be dated back to the 1960s and the first major study is by Fisher and Kaysen (1962) who studied US electricity demand. Taylor (1975) surveys the econometric literature on the demand for electricity with the greatest focus on the residential sector. He discusses the proper way to model decreasing block tariffs so as to take account of the nonlinear budget constraints and suggests that researchers should include "both a marginal and an average price as predictors in the demand function."[1] He also makes a distinction between short- and long-run demand.[2] Based on the reviewed literature, he concludes that price elasticity of demand (in absolute value) is generally smaller in the short run than in the long run, that income elasticity of demand is detected but with mixed results in the long run, and that cross-price elasticity[3] of demand is subject to a lack of evidence in the short run and is vague in the long run. Taylor cautions that there may be possible estimation bias due to potential/possible modeling deficiencies and points out that "the coefficients and elasticities must be interpreted for what they are, as representing effects for the aggregates involved, rather than effects for individual consumers"[4] since all data employed in surveyed studies involve certain levels of aggregation.

Aigner and Poirier (1979) provide a succinct summary of the studies on time-of-day (TOD) electricity demand under a time-of-use pricing schedule. All empirical work included is conducted on DOE-sponsored pricing experiments. The analysis is mostly based on the

[1] Op. cit., 79.

[2] Taylor refers to short-run demand as the condition that the stock of electricity-consuming appliances remains fixed, while long-run demand takes the stock as variable.

[3] Substitute goods for electricity include fuel, natural gas, oil, etc.

[4] Op. cit., 104.

residential sector with an assessment of the related economic theory and econometric methods. They conclude the existence of a short-run price response qualitatively but are cautious about interpreting and accepting the results quantitatively because of the inconsistent results generated by different authors. Bohi (1981) studies elasticity for different forms of energy such as electricity, natural gas, petroleum, and coal. He discusses methodologies for energy demand estimation and surveys the literature based on energy type. Bohi categorizes studies of electricity demand according to whether they study residential, commercial, and industrial demand and further categorizes them according to modeling techniques used. He points out that estimated results generally vary with the type of model and the type of data. He also evaluates the estimation procedures at great length to gain insight into modeling techniques and recommends careful interpretation of the results' magnitude in the surveyed literature.

With the development of econometric techniques and theoretical scope, more insights are gained from estimation of the demand for electricity. Thus, in the following parts, I attempt to evaluate and analyze empirical works on electricity demand, especially literature published after 2000. A selection of papers which are representative of the literature (based on the author's judgment) are included. First, in relatively more detail, I summarize seven studies variously studying the industrial, commercial, and residential sectors, and the wholesale market, and then present summary results of 14 papers in Table 2.9.

2.1 Reviews

2.1.1 Patrick and Wolak (PW 2001) and Taylor, Schwarz and Cochell (TSC 2005)

Both articles apply the same methodology to investigate own-price elasticity and intertemporal substitutions in electricity consumption within the day based on the real-time pricing (RTP) industrial customer level data. The data set employed in PW is from the English and Welsh electricity market (April 1, 1991 to March 31, 1995) and in TSC from the Duke Energy Hourly Pricing program (1994 to 2001 summer months (Jun. to Sep.)).

PW assume that the customer minimizes their expected variable production costs for the next day subject to certain constraints. Cost minimization is due to the assumed pre-determined output level that the customer must meet in real production. They treat each load-period[5] electricity consumption within the day as different commodities and thereby model them as inputs into the customer's daily variable cost function. The cost function is assumed to take the Generalized MacFadden (GM) form due to its second-order flexibility with the global curvature restrictions imposed by economic theory and possible existence of both substitution and complementary electricity effects among load periods. The demand equation, as a consequence, is derived by Shepard's lemma.[6]

Electricity demand E_{id} in PW for each customer in load period i and day d is given by,

$$E_{id} = [\frac{1}{PZ_d} \sum_{j=1}^{48} c_{ij} PE_{jd} + b_{it}] Y_d + a^*_{it} + \sum_{k=1}^{2} k_{ij} X_{jd} + d'_i W_{id} + u_{id} \tag{2.1}$$

Where t refers to fiscal year t; PZ_d is the materials and fuels composite price index for day d; PE_{jd} is the price for electricity; X_{jd} is the customer's capital stock holdings and the level of employment, which is constant within each fiscal year t; W_{id} is a vector of the average hourly temperature in degrees centigrade and sun intensity; u_{id} is the error term, representing the unobservable portion of the customer's conditional variable cost function, and $u_d = \{u_{1d}, ..., u_{48d}\}$ is the random vector with mean 0 and covariance matrix Ω, which is distributed independently across days. Let $a_{it} = a^*_{it} + \sum_{j=1} 2k_{ij} X_{jd}$ be the fiscal year t constant term for each customer.

TSC adopt the same demand model as PW, but they take it a step further to account

[5]In the English and Welsh electricity market, there are 48 load periods (half-hour intervals); In the US electricity market, there are 24 load periods (one-hour intervals).

[6]Patric and Wolak (2001) have a detailed discussion of GM, Generalized Leontief, and translog cost functional forms and GM specifications.

for time-series correlation and heteroscedasticity of the error term by modeling u_d as an AR(m)-EGARCH(p,g) process.

Matrix $[c_{ij}]$ controls globally concavity, which is positive definite. Other time-varying coefficients, a, b, and d, are modeled as Fourier series with basic daily frequency in PW and basic frequency of the length of the yearly sample periods in TSC.[7]

Since output Y_d in both PW and TSC are assumed to be pre-determined by a customer's planning schedule and to fluctuate based on some periodic process, it thus can be modeled by certain proxies. PW mention two ways to proxy the output. One way is to utilize dummy variables:

$$Y_d = \mathrm{DAY} + \mathrm{MONTH} + \mathrm{YEAR} \tag{2.2}$$

where DAY represents the day-of-the-week indicator variables, MONTH represents the month-of-the-fiscal-year indicator variables, and YEAR is a set of fiscal year indicator variables.

The other way is use a frequency series which is applied to the estimation model:

$$Y_d = (1 + \sum_{j=1}^{Nw} w_j \cos(j\tau d) + w_{j+Nw} \sin(j\tau d) + h_d(j))\lambda_t \tag{2.3}$$

where $\tau = \frac{2\pi}{366}$ is the basic yearly frequency (leap year); $h_d(j)$ is a dummy variable taking a value 1 if d is a national holiday and 0 otherwise; λ_t is dummy variable for fiscal year t ($t = 1, ..., 4$); Nw is the number of harmonious frequencies; and 1 is the normalized constant term because Y_d is not identified separately from its coefficient.

TSC model output in the same way as PW do at first with elaboration as,

$$Y_{dyk} = DOW_{dyk} * Week_{dyk} * Holiday_{dyk} \tag{2.4}$$

[7]Please refer to PW and TSC for details of coefficient specification.

where

$$\begin{aligned} DOW &= 1 + a_1 * \text{Sunday} + a_2 * \text{Monday} + ... + a_6 * \text{Saturday} \\ Week &= 1 + b_1 * \text{Week}_1 + ... + b_{(x)}\text{Week}_{(x)} \\ Holiday &= 1 + c_1 * \text{July 4th} + c_2 * \text{Labor day} \end{aligned}$$

where y is the year index and k is the customer's index. a, b, and c are restricted to be less than 1 to guarantee that they have a positive value.

In PW, identification is justified by the market price-setting process in England and Wales. TSC do not, however, mention identification conditions.

PW further simplify the computational complexities by grouping customers in the same Business Industry Classification (BIC) code based on the assumption that they have the same parameters in their demand model. The grouped data version is given by,

$$\bar{E}_{id} = \frac{1}{M}\sum_{k=1}^{M} E_{idk} = [\frac{1}{PZ_d}\sum_{j=1}^{48} c_{ij}PE_{jd} + b_{id}]Y_d + a_{it} + d_i' W_{id} + \frac{1}{M}\sum_{k=1}^{M} u_{idk} \tag{2.5}$$

where Y_d is the average daily output of M firms within a certain BIC code.

Both PW and TSC estimate their models by year. Nonlinear SUR estimation techniques are adopted due to cross-equation restrictions for the load-period equation model in PW equation 2.1 and in TSC. Equation 2.5 is estimated by quasi-maximum likelihood using a 48-dimensional multivariate normal density as the objective function for a single observation in PW.

Table 2.1 presents the implied own- and cross-price demand elasticities for five BIC codes in PW.

Compared with PW, TSC in general detect larger price responses. Based on average price elasticity of hours 14-21, Table 2.2[8] shows the top ten price response individual customers

[8]Op. cit., 245.

in TSC. Among them, six customers have their own self-generators and two (steel and electrodes) have arc-furnaces, which allow discrete production processes. Dependence of grid generation and flexibility of production are factors affecting price responses. Furthermore, seven out of 51 customers' average absolute value price elasticities exceed 0.2 of hours 14-21 and the largest hourly value is 0.775. Aggregate elasticities demonstrate the importance of generators and arc-furnaces in facilitating customers' price responses in Table. 2.3[9].

For intra-day cross-price elasticities, TSC show general complementarity in adjacent hours in the afternoon, but substitution relationships during other hours. They also find that "customers with generators appear to demonstrate more complementary behavior in the afternoon when prices are high implying that there is limited flexibility to adjust generation on an hour to hour basis."[10] A requirement for block hour decisions precludes the potential for substitution. TSC argue that "in PW, the steel tubes industry shows complementarity in adjacent hours, which is hypothesized to be consistent with little ability to shut down production. In our study, we find complementarity for customers with self-generation. We attribute that finding to the fact that it only makes sense to use self-generation for at least several consecutive hours."[11]

2.1.2 Boisvert, Cappers, Goldman, Neenan, and Hopper (BCGNH 2006)

This study focuses on intensity of electricity price responsiveness within a day by industrial customers, including manufacturing, education/government, commercial/retail, health care, and public works consumers. Data are from the Niagara Mohawk Power Corporation (NMPC) encomprising 119 RTP customers in the summer months in 2000 to 2004.

BCGNH group hourly electricity demand by peak and off-peak periods of the day and

[9] Op. cit., 248.
[10] Op. cit, 249.
[11] Op. cit., 249.

model them as two inputs in the customer's daily production function, which is assumed to be separable among electricity and other inputs. The cost function is assumed to take the form of Generalized Leontif (GL) as the following,

$$C = E(d_{pp}P_p^{1/2}P_o^{1/2} + d_{po}P_p^{1/2}P_o^{1/2} + d_{op}P_o^{1/2}P_p^{1/2} + d_{oo}P_o^{1/2}P_o^{1/2}) \tag{2.6}$$

where E is the effective electricity input depending on peak and off-peak period electricity usage, K_p and K_o (kWh), respectively; P_p and P_o (\$/kWh) are average prices for peak and off-peak periods; and the restriction $d_{po} = d_{op}$ is imposed.

Since BCGNH investigate two electricity commodities that function as substitutes, they do not need to accommodate both positive and negative elasticities as in PW and TSC.

By Shepard's lemma, electricity demand for peak and off-peak electricity is given by,[12]

$$K_p = \frac{\partial C}{\partial P_p} = E(d_{pp} + d_{po}(\frac{P_o}{P_p})^{1/2}) \tag{2.7}$$
$$K_o = \frac{\partial C}{\partial P_o} = E(d_{oo} + d_{op}(\frac{P_p}{P_o})^{1/2}) \tag{2.8}$$

Accordingly, the Hicks-Allen partial elasticity of substitution (the percentage change in factor intensities as the inverse price ratio changes by one per cent, holding effective electricity constant) is derived as,

$$\sigma_{op} = \frac{Cd_{op}(P_pP_o)^{-1/2}}{2Ea_pa_o} \tag{2.9}$$
$$\sigma_{pp} = -\frac{Cd_{po}(P_o^{1/2}P_p^{-3/2})}{2Ea_p^2} \tag{2.10}$$
$$\sigma_{oo} = -\frac{Cd_{op}(P_p^{1/2}P_o^{-3/2})}{2Ea_o^2} \tag{2.11}$$

where $a_p = \frac{K_p}{E}$, $a_o = \frac{K_o}{E}$ are input-output demand equations.

Certain restrictions on d_{pp}, d_{op}, and d_{oo} are needed for a well-behaved GL function.

[12]In the paper, there is a typographical error in the peak and off-peak demand equations on p.14.

BCGNH list three cases of substitution effects. (1) Peak and off-peak electricity are substitute inputs, $\sigma_{op} > 0$, when d_{pp}, d_{op}, and d_{oo} are non-negative. (2) Peak and off-peak electricity are perfect complements, used in fixed proportions, $\sigma_{op} = 0$, when $d_{op} = 0$ ($d_{pp} \geq 0$ and $d_{oo} \geq 0$ is required). (3) Peak electricity demand are substitutes up to a threshold P_p^*/P_o^* at which point demand from the grid is zero when $d_{pp} < 0$, $d_{op} > 0$, and $d_{oo} > 0$. The corresponding condition is $P_o/P_p < d_{pp}^2/d_{oo}^2$.

Since E is unobservable, the estimation model is given by the ratio of two input-output demand equations:

$$\ln(\frac{a_p}{a_o}) = \ln(\frac{k_p}{k_o}) = \ln\{\frac{d_{pp} + d_{po}(\frac{P_o}{P_p})^{1/2}}{d_{oo} + d_{op}(\frac{P_p}{P_o})^{1/2}}\} \tag{2.12}$$

The estimation model is further elaborated for customer f and weekday observation index t as,

$$\begin{aligned} \ln(K_{p,t,f}/K_{o,t,f}) = \quad & w_f W_{t,f} + \ln[h_{p,h} H_{t,f} + d_{pp,f} + d_{po,f}(P_{o,t,f}/P_{p,t,f})^{1/2}] \\ & - \ln[h_{o,f} H_{t,f} + d_{oo,f} + d_{op,f}(P_{p,t,f}/P_{o,t,f})^{1/2}] + \varepsilon_{t,f} \end{aligned}$$

Where $W_{t,f}$ is the intercept shifter, measuring cooling degrees (the difference between the average peak period temperature and 65 degrees F); $H_{t,f}$ is a dummy variable taking the value 1 when the average temperature heating index (THI) constructed from temperature and dew point values is greater than or equal to 85 during peak periods and 0 otherwise; and $\varepsilon_{t,f}$ is a random error. Restrictions are $d_{po,f} = d_{op,f}$, $h_{p,f} = h_{o,f}$. BCGNH normalize coefficients by $d_{oo,f} + d_{pp,f} + d_{op,f} + d_{po,f} = 1$ to reflect a unit isoquant for effective electricity.

GCGHN point out that "determination of peak and off-peak periods is an empirical question driven by prices and the circumstances by which customers use and value electricity."[13] They compare three alternative peak-period definitions (12-5pm, 1-5pm, and 2-5pm) and select the one with the highest elasticity of substitution, which is 2-5pm.

Estimation is done for each customer by using the full information maximum likelihood

[13] Op. cit., 11.

(FIML) method.

The results show that the price response is heterogeneous for customers even with the same business classification. The average price elasticity of substitution is 0.11. In general, industrial customers are the most responsive to price, followed by government/education, and commercial/retail customers. On the other hand, the latter two groups are more responsive to price than industrial customers on hot days and when the peak-to-off-peak price ratio is high. For all customers, there is no evidence to support the following two hypotheses: (1) price responsiveness is reduced when operations are at, or near, their peak electricity usage; and (2) the price response is primarily driven by customer size. In detail, "about one-quarter of customers appear to use peak and off-peak electricity in fixed proportions; the remaining firms have positive elasticities of substitution, indicating that peak and off-peak electricity are substitute inputs. A large group of customers are only slightly responsive (elasticities of substitution of less than 0.05). About 20% are moderately responsive, with elasticities of substitution between 0.05 and 0.10. Finally, the 10% of customers in the high and very high groups (elasticities of substitution greater than 0.10) provide 75-80% of the sample customers' aggregate demand response."[14] In the high-response group (four government/education and ones commercial/retail customer), BCGNH identify a threshold price ratio range from 7:1 to over 100:1 that customers curtail their grid-supplied peak electricity usage.

Since customer-specific characteristic information is limited in their data set, they point out that "It is only through knowing something about the specific customer characteristics or circumstances that one can understand why certain customers in a particular business class are price responsive and others are not."[15]

[14]Op. cit., 23.
[15]Op. cit., 29.

2.1.3 Reiss and White (RW 2005) and Ito (2012)

Both RW and Ito, with different data sets, investigate the demand response with respect to the electricity tariff, which is increasing stepwise, of Californian residential households. The key disparity between the two studies is the assumption of household rationality. In RW, households are assumed to be perfectly rational, while, in Ito, households are assumed to be characterized by bounded rationality. It matters whether an assumption of bounded rationality is made or not because it impacts upon the household's responses to perceptions of marginal or average prices and thus impacts the credibility of the study.

RW (2005)

Data are obtained from the Residential Energy Consumption Survey (RECS) of the US Department of Energy, which contains detailed information on household appliance holdings and dwelling characteristics, annual electricity consumption and expenditure, and weather information from 1993 to 1997 and supplies this information in respect of 1307 Californian households. Price data are gathered by matching rate schedules to each observation in the survey.

During the study periods, the electricity price schedule is a two-tier increasing block tariff. In standard theory, electricity consumption decisions of utility-maximizing households depend on the entire tariff structure rather than average price or any single marginal price in which perfect household rationality is required.

In order to explicitly model household electricity demand with a nonlinear budget constraint, RW define $x(p, y, z, \varepsilon; \beta)$ to be the ordinary demand with constant price p, household income y, observed consumer characteristics (demographic and dwelling structure characteristics, appliance attributes, and weather information) z, unobservable characteristics ε, and parameter sets β. The budget constraint is linearized at the optimal consumption level. $x(\cdot)$ is assumed to be strictly increasing in ε and strictly decreasing in p. A reduced form of household demand is presented by utilizing $x(\cdot)$,

$$x^* = \begin{cases} x(p_1, y, z, \varepsilon; \beta) & \varepsilon < c_1 \\ \bar{x} & c_1 < \varepsilon < c_2 \\ x(p_2, y_2, z, \varepsilon; \beta) & \varepsilon > c_2 \end{cases} \tag{2.13}$$

where $y_2 = y + \bar{x}(p_2 - p_1)$ and c_j is the solution to $x(p_j, y_j, z, c_j; \beta) = \bar{x}$ with $y_1 = y$. $c_1 < c_2$ provided that income effects are quite small.

Assuming ε is distributed as $N(0, \sigma^2)$ and enters equation (2.13) additively, the expression for x^* can be simplified as,

$$E(x^*|\cdot) = [x(p_1, y, z; \beta) - \sigma\lambda_1]\Phi_1 + \bar{x}(\Phi_2 - \Phi_1) + [x(p_2, y_2, z; \beta) + \sigma\lambda_2](1 - \Phi_2) \tag{2.14}$$

where Φ_j is the standard normal distribution evaluated at $c_j(\beta)/\sigma$, ϕ_j is the normal density at $c_j(\beta)/\sigma$, and $\lambda_1 = \phi_1/\Phi_1$, $\lambda_2 = \phi_2/(1 - \Phi_2)$.

RW point out that the mis-match between annual aggregate household consumption data and the monthly electricity bill might cloud the true household price response. This is because the aggregate data confound not only the price effect but also the effects of other factors, such as the weather. To clarify the direct price effect, RW rely on the monthly observable variables to control the effects of other factors. Let $x_t(w_t)$ represent consumption at month t conditioned on controls w_t and annual electricity expected demand is thereby expressed as,

$$E(x^a|w_1, w_2, ...w_12) = \sum_{t=1}^{12} E(x_t^*|w_t) \tag{2.15}$$

Corresponding ordinary demand $x_t(w_t)$ is then specified as the sum of the demand for eight different appliance categories (baseline electricity use (appliances that are universally owned); electric space heating; central air conditioning; room air conditioning; electric water heating; swimming pool; additional refrigerators and freezers; and other appliances).

$$x_t = \sum_k d_{k,t}\alpha_k p + \sum_k d_{k,t}\gamma_k y + \sum_k d_{k,t} z_k' \delta_k + \sum_k d_k \varepsilon_k \tag{2.16}$$

where k refers to the number of distinct appliance categories.

Combining equations (2.14) to (2.16) forms the basic estimation model with an implicit assumption that the household does not make appliance purchasing decisions. Three problems are addressed by RW's modeling techniques: (1) identification is established due to the increasing stepwise tariff with constant-price segments; (2) different price and income effects for each category of appliances are accounted by utilizing micro-level appliance stock information in the model; (3) the price effect is identified from annual aggregate consumption by controlling monthly varying variables.

RW apply the generalized method of moments (GMM) procedure by matching first and second moments to estimate the model.

Table 2.4[16] presents the results of estimated marginal effects for different categories of appliances. Price effects are heterogeneous across appliances and most income effects are not significant at the 1% significant level due to the fact that the analysis is conditional on households' appliance stocks and that income effects generally occur through households' choices of appliances rather than through utilization behavior. The mean annual electricity price elasticity for the households in the data is around -0.39. Average annual household price and income elasticities[17] focused on the demand sensitivity in the margin based on different appliance household groups are shown in Table 2.5.[18] RW also apply the OLS method using household average price to estimate the model as a way to substantiate their modeling techniques. Most price elasticities in the OLS method are positive due to the failure to address the simultaneity problem.

[16] Op. cit., 867.

[17] Elasticities should be interpreted as short-run based on Taylor (1975)'s definition, since RW implicitly assume constant household holdings of appliance stocks.

[18] Op. cit., 868.

RW also estimate the price elasticities by different household groups based on household annual income and annual electricity consumption. Conventional wisdom is confirmed that households with medium-to-high incomes are less sensitive to electricity prices than households with lower incomes, although the differences are small. Furthermore, price elasticities are lower for households with high amounts of electricity usage. This appears to conflict with the results that electricity-intensive appliances, like electric space-heating/cooling systems, have greater price sensitivity *ceteris paribus*. RW argue that "this inverse relationship reflects both a weak correlation between household income and ownership of electric space-heating/cooling systems, and the fact that households tend to substitute toward more price-inelastic electricity uses as income rises."[19]

Ito (2012)

The research question of Ito (2012) is to uncover which perceived price consumers respond to under nonlinear price schedules in the Californian electricity market. Standard theory predicts consumers respond to the marginal price under a non-linear budget set. This requires two implicit assumptions: (1) consumers are engaged with no uncertainty; and (2) they are fully aware and understand the nonlinear pricing structure. Borenstein (2009) finds that customers are less likely to respond to their marginal price with uncertainty, but is inconclusive about whether they respond to the expected marginal price or the average price. Intuitively, if the cognitive cost of understanding the nonlinear price schedule is substantial, consumers may respond to the average price rather than the marginal price.[20] I mainly focus on the regression analysis part in the paper.

Data are obtained from Southern California Edison (SCE) and San Diego Gas & Electric (SDG&E) in California on a panel data set of household-level monthly electricity billing

[19] Op. cit., 871.

[20] In the 2010 version of this study, Ito calculates a consumer's utility gain from responding to marginal not average price by assuming a simple quasi-linear utility function and constant price elasticity equal to -0.5. The result is about $2 on average, which is about 3% of the average consumer's electricity bill.

records from 1999 to 2007. Two utility companies adjust their price schedule multiple times independently, providing substantial spatial and temporal price variation in the sample period.

The research design belongs to the category of Regression Discontinuity Design by exploiting a spatial discontinuity in electricity utility service areas. In detail, the study focuses "on the territory border of SCE and SDG&E in Orange County because this is the only border in populated areas that also does not correspond to city or county boundaries."[21] Ito substantiates the balance of consumers' demographics and housing characteristics across the border by matching census data. Thus, the application of consumers on the other side of the utility border as a control group is legitimate.

The regression model is as follows,

$$\Delta \ln x_{it} = \beta_1 \Delta \ln mp_{it} + \beta_2 \Delta \ln ap_{it} + f_t(x_{it_m}) + \gamma_{ct} + \delta_{bt} + u_{it} \tag{2.17}$$

where $\Delta \ln x_{it} = \ln x_{it} - \ln x_{it_0}$, in which it_0 is the previous year's corresponding billing month, as for $\Delta \ln mp_{it}$ and $\Delta \ln ap_{it}$, γ_{ct} is the city-by-time fixed effects, δ_{bt} represents the billing-cycle-by-time fixed effects, $t_m = t - 6$ is the middle period between t and t_0, and x_{it_m} is the corresponding consumption level.

First difference eliminates household-by-month fixed effects since it is impractical to add consumer-by-time controls. The price instruments are $\Delta \ln mp_{it}^{PI} = \ln mp_t(x_{it_m}) - \ln mp_{t_0}(x_{it_m})$ and $\Delta \ln ap_{it}^{PI} = \ln ap_t(x_{it_m}) - \ln ap_{t_0}(x_{it_m})$. To make the instrument valid, that is, the error term u_{it} is uncorrelated with x_{it_m}, Ito uses $f_t(x_{it_m})$ to control the confounding factors such as underlying distributional changes in consumption. He specifies $f_t(x_{it_m})$ by non-parametric controls: defining group dummy variables by $G_{j,t} = 1\{x_{j,t_m} < x_{it_m} < x_{j+1,t_m}$ for each percentile of consumption in t_m, in which $G_{j,t}$ equals to 1 if x_{j,t_m} falls between j and $j+1$ percentiles.

[21]Op. cit., 7.

The model does not suffer multicollinearity problems due to the fact that utilities change each of the five tier rates differently over time, which breaks down the typical high correlation between marginal prices and average prices.

Consumers' price responses can be tested in the model. The null for consumers responding to marginal price is $H_0 : \beta_2 = 0$, while the null for responding to average price is $H_0 : \beta_1 = 0$. The testing method belongs to encompassing tests. Table 2.6 presents the estimation results of model (2.17). Columns 1 to 3 show that "once the average is included, adding marginal price does not statistically change the effect of average price, and the effect of marginal price becomes statistically insignificant from zero."[22] Columns 4 to 6 provide the results with one-month lagged price, since consumers receive electricity bills at the end of a monthly billing period. The main results still hold with this modification. Table 2.7 investigates the point at which consumers respond more to the lagged price than the contemporaneous price. Column 1 shows that once the lagged prices are controlled for, the effect of contemporaneous price is statistically insignificant from zero. Columns 2-5 present the medium-long-run price elasticity by regressing with the average lagged price for one to four months instead. Ito then examines the possibility of consumers responding to the *expected* marginal price by the same method, as listed in Table 2.8. Still, it is substantiated that consumers respond to the average price. Note that the price elasticity estimated in this study is smaller than in RW.

2.1.4 Allcott (2011)

Allcott studies residential household electricity demand based on household level data obtaining from the energy-smart pricing plan operated by the Center for Neighborhood Technology (CNT) in Chicago since 2003. Household participants in the program are randomly assigned to control and treatment groups by CNT. Data properties are tested by Allcott and the two groups in the experiments are ascertained to be well balanced.

[22]Op. cit., 18.

Allcott defines an *ad hoc* indirect utility function $V(\mathbf{P}, w_i)$ with respect to household wealth, w_i, and electricity price vector, $\mathbf{P}$,

$$V_i(\mathbf{P}, w_i) = w_i - \sum_d (\sum_{h=1}^{24} P_{hd}(\frac{1}{2}\eta_{ih}^D P_{hd} + (\eta_{is}^A - \eta_{ih}^D)\bar{P}_{hs} + \eta_{ig}^{HP} HP_d + \eta_{is}^T T_i + \varepsilon_{ihd})) \quad (2.18)$$

where $\bar{P}_{hs}$ (cents/kWh) is the average price for hour h in season (summer and non-summer) s; P_{hd} is the hourly electricity price; HP_d is the high price alert day indicator function; T_i takes a value of 1 if the household is in the treatment group, and 0 otherwise; g denotes a group of hours: early morning (6-10am), morning (10am-2pm), afternoon (2-6pm), and evening (6-10pm); and ε_{hd} is the unobservable demand shock in hour h of day d.

By Roy's identity, household i's electricity demand function, q_{ihd} (W), is derived as,

$$q_{ihd} = \eta_s^A \bar{P}_{hs} + \eta_h^D (P_{hd} - \bar{P}_{hs}) + \eta_g^{HP} HP_d + \eta_s^T T_i + \varepsilon_{ihd} \quad (2.19)$$

This model captures three types of demand response to time-varying prices: (1) the response to the average seasonal hourly price shape, captured by η_s^A; (2) the real-time response when the instantaneous hourly price deviates from the average seasonal price, captured by η_h^D; and (3) the response to High Price Alert Days, captured by η_g^{HP}.

Allcott points out that the interpretation of η_h^D should incorporate intra-day substitution price effects since related parameters are not included in the model due to the high correlation among hourly prices.

Demand parameters are estimated by pooling across all hours of the experiments in May to December, 2003 (weekend effects are not estimated separately due to the lack of data). The model is set in the average treatment effect (ATE) framework based on the potential outcome approach developed by Rubin (1974, 1978) which views causal effects as comparisons of potential outcomes defined on the same units to deal with the simultaneity problem.[23]

[23] Simultaneity is caused by the fact that the unobserved demand factor ε affects the equilibrium market

In expectation, control and treatment groups have the same unobserved demand factor, ε, due to the randomness of the experiment, and therefore, any difference in demand can be attributed to the price effect on demand.

The ATE is given by,

$$\tau_{hd} = E[\eta_s^A(\bar{P}_{hs} - \bar{P}_s^{T=0}) + \eta_h^D(P_{hd} - \bar{P}_{hs}) + \eta_g^{HP}HP_d + \eta_s^T|h, d] \tag{2.20}$$

Elaborating the model by including control variables, the estimating equation is,

$$\begin{aligned} q_{ihd} =& \{\eta_s^A(\bar{P}_{hs} - \bar{P}_s^{T=0}) + \eta_h^D(P_{hd} - \bar{P}_{hd}) + \eta_g^{HP}HP_d + \eta_s^T\}T_i \\ &+ \{\alpha_i(\bar{P}_{hs} - \bar{P}_s^{T=0}) + \alpha_2(P_{hd} - \bar{P}_{hd}) + \alpha_3 HP_d^{aft} + \alpha_4 HP_d + \alpha_5\}X_i + \zeta_{hd} + \epsilon_{ihd} \end{aligned}$$

where X is the set of controls, including pre-program average hourly consumption (monthly electricity bills from May in 2002 are used to construct pre-program baseline electricity consumption), household size, and log(income); HP_d^{adf} is the dummy for an afternoon hour of a High Price Alert Day; ζ_{hd} is the fixed effect for hour h of day d; and ϵ_{ihd} is the error term.

Estimation applies the standard fixed effect method. One potential limitation of the model is that only the correlation between average treatment effects and prices is estimated since there are no observations of the household behaviors that underlie the treatment effects or when those behaviors occurred. Allcott argues, “it does not identify the frequency at which behavioral changes occurred, i.e. whether the responses were day-to-day, short term adjustments to air conditioner settings or long-term adjustments to thermostats each season. It similarly does not identify whether the effects were produced by long-run changes to energy-using capital stock versus short-run changes to the usage of the capital stock.”[24]

The estimation results show that most price marginal coefficients are statistically signif-

price by shifting the aggregate market-level demand curve.

[24] Op. cit., 832.

icant. The average price coefficients, $\hat{\eta}_s^A$, are -17.4 and -21.8 W/(cents/kWh) for summer and non-summer seasons, respectively and the deviation coefficients η_h^D are on average -12 W/(cents/kWh). The overall price elasticity of demand is about -0.1.

No statistical evidence supports net load shifting from high to low price hours. Households, therefore, mainly react to high prices with conservation. Allcott is cautious to extend this conclusion to say there are no substitution effects of household consumption across hours since no substitution parameters are included in the model. He therefore only concludes that "*on net*, RTP causes households to significantly reduce consumption on the average afternoon and does not cause significant increases in average consumption at night."[25]

Allcott also admits that it is not appropriate to extrapolate his results to the general population since the households in the RTP experiment may not be representative due to self-selection issues.

2.1.5 Fezzi and Bunn (FB 2010)

FB study the hourly aggregate day-ahead data from the Pennsylvania, New Jersey, and Maryland (PJM) wholesale electricity market from Apr. 1, 2002 to Aug. 30, 2003 and point out that "most research on measuring the demand response to price has focused upon micro survey data, at company or individual level, but despite clearly giving behavior insights, it does not help directly in specifying aggregate elasticities for spot and forward trading markets."[26]

FB adopt the structural asymmetric vector error correction modeling technique and incorporate three types of demand responses to price changes: instantaneous responses; equilibrium[27] (long-run) responses; and responses through an error-correction mechanism.

[25] Op. cit., 833.

[26] Op. cit., 829.

[27] Since the analysis employs high-frequency (daily) data, the equilibrium refers to the condition that the system will converge over days or weeks, in the absence of perturbations. Therefore, the long-run (different from Taylor (1975)'s definition) corresponds to a short-period window, instead of the situation when electricity-usage appliances are taken as variables.

Demand and supply function specification is given by,

$$\text{Demand:} \qquad Q(P, a, b, v_d) = P^{\lambda_1}\phi(a, b)v_d \tag{2.21}$$

$$\text{Inverse Supply:} \quad P(Q, P_f, C, t_r, v_s) = \beta_0 Q^{\beta_1} P_f^{\beta_2} C^{\beta_3}\eta(t_r)v_s \tag{2.22}$$

where Q is electricity quantity, P is the electricity price, $\phi(\cdot)$ is a generic function taking account of the nonlinear relationship between demand and weather conditions a and aggregate usage behavior (work patterns and living habits) b, λ_1 is the price elasticity of demand, P_f is the fuel price, C is the excess capacity available in the system, t_r is the system operation technical rationale, $\eta(\cdot)$ is a nonlinear function capturing the effect of technical constraints, and v_d and v_s are random errors.

$\phi(a, b)$ is modeled as,

$$\phi(a, b) = \lambda_0(\lambda_f^{d_f} \prod \lambda_i^{d_i})(\lambda_t^{d_t} \exp[\lambda_c t_{cold} + \lambda_t t_{hot}]) \tag{2.23}$$

where d_f is a dummy variable of national holidays, d_i are dummies for each workday of the week (Monday to Friday), t_{cold} (t_{hot}) is equal to the temperature when the temperature is lower (higher) than the given threshold and 0 otherwise, and d_t is a dummy variable set to 1 when the temperature is above the threshold and 0 otherwise.

FB take the log of both sides of the demand and supply functions and specify them in an asymmetric error correction form by allowing for the possibilities of asymmetries in demand response. Using the smaller case to represent variables' logarithm form for Q, P, P_f, and C,

$$\begin{aligned}\Delta q_t =& \theta_{21}\Delta p_t + \delta_{21}\Delta t_{cold,t} + \delta_{22}\Delta t_{hot,t} + \alpha^{+}_{Sq} I_{St}\beta_s w_{S,t-1} + \alpha^{-}_{Sq}(1 - I_{St})\beta_S w_{S,t-1} \\ &+ \alpha^{+}_{Dq} I_{Dt}\beta_D w_{D,t-1} + \alpha^{-}_{Dq}(1 - I_{Dt})\beta_D w_{D,t-1} + \varphi d_{j,t} + \sum_{i=1}^{k} \varsigma'_i \Delta w_{D,t-i} + u_{D,t} \\ \Delta p_t =& \theta_{11}\Delta q_t + \delta_{11}\Delta c_t + \delta_{12}\Delta p_{f,t} + \alpha^{+}_{sp} I_{St}\beta_s w_{s,t-1} + \alpha^{-}_{Sp}(1 - I_{St})\beta_s w_{s,t-1} \\ &+ \phi d_{j,t} + \sum_{i=1}^{k} \vartheta'_i \Delta w_{S,t-i} + u_{St}\end{aligned}$$

where $d_{j,t}$ is a vector of dummy variables containing $d_{i,t}$ and $d_{f,t}$; $w_d = [p_t, q_t, t_{hot}, t_{cold}, d_t, 1]$ and $w_s = [p_t, q_t, c, p_f, 1]$ are vectors containing variables influencing supply and demand respectively; β_D and β_S are vectors of coefficients for w_d and w_s in the demand and supply long-run co-integration equations; I_{St} (I_{Dt}) is the indicator variable for positive disequilibrium in supply (demand) side, and equals 1 if $\beta_S w_{S,t-1} > 0$ ($\beta_D w_{D,t-1} > 0$) and 0 otherwise; and k is the number of lags to eliminate serial correlation in the error terms.

Regression models are for each hour of the day[28] and thus the basic data frequency is daily. Weekend data are excluded from the analysis for two reasons: (1) the electricity demand profile is different on weekends, and so are the demand-supply interactions; (2) many energy markets are closed during the weekends. Two hours, 7pm and midnight, which represent peak and off-peak demand behaviors, are selected for analysis.

Unit roots are detected in the demand, price, and temperature series and model estimation follows Johansan procedures and applies full information maximum likelihood (FIML) for short-run error correction models. The results indicate that both the short- and long-run price elasticities of demand are insignificant, which can be considered as vertical. However, demand response to past positive disequilibrium in the long-run supply function disequilibrium (price is higher than in the equilibrium) through error-correction term, α^{+}_{Sq} equals -0.085 and -0.037 for 19pm and 24pm with 5% significance level respectively, but not to past

[28]FB implicitly assume the hourly electricity demand market is independent.

negative supply disequilibrium. FB argue that "the demand function, even though vertical, does adjust to past high prices shifting to the left according to the adjustment coefficients."[29]

2.1.6 Summary of Other Works

Tables 2.9 and 2.10 present summaries of other empirical works. The square bracket means the value is not significant at the 10% significance level.

2.2 Evaluation and Critique

Firstly, based on the above studies, electricity price elasticity of demand in the short run is generally inelastic, while it is more responsive in the long run, but is still quantitatively inconclusive, even for the same sector and country. For example, Karmerschen and Porter (2004), Dergiades and Tsoulfidis (2008), and Alberini and Filippini (2011) all adopt US residential sector data at the aggregate level. KP and DT find long-run elasticity, in its absolute value, is approximately unitary, but AF show a much smaller elasticity. Price elasticity is also demographic-, sector-, and business-specific; for example, the Swiss residential sector (Fillippini, 2011) and the Turkish industrial sector (Dilaver and Hunt, 2011) have the highest and lowest long-run elasticities respectively in its absolute value among included studies, and price elasticities are heterogeneous among different BIC code industries (Patric and Wolak, 2001 and Taylor, et. al., 2005). Compared to aggregate low-frequency data, high-frequency and micro-level data generally suffer less vulnerability to measurement error and thus convey finer stories about demand behavior. Price has intra-day substitution and complementarity effects — customers exhibit electricity-usage conservation during peak hours when prices are high, and may shift demand to off-peak hours when prices are low. Some studies investigate price elasticity determination to a certain extent. Price elasticity can be dependent upon

[29] Op. cit., 848.

weather, time (peak/off-peak), and technology[30] in general, and for residential households, dependent upon income and usage-level. Studies on treatment and control groups in the experiments of RTP and CPP substantiate the conclusion that customers' demand-response with respect to price can be stimulated with proper pricing policies and facilitating technologies[31] However, their effects are possibly customer- and period-dependent. For example, the steel industry, as in Patric and Wolak (2001), has restricted the ability to adjust production when facing high prices, and Reiss and Wright (2007) investigate the question whether customers can learn to be conservative in energy usage.

Secondly, I discuss some aspects of model specification and estimation. (1) Hourly electricity demand is possibly not independent, especially for the wholesale market and the industrial and commercial sectors. Customers may make electricity-appliance usage decisions for block periods and hourly-utilization behaviors could be correlated. Note that even if the price and weather conditions are the same, the same electricity usage would not occur as it is influenced by its own recent history. Treating hourly electricity independently, for instance, Fezzi and Bunn (2010), or modeling demand without controlling past hour utilization behaviors might overlook the relevant information contained in previous hours, and these effects may confound in the model included interested price variables and thus bias the estimates. Moreover, if the errors are autocorrelated due to this omission, dynamic modeling will lead to more bias. Wolak and Patrick (2001) and Taylor et. al. (2005) adopt the same modeling techniques except that Taylor et. al. (2005) control the unobserved error terms by an AR-EGARCH process. With better utilization of previous hours' information, Taylor et. al. (2006) estimate higher price elasticities in absolute value. This comparison, however, requires an assumption that the differences in customers and time periods in these two papers are negligible. (2) There needs to be more investigation as to whether modeling

[30] Technology refers to equipment, like on-site generators, that facilitates electricity demand adjustment on grid and that reduce the cost to customers of acquiring knowledge of real-time prices.

[31] Costa and Kahn (2010) show that electricity conservation "nudges" promote conservation, but depend heterogeneously on ideology.

should be linear or nonlinear and whether variables should be expressed in log form or otherwise. Whether log forms of variables are used or not generally depends on the preference of researchers. Both methods distort price elasticity to some extent. In a linear level model, when prices are low, price elasticity might be over-estimated when prices are low, and when prices are high, they might be under-estimated. Zarnikau (2003) and Xiao, Zarnikau, and Damien (2006) discuss the model selection based on statistical methods but their results are still inconclusive. (3) There is a possible multicollinearity issue with high frequency data in regressions. Some demand models include hourly lag and/or leading prices in level in the model under RTP. Hourly prices, however, are generally highly correlated. High correlation might distort the estimates and result in low credibility. In Wolak and Patric (2001) and Taylor et. al. (2005), both regressions are conducted in this way without mentioning the possible multicollinearity problems. (4) In the residential sector with a nonlinear pricing system, Ito (2012) empirically examines, using rich, micro-level data, to what price households respond: average price, expected marginal price, or marginal price. In contrast to standard economic theory, he concludes that households respond to average price possibly due to cognitive cost. Reiss and Wright (2005), however, present problematic estimates by using simple OLS with average price used as the basis of comparison for the results of OLS with marginal price. That said, however, RW's results cannot be used to contradict Ito's. This is because the use of simple OLS in RW neglects the endogeneity problem and thus the estimates of the price effect are biased, which is something acknowledged in RW. Since RW does not provide the IV results with average price, the results based on marginal price are not well substantiated. In all, the potential specification or estimation flaws should not be reason to dismiss the results. They are all attempts to uncover the truth of the reality, but researchers need to approach all research results with a skeptical mind and an appreciation of the limitation of each study. Also as Granger (1997) points out, "the actual economy appears to be very complicated, partly because it is the aggregation of millions of non-identical, non-independent decision-making units ,.., the modeling objective has thus to be limited to providing an adequate or satisfactory approximation to the true DGP. Hopefully, as modelling technology improves and data increases, better approximation will be achieved, but actual convergence to the truth is highly unlikely."[32]

Also, there are still many veils obscuring customers' demand responses with respect to price. Herter (2007), unlike econometric empirical demand studies, adopts exploratory analysis making use of descriptive statistics and graphics, and discovers that high-use customers respond significantly more in terms of electricity usage reduction than do low-use customers, while low-use customers save significantly more in terms of percentage reductions in annual electricity bills than do high-use customers based on the data from the California Statewide Pricing Pilot of 2003-04 with CPP. These results contradict those of Reiss and Wright (2005) and challenge the price targeting strategies with respect to high-use customers. To fully understand customers' demand responses, it is necessary to employ a finer data set and researchers need to conduct in-depth investigations. It will also be interesting for researchers and policy makers to know how variations in household composition, such as the number of occupants, their ages, or other characteristics in the residential sector, industrial/businese type, size, or available technologies in the industrial and commercial sectors, and geographic, socio-demographic, or cultural[33] variations influence price effects.[34]

[32]Op. cit., p.169.

[33]Luzenhiser (1992) supported the idea that cultural forms (lifestyles) affect patterns of energy consumption, pointing out that individuals make choices, including energy consumption choices, based on cultural contexts.

[34]These factors provide useful information for tariff design, energy-efficient policy subsidies, and so on.

2.3 Tables

Table 2.1: PW results summary

BIC	Own-price elasticity	Cross-price elasticity
17000 (Water supply)	Most price-responsive industry among five analyzed industries. Morning and early evening have relatively high responses, and the largest occurs at 9:30-10am around -0.27. 2:30-6pm is uniformly small around -0.02.	Most of the possibility of substitutions detected in adjacent load periods.
22200 (Steeltube)	Generally very small (in absolute value) in the range of (-0.002,0). The largest occurs during the late afternoon (5-6pm) around -0.008.	Consistently negative cross-price effects (complementary effects) during early morning and in the afternoon of adjacent load periods. Slightly positive (substitution effect) after 5pm for about one and half hours.
22460 (Copper, brass alloys)	Relatively large around noon, early afternoon, and early evening (6-8pm) around -0.05, other hours around -0.02.	Most of the substitution possibilities are around early evening across adjacent hours, and with a lower amount early in the daylight periods.
24890 (Ceramic goods)	Relatively constant and small (in absolute value) and in the range of (-0.05, 0) for most hours, largest occurs around 5:30 pm with the value around -0.013	Slightly negative of the close adjacent hours early in the working day, major substitution possibilities in adjacent hours occur in the late afternoon and early evening.

Table2.1(cont.)

BIC	Own-price elasticity	Cross-price elasticity
31600 (hand tools&metal goods)	Relatively constant and small (in absolute value) and in the range of (-0.005, 0) for most hours, largest occurs 4-6pm around -0.025	Relatively large substitution possibilities of adjacent hours during early morning and early evening.

Table 2.2: Summary own-price elasticities for each customer

Industry	Generator	Hour 14 to 21	Hour 14	Hour 18	Hour 20
University	Y	-0.376^{***}	-0.349^{***}	-0.519^{***}	-0.369^{***}
Textiles	N	-0.353^{*}	-0.405^{*}	-0.339^{*}	-0.334^{*}
Paper	Y	-0.352^{***}	-0.775^{***}	-0.340^{***}	-0.119^{***}
Textiles	Y	-0.333^{***}	-0.332^{***}	-0.353^{***}	-0.313^{***}
Glass tempering	N	-0.316^{***}	-0.069^{***}	-0.513^{***}	-0.390^{***}
Steel	N	-0.299^{***}	-0.426^{***}	-0.302^{***}	-0.206^{***}
Electrodes	N	-0.208^{***}	-0.355^{***}	-0.215^{***}	-0.138^{***}
Laboratory	Y	-0.168^{***}	-0.111^{***}	-0.225^{***}	-0.156^{***}
University	Y	-0.163^{***}	-0.104^{***}	-0.270^{***}	-0.177^{***}
Chemical fiber	Y	-0.106^{***}	-0.110^{***}	-0.120^{***}	-0.080^{***}

Table 2.3: Aggregate own-price elasticities[a]

Group	Complete group	Generator	Arc furnace	No generator or arc furnace
All customers	-0.155 (51)	-0.269 (10)	-0.269 (2)	-0.029 (39)
Textiles	-0.212 (23)	-0.333 (1)	n.a. (0)	-0.038

[a] The number in the bracket represents the number of customers in that group.

Table 2.4: Estimated marginal effects[a]

Explanatory variable	Effect on kWh consumed per month for: Baseline use	Elec. space heating[b]	Central air cond.[c]	Room air cond.	Elec. water heating	Swimming pool
Price(cents/kWh)	0.4	−37.8	−22.5	−63.4	-34.0	-27.5
	(3.7)	(14.8)	(21.3)	(31.1)	(9.5)	(18.4)
Income ($1000)	0.4	16.2	9.1	21.6	-32.8	6.3
	(2.3)	(13.0)	(10.6)	(20.8)	(7.5)	(9.8)

[a] Asymptotic standard errors in parentheses.
[b] Heating-season months only.
[c] Cooling-season months only.

Table 2.5: Price and income elasticities for Californian households[a]

	Price		Income	
Mean elasticities of electricity demand	GMM method	OLS method	GMM method	OLS method
All households	-0.39	0.16	-0.00	+0.00
Households with:				
Electric space heating	-1.02	-0.46	-0.00	0.01
No electric space heating	-0.20	0.35	-0.00	-0.00
Central or room air conditioning	-0.64	0.05	0.02	0.02
No air conditioning	-0.20	0.24	-0.01	-0.01
No electric space heating nor air conditioning	-0.08	0.39	-0.01	-0.01

[a] Price elasticity measures annual electricity consumption resulting from a 1% increase in marginal price in each month of the year, holding appliance stocks fixed.

Table 2.6: Marginal price vs. average price

	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta \ln(\text{Marginal Price}_t)$	-0.040		-0.005			
	(0.004)		(0.011)			
$\Delta \ln(\text{Average Price}_t)$		-0.055	-0.061			
		(0.005)	(0.014)			
$\Delta \ln(\text{Average Price}_{t-1})$				-0.052		0.003
				(0.004)		(0.012)
$\Delta \ln(\text{Average Price}_{t-1})$					-0.075	-0.079
					(0.005)	(0.014)

Table 2.7: Lagged responses and medium-long-run price elasticity

		Medium-long-run responses			
		1 month	2 month	3 month	4 month
	(1)	(2)	(3)	(4)	(5)
$\Delta \ln(\text{Average Price}_t)$	0.002				
	(0.006)				
$\Delta \ln(\text{Average Price}_{t-1})$	-0.045				
	(0.008)				
$\Delta \ln(\text{Average Price}_{t-2})$	-0.036				
	(0.007)				
$\Delta \ln(\text{Average Price}_{t-3})$	-0.013				
	(0.006)				
Δ ln(Average of Lagged		-0.075	-0.093	-0.099	-0.101
Average Prices)		(0.005)	(0.005)	(0.005)	(0.005)

Table 2.8: Expected marginal price vs. average price

	(1)	(2)	(3)	(4)
$\Delta \ln(\text{Expected Marginal Price}_t)$	-0.043	-0.009		
	(0.004)	(0.011)		
$\Delta \ln(\text{Average Price}_t)$		-0.057		
		(0.014)		
$\Delta \ln(\text{Expected Marginal Price}_{t-1})$			-0.055	0.005
			(0.004)	(0.011)
$\Delta \ln(\text{Average Price}_{t-1})$				-0.083
				(0.014)

Table 2.9: Summary

Author	Data type	Methodology	Sector	Results	
				Elasticity	
				SR	LR
Kamerschen and Porter (2004)	Annual aggregate (1973-1998 US)	Simultaneous structure model, 3SLS	Residential Industrial Aggregate		-0.9325 [-0.3399] [-0.1348]
Narayan and Smyth (2005)	Annual aggregate (1969-2000 Australia)	Bounds testing procedure to cointegration, ECM	Residential	-0.2631	-0.5409
Lijesen (2007)	9am-6pm hourly (peak) load (2003 Netherlands)	Static Model 2SLS	Aggregate	(-0.0014,0.0043)	
Dergiades and Tsoulfidis (2008)	Annual aggregate (1965-2006 US)	Bounds testing procedure to cointegration, ECM	Residential	-0.386	-1.0652
Sa'ad (2009)	Annual aggregate (1973-2007 Korea)	Structural time-series model (STSM)	Residential		-0.5
Jamil and Ahmad (2011)	Annual Aggregate (Pakistan)	Johansen cointegration and VECM	Aggregate Residential Commercial Manufacturing Agricultural	[-0.07] [0.07] [-0.31] [-0.06] [-0.19]	-1.27 -1.22 -1.68 -1.22 -0.51
Alberini and Filippini (2011)	Annual aggregate (1995-2007 48 US states)	Kiviet corrected LSDV and the Blundell-Bond GMM	Residential	(-0.15,-0.08)	(-0.43,-0.73)
Dilaver and Hunt (2011)	Annual aggregate (1960-2008 Turkey)	Structural time-series model (STSM)	Industrial		-0.1609
Lee and Chiu (2011)	Annual aggregate (1978-2004 24 OECD countries)	Panel smooth transition regression model	Aggregate		-0.229
Filippini (2011)	Annual city aggregate (2000-2006 22 cities Swiss)	LSDV, Kiviet corrected LSDV	Residential	(peak) -0.778 (off-peak) -0.652	(peak) -2.266 (off-peak) -1.652

Table 2.10: Summary

Author	Data type	Methodology	Sector	Results
Boisvert et al. (2004)	Peak/ off-peak daily customer-level aggregate (1998-2001 summer CSW customers)	Generalized Leontief (GL) cost specification	Industrial and commercial customers	Substitution elasticity ranges from 0.10 to 0.18
Wolak (2006)	CPP houly household-level (Jun. 1, 2005 to Oct. 14, 2005, Anaheim)	DID, nonparametric condition mean estimation framework	Residential	On average, households in the treatment group consumed 12% less electricity during the peak hours of CPP days than customers in the control group
Chio, Sen, and White (2011)	RTP houly aggregate (2005-2008 Ontario)	Non-linear SURE based on Generalized Leontief (GL) cost specification	Industrial	Substitution elasticity (peak/off-peak load) is 0.02 to 0.07
Faruqui and Sergici (2011)	RTP hourly household-level (2008, 2009 summer, SEP, BGE)	CES demand model, fixed effect panel	Residential	Substation elasticity (peak/off-peak load) is weather- and technology-dependent, and the value based on average weather is -0.096. Price elasticity of daily aggregate data is -0.039, which is weather-dependent.

Chapter 3

Korean Electricity Market Brief

The Korean electricity market restructuring began in April, 2001. Prior to the reform, Korean Electric Power Corporation (KEPCO), as a state monopoly, controlled electricity generation, transmission/distribution, and retail. The original restructuring plan included three phases:

- Generation competition: divesting the generation sector from KEPCO by splitting KEPCO into six generating companies (Gencos) to initiate competition in the wholesale market. Private generating companies are allowed to enter the market.
- Wholesale competition: divesting the distribution sector from KEPCO to introduce competition to the distribution sector and the transmission system serves as a common carrier.
- Retail competition: privatizing distribution.

The phase 2 and 3 reforms were put on a hold in 2004 due to the economic recession, the changing political environment,[1] and the devastating effect of the California electricity crisis. KEPCO still functions as the main utility in the market and its retail tariffs are subject to government supervision.

In the first phase, Korea Power Exchange (KPX) was created by the government as an independent system operator to take charge of the wholesale market. KPX operated a cost-based pool (CBP) to determine real-time wholesale equilibrium prices, equating electricity

[1] President Roh came into power in 2003 and his government has progressively backed away from the market reforms initiated by his predecessor, Kim Dae-jung.

demand and supply on an hourly basis.[2] The pool selling price (PSP) equals the sum of system marginal price (SMP),[3] capacity price (CP), and others including constraints and ancillary service costs. SMP accounted for a large portion of the PSP; for example, in 2008, SMP accounted for of 70% of the PSP on average. The unique characteristic of this market is that the market price is set by actual variable cost, as opposed to the common method used in other countries' electricity markets using the bidding price. In operation, SMP is published at 3pm by KPX in the trading day ahead, and the PSP is known to the wholesale buyers in the settlement statements after the trading day.[4]

There is no regional trading market since the power system is operated on a single national grid. All generators over 1 MW capacity are required to participate in the wholesale market for trading. District electricity businesses, the utility KEPCO, and large customers (above 500MVA) are allowed to enter the wholesale market as buyers. From 2006 to 2010, the percentages of the total electricity amount traded in the wholesale market were 97.2%, 96.9%, 96.9%, 97.7%, and 96.9%. The rest of the electricity load[5] is traded by pre-determined power purchase agreements (PPA). Since the vast majority of the volume of the electricity load is traded in the wholesale market, the aggregate electricity load of the whole economy is a good representation of the load in the wholesale market.

District electricity businesses arose from the private sector, which built gas-turbine generators in urban residential complexes and competed with KEPCO to sell power to those residential complexes. In total, 22 district suppliers were approved across 28 districts from October 2004 to July 2008, and the total capacity approval reached 1512.25MV. In 2008,

[2]Generators' costs are evaluated by the Generation Cost Assessment Committee (GCAC). The GCAC assesses each generator's variable cost on a monthly basis and fixed cost on a yearly basis.

[3]SMP is determined by a Price Setting Schedule (PSS) system, which, after considering forecast demand, cost per power-generator, bidding information, and availability, calculates SMP in order to meet power system demand at minimal cost.

[4]The preliminary settlement is issued within nine business days after the trading day, and the final settlement is within 22 business days by KPX.

[5]The definition of electricity load refers to the sum of the electricity requirements of all consumers on a network or the amount of electricity consumed by some particular consumers. In this dissertation, I will use electricity load and consumption interchangeably.

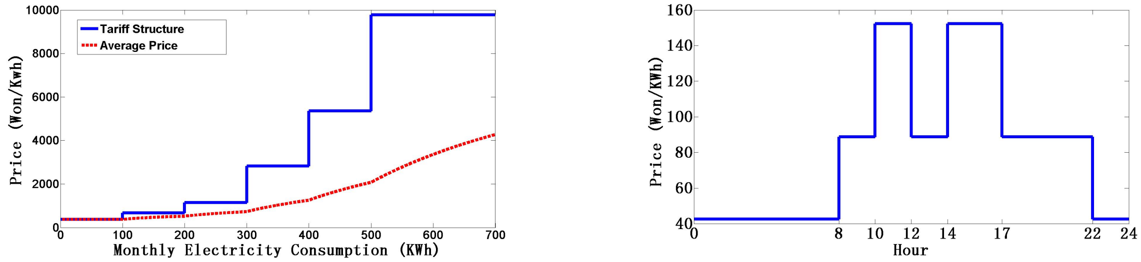
however, rising oil prices eroded their profit, discouraging further activities. Suppliers gradually withdrew from the districts from 2008 onwards. District electricity businesses can also join the wholesale market as buyers and sell their leftover power to KEPCO at the market price.

KEPCO, as a national utility, has a legal obligation to provide electricity to its retail customers. There are six sectors in the retail market: residential, commercial, industrial, educational, agricultural, and street lighting.[6] From 2006 to 2010, based on monthly data, the average percentages of each sector's consumption were 20.14%, 22.5%, 52.8%, and 4.56% for the other three sectors in KEPCO. Electricity purchased by KEPCO accounts for about 90% of the total consumption on average. Although KEPCO can relieve system stress by mandating its large customers to reduce their electricity usage (direct load control) when the price in the wholesale market is high, its greatest price response is restricted by the unresponsive retail customers. The retail prices set by KEPCO are predetermined, and thus are not directly linked to the fluctuations in the wholesale market price on an hourly basis. Specifically, in the residential sector, the retail tariff is set as increasing stepwise, such that the marginal price is a step function of the consumption in each billing cycle, as shown in Figure 3.1 (a). In the industrial and commercial sectors, the tariff is Time Of Use (TOU)[7], which is designed for three pricing intervals within the day and is differentiated seasonally. Table 3.1 lists pricing periods for off-, mid-, and on-peak hours, and Figure 3.1 (b) presents a typical tariff structure in a summer day. A major price difference exists with peak rates, where on-peak rates can be 403% higher than off-peak rates in a sampling period. KEPCO offers several different TOU pricing options for industrial and commercial customers.

[6]KEPCO on its website defines industrial service as applying to "customers using electricity for mining, manufacturing, gas production and supply, water supply defined by the Water Supply and Waterworks Installation Act; and electric railroads," and commercial service as applying to customers excluding residential, educational, industrial, agricultural, and street lighting customers.

[7]The term "time of use" (TOU) denotes that prices vary depending upon when the electricity is used.

3.1 Figures and Tables



(a) (b)

Figure 3.1: (a) Residential tariff structure in 2006; (b) A typical industrial TOU structure in summer in 2006. Source: KEPCO.

Table 3.1: TOU pricing periods

Classification	Spring, summer, and fall[a]	Winter[b]
Off-peak load	23pm ∼ 9am	23pm ∼ 9am
	9am ∼ 11am	9am ∼ 10am
Mid-peak load	12pm ∼ 13pm	12pm ∼ 17pm
	17pm ∼ 23pm	20pm ∼ 22pm
	11am ∼ 12pm	10am ∼ 12pm
On-peak load	13pm ∼ 17pm	17pm ∼ 20pm
		22pm ∼ 23pm

Source: KEPCO.
[a] Spring is from Mar. 1st to Jun. 30th, summer from Jul. 1st to Aug. 31st, and fall from Sep. 1st to Oct. 31st.
[b] Winter is from Nov. 1st to Feb. 28th.

Chapter 4

Preliminary Data Analysis

In this chapter, I investigate the underlying features of data both graphically and statistically. I first present graphical illustrations of the time-series process of the data, the autocorrelation function (ACF), and spectral plots in Section 4.1. Then, in Section 4.2, I apply statistical tests to check whether the series contain unit roots.

4.1 Graphical Description

I first analyze aggregate hourly consumption in Figures 4.1 to 4.6. There are predictable patterns of the load curve[1] across days, weeks, and years.

Figure 4.1 provides some insights into the electricity consumption. All load curves are approximated by a step function defined over the hours in a day. Sub-figure (a) illustrates daily average consumption. A small positive trend is detectable as well as yearly and seasonal cycles. Sub-figure (b) displays hourly consumption in 2007. Weekly cycles of electricity consumption become much clearer. Sub-figures (c) and (d) present hourly consumption at 3am (off-peak load hour) and 12pm (peak load hour) respectively. The magnitude of the 12pm load is higher than the 3am load in general and the 12pm load shows more variations, indicating heterogeneity in hourly consumption.

Figure 4.2 contains plots of different averages of electricity consumption. Sub-figure (a) presents monthly average consumption. Consumption is higher in winter and summer because of the need for heating and cooling services provided by electricity. Sub-figures (b)

[1] A load curve is a chart showing how much electricity customers utilize during a given period of time.

to (f) present average hourly consumption in general and respectively for four seasons. The daily load cycle becomes apparent. The lowest electricity consumption is at around 4am and the highest at 12pm except in summer when the highest is at 3pm. The load curves are generally similar for the four seasons except in the evening hours of 8pm to midnight when consumption is flatter in spring and fall. In summer, with the cooling-down temperature at night, air conditioning usage is generally reduced, which causes electricity usage to decrease, while on the contrary, in winter, cold nights require more heating, causing electricity usage to increase.

Figure 4.3 provides further insights into hourly average consumption on different types of days. Sub-figures (a) and (b) present hourly average consumption on holidays and weekends. I categorize the important holidays based on the criterion that the aggregate electricity daily load curve has the distinct feature that the usage is decreasing in the morning and remains low until the afternoon. Table 4.1 lists the holidays in terms of whether they are important or normal holidays. The load curve of Sunday consumption has a similar shape as that of important holidays. This is probably because people tend to wake up late on these days and most firms also stop operating. The daily load curves of normal holidays and Saturdays have similar shapes as those of general days, but with a lower magnitude during the day. Sub-figure (c) plots the average hourly consumption on Mondays and the other work days. Consumption on early Monday morning is lower than that on other days and it converges to the same level of the other days during the daytime. Sub-figure (d) illustrates the non-linear relationship between electricity consumption and temperature. The break point of the curve is around 65° F.

Figure 4.4 plots the autocorrelation function for the hourly load series both in level and in first difference. The level series are highly correlated and will not die out in long periods of time, while autocorrelation is greatly reduced by first difference. Also the plots show a strong daily pattern in both level and first-difference data. Figure 4.5 provides a summation of hourly electricity consumption cycles by spectral plots. Sub-figures (a) and

(b) are the periodograms[2] of hourly and first-difference hourly consumption for all sample periods. In sub-figure (a), low frequency cycles, for instance seasonal, yearly, etc., are clearly visible as well as daily and half-day cycles, indicating that low and high frequencies are both important determinants of the sample variance of hourly electricity demand. In sub-figure (b), low frequency cycles are overshadowed by daily and its harmonious high frequency cycles in first-difference series. Sub-figures (c) and (d) are the periodograms of hourly and first-difference hourly consumption at 12pm, which illustrate the same results as in sub-figures (a) and (b) and the weekly cycle is also observable. The first-difference series eliminates seasonal and yearly cycles and its variance is mainly from weekly, daily, and within-day fluctuations.

For hourly SMP, Figure 4.6 plots the hourly average price during a day, which has a similar shape as the hourly average load plot in Figure 4.2. SMP is more volatile than consumption: the maximum ratio of the highest to lowest SMP within a day is 7.4989 with an average of 2.1869, while the maximum ratio of the highest to lowest daily load within a day is 1.8291 with an average of 1.2906. Two possible reasons might explain the high volatility in SMP: one is that not all customers respond to SMP in the wholesale market; most customers are engaged in the retail market, which means that demand in the wholesale market is quite inelastic to SMP; and the other reason is that SMP can increase rapidly when supply is constrained (hourly prices are high when capacity is expected to be tight).

Next, I present the graphical figures in the residential, industrial and commercial sectors in the retail market. In Figure 4.7, sub-figures (a) and (b) illustrate the time-series pattern of monthly consumption, and sub-figures (c) and (d) plot the monthly average real price in these three sectors. All consumption series contain a visible upward time trend. The

[2] A periodogram is defined as,

$$I(w_j) = N^{-1}|\sum_{t=1}^{N} x_t e^{-2\pi i w_j t}|^2, \quad \text{j=0,1,...,N-1}$$

where $w_j = j/N$. If the data contain some strong periodic components, then the periodogram values corresponding to those frequencies (or near those frequencies) will be large.

residential consumption series contain a clear yearly cyclical pattern. Consumption is quite volatile within a year compared to the consumption series in the other sectors, and this variation seems to be increasing over time. The highest consumption during a year is in winter, probably due to electrical heating usage. For industrial consumption series, there seems to be no clear cyclical patterns. For commercial consumption series, the highest consumption within a year is in summer before 2007, while the consumption in winter caught up to summer consumption after 2007. The yearly cycle is also visible in the plot. The real price plot in sub-figure (c) in the residential sector has a similar shape as the plot of the consumption series. In the industrial and commercial sector, the highest real average price during a year is in summer, since TOU prices are highest in summer. Moreover, the real price in the commercial sector actually contains a decreasing trend before 2001.

In Figure 4.8, sub-figures (a) and (b) plot the series of average household income and Gross Domestric Product (GDP) in real terms. They both contain a steady increasing trend, but the household income series are more volatile than the GDP series. In Figure 4.10b, sub-figures (a) and (b) plot the natural-gas real price for industrial and residential customers, and the high-sulfur fuel oil real price for industrial customers in real terms. All price series have an increasing time trend. The natural-gas prices for industrial customers are cheaper than the prices for households. Figure 4.10 presents the scatter plot of the monthly consumption against average temperature in these three sectors. In the residential sector, there seems to be a non-linear relationship, with the breaking point around 75° F. Below 75° F, electricity consumption generally decreases when the temperature increases, while above it, consumption begins to change in the same direction. In the industrial and commercial sectors, however, there are no clear patterns between electricity consumption and temperature.

Figures 4.11 and 4.12 present the autocorrelation functions and spectral plots for all three sectors. In residential sector, both the ACF and spectral plot in the residential sector show a clear yearly cyclical pattern. In the industrial and commercial sectors, both ACF plots die out after about a year. There are also spikes in the spectral plot in the commercial sector corresponding to yearly and its harmonious cycles and the half-year cycle is the strongest. No yearly cycle is detected in the industrial sector.

4.2 Unit-Root Test

In this section, I test whether the time-series data contain a unit root based on two statistical tests: the KPSS stationarity test by Kwiattowski et al. (1992) and the Dickey-Fuller-Generalized Least Squares (DF-GLS) unit root test by Elliott, Rothenberg, and Stock (1996). Tables 4.2 to 4.3 list results for hourly series, i.e., series of electricity consumption, system marginal price, and temperature in the wholesale market, and monthly series in the residential, industrial, and commercial sectors respectively.

In Table 4.2, for both the consumption and SMP series, the KPSS test can reject the stationarity null hypothesis at the 1% significance level, and the DF-GLS test cannot reject the unit root null hypothesis at the 10% level. For the temperature series, the results are weak in the sense that the KPSS test can reject the stationarity null at the 5% level and the DF-GLS test cannot reject the unit root null hypothesis at the 1% level, but not at the 5% level. For all three first-difference series, the results favor stationarity, i.e., the KPSS test cannot reject the stationarity null at the 10% level and the DF-GLS test can reject the unit root null at the 1% level. In Tables 4.3 and 4.4, for all monthly series in the residential, industrial and commercial sectors, once the yearly effects are removed, the KPSS test cannot reject the stationary null at the 10% level, and the DF-GLS test can reject the unit root null at the 1% level. In summary, I can conclude that the unit root hypothesis is rejected in favor of the stationarity hypothesis for the monthly and quarterly series and first-difference hourly and daily series.

With a little more contemplation, the results for the consumption series in different observation frequencies seem to suggest that the series become stationary by temporal aggregation,[3] which apparently is in contradiction with the statistical theory. A unit root is a characteristic of a stochastic process, and thus should not be affected by observation frequency, which indicates that accepting the unit root null hypothesis in high-frequency series would imply that a lower-frequency series would conform to a similar process. In empirical practice, however, some studies notice similar consequences as mine in relation to temporal aggregation.[4] Chen and Tran (1994) argue that this temporal aggregation issue with unit roots may arise when the series "have a long mean-reversion component that the conventional unit root tests are not powerful enough to detect in a short span sample." Fujihara and Mongoue (1994) mention an explanation by Sims that "high frequency data may contain a stationary autocorrelation coefficient close to one, making it difficult for classical test procedures to discriminate between a unit root and a large but stationary autocorrelation coefficient." They suggest readers should be cautious in their acceptance of a unit root. Pierse and Snell (1995) explored the relationship between power and sampling frequency, and show that, based on a Monte Carlo simulation, a lower frequency increases the power of the unit root test, and only small-span increases are generally required to maintain power when reducing sampling frequency. Therefore, empirically, the unit root results for lower frequency data are generally more trustworthy.

In my data set, hourly data are all highly autocorrelated and will not die out over long periods of time, as shown in Figure 4.4, and thus the tests may have limited capacity to distinguish them from a unit root process as Sims implied. Further, since temporal aggre-

[3]I apply the two tests for the monthly consumption series in the wholesale market, and the results also favor stationarity.

[4]For example, Pierse and Snell (1993) re-estimate the consumer behavior model in Molana (1991), and the null hypothesis that consumption and wealth are not cointegrated is convincingly rejected in the annual data, although it cannot be rejected by Molana with quarterly data. Chen and Tran (1994) investigate the process of the real exchange rate. They show that "the standard augmented Dickey-Fuller tests can not reject the null hypothesis of nonstationarity for all the monthly data, but reject it for all the annual data except the real exchange rate of Pound," while the revised KPSS tests "suggest that we can only reject in the monthly data the null hypothesis of stationarity for yen. And we can not reject the null hypothesis of stationarity for the annual data." Fujihara and Mougoue (1994) point out that the unit root test may produce conflicting evidence that "high frequency spot exchange rate (e.g. daily and weekly) appear to possess unit roots while low frequency (e.g. monthly and quarterly) spot rate may not."

gation would smooth out the series and reduce autocorrelation, my results for the monthly series thus are more convincing. On the other hand, however, according to Blough (1992), "given the low power of unit root test, an acceptance of the null that a variable is difference-stationary compared to stationary could be taken as an indication of a spurious regression rather than as firm evidence of difference stationarity itself." Since spurious regression invalidates standard statistical inference, it is important to guarantee appropriate inferential procedures in model specification. On this point, despite limited interpretation based on the tests, there is strong evidence that all first-differenced hourly series are stationary, so modeling the high-frequency series in first difference to avoid spurious regression seems to be more correct.

Summary

I aware that, in empirical practice, statistical properties can only facilitate model specification, but should not dictate model selection, as Belsley and Welsch (1988) assert that we should not try to "model without understanding the non-statistical aspects of the real-life system you are trying to subject to statistical analysis. Statistical analysis done in ignorance of the subject matter is just that – ignorant statistical analysis."[5]

For hourly series, the statistical tests suggest modeling them using first differences to avoid spurious regressions. Also, using first differences of the corresponding consumption series could get rid of complicated low frequency cycles, as shown in the spectral plots in Figure 4.5, and thus this data arrangement would facilitate controlling cyclical effects in empirical analysis. Moreover, technically, the use of first differences can also help reduce the multicollinearity problem if lagged consumption variables are used as regressors in regressions. Combining these results, I conclude that modeling the hourly demand model in first differences is preferable.

[5] Op. cit., 447.

For the monthly and quarterly series related to the residential, industrial, and commercial sectors, once the yearly effects are controlled, the series are stationary. Further, the cyclical pattern in consumption is mainly in the form of a yearly cycle, which can be controlled by use of a monthly or seasonal dummy. Therefore, I would model the demand in level.

4.3 Figures and Tables

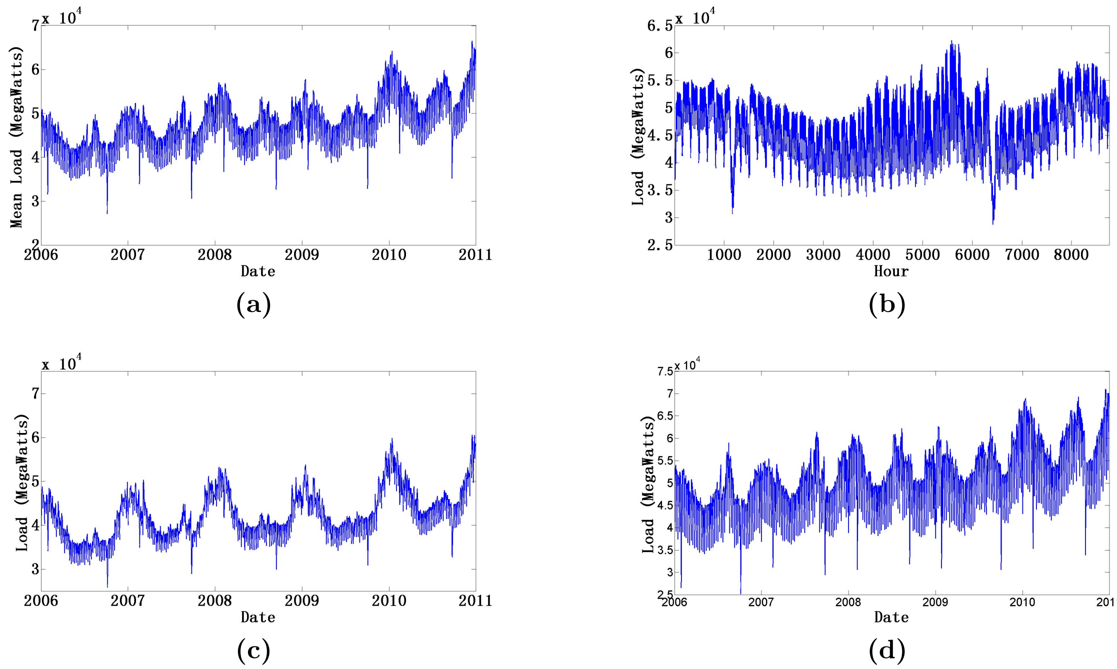


Figure 4.1: Data description of Korean electricity load in the wholesale market from Jan. 1, 2006 to Dec. 31, 2010: (a) Daily average load; (b) Hourly load in 2007; (c) 3am daily load; (d) 12pm daily load.

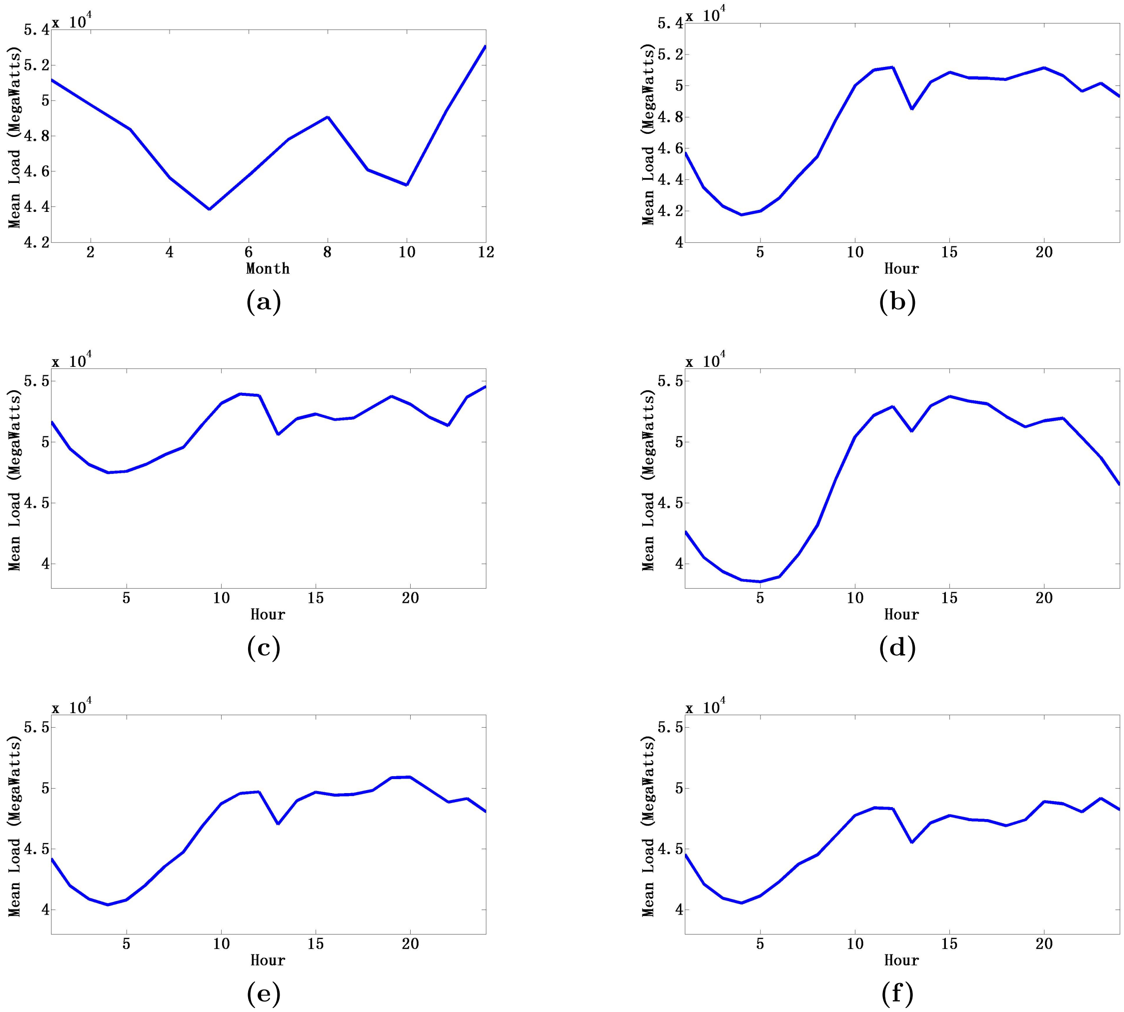


Figure 4.2: Data description of Korean electricity load in the wholesale market from Jan. 1, 2006 to Dec. 31, 2010: (a) Monthly average load; (b) Hourly average load; (c) Winter (Nov. 1st to Feb. 28th) hourly average load; (d) Summer (Jul. 1st to Aug. 31st) hourly average load; (e) Fall (Sep. 1st to Oct. 31st) hourly average load; (f) Spring (Mar. 1st to Jun. 30th) hourly average load.

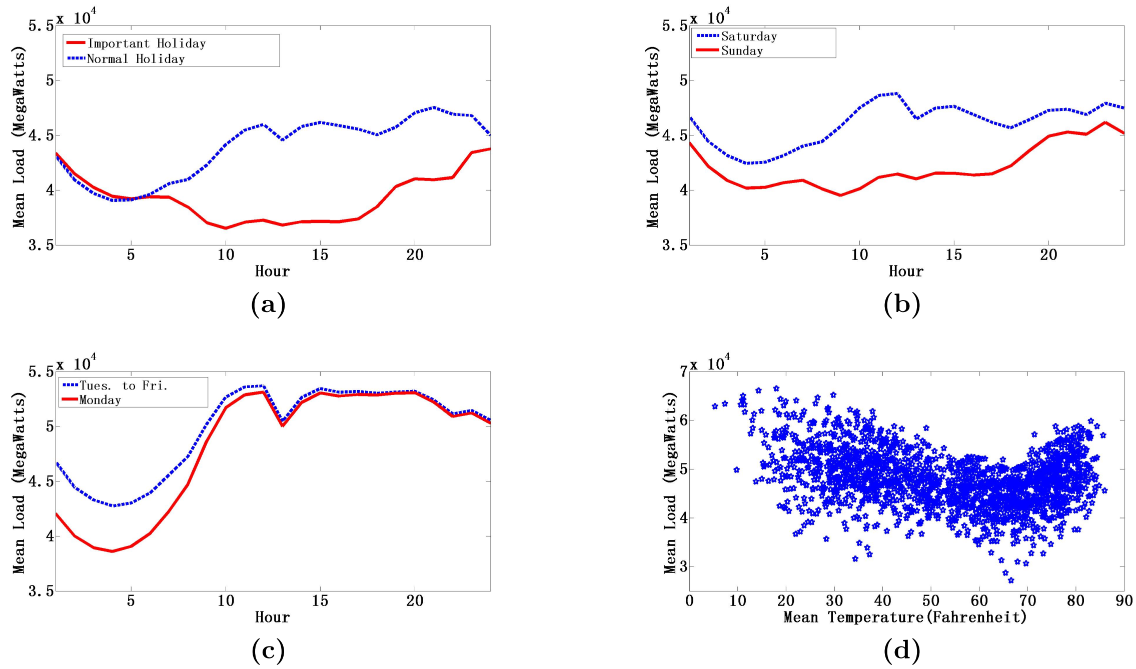


Figure 4.3: Data description of Korean electricity load in the wholesale market from Jan. 1, 2006 to Dec. 31, 2010: (a) Hourly average load (holiday); (b) Hourly average load (weekend); (c) Hourly average load (workday); (d) Daily average load versus average temperature.

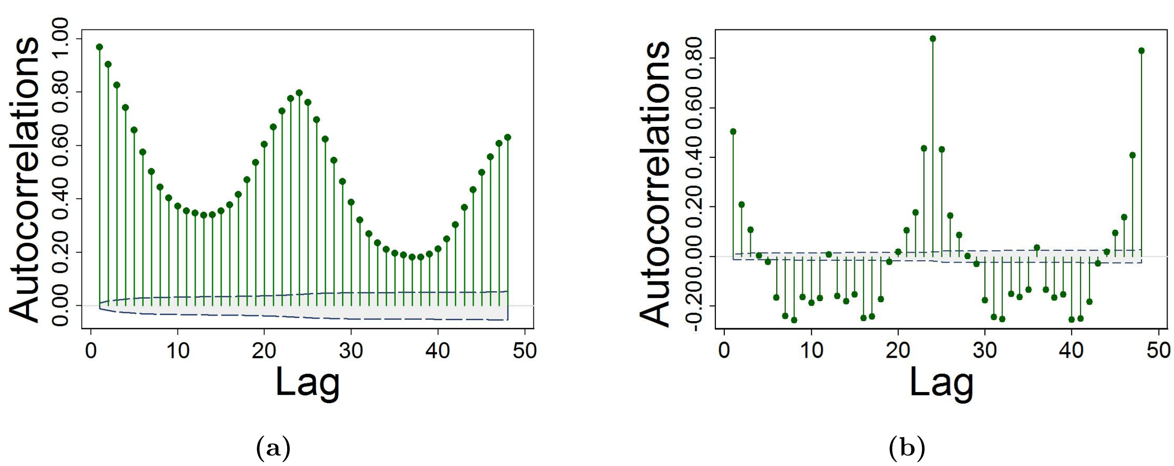


Figure 4.4: Data description of Korean electricity load in the wholesale market from Jan. 1, 2006 to Dec. 31, 2010: (a) ACF plot in level; (b) ACF plot in first difference. Barlett's formula for MA(q) 95% confidence interval.

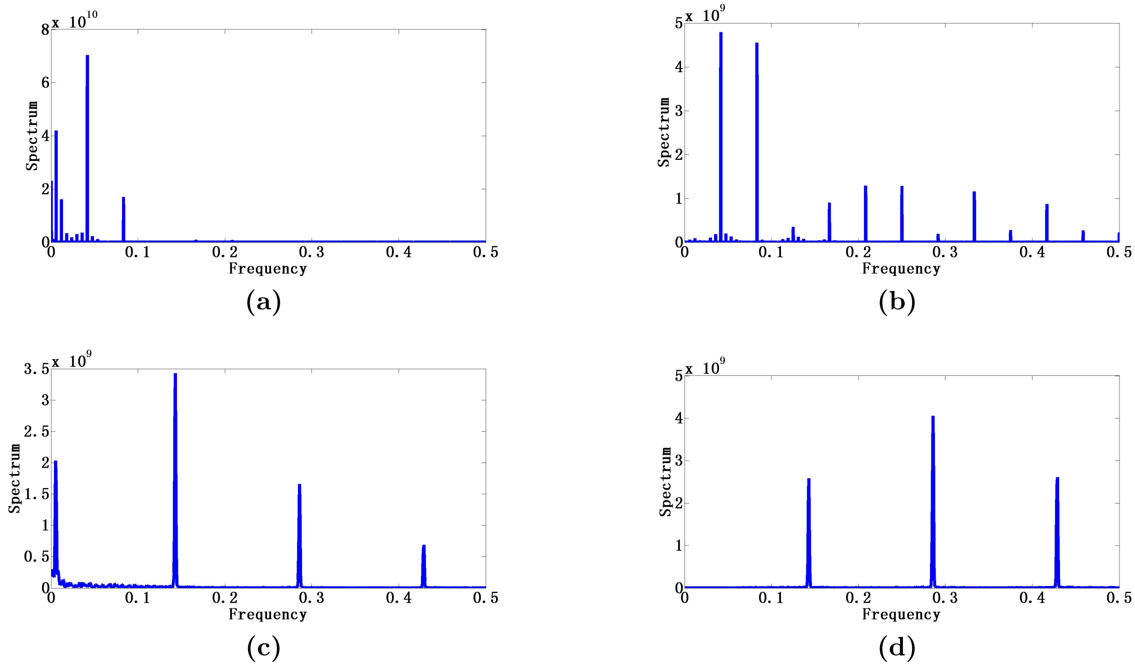


Figure 4.5: Data description of Korean electricity load in the wholesale market from Jan. 1, 2006 to Dec. 31, 2010: (a) Smoothed periodogram of hourly load; (b) Smoothed periodogram of first-difference hourly load; (c) Smoothed periodogram of 12 pm hourly load; (d) Smoothed periodogram of first-difference 12pm hourly load.

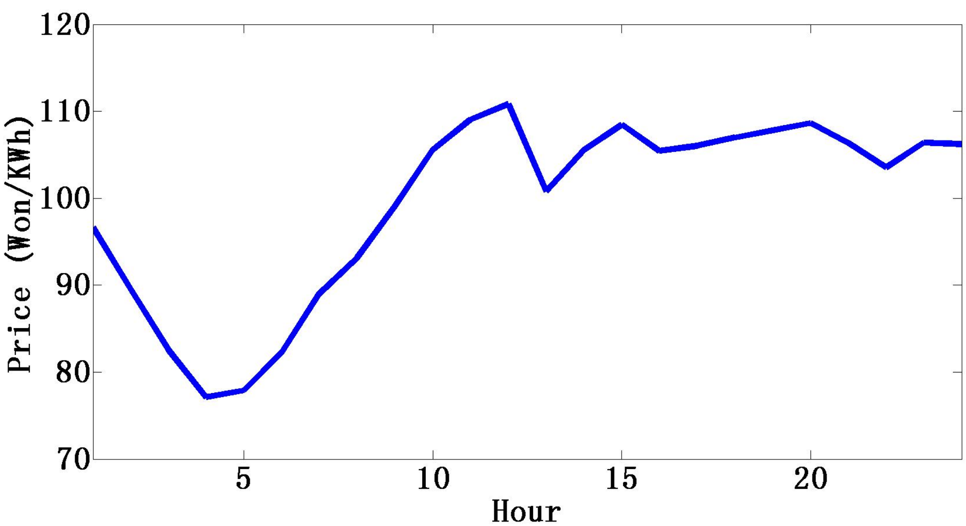
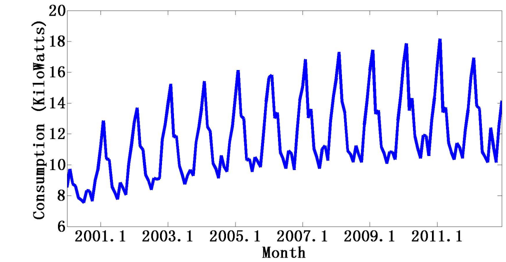
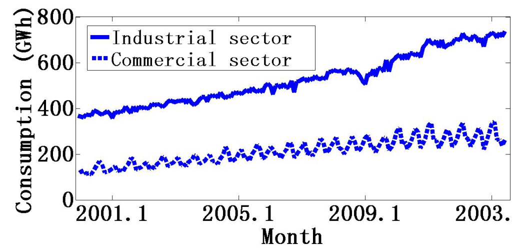


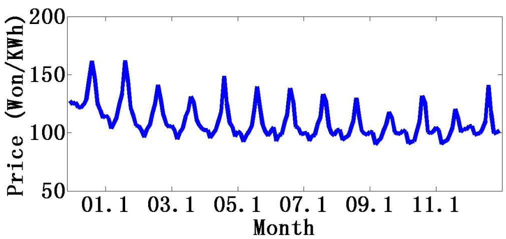
Figure 4.6: Hourly system marginal price (SMP) in the wholesale market from Jan. 1, 2006 to Dec. 31, 2010.



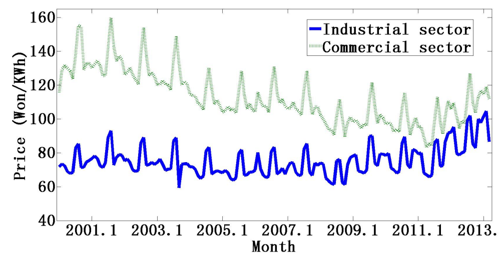
(a)



(b)

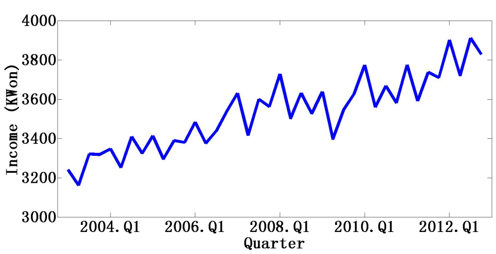


(c)

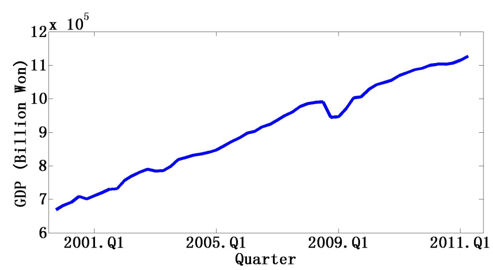


(d)

Figure 4.7: Data description of Korean electricity load and price in retail market: (a) Monthly consumption per household per day from Jan. 2000 to Dec. 2012 in the residential sector; (b) Monthly consumption per day from Jan. 2000 to Mar. 2013 in the industrial and commercial sectors; (c) Monthly average real price from Jan. 2000 to Dec. 2012 in the residential sector; (d) Monthly average real price from Jan.2000 to Mar. 2012 in the industrial and commercial sectors. Base year = 2010.

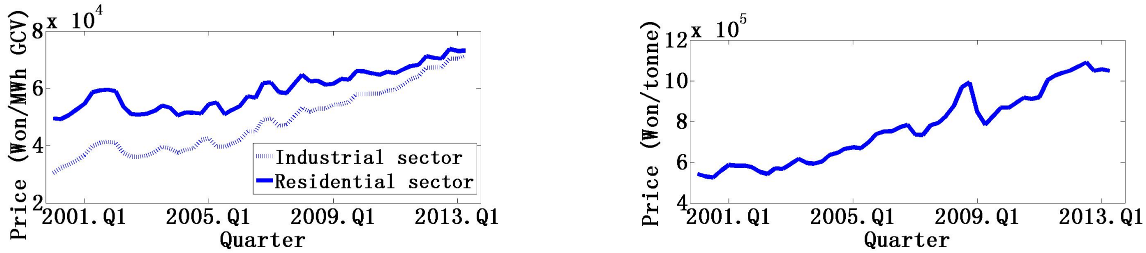


(a)



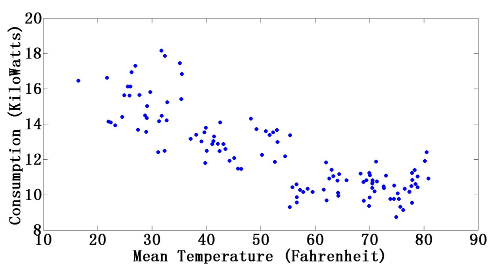
(b)

Figure 4.8: (a) Quarterly real average household income from Q1, 2003 to Q4,2012; (b) Quarterly real GDP from Q4, 1999 to Q2, 2013. Base year =2010.

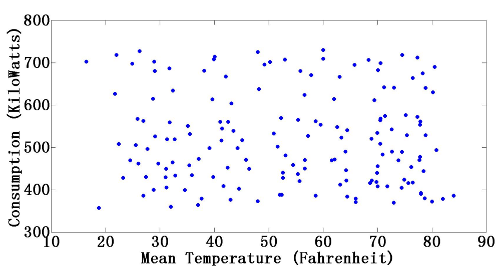


(a)

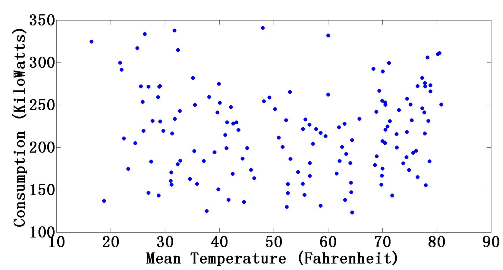
Figure 4.9: (a) Quarterly real average price of natural gas for the residential and industrial customers from Q4, 1999 to Q1,2013; (b) Quarterly real average price of high-sulfur fuel oil for the industrial customers from Q4, 1999 to Q1,2013. Base year =2010.



(a)

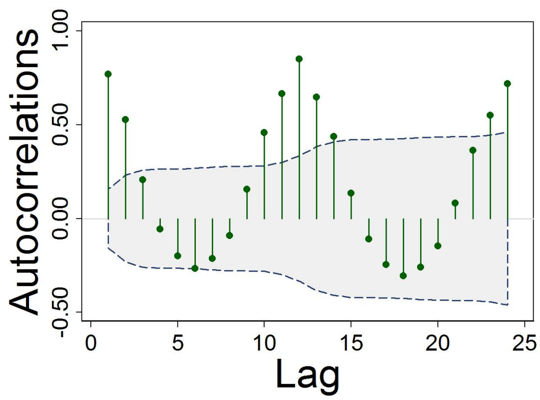


(b)

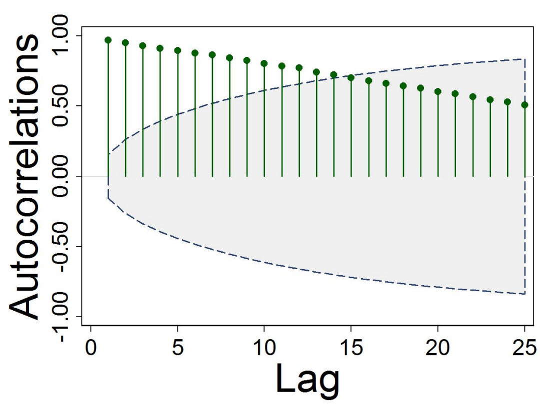


(c)

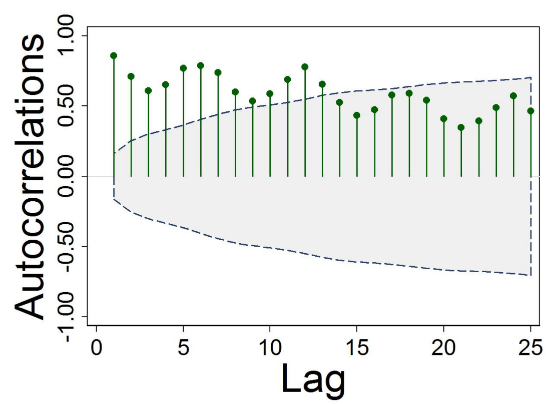
Figure 4.10: (a) Plots of monthly consumption against average temperature in the residential sector; (b) Plots of monthly consumption against average temperature in the industrial sector; (c) Plots of monthly consumption against average temperature in the commercial sector.



(a)



(b)



(c)

Figure 4.11: Data description of Korean electricity load in the retail market: (a) ACF plot in the residential sector; (b) ACF plot in the industrial sector; (c) ACF plot in the commercial sector. Barlett's formula for MA(q) 95% confidence interval.

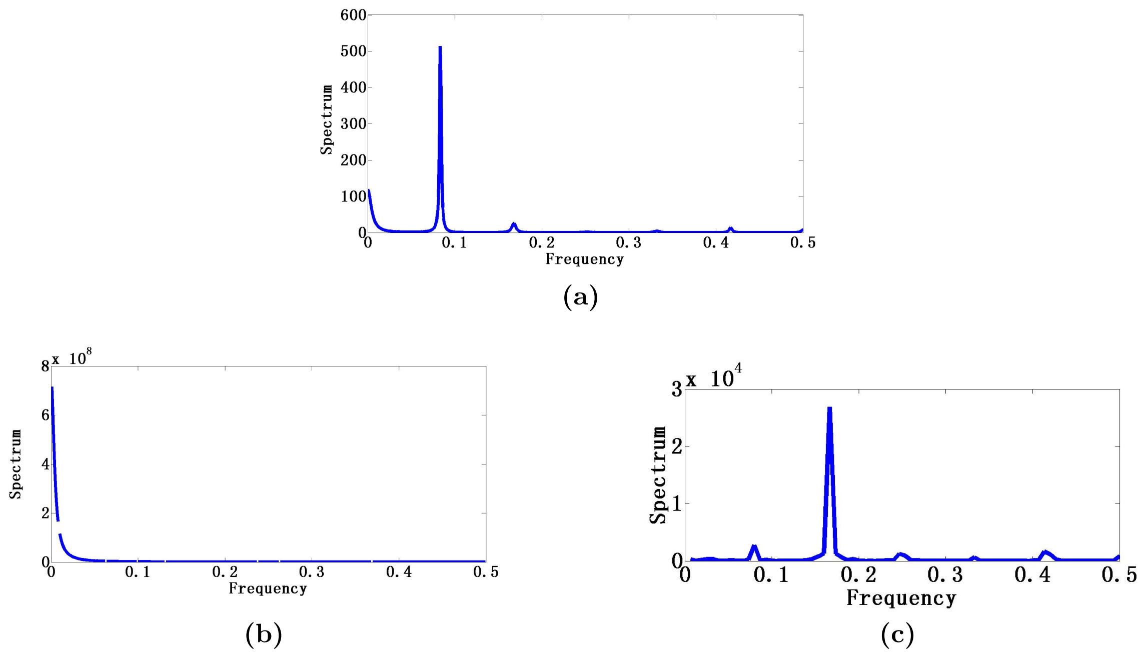


Figure 4.12: Data description of Korean electricity load in the retail market: (a) Smoothed periodogram in the residential sector; (b) Smoothed periodogram in the industrial sector; (c) Smoothed periodogram in the commercial sector.

Table 4.1: Important and normal holidays

Normal Holiday	Independence Day, Children's Day, Memorial Day, Buddha's Birthday, Constitution Day, Liberation Day, National Foundation Day
Important Holiday	New Year's Day, Spring Festival, Moon Festival, Christmas Day

Source: http://english.visitkorea.or.kr.

Table 4.2: Stationarity test (KPSS) and unit root test (DF-GLS) for the series associated with the wholesale market

	Levels(trend)		First differences (no trend)	
	KPSS	DF-GLS	KPSS	DF-GLS
Consumption	1.0303^{***}	-2.4543	0.0013	-9.8654^{***}
SMP	3.0961^{***}	-2.3472	0.0004	-15.1999^{***}
Temperature	0.1818^{**}	-3.2106^{**}	0.0083	-8.1242^{***}

Tests applied to de-seasonalized data using hourly, daily, and seasonal dummies.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

Table 4.3: Stationarity test (KPSS) and unit root test (DF-GLS) for the series associated with the residential sector

	Levels	
	KPSS	DF-GLS
Consumption	0.0171	-8.658^{***}
Average price	0.0131	-5.963^{***}
Income	0.0393	-5.527^{***}
Natural gas price	0.454	-3.005^{***}

Tests applied to the series getting rid of yearly effects by yearly dummy.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

Table 4.4: Stationarity test (KPSS) and unit root test (DF-GLS) for the series associated with industrial and commercial sectors

	Levels	
	KPSS	DF-GLS
Industrial consumption	0.0357	-5.604^{***}
Commercial consumption	0.0604	-8.159^{***}
Industrial average price	0.0131	-5.604^{***}
Commercial average price	0.0702	-8.159^{***}
Temperature	0.0100	-7.912^{***}
GDP	0.0593	-5.106^{***}
High-sulfur fuel oil price	0.0351	-5.909^{***}
Natural-gas price	0.101	-4.507^{***}

Tests applied to the series getting rid of yearly effects by yearly dummy.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

Chapter 5

Electricity Demand in the Wholesale Market

Electricity demand is a derived demand, resulting from the aggregate stock of electrical appliances and corresponding agents' habits, schedules, experiences, and other utilization behaviors. If one is looking at electricity demand in a reduced from, the price of electricity and its substitutes, such as oil and natural gas, GDP, and weather are theoretically sound factors that affect electricity demand through consumers' utilization behaviors and appliance stocks. Except for these factors, based on the data analysis in Section 4.1, I conjecture that there are also certain regularities governing hourly aggregate electricity utilization behaviors:

1. There is a periodic and cyclical stable utilization routine,[1] which implies that consumers tend to do the same thing daily, weekly, monthly, or yearly.

2. There is a high dependency of hourly utilization behaviors. The previous hour's consumption level contains information about current utilization behaviors. This dependency has two channels: (1) Current utilization behaviors might be affected by the adjacent previous hours' usage, defined as "usage persistency". For instance, consumers tend to keep using an appliance once they open it or rearrange/shift its usage to later hours; (2) Current utilization behaviors might be affected by consumers' expected utilization behavior; for example, if consumers plan to use some appliance at a later hour, they may open it before the actual usage.

3. Current utilization behaviors might also be influenced by previous daily or weekly

[1] In the residential sector, Borenstain 2009 points out that the typical consumer might set behavior consumption rules ex ante, for example set thermostats at a certain level, or turn off lights in unused rooms. In industry, business, or other sectors, managers generally set down production plans in advance.

consumption which deviates from the periodic routine. In other words, previous daily or weekly consumption contains information about current utilization behaviors which is not captured by the periodic routine.

Both 2 and 3 indicate that, at the same hour with the same price on the same type of day, the same electricity consumption might not occur as it is influenced by its past history.

The difficulty of modeling high-frequency electricity data, compared to low-frequency data, lies in controlling these complex regularities. Empirically, these regularities are correlated with price variables through aggregate electricity consumption, and thus if they are not properly controlled, their effects would be confounded in the price coefficients, causing omitted-variable bias. To avoid this misidentification, my strategy is to take advantage of trigonometric functions, which is a suitable method to model cyclical phenomena, and lagged consumption variables as utilization regularity proxies in demand modeling.

5.1 Empirical Model

In Chapter 4, I conclude that modeling the demand equation in first differences is the better approach to avoid spurious regressions and to control cyclical pattern in a practical way. In an ad hoc way, I assume a linear demand functional form as a starting point and model the wholesale demand,[2] with the consideration of the statistical facts that first-differenced

[2]There is no consensus in the literature about the most appropriate functional form of electricity demand estimation, and most of the studies adopt a linear or logarithmic form in an ad hoc way. In demand estimation, Zarnikau (2003), based on residential cross-sectional household level quarterly data, finds that linear representation is superior to log-linear and translog share equation functional forms by a frequentist non-parametric bootstrapping test. Xiao, Zarnikau, and Damien (2007) conclude the model preference order as translog, log-linear, and linear functional form by a Bayesian approach based on the same data set. They remark that "model selection remains a very difficult task and different model selection criteria may steer one towards different model choices." In practice, we should keep in mind that, "letting the data speak for themselves" through econometric methods is helpful in model specification, but that its role should not be exaggerated over economic theory and underlying reality, which are the foundation of, and guiding force in, model specification. On the other hand, the demand modeling strategy can also be structural and explicitly incorporate information about aggregate electricity-consuming appliance stocks, utilization rate levels, and purchasing decisions. This is, however, hard to formalize directly and is beyond the scope of my data set.

hourly series are stationary, as the following,

$$\Delta y_t = \alpha + \gamma \Delta p_t + \sum_{i=1}^{m} \beta_i \Delta p_{t-i} + \Delta T_t' \theta + + D_t' \delta + X_t + \epsilon_t \tag{5.1}$$

where y_t is the aggregate consumption in unit MWh in the wholesale market at hour t; p_t is the wholesale market system marginal price (SMP) in thousand Won per MWh at hour t; T_t is the vector, including $\{HDD_t, HDD_t^2, CDD_t, CDD_t^2\}$, HDD_t (CDD_t) is equal to the temperature in degrees F when the daily average temperature is lower (higher) than 65° F and 0 otherwise at hour t[3]; D_t is the vector containing dummy variables for weekday, weekend, holidays, and years, which are used to control possible structure shifts listed in Table 5.1; X_t represents utilization regularities of y_t over time, at hour t; ε_t represents an unobservable random disturbance; and m is the largest number of price lag.

X_t is specified as the following,

$$X_t = \sum_{i=1}^{n_1} a_i \Delta y_{t-i} + \sum_{i=1}^{n_2} (b_i \Delta y_{t-24*i} + c_i \Delta^2 y_{t-24*i}) + \sum_{i=1}^{n_3} (\eta_i^1 \sin(2\pi \frac{i}{168} t) + \eta_i^2 \cos(2\pi \frac{i}{168} t)) \tag{5.2}$$

Since utilization cycles are not strictly sinusoidal-like, trigonometric functions with a basic weekly frequency and all its harmonious frequencies are included, and thus $n_3 = 84$. Further since trigonometric functions cannot capture nonlinearity within one hour, I apply second-order differences of daily lagged consumptions to capture the non-linearity, especially the tendency changes of consumption, within one hour. Lagged consumption $\{y_{t-i}\}_{i \leq n_1}$ are used to capture the lag inter-temporal usage effects and $\{y_{t-24*i}\}_{i \leq n_2}$ captures previous daily or weekly influences; $n_1 \in [1, N_1]$, $n_2 \in [1, N_2/24]$, $N_1 (\leq 23)$ and $N_2 (\leq 15)$ are the largest number of hourly and daily lags respectively.[4]

I impose one assumption for model (5.1) that the disturbance does not contain auto-

[3] Recall the relationship between temperature and electricity demand, which is kinked at around 65° F.

[4] I do not specify the behavioral and equilibrium conditions from which mapping of the observable and unobservable explanatory variables into the observable dependent/variables might have been derived. The reduced form suffices to analyze many situations as long as the underlying structure stays stable.

correlation. This assumption guarantees that I can treat the lag consumption variables as pre-determined exogenous variables.

GDP and the price of electricity substitutes, i.e., oil and natural gas, are excluded in the model for two reasons. Firstly, substitution and GDP effects might both involve changing the type of appliance. Given the short response time in the model, these effects would be nil. Secondly, these data are available on a quarterly basis, which do not match the data frequency in the model. Further, since interpolation of quarterly data to hourly data is not practical, much variation will be lost by this conversion.

γ represents the price sensitivity for market customers, who respond to the market price. β captures hourly average direct lagging price substitution or complementary effects.[5] The size of γ and β are affected by the size of responsive customers in the market in the sense that if the ability of customers to respond to market price increases, then the price effects in absolute value will increase.

In practice, I apply aggregate hourly electricity consumption of the whole economy as y_t, since it is a good representation of the total load in the wholesale market[6] and SMP as market price, based on the assumption that wholesale market buyers respond to SMP rather than PSP. The rationale behind this assumption is that PSP is generally known to customers after the trading day and SMP accounts the majority, around 70%, of PSP. The rest of PSP includes the capacity price (CP), and other constraints and ancillary service costs, and this cost information is unavailable to customers beforehand. Therefore, customers might use SMP as a proxy for the true price and thus respond to it. Moreover, SMP is applied in its nominal level, because data are in hourly frequencies, and it might be more realistic to assume that consumers respond to the nominal price rather than the real price.[7] The

[5]The substitution effect refers to electricity consumption shift to later hours due to price increases and the complementarity effect refers to a price increase that has the same effect as current and later hours.

[6]The average yearly ratio of electricity supplied from the market is 97.1%. The rest of the supply is from PPA, which is small and not influenced by SMP. This provides the basis that aggregate demand is well suited for analyzing the wholesale market and the proportion effect is very low.

[7]For example, Boisvert et.al. (2007) apply the nominal level of electricity prices in their daily electricity demand model.

Sampling period is from 2006 to 2010.

Since economic theory offers little guidance about the lag lengths, I will select the included lag variables based on the *general-to-specific* principle, which is to start with a larger model and then to simplify it by removing insignificant variables. I impose an ad hoc restriction for the largest number of lags in the model for considerations of parsimony, i.e., $n_1 \in [1, 23]$, $n_2 \in [1, 15]$, and $m_1 \in [1, 48]$. I hope that, by using first differences and including a variety of utilization proxies as well as structure shift binary covariates, I diminish the possibility that spurious correlations will contaminate estimates of the effect of price on electricity demand in the wholesale market.

Finally, instead of trigonometric functions, binary variables, i.e., hourly, workday dummies, and the interactions among them, could also control the cyclical behaviors. Under binary dummies, periodic routines are approximated by a step function. Compared with trigonometric functions, binary variable proxies are less flexible and lack the power to track non-linearity information in the data. Nonetheless, I will also provide results with binary variable proxies as a comparison.

5.2 Identification

5.2.1 Theoretical Foundations

The most important issue in the model is whether the price coefficient γ is identifiable. I establish identification in a simultaneous equation system. I consider that the electricity equilibrium price (SMP) and quantity are jointly determined by the interaction of electricity supply and demand. This setup requires the rationality assumption of KPX, in the sense that they take the demand response into consideration such that there is no systematic deviation in SMP determined in the previous trading day and the potential price at real time. I believe this is a reasonable assumption provided that KPX would revise their forecasting if there were systematic errors. I also assume that some variables available to KPX are not available to me.

In this demand and supply system, the electricity supply curve is relatively stable whereas the demand curve is quite volatile. The supply shifters such as cost adjustment, maintenance, outage, ramping, and other technological constraints are generally stable or with limited effects. Table 5.2 lists the coefficients of variation for hourly price, temperature, load, and available capacity of base, LNG, and oil generators in 2007 as an illustration. The coefficients of variation are much smaller for generators' capacity compared with other variables. Further by checking the generator cost data in 2007, I find that out of 82 LNG and oil plants, only 12 plants adjusted their cost slightly one time during 2007.[8] Therefore, there are no suitable supply-side shifters to facilitate identification.

Exclusion conditions (zeros restrictions) are generally the standard method to establish identification of the demand curve in a simultaneous system in common practice. If the exclusion conditions are satisfied in the demand equation, the demand equation can establish identification and one could adopt exogenous supply shifters, straight-forward instrumental variables (IV), to instrument endogenous price in the demand equation to correct simultaneous bias and obtain consistent estimates. It is a popular method, but not the only approach to guarantee identification in the sense that no transformation of the structural parameters yields the same reduced form in the system. "The popularity of zeros restriction as a means of identifying equations," according to Kennedy (1988, p.170), "probably stems from the fact that this method is easier to apply and has been given formal mathematical treatment." In essence, the identification problem with simultaneous equation systems is a mathematical problem and it is concerned with the question of the possibility or impossibility of obtaining meaningful estimates of the structural parameters (Kennedy 1998, p.159). Except for exclusion conditions, several other forms of restrictions suggested by extraneous information and/or economic theory in the systems could also be sources of identification.[9] A few of

[8]The generator cost function takes the quadratic form: $C = a^2q + bq + c$. For all 82 LNG and oil generators, the mean (median) proportion changes, in absolute value, of total cost and marginal cost change when producing 1 kwh are 31.38%(15.81%) and 15.45%(8.21%), respectively.

[9]I should also be aware that "it is not true that if an equation is under-identified, I have to give up all

these are, according to Kennedy (1998, p.159):

- extraneous estimates of parameters;
- knowledge of exact relationships among parameters;[10]
- knowledge of functional form;
- knowledge of relative variances of disturbances;[11]
- knowledge of zeros correlations between disturbances in different equations.

Thus, to establish identification in this case, I show that non-linearity in the supply equation can help identification of the demand equation.[12] Figure 5.1 (a) presents a typical hourly aggregate supply curve of LNG and oil generators in the wholesale market. The Appendix provides both informal and formal identification proofs. I also extend the results to the identification[13] of the nonlinear demand equation with respect to price.

Intuitively, the role of exclusion conditions in a linear system is to provide restriction on the relationship between the slope and the horizontal shift variables in the demand equation, while non-linearity in the supply curve, in essence, has the same role to provide this restriction. As a simple illustration, instead of the nonlinear supply curve, I could instead think of a linear supply curve with arbitrary shifters, which would provide the same equilibrium results as the original system, as shown in Figure 5.1 (b). Alternatively, the exclusion condition and the non-linear supply curve have the same role in terms of ruling

hopes of estimating it" (Maddala (1992), p.386)). Blind adherence to the standard rule without investigating the properties of the system and data is more likely to reflect dogmatic faith than the scientific approach.

[10]This knowledge includes homogeneous/non-homogeneous within equation linear restrictions, cross-equation restrictions, and nonlinear restrictions. The exclusion condition is a special case.

[11]For instance, in the demand and supply system with no exogenous variables in either the supply or demand equations, if the variance of the disturbance terms in the supply equation is known to be much larger than the variance of the disturbance terms in the demand equation, that is the demand curve is relatively stable whereas the supply curve is quite erratic, the data observations could still trace out a demand curve without the application of IV. Christ (1966), Johnston (1984), and Maddala (1977, pp.226-231) discuss identification based on the restrictions on the knowledge of the disturbances in the system.

[12]Reiss (2005) points out that the demand function is identifiable with a stable two-tier increasing block tariff schedule.

[13]Identification here refers to identify the first-order derivative of demand with respect to price.

out the possibility that the equation under consideration cannot be distinguished from a linear combination of all equations in the system.

5.2.2 Choices of IVS

To correct simultaneous bias, proper IVs are needed. Without an instrument, there will be an upward bias of the estimated price coefficient. Based on the literature, possible choices of an instrument for price come from supply-side information, including base generator availability (Wolfram 1999), fuel (natural gas) price, excess capacity availability in the system (Fezzi and Bunn 2010). As discussed in the identification section, with a relatively more stable supply curve in the system, supply shifters would not be legitimate IVs. I provide some statistical results to verify this claim. For base generator availability, its coefficient of variation, 5.4, in Table 5.2, is much lower than that of SMP, 20.82, indicating its potentially low explanatory power of the variation of hourly SMP.[14] Further, an F-test applied in the first stage 2SLS confirms that base generator availability is a weak IV, as the F-statistic is 0.4281, which is much smaller than the conventionally acceptable level 10.[15] For fuel (natural gas) price, intuitively, if the fuel price increases, the electricity price will increase correspondingly. In the Korean electricity market, however, the generator cost function does not adjust instantaneously with the fuel price, but rather does so with an adjustment lag due to the fact that generator variable cost is generally approved by the cost-evaluation committee on a monthly basis. This fact breaks down the contemporaneous correlation of fuel market price and electricity price. Further, the data of fuel price are available on a quarterly basis, providing limited variation for hourly price changes. Thus, the lag of the fuel price is also ruled out as an IV candidate. For excess capacity availability on the system,

[14]Conley (2008) points out that IV methods exploit only the portion of the variation in the endogenous variable induced by shifting the instrumental variable, and inference tends to be imprecise. This problem is also exacerbated when the instruments are only weakly related to the endogenous variables and induce only small amounts of variation, resulting in large standard errors and biased estimates.

[15]Stock and Watson (2003) show that if there is only one endogenous variable, the first-stage regression in 2SLS can test for the significance of the instruments using an F-test by a rule of thumb criterion: if the F-statistic is less than 10, then the set of instruments is weak.

it still suffers an endogeneity problem since excess capacity depends on the level of demand.

According to the time-series nature of the data and its inherent cyclical patterns, monthly[16] and yearly lag (365 days) prices can serve as possible IV candidates with the no autocorrelation assumption. The rationale is that, since aggregate consumption contains monthly and yearly cyclical patterns, and the supply curve is relatively stable, the monthly and yearly lag prices are thus correlated with the current price, enabling the prediction power of the current price. Lijesen (2007) and Alberini and Filippini (2010), for instance, also apply monthly lag prices as IVs in their electricity demand model. Morevoer, yearly and monthly lag prices both have their own merits in the sense that a yearly lag provides a better match for holidays with the Western calendar, and a monthly lag provides a better match for a weekly usage pattern. The leap year in 2008 and the Lunar calendar holidays in Korea would, however, weaken the explanatory power of a yearly lag as an instrument. Again, I apply the standard F-test in the first-stage regression to test the validity of each IV. The F-statistics are 46.72 and 429.24 for a one-year and one-month lag respectively. Both of them are much greater than the conventional level 10, although the strength of the yearly lag IV is lower. The results indicate both IVs are proper instruments for the current price variable.

5.3 Estimation Results and Implications

I first select the number of lags of price and consumption variables based on the *general to specific* criterion for model (5.1) by GMM with a one-month lag price IV. Both hourly price and consumption lags within 23 hours, $m = 23$ and $n_1 = 23$, daily consumption lags within a week, and a two-week consumption lag, $n_2 = [1 : 7, 14]$, are included.

Table 5.3 lists some estimation results of model (5.1). Column (1) contains the results without instruments estimated by OLS, and columns (2)-(4) are the results with different

[16] I normalize a month as being 28 days.

instruments estimated by GMM. The high R^2 values (> 0.9) on the regression indicate that the control variables comprise a relatively comprehensive set and explain a large amount of the variation in electricity demand. As expected, there is an upward bias of the price coefficient without IV. Further, the estimated price coefficients with IVs are all significant at the 5% level and are consistent with the economic theory that as electricity prices rise, the quantity of electricity demanded will fall, ceteris paribus.

Before interpreting the results, I present some diagnostic test statistics in Table 5.4 to check the models' performance. Columns (2)-(4) correspond to the models in the same columns in Table 5.3. Firstly, the Cumby-Huizinga Test detects no autocorrelation in the error term, since its statistic is insignificant at the 10% level. This result provides empirical support for the no autocorrelation assumption and legitimatizes the use of monthly and yearly lag price variables as instruments. Secondly, these instruments are needed based on the endogeneity test. The test statistics are all significant at the 1% level, suggesting the price variable suffers from an endogeneity problem and both instruments are not redundant. Thirdly, the Pagen Hall Heteroscedasticity test rejects the null of homogeneity in the error terms at the 1% level, supporting the use of the GMM method since it is more efficient when the error terms are heteroscedastic than the IV method. Fourth, for the Hansen J test, I could accept the null hypothesis that the error terms are uncorrelated with the instruments when both one-month and one-year lag price IVs or one- and two-month lag price IVs are used to instrument price variable at the 1% or 10% levels. This does, however, cast some doubts when monthly and yearly lag IVs are used together. This is because the yearly lag price IV is weaker than the monthly lag price IV in the aspects of matching weekly consumption patterns. Further, in respect of the Hansen J test, I conclude that the monthly lag price IVs are a better choice. Moreover, the results with different IVs are not systematically different from each other. Therefore, I will present my interpretations based on the model with one-month lag price instruments in column (2) in Table 5.3. Figure 5.3 provide CUSUM test plots of this model, suggesting that the model is stable within a 95%

confidence band.

Table 5.5 presents the corresponding instantaneous price elasticities, calculated at the mean consumption and prices level for each year. The high average prices in 2008 and 2010 are due to increased global fuel prices. For example, the market price went up by 11% between the end of 2009 and August 2010 due to persistently high fuel prices. The resulting real-time elasticity is highly inelastic, -0.0034 on average. This result first confirms the theory that consumers' demand for electricity is less sensitive to price changes than the demand for many other commodities, and second it aligns with the real-time estimates in Lijesen (2006), -0.0043, based on wholesale data in 2003 in the Netherlands. The long-run elasticity based on the whole sampling period is -0.0640.

Real-time price elasticities for each hour within a day are calculated and presented in Table 5.6 and Figure 5.2 (a). The highest price response is around 12pm when consumption is in general the highest, and the lowest price response is at midnight.

Table 5.8 presents estimated price coefficients with different cyclical regularity specifications. Binary dummies are also applied as utilization routine proxies, i.e., hourly, weekday, and their interaction dummies. Column (1) contains the results of the basic model in column (2) in Table 5.3. Columns (2) and (3) suffer autocorrelation based on the Cumby-Huizinga autocorrelation test, indicating not only that the lagged price variable is not a suitable IV, but also that the lagged consumption and price variables in the model are endogenous. This result also indicates that the second-order difference consumption lag terms sufficiently control the nonlinearity within one hour. This autocorrelation problem does not, however, bias the estimated real-time price coefficients, since they are of similar magnitude and are not statistically different from the results in columns (1) and (4). One justification for this is that, although autocorrelation exists, the correlation between monthly lagged price and the current residue is quite low, and thus the validity IV conditions are not jeopardized so much as to bias the estimates. Autocorrelation does, however, bias long-run price elasticity through the biased estimates of the lagged variables. I calculate long-run elasticity to illustrate the bias. Long-run price elasticities in columns (1) and (4), with no autocorrelation, have similar magnitude, while columns (2) and (3) are more than twice as large as in models (1) and (4). Thus, autocorrelation leads to an upward bias of long-run price elasticity in absolute value.

For more comparisons of price elasticities, I also regress the model with a CPI deflated price variable where 2010 is the base year, and a log-log specification,[17] respectively. Firstly, with a CPI-deflated price, the estimated coefficient of the price variable is -1.8278 (0.7236), significant at the 5% level, and is not significantly different from the results with nominal price. The corresponding real-time price elasticity calculated at the sample mean is -0.0034, and the long-run price elasticity is -0.0640, the same as the basic results in column (2) of Table 5.3 when reserving four decimals. Secondly, a log-log specification for the wholesale market demand is applied based on model (5.1). All variables, except for the dummies and trigonometric function, are in log form. The estimated real-time price elasticity is -0.0027 (0.0014), with a p-value of 0.0511. Long-run price elasticity is -0.0478. Still, the results are not statistically different from those of the linear specification.

Next, all estimated lag price coefficients are significant at 1%, as shown in Table 5.9 Panel A. I calculate inter-temporal cross-price elasticities by the formula,

$$\frac{\partial y_i}{\partial p_j} = \frac{\partial y_i}{\partial p_j} + \sum_{k=1}^{i-j} \frac{\partial y_i}{\partial y_{i-k}} * \frac{\partial y_{i-k}}{\partial p_j} \tag{5.3}$$

where $j < i$ and $\frac{\partial y_i}{\partial y_{i-k}}$ incorporate both direct and indirect lag consumption effects. The first part of the formula represents the direct cross-price effect, and the second part is the indirect price effect through demand persistency effects.

Table 5.7 Panel A shows cross-price elasticity within 23 lag hours based on the whole sample. All cross-price elasticities, except for 23 lag hour with a very small positive magnitude, are negative, indicating major complementary effects. Figure 5.2 (b) illustrates the

[17] Allcote (2011) points out that linear specification tends to under-estimate price elasticity when prices are high, while log-log specification tends to over-estimate price elasticity when prices are low.

cross-price elasticities graphically. The plot shows that as the lag periods move further away, the magnitude of the cross-price effects weakens. The highest complementarity effects are in adjacent hours, and cross-price elasticities with respect to the previous five hours are higher than the real-time price elasticity in absolute value, indicating consumers are more responsive to previous adjacent hours' prices than the current price, probably because customers need some response time and tend to adjust usage in block hours. I conjecture that the reason for a lack of statistical evidence for a consumption shift might be that (1) SMP is not high enough in the sample to induce the incentive of consumption shift, and therefore the reduced electricity consumption might only involve some infra-marginal usage of the electricity, not the real operational work, i.e., lighting and air-conditioning, and these reductions will not shift electricity usage to later hours, and it will probably get back to its original level after some hours; (2) it is possible that some individual consumers do exhibit a consumption shift, but aggregation masks this behavior.

Moreover, except for the interpretation of own- and cross-price elasticities, the regression results in column (2) in Table 5.3 also convey other implications.

Firstly, Figure 5.4 presents demand impulse response within the previous 23 hours, with the assumption that all prices remain the same. Note that the effects of lagged consumption are positive but with a decreasing trend within the previous three hours, and the effects become negative afterwards and the effects gradually converge to zero. Usage persistency or expectations might be the factors causing previous consumption to influence current consumption. With aggregate data, however, these two factors are confounded in the data and thus cannot be distinguished. Therefore, I will not go further into explanations of the behavior. Thus, I conjecture that omitting the lagged consumption variable might introduce bias in the coefficients of the price variables. To illustrate this, I run the regression without the lagged consumption variables within 23 hours. The estimated current and lagged price coefficients are listed in Table 5.9 Panel B and the corresponding price elasticities are listed in Table 5.7 Panel B. Compared to the original results, omitting lagged consumption variables tends to overestimate price elasticities in absolute value.

I conclude that hourly electricity demand is not independent, and current consumption is influenced by the previous hour's consumption and price level. Table 5.10 lists a selection of three hours' residual correlation coefficients when I regress hourly demand in a day separately in an SUR framework. Based on Breusch and Pagans (1980) LM test,[18] I could reject the null that each of the two equations are not correlated at any conventional significance level, substantiating the dependence of hourly demand. Therefore, models treating each hour of demand independently neglect the information of hourly dependency, and the corresponding price coefficients are thus less efficient and probably biased, and so are the total cross-price elasticities. The size and the direction of the bias depend on the magnitude of the correlation between price variables with lag demand variables and the size of lag demand effects.

Secondly, temperature effects are not symmetric. Temperature has a significant negative linear effect on electricity usage in heating degree days while the quadratic effect is insignificant. Temperature has both significant linear and quadratic effects in cooling degree days. When the temperature falls below 65° F, the higher temperature helps reduce electricity usage, while when the temperature rises above 65° F, electricity consumption increases with higher temperatures, ceteris paribus. The temperatures effect in cooling- and heating-degree days are heterogenous based on the Wald test. The Wald statistic for linear coefficients of the two types of days is 44.47, and for quadratic coefficients is 49.32; both are significant at the 1% level. The corresponding temperature elasticities, calculated at the sample mean, are 0.0153 and -0.0109 for cooling- and heating-degree days, respectively.

Thirdly, on average, important holidays have larger effects than normal holidays in reducing electricity consumption, and this relationship holds true for Sunday and Saturday but with a relatively smaller magnitude in absolute value, and the reduction on Saturday is not significantly different from zero. I apply the Wald test to check whether their magnitudes are also statistically different, and this is shown in Table 5.12. The reduction effects on Sundays

[18] The test statistic is $\lambda = N\rho_{ij}^2$, where N is the sample size and ρ_{ij} is the correlation.

and normal and important holidays are not statistically different. Saturday effects are statistically different to others. On the day before an important holiday, there is a statistically significant reduction in consumption, while, on the contrary, on the day after an important holiday and Monday, ceteris paribus, there are demand increases. The reason for these results might be that people tend to relax, leave the office, and shut down production earlier than normal workdays on days before important holidays, while they need to resume work, and get back to normal production after the holiday, which causes demand to increase when controlling for lagged consumptions. The effects of the day after important holidays and Mondays are statistically different; the test statistic is 4.68, which is significant at the 5% level, using the Wald test. Table 5.11 lists the reduction/increasing consumption percentage compared with average hourly consumption level. All yearly dummies are not statistically significant. Thus, there is no evidence of a structural shift due to socio-demographic factors, such as age distribution and living standard within my sample periods.

5.4 Hypothesis Testing

In this section, based on model (5.1), I investigate possible heterogeneity of the price coefficients, which might be affected by different factors, for example, temperature, size of the customers in the market, and different hours of the day. To do so, I add interaction terms between factor indicators and price variables to allow the estimated coefficients to vary under different situations, and I then use hypothesis tests to determine whether price responses are significantly different in each scenario. In detail, I propose the following three hypotheses:

- The absolute value of the price coefficient is higher in cooling-degree days than in heating-degree days, which is suggested by Boisvert, et. al. 2006.
- The absolute value of the price coefficient is higher with district electricity businesses.
- The absolute value of the price coefficient is higher in peak hours.

5.4.1 Hypothesis 1: The absolute value of the price coefficient is higher in cooling-degree days than in heating-degree days.

Table 5.13 shows the estimation results regarding price coefficients. The estimated coefficient of the interaction term is highly insignificant. Thus, statistically, I can reject the hypothesis that the price coefficient is higher in cooling-degree days. The corresponding real-time price elasticities, based on the sample mean, are -0.0032 and -0.0040 for cooling- and heating-degree days, respectively, indicating that customers, in aggregate, tend to be less responsive in summer. Economically, these results might be because (1) electricity is more valuable in the summer due to hot weather; and (2) there are fewer substitutes in summer than in winter; for instance, in summer, fans are substitutes for air-conditioners, but they both use electricity, while in winter, heating equipment includes gas boiler systems, electric heaters, electric heating beds, etc.[19]

In contrast to the Korean residential sector, Boisvert et. al. (2006) show that price elasticity is higher in cooling-degree days than in heating-degree days in absolute value based on PJM data in the US residential sector. Therefore, caution is required when generalizing empirical results to different geographic and demographic situations.

5.4.2 Hypothesis 2: The absolute value of the price coefficient is higher with district electricity businesses.

District electricity businesses emerged in 2004, and in total 22 district suppliers were approved across 28 districts from October 2004 to July 2008. I therefore add the interaction term of the price variable and the dummy variable indicating Aug. 2008 to 2010. As a

[19]Taking households for example, Jeong et.al. (2011, pp.??) point out that "In 2005 Korean statistics, 77.76% of households in Seoul were using gas boilers as their main heating system. Among the HHs that can choose their heating system, 96.35% were using gas boilers. 6% of all HHs in Korea were using electric heaters. The electric heating bed market in Korea reached 100 billion won in 2010, amounting to 21% of the market for every kind of bed, which became a major type of electric equipment and furniture in Korean households."

comparison, I also add the interaction term of the price variable and the dummy variable indicating 2008 to 2010, because district electricity businesses began to withdraw from the market in 2008 due to rising oil prices. Results are listed in Table 5.14. Clearly, I cannot reject the null of the hypothesis at the 5% significant level. The magnitude of the estimated price coefficient in column (2) is greater than that in column (1) in absolute value. District electricity businesses do help increase the overall price sensitivity in the wholesale market. The corresponding instantaneous elasticity is -0.0075, calculated at the sample mean, during 2007 to July, 2008, and -0.0034 for the rest of the sample period in column (1), and -0.0121 in 2007 and -0.0035 from 2008 to 2010 in column (2). Therefore, the greater the size of the responsive consumers, the larger the price effects.

5.4.3 Hypothesis 3: The absolute value of the price coefficient is higher in peak hours.

Generally, which hours are peak hours is an empirical question. Thus, in practice, I select five possible time ranges in a work day: (1) 11am-8pm, (2) 2-7pm, (3) 2-5pm, (4) the peak period in the tariff structure in the industrial and commercial sectors, and (5) the peak and mid period in the industrial and commercial tariff schedule in Table 3.1. Table 5.15 lists the average price level in these candidate periods. The first and second rows are the average price corresponding to the candidate peak and the rest time periods. The average price is higher in all five candidate peak periods. I add the interaction term of the price variable and each candidate peak-hour dummy, respectively, in the model. Table 5.16 presents the results for each period candidate. Statistically, for all five candidates, I reject the null hypothesis since there is no evidence that consumers are more responsive in peak hours. The corresponding instantaneous price elasticity, calculated at the sample mean, is -0.0031 and -0.0032 for off-peak and peak periods, respectively, based on column (5).

5.5 Smooth Transition Model

In this section, I generalize the basic linear regression model to a smooth transition model. This model could help to test the conjecture that the demand curve is kinked when the price is high. The rationale behind this conjecture is that KEPCO can relieve the system stress by mandating its large customers to reduce their electricity usage (direct load control) when the market price is high, which is known as interruptible services. Furthermore, large customers in the market could withdraw from the wholesale market when the price is high by starting their on-site generators. These two activities suggest that the wholesale market demand curve might be flatter when the price is high. Thus, there might exist a threshold price above which the demand response increases. To model this situation, based on model (5.1), suppose there exists a cutoff price a such that,

$$\Delta y_t = \begin{cases} \gamma\Delta p_t + \alpha + \sum_{i=1}^{k}\zeta_i\Delta p_{t-i} + \Delta T_t'\theta + D_t'\delta + X_t + \varepsilon_t, & p_t \leq a \\ \gamma_2\Delta p_t + \alpha_2 + \sum_{i=1}^{k}\zeta_i\Delta p_{t-i} + \Delta T_t'\theta + D_t'\delta + X_t + \varepsilon_t & p_t > a \end{cases}$$

$$\Rightarrow \Delta y_t = (\gamma\Delta p_t + \alpha + \sum_{i=1}^{k}\zeta_i\Delta p_{t-i} + \Delta T_t'\theta + D_t'\delta + X_t) + (\widetilde{\gamma}\Delta p_t + \widetilde{\alpha})1_{p_t>a} + \varepsilon_t \qquad (5.4)$$

where $\widetilde{\gamma} = \gamma_2 - \gamma$ and $\widetilde{\alpha} = \alpha_2 - \alpha$.

Suppose customers adjust grid demand smoothly on two states. It is possible that w-holesale buyers may not be able to perfectly fine-tune their usage. The indicator function can be elaborated by a logistic transition function, $G(p_t;\eta,a) = 1/(1+\exp^{-\eta(p_t-a)})$, such that,

$$\Delta y_t = (\gamma\Delta p_t + \alpha + \sum_{i=1}^{k}\zeta_i\Delta p_{t-i} + \Delta T_t'\theta + D_t'\delta + X_t) + (\widetilde{\gamma}_2\Delta p_t + \widetilde{\alpha}_2)G(p_t;\eta,a) + \varepsilon_t \qquad (5.5)$$

η determine the smoothness of the transition from one stage to the other, and there are two limiting cases: when $\eta \to \infty$, $G(p_t;\eta,a)$ tends to be an indicator function; and when $\eta = 0$,

$G(p_t;\eta,a)=1/2$, the model collapses to a homogeneous linear regression model.

In practice, since p_t is endogenously determined, for simplicity, I instead apply p_{t-1} in the indicator function due to their high correlation.

Before estimating the model, I should first test whether, empirically, the kink point is relevant or not. The corresponding hypothesis is presented by,

$$
\begin{aligned}
H_0 :& \widetilde{\gamma}_2 = \widetilde{\alpha}_2 = 0 \\
H_a :& \widetilde{\gamma}_2 \neq 0 \text{ or } \widetilde{\alpha}_2 \neq 0
\end{aligned}
$$

Because model (5.4) is unidentified under the null, I instead adopt model (5.5) and revise the hypothesis as,

$$
\begin{aligned}
H_0 :& \eta = 0 \\
H_a :& \eta > 0
\end{aligned}
$$

Although model (5.5) is still unidentified under the null, by the Talor expansion of $G(\cdot)$ under the null, an auxiliary regression model can be derived for testing.

The Taylor expansion of $G(\cdot)$ under the null is,

$$
\Delta y_t = \ (\gamma^*\Delta p_t + \textstyle\sum_{i=1}^{k}\zeta_i\Delta p_{t-i} + \Delta T_t'\theta + D_t'\delta + X_t) + p_{t-1}(\widetilde{\gamma}_2^*\Delta p_t) + p_{t-1}^2(\widetilde{\gamma}_2^{**}\Delta p_t) + u_t
$$

where $u_t = e_t + R(p_{t-1};\eta,a)$, and γ^*, $\widetilde{\gamma}_2^*$ and $\widetilde{\gamma}_2^{**}$ are multiples of η.

The hypothesis is further revised as,

$$
\begin{aligned}
H_0 :& \widetilde{\gamma}_2^* = \widetilde{\gamma}_2^{**} = 0 \\
H_a :& \widetilde{\gamma}_2^* \neq 0 \text{ or } \widetilde{\gamma}_2^{**} \neq 0
\end{aligned}
$$

I adopt test statistics $\mathit{\Gamma}$ from Massacci (2011) to account for endogenity, which is a revised

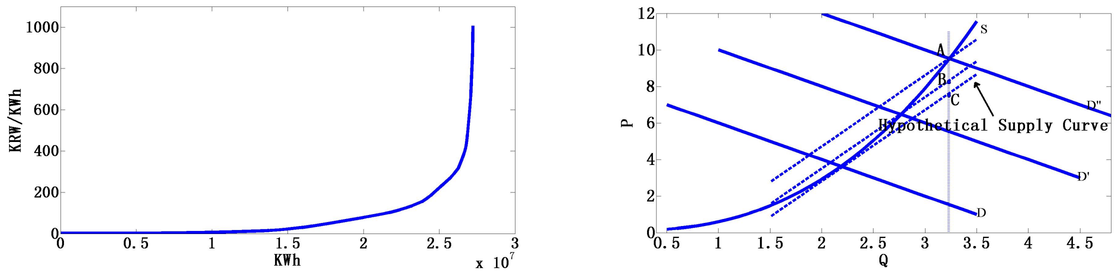
LM test. This test statistic is equal to 13.4945, which is slightly greater than $\chi^2_{.99,4} = 13.2676$, and thus the null is rejected at the 1% significance level and a nonlinear specification is suitable based on the test.

I am aware, however, that there might be some possible problems with the test. As pointed out by Massacci (2011), the power of the test tends to be lower the smoother the transition function. It also depends on the magnitude of the nonlinear effect, and it tends to be higher when the slope coefficient varies across regimes than when it does not. Moreover, the test requires correct specification of the model.

Parameters η and a in the indicator function are estimated based on a grid search based on the criterion of minimized mean square error and given η and a, the coefficients are estimated by iterated GMM. In order to maintain identification, I assume η is large in the grid search method. Based on grid search, the transition is sharp as $\hat{\eta} = 205.7$, which reduces the transition function to an indicator function. $\hat{a} = 32.6785$, which is about the 1.5th percentile of the price distribution. The price coefficient in the first regime (price less than 32.6785) is -6.2971 (2.5593), and the price coefficient with the dummy in the second regime is 4.7255 (2.3211); both estimates are significant at the 5% level. Thus, the price coefficient in the second regime is -1.5716 with the corresponding instantaneous elasticity -0.0034; the same as the estimates in the linear case with four decimals. Moreover, the number of the data are limited at low prices, and thus this kink result does not provide much credibility.

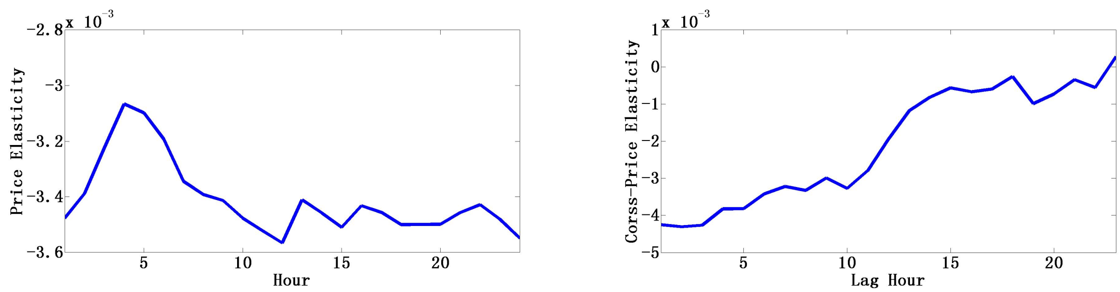
I further test the model if there exists more regime. I cannot, however, reject the null that there is no more kink point. The conjecture that the demand curve is kinked when the electricity price is high cannot be established.

5.6 Figures and Tables



(a) (b)

Figure 5.1: Identification illustration



(a) (b)

Figure 5.2: (a) Electricity real-time price elasticity for each hour within a day; (b) Cross-price elasticity for lagged hours.

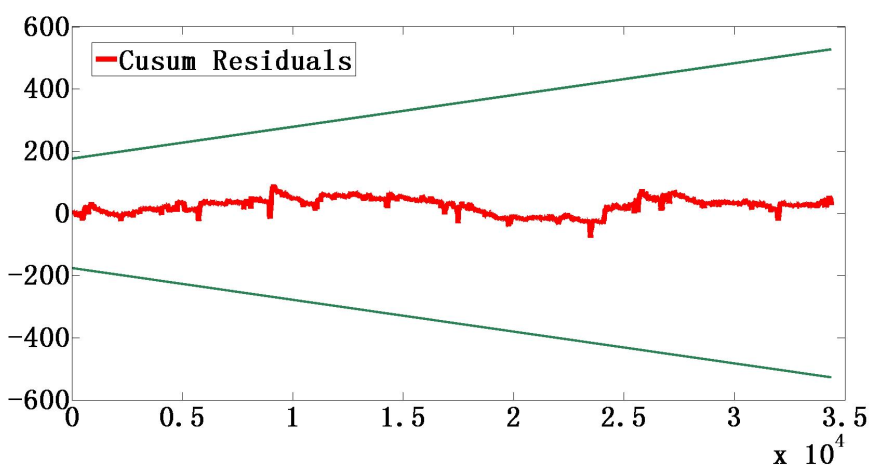


Figure 5.3: CUSUM test for models (5.1) in column (2) in Table 5.3.

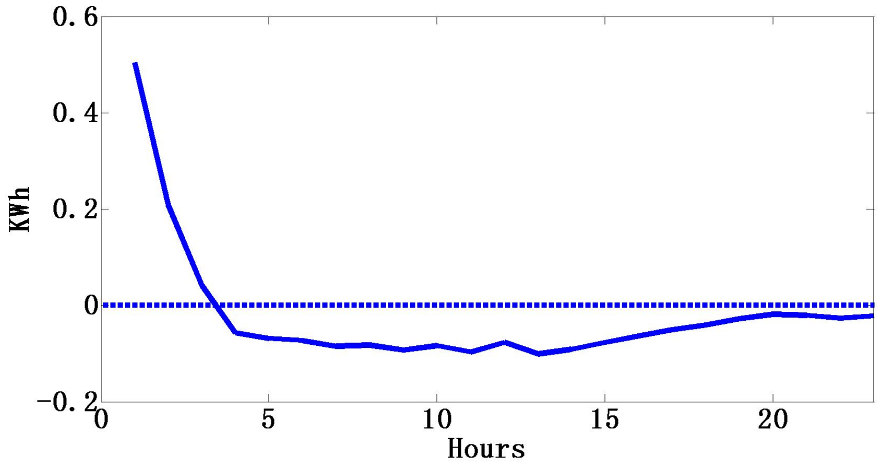


Figure 5.4: Consumption impulse response plot

Table 5.1: Dummy variables for calendar effect

D_t^1	Saturdays (if not a holiday)
D_t^2	Sundays (if not a holiday)
D_t^3	Holiday (Independence Day, Children's Day, Memorial Day, Buddha's Birthday, Constitution Day, Liberation Day, National Foundation Day)
D_t^4	Holiday (New Year's Day, Spring Festival, Moon Festival, Christmas Day)
D_t^5	Days before important holidays (if not weekend)
D_t^6	Days after important holidays (if not weekend)
D_t^7	Mondays (if not a holiday or days before and after important holiday)
D_t^8	Year indicators

Table 5.2: Coefficient of variation

	Coefficient of variation $100 * s/\bar{x}$
SMP	20.82
Temperature	35
LNG, oil generator capacity	3.5
Base generator capacity	5.4
Load	12.19

Source: Author's own calculation based on data from 2007.

Table 5.3: Estimates of model 5.1

	(1)		(2)		(3)		(4)	
Δp_t	0.2582**	(0.1316)	−1.5999**	(0.6859)	−2.0202***	(0.6679)	−1.4771***	(0.5923)
ΔHDD_t	−11.9286***	(4.2593)	−12.6952***	(4.3000)	−13.3244***	(4.2977)	−12.5164***	(4.2692)
ΔHDD_t^2	-0.0104	(0.0442)	−0.0024	(0.0446)	0.0028	(0.0446)	−0.0043	(0.0443)
ΔCDD_t	−28.8779***	(3.5077)	−29.5158***	(3.5329)	−30.0194***	(3.5316)	−29.3955***	(3.5156)
ΔCDD_t^2	0.2622***	(0.0297)	0.2668***	(0.0298)	0.2701***	(0.0298)	0.2661***	(0.0297)
D_t^1	-48.2204	(46.0964)	−42.8501	(45.8959)	-42.4767	(45.9053)	-42.3836	(45.8748)
D_t^2	−141.8640***	(39.8780)	−138.0514***	(39.9777)	−136.9715***	(40.0347)	−137.9790i***	(39.9599)
D_t^3	−110.9556***	(28.0600)	−106.1475***	(27.9496)	−105.7374***	(27.9543)	−106.1442***	(27.9483)
D_t^4	−166.8648***	(39.1226)	−162.6396***	(38.9215)	−157.9896***	(38.8745)	−162.3632***	(38.9153)
D_t^5	−93.5442***	(30.1226)	−93.9165***	(31.4170)	−92.6536***	(30.4383)	−93.8401***	(30.4090)
D_t^6	304.6848***	(35.4851)	309.5426***	(35.6572)	312.3413***	(35.7007)	309.1132***	(35.6192)
D_t^7	143.7125**	(66.7958)	147.5915**	(66.4501)	146.2062**	(66.4394)	147.6196**	(66.4531)
$D_t^8(2008)$	0.3393	(4.8766)	0.6521	(4.8731)	0.5801	(4.8787)	0.6555	(4.8716)
$D_t^8(2009)$	0.0833	(4.9196)	0.2338	(4.9110)	0.2074	(4.9147)	0.2556	(4.9096)
$D_t^8(2010)$	-0.0725	(5.0872)	0.0136	(5.0835)	-0.2142	(5.0879)	0.0248	(5.0820)
IV1[a]			√		√		√	
IV2[b]					√			
IV3[c]							√	
R^2	0.9685		0.9683		0.9680		0.9683	

Source: Author's estimations using the data from 2006 to 2010.
Standard errors in parentheses are heteroscedastic robust.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.
[a] IV1: Monthly (28 days) lagged electricity price.
[b] IV2: Yearly (365 days) lagged electricity price.
[c] IV3: Two-monthly (56 days) lagged electricity price.

Table 5.4: Diagnostic test for the model in columns (2) - (4) in Table 5.3

	(2)	(3)	(4)
Endogeneity test statistics	7.787***	12.314***	9.213***
Hansen J statistics	-	6.465**	0.127
Cumby-Huizinga test statistics[a]	5.9895	6.0914	6.1986
Pagen Hall test statistics	6271.228***	6263.556***	6275.194***
Ramsey Reset test statistics	0.43	0.51	0.42

Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

[a] Test parameters are set as $q = 0$, $s = 3$.

Table 5.5: Real-time price elasticity[a]

	Elasticity	Mean load (MWh)	Mean price(Kwon/MWh)
2007	-0.0030	44067	82.1062
2008	-0.0040	47701	117.9824
2009	-0.0033	49043	102.5536
2010	-0.0035	53571	116.8262
07-10	-0.0034	49094	104.8761

[a] Price elasticities are calculated based on the sample mean based on the results in column (2) in Table 5.3 and four decimals are reserved for presentation.

Table 5.6: Real-time price elasticity[a]

Hour	1	2	3	4	5	6	7	8	9	10	11	12
	-0.0035	−0.0034	-0.0032	-0.0031	-0.0031	-0.0032	-0.0033	-0.0034	-0.0034	-0.0035	-0.0035	-0.0036
Hour	13	14	15	16	17	18	19	20	21	22	23	24
	-0.0034	-0.0035	-0.0035	-0.0034	-0.0035	-0.0035	-0.0035	-0.0035	-0.0035	-0.0034	-0.0035	-0.0035

[a] Price elasticities are calculated based on sample mean based on the results in column (2) Table 5.3 and four decimals are reserved for presentation.

Table 5.7: Cross-price elasticities[a]

A. Model 5.1												
Lag hour	1	2	3	4	5	6	7	8	9	10	11	12
	-0.0043	-0.0043	-0.0043	-0.0038	-0.0038	-0.0034	-0.0032	-0.0033	-0.0020	-0.0033	-0.0028	-0.0019
Lag hour	13	14	15	16	17	18	19	20	21	22	23	
	-0.0012	-0.0008	-0.0006	-0.0007	-0.0006	-0.0003	-0.0010	-0.0007	-0.0003	-0.0006	0.0003	
B. Model 5.1 without lag consumption variables within a day												
Lag hour	1	2	3	4	5	6	7	8	9	10	11	12
	-0.0039	-0.0037	-0.0045	-0.0045	-0.0043	-0.0041	-0.0044	-0.0044	-0.0049	-0.0047	-0.0052	-0.0048
Lag hour	13	14	15	16	17	18	19	20	21	22	23	
	-0.0039	-0.0037	-0.0033	-0.0035	-0.0035	-0.0032	-0.0037	-0.0034	-0.0032	-0.0034	-0.0023	

[a] Price elasticities are calculated based on sample mean based on the results in column (2) in Table 5.3 and four decimals are reserved for presentation.

Table 5.8: Results with different cyclical regularity specifications

	(1)	(2)	(3)	(4)
Δp_t	-1.5999^{**}	-1.4660^{*}	-1.8483^{**}	-1.9399^{***}
	(0.6859)	(0.7718)	(0.7629)	(0.6882)
Long-run price elasticity	-0.0640	-0.1788	-0.2094	-0.0804
Trigonometric term	$\surd$	$\surd$		
Binary dummy			$\surd$	$\surd$
Second-order difference term	$\surd$			$\surd$
Cumby-Huizinga test statistics[a]	5.9895	47.6411^{***}	158.7031^{***}	6.7946

Source: Author's estimations using data from 2006 to 2010.
Instrumental variable: one month lagged price variable.
Standard errors in parentheses are heteroscedastic robust.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.
[a] Test parameters are set as $q = 0$, $s = 3$.

Table 5.9: Estimated price coefficients

A. Model 5.1												
Lag hour	0	1	2	3	4	5	6	7	8	9	10	11
	-1.5999^{***}	-1.1853^{***}	-1.0856^{***}	-1.1348^{***}	-1.0332^{***}	-1.1622^{***}	-1.0124^{***}	-1.0720^{***}	-1.2061^{***}	-1.0854^{***}	-1.3322^{***}	-1.1107^{***}
	(0.6859)	(0.3280)	(0.3032)	(0.3056)	(0.3029)	(0.2882)	(0.2912)	(0.3062)	(0.3201)	(0.2988)	(0.3427)	(0.3052)
Lag hour	12	13	14	15	16	17	18	19	20	21	22	23
	$-0.8620{*}{*}{*}$	-0.7865^{***}	-0.8272^{**}	-0.7925^{***}	-0.9057^{***}	-0.8486^{***}	-0.7169^{**}	-1.1316^{***}	-0.8233^{***}	-0.7239^{***}	-0.9182^{***}	-0.4828^{**}
	(0.2938))	(0.3033)	(0.3393)	(0.2962)	(0.2963)	(0.3070)	(0.3126)	(0.2836)	(0.2755)	(0.2558)	(0.2567)	(0.2136)
B. Model 5.1 without lag consumption variables within a day												
Lag hour	0	1	2	3	4	5	6	7	8	9	10	11
	-2.1221^{***}	-1.3722^{***}	-1.7186^{***}	-2.1269^{***}	-2.1099^{***}	-2.0162^{***}	-1.9154^{***}	-2.0701^{***}	-2.2806^{***}	-2.1998^{***}	-2.4308^{***}	-2.2401^{***}
	(0.8006)	(0.3865)	(0.3591)	(0.3642)	(0.3617)	(0.3376)	(0.3495)	(0.3687)	(0.3904)	(0.3631)	(0.4052)	(0.3509)
Lag hour	12	13	14	15	16	17	18	19	20	21	22	23
	$-1.9247{*}{*}{*}$	-1.8262^{***}	-1.7242^{**}	-1.5441^{***}	-1.6441^{***}	-1.6606^{***}	-1.4776^{***}	-1.7095^{***}	-1.6007^{***}	-1.4951^{***}	-1.5836^{***}	-1.0801^{***}
	(0.3451))	(0.3457)	(0.3974)	(0.3487)	(0.3507)	(0.3648)	(0.3722)	(0.3438)	(0.3241)	(0.3059)	(0.3124)	(0.2518)

Source: Author's estimations using the data from 2006 until 2010 by GMM with one month lag price instrument.
Instrumental variable: One month lagged price variable.
Standard errors in parentheses are heteroscedastic robust.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

Table 5.10: Correlation of coefficients of the equations in SUR model

	3am	4am	11am
3am	1		
4am	0.9532	1	
11am	0.2317	0.3243	1

Source: Author's estimations using data from 2006 to 2010.

Table 5.11: Wald test of the coefficients of the holidays and weekend dummies

D_t^1 v.s. D_t^2	D_t^3 v.s. D_t^4	D_t^1 v.s. D_t^3	D_t^1 v.s. D_t^4	D_t^2 v.s. D_t^4	D_t^2 v.s. D_t^4
4.25**	2.03	3.12*	6.66***	0.93	0.44

[a] Wald tests are applied to the results in column (2) in Table 5.3.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

Table 5.12: Percentage of consumption reduction/increasing on different types of the day compared with average hourly consumption

D_t^2	D_t^3	D_t^4	D_t^5	D_t^6	D_t^7
−0.29%	−0.22%	−0.34%	−0.20%	0.65%	0.31%

Source: Author's own calculation based on the results in column (2) in Table 5.3.

Table 5.13: Estimated price coefficients for hypothesis 1

Δp_t	Cooling-degree day dummy$*\Delta p_t$
−1.6182*	0.1250
(0.6394)	(0.9468)

Source: Author's estimations using the data from 2006 until 2010.
Instrumental variable: One month lagged price variable.
Standard errors in parentheses are heteroscedastic robust.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

Table 5.14: Estimated price coefficients for hypothesis 2

(1)		(2)	
Δp	2008.8-10 dummy$*\Delta p_t$	Δp	2008-10 dummy$*\Delta p_t$
-3.8163^{***}	2.2998^{**}	-6.7672^{***}	5.2151^{**}
(1.2682)	(0.9940)	(2.3094)	(2.0555)

Source: Author's estimations using data from 2006 to 2010.
Instrumental variable: One month lagged price variable.
Standard errors in parentheses are heteroscedastic robust.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

Table 5.15: Average price for peak hour candidates

	(1)	(2)	(3)	(4)	(5)
Peak	116.1858	116.3560	115.9056	116.5405	115.2347
Off	99.9425	102.3178	103.3507	102.7870	97.2127

Source: Author's own calculation based on data from 2006 to 2010.

Table 5.16: Estimated price coefficients for hypothesis 3

	(1)	(2)	(3)	(4)	(5)
Δp_t	-1.4588^{**}	-1.2410^{*}	-1.6647^{**}	-1.4654^{**}	-1.4669^{**}
	(0.6929)	(0.7284)	(0.7410)	(0.6622)	(0.6692)
Peak dummy$*\Delta p_t$	-0.4652	-1.2401	0.2212	-1.0629	-0.4464
	(0.8731)	(0.8779)	(0.7981)	(1.1246)	(0.8780)

Source: Author's estimations using data from 2006 to 2010.
Instrumental variable: one month lagged price variable.
Standard errors in parentheses are heteroscedastic robust.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

Chapter 6

Electricity Demand in the Retail Market

6.1 Residential Sector

6.1.1 Basic Empirical Framework

In the residential sector, the tariff has an increasing-step pricing schedule with a monthly billing cycle. Based on the results from Ito (2012), I assume that average household consumption responds to the monthly average price. One drawback of the aggregated data is that aggregation masks the billing cycle of different households. Empirically, I apply the calender month as the billing month, and consider the following estimating equation with a log-log specification:

$$\log(y_t) = \mu + \gamma_1 \log(p_t) + \gamma_2 \log(p_{t-1}) + \alpha \log(I_t) + \beta \log(p_t^{gas}) + \theta_1 \log(CDD_t) + \theta_2 \log(HDD_t) + D_t'\delta + \epsilon_t \quad (6.1)$$

where y_t is the average daily consumption in unit KWh per household at month t; p_t is the average electricity price among all households in Korean won per KWh at month t; p_t^{gas} is the average natural gas price for households in Korean won per MWh GCV at month t; I_t is the average income per household in million Korean won at month t; CDD_t and HDD_t are defined as cooling- and heating-degree days; and CDD_t (HDD_t) equals the monthly average temperature in degrees F when it is greater (smaller) than 75° F[1] and zero otherwise at month t. D_t is the vector of dummy variables, including a yearly dummy, winter and summer vacation dummies (Jan., Feb., Dec., and August), and a national holiday dummy

[1]The choice of the breaking point, 75 degrees F is based on Figure 4.10, which plots the relationship between monthly consumption per household per day and average temperature. The breaking point of the plot is around 75 degrees F.

which indicates a specific month that has at least three days of national holidays;[2] and ϵ_t is the unobservable random disturbances. All nominal price and income data are deflated by the CPI with 2010 = 100. Income and natural gas price data are interpolated by cubic spline to monthly frequency.

Electricity demand is a derived demand, depending on households' utilization behaviors and electrical appliance stocks. Changes in utilization behaviors and appliance stocks can be caused by two channels. One is induced by household income and the price of electricity and its substitutes. The other is induced by other exogenous factors, such as households' lifestyles. In this sense, model (6.1) could be viewed as a reduced form of electricity demand model. The price and income coefficients not only capture the household utilization behaviors but also absorb the effects of any adjustments to appliance stocks induced by them on electricity consumption. Temperature, holiday, and other dummies are used to control the exogenous factors. In Section 6.1.2, I will introduce a Structural Time Series Model (STSM) to better control the changes in utilization behaviors and appliance stocks induced by exogenous factors.

The identification of the electricity price coefficients is substantiated by the nonlinear price variations in the system. These variations result from the six increasing-step tariff schedule in the residential sector, as shown in Figure 3.1, and monthly household consumption fluctuations. Because of these two factors, the contemporaneous monthly average price is endogenously determined. An instrumental variable (IV) is thus needed to correct the simultaneous bias asymptotically.

In practice, the price variable is computed as monthly average revenue from the aggregate revenue and consumption data provided by KEPCO. I am concerned that this arrangement might introduce measurement error in the model for two reasons: (1) aggregate revenue data might include other miscellaneous payments, such as delivery charges, and thus average

[2]National holidays includes New Year's Day, Spring Festival, Moon Festival, Christmas Day, Independence Day, Children's Day, Memorial Day, Buddha's Birthday, Constitution Day, Liberation Day, and National Foundation Day.

revenue would serve as a proxy of the true consumption weighted average price; (2) these data might involve calculation errors, since KEPCO points out in its release that "the sales figures of KEPCO have been prepared based on interval estimates of KEPCO for your convenience only, and have been neither audited nor reviewed by KEPCO's independent accountants, Samjung & Co. or any other accountants." Therefore, measurement error would be another source of an endogeneity problem in the model. In detail, the difference between the true average price and average revenue is absorbed in the residuals and this could possibly lead to bias in the electricity price coefficients, and this bias requires not only an instrument for the current average but also for the lagged price used in the model. Standard econometric theory indicates that classical measurement error[3] leads to a bias towards zero (Greene, 2008, p. 325). Therefore, if the measurement error is solely caused by reason (2), then the price coefficients tend to underestimate the price elasticity in absolute value. If, however, the measurement error is also caused by reason (1), the direction of bias would depend on the correlation between the measurement error and the electricity average price variable.

To solve the endogeneity problem, I assume that there is no autocorrelation in ϵ_t. This assumption allows me to apply the yearly lagged average price as possible IVs for price in the model. The underlying IV rationale is that y_{t-12} is correlated with y_t due to the yearly consumption cycle but not ϵ_t, and with a relatively stable tariff schedule, $p(y_{t-12})$ can be a good predictor of $p(y_t)$. To make sure the proposed instruments are valid, I apply a standard F-test in the first stage regression regarding the two endogenous electricity price variables and the Cragg-Donald Wald F-test. The standard F statistics are 86.39 and 43.15 for the current and lagged price, respectively. As a rule of thumb, an F-statistic over 10 is required to suggest IVs are sufficiently strong,[4] (Staiger and Stock 1997, p.557) and both the F statistics satisfy this criterion. The Cragg-Donald Wald F statistic[5] is 11.77, which

[3]Classical measurement error refers to the error that has a mean of zero and is independent of the true regressor and the economic error disturbance in the regression.

[4]Staiger and Stock (1997) show that the weak instruments problem can arise even when the first-stage t- and F-tests are significant at conventional levels in a large sample.

[5]The Stock-Yogo weak ID test critical values for two endogenous variables and two instruments: 10%

is greater than the 10% maximal IV size, suggesting that I can reject the null hypothesis that the equation is weakly identified based on this criterion, and thus the proposed IVs are relevant.

Furthermore, I conjecture that, in aggregate, households respond to the previous monthly price rather than the current price. This is because households generally know their exact monthly bill at the end of the billing cycle without special effort, but they still need to struggle to keep track of their usage within the billing cycle to estimate their current average monthly price. The gain associated with tracking might be far less than the cost of effort, the so-called cognitive cost. Therefore, households would tend to respond to the easily obtained previous monthly bill to adjust their current electricity usage. KEPCO, however, offers an on-line application that calculates the electricity payments for households automatically once they have provided the total electricity usage. This application can be viewed as a way to reduce a household's cognitive cost. Accordingly, against this background, it is not obvious to which price households do respond: lagged or contemporaneous. Therefore, I will apply an encompassing test to examine which price should be used to estimate the price effect. An encompassing test is used to investigate if one model encompasses an alternative model. In my case, if households respond to the monthly lagged price rather than the contemporaneous average price, I expect $\gamma_1 = 0$ because the contemporaneous price would not affect demand conditional on the effect of lagged prices. Distinguishing current and lag prices is instrumental for model specification: the price elasticity would be biased based on a model with only contemporaneous price if households actually respond to lagged prices.

Empirical Results and Implications

Table 6.1 presents the estimation results of the variables of interest in model (6.1). Results from the linear specification are also listed. Columns (1) and (3) are the results without

maximal IV size 7.03; 15% maximal IV size 4.58; 20% maximal IV size 3.95; 25% maximal IV size 3.63. Source: Stock-Yogo (2005).

IVs estimated by OLS, and columns (2) and (4) are the results with IVs estimated by GMM. In columns (2) and (4), the estimated current electricity price coefficient is highly insignificant while the estimated lag electricity price coefficient is significant at 5% with a theoretically sound sign. The corresponding price elasticities are -0.2720 and -0.4585 (the latter is calculated at the sample mean) respectively, which are quite inelastic. Although the magnitude of the price elasticity in absolute value is larger in the linear specification, they are not significantly different. The estimated income effects are with the expected signs and significant at the 5% level in all columns. The corresponding income elasticities are 1.1043 and 1.4554 (the latter is calculated at the sample mean) in columns (2) and (4) respectively. Thus, households, in aggregate, are more responsive to changes in income than to changes in electricity prices.

The estimates of substitution effects, however, are all with unexpected signs and significant at the 5% level. These results violate the standard economic theory that substitution effects should be positive. I conjecture that some decreasing trend effects might confound these estimates, leading to negative substitution effects and will further investigate this effect. I will investigate this conjecture in Section 6.1.2.

The cooling- and heating-degree days effects are all negative and significant at the 1% level, indicating that households tend to reduce electricity (i.e., electrical heating) usage when the temperature increases. As seen in Figure 4.10, temperature seems have nonlinear effects on electricity consumption, and when the temperature is greater than the breaking point, temperature and consumption seem to change in the same direction. I therefore apply the Wald test to test whether temperature has heterogeneous effects in these two types of days, with the null hypothesis being $\theta_1 = \theta_2$. The test statistics are 0.39 and 0.42 for the model in columns (2) and (4) respectively, suggesting that it is possible to accept the null hypothesis of homogeneous temperate effects.[6] The temperature elasticity is around -0.32 for both cooling- and heating-degree days and the two specifications. The reason for not

[6]The corresponding p-values are 0.5315 and 0.5180.

detecting nonlinear effects might be the small sample size with the high temperatures; there are only three data points when the temperature is greater than 80 degrees F in the sample.

The sign for the holiday dummy is positive and significant at the 5% level in all columns. This result implies that most people are off-work during holidays and then spend more time at home, causing electricity usage to increase. The estimates indicate that households tend to increase their daily electricity use by about 6% or around 0.65 KWh in the month with at least three days of national holidays. I also test the effects with other numbers of holidays during a month: the month with at least one and two days of national holidays, respectively. The coefficient of the former is -0.0145 (0.0159), and the latter 0.0153 (0.0145); both are highly insignificant, indicating that months with fewer than three days of holidays do not significantly affect households' utilization behaviors in aggregate. Generally, the most important holidays, such as Lunar New Year and Moon Festival, give people three days off. Families tend to get together and celebrate and this may well lead to the increase in electricity usage.

Moreover, the significance of the lagged price coefficient is in accordance with my conjecture that households respond to monthly lag electricity price. To further confirm this conjecture, I apply an encompassing test for these two model specifications, shown in Table 6.2. All estimations are with instruments. For model (6.1), column (1) lists the results that only the contemporaneous price is included in the regression, and the corresponding price elasticity is significant at the 1% level. However, the encompassing test in column (3), including both the lagged and contemporaneous prices, the same as the results in column (2) in Table 6.1, shows that the current price becomes highly insignificant when the lagged price is in control. Columns (2) and (3) indicate that, once the lagged price is included, adding contemporaneous price does not significantly change its effect. The results for the linear model in columns (4)-(6) yield similar conclusions. Therefore, statistically, there is empirical evidence that households, in aggregate, respond to monthly lagged price rather than the contemporaneous price.

Table 6.3 presents the statistics for different diagnostic tests. Columns (1)-(6) are corresponding to the models as in the same columns in Table 6.2. First, all statistics of the endogeneity test are significant at the 10% level, suggesting that the current and lagged price variables suffer from an endogeneity problem and the instruments are not redundant. Second, for the Cumbly-Huizinga autocorrelation test,[7] all statistics are not significant at the 5% level except in column (5), which indicates that I cannot reject the null at the 10% level. The log-log specification seems more robust to autocorrelation. Generally, autocorrelation is not a big concern in the model, justifying the use of yearly lags as instruments in the model. Fourth, all the statistics of the Pagan-Hall heteroscedasticity test suggest that it is possible to be unable to reject the null of homoscedasticity.

I also apply the Hansen J test to test the exogeneity of the IVs: a rejection of the null for the Hansen J test would cast doubt on the validity of the instruments. This test served as a supplement to the autocorrelation test, to make sure the orthogonal conditions hold for the instruments. Since the Hansen J test cannot be applied in the exact identification cases, instead, I apply yearly lag IVs for the both current and lagged price variables to instrument the models with only one price variable, as the models in columns (1)-(2) and (4)-(5) in Table 6.2. Results are listed in Table 6.4. Columns (1)-(2) and (4)-(5) correspond to the models in the same columns in Table 6.2. I cannot reject the null hypothesis for the test statistics in columns (2) and (5) at the 10% level, in column (1) at the 5% level, and column (4) at the 1% level. The results in columns (1) and (4) are weaker than the results in the other columns, but are acceptable. Still, the validity of the IVs is better in the log-log specification.

Generally, the results in model (6.1) are robust to linear specification. Figure 6.1 provides CUSUM test plots for model (6.1) in column (3) in Table 6.2. The CUSUM test is used to check the stability of the coefficients over time and the shape of the two plots suggests that the model is stable within a 95% confidence band.

[7] $s = 1$, $q = 3$.

Furthermore, I introduce a variation in model (6.1). Homogenous temperature effects allow me to replace the $\log(CDD_t)$ and $\log(HDD_t)$ variables by their combination, denoted as $\log(Temp_t)$. I then add the $\log(Temp_t)^2$ variable in the model to test whether there exist quadratic temperature effects on household electricity usage by testing whether the coefficient for $\log(Temp_t)^2$ is significantly different from zero. Results are presented in Table 6.5. The estimated quadratic effects are statistically insignificant, implying that temperature affects household electricity usage linearly in the log-log specification.

Last but not least, the empirical model should be interpreted as capturing the equilibrium demand relationship, and thus the price elasticity is described as long-run elasticity with a couple of months' window in the time-series viewpoint.

6.1.2 Structural Time-Series Model

In this section, I relax the basic model specification by allowing the intercept coefficient in the basic model to evolve stochastically over time under the framework of a structural time-series model, as a way to better control electricity usage trend effects associated with possible exogenous factors not controlled in the model. On the other hand, Leamer (1983) advocates conducting sensitivity analysis to help diagnose misspecification, and the "fragility" of regression coefficient estimates is indicative of specification error. Therefore, the study in this section could be also viewed as a simple sensitivity analysis in line with Leamer (1983) to check the robustness of the estimates.

I revise model (1) as the following,

$$\log(y_t) = \mu_t + \gamma_1 \log(p_t) + \gamma_2 \log(p_{t-1}) + \alpha \log(I_t) + \beta \log(p_t^{gas}) + \theta \log(Temp_t) + D_t'\delta + \epsilon_t \quad (6.2)$$

where μ_t is now called the stochastic trend component.

μ_t is then expressed as

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \quad (6.3)$$

$$\beta_t = \beta_{t-1} + \zeta_t \quad (6.4)$$

where $\eta \sim N(0, \sigma^2_\eta)$ and $\zeta_t \sim N(0, \sigma^2_\zeta)$. Equations (6.3) and (6.4) represent the level and the slope of the trend, respectively. The exact form of the trend depends on the size of the variances (referred as the hyperparameters), σ^2_η and/or σ^2_ζ. In other words, η_t shifts the level of the trend up and down and ζ_t allows the slope to change. Model (6.1) in the last section is actually a special case here as both σ^2_η and σ^2_ζ equal zero.

μ_t can be interpreted as a proxy capturing the underlying structure or Underlying Energy Demand Trend (UEDT), as it is defined in the literature, in the residential sector. Structural factors include living standards and dimensions, household demography, lifestyles, habits, energy usage standards, etc., which can be assumed to be exogenous factors that influence household electricity consumption. For instance, a higher standard of living might involve more usage of electrical appliances, and thus increase electricity consumption, while, on the other hand, environmental concerns involve adopting efficient appliances, and thus reduce electricity consumption. Generally speaking, compared to model (6.1), stochastic μ_t could better control the effects of the exogenous factors imposed on household electricity usage either through changes in appliance stocks or utilization behaviors.

Since price variables are endogenous, I apply the LIML estimation method in STAMP 8.3. The results are presented in Table 6.6 with four different structure specifications. The presented component variables are estimated at the end period. The slope estimates, $\hat{\beta}_t$, are all negative, indicating a downward consumption trend. The estimated coefficients of contemporaneous price, except for the *fixed level/stochastic slope* specification in column (1), are highly insignificant, while the estimated coefficients of the lagged price are all significant at the 1% level, with expected signs for all specifications. The estimated price elasticities are generally more responsive than the results of model (6.1) in column (2) in Table 6.1, but are not significantly different.

The estimated coefficients for natural gas are all insignificant once the underlying demand trend is controlled, confirming my conjecture that the significant negative estimated substitution coefficients confound the trend effects for all results in the last section. I conclude that natural-gas price does not statistically affect household electricity usage. Generally, changing the type of energy usage requires a change in the type of appliance. I conjecture that the lack of evidence of substitution effects indicates that the gain from using a relatively cheaper form of energy is less than the cost of adjusting the type of appliance.

The rest of the estimated coefficients are all with similar size and sign as the results in the last section. Thus, except for natural gas, the estimated coefficients of model (6.1) are robust to STSM specification.

Table 6.7 provides statistics of the diagnostic tests. Based on the AIC and BIC criteria, models with *stochastic level/fixed slope* and *stochastic level/stochastic slope* specifications are preferred to the other two. Moreover, the results in these two specifications in Table 6.6 are the same when reserving four decimals except for the price estimates, which are not significantly different. Based on the criterion of parsimony, *stochastic level/fixed slope* is further preferred. Further, for the *stochastic level/fixed slope* specification, autocorrelation and heteroscedasticity are not detected, and the normality standard can be accepted based on the Bowman-Shenton test at the 10% level. The corresponding price and income elasticities are -0.2926 and 1.0930 respectively. Figure 6.2 (a) plots the UEDT from the *stochastic level/fixed slope* model over time throughout the sample period. The UEDT plot seems quite deterministic with a downward-sloping trend. The determinacy can be explained by the estimates of the hyperparameters, $\hat{\zeta}_t$ and $\hat{\eta}_t$, which are nil even when reserving six decimals. Thus, the extremely small variance leads to a trend that is close to deterministic. Figure 6.2 (b) illustrates the Cusum test based on the same model. The shape of the figure suggests that the model is generally stable. I also check the results if I include a time trend in model (6.1), the estimated substitution elasticity is -0.2095 (0.2031), which is insignificant, and the other estimates are not statistically different from the results in column (2) in Table 6.6.

Sa'ad (2009) also suggests that a *stochastic level/fixed slope* specification is the most preferred model based on his dataset[8] with a downward-sloping UEDT plot. He thus concludes that "the improved energy efficiency of household appliances and equipment outweigh the effects of the structural factors and households' lifestyles during the study period"[9] is also confirmed. The estimated UEDT slope at the end period in my result, -0.00002, is, however, much smaller than his, -0.0093, and the UEDT plot is much more deterministic than his. I further investigate why his results are more stochastic with a large absolute value of the slope estimate. Since he does not include a temperature variable in the model, I drop the temperature variable in my model and repeat the STSM exercise. Still, the *stochastic level/fixed slope* model is the most preferred, and the resulting UEDT at the end period is -0.0092, and the UEDT plot in Figure 6.3 (a) becomes more stochastic, similar to Sa'ad (2009). I therefore conjecture that omitting the temperature variable is what leads him to obtain a slope estimate with a larger absolute value and a more stochastic UEDT plot.

Finally, I run the regression of model (6.2) without the income variable to check what will happen to the UEDT plot. Omitting the income variable would lead to omitted-variable bias in the model. The increased electricity consumption due to income effects would be partly absorbed by the price coefficients. Thus, the estimated price coefficient can be viewed as an upper bound of the absolute value of the price elasticity. I adopt the *stochastic level/fixed slope* specification, and the results are shown in Table 6.8. The estimated price elasticity is -0.4432, which is significant at the 5% level. The estimated slope becomes positive and the UEDT plot shows an upward-sloping trend in Figure 6.3 (b). Part of the income effects are also confounded in the UEDT estimates.

[8]He uses 1973-2007 annual data relating to the Korean electricity residential sector.
[9]Op. cit., 5473.

6.1.3 Figures and Tables

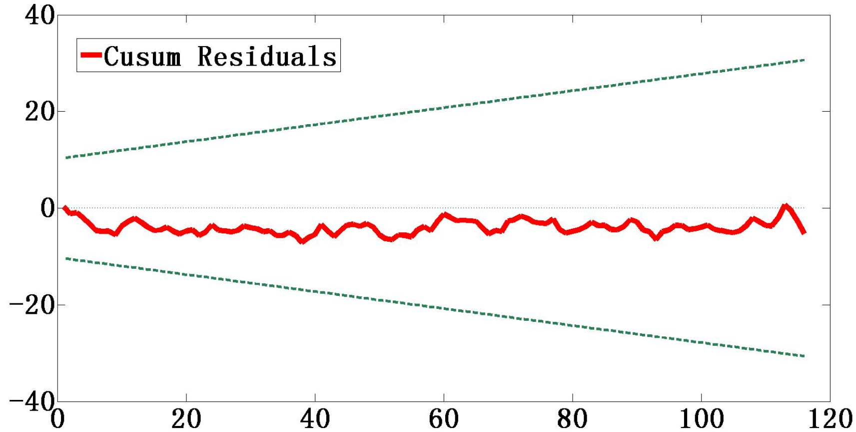
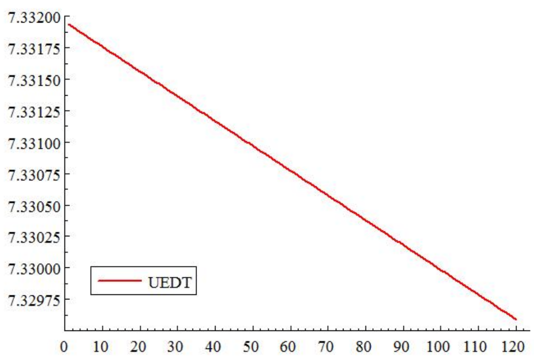
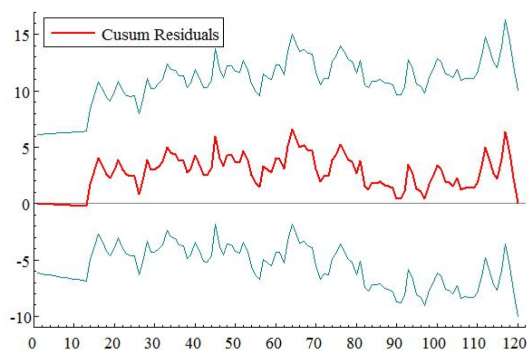


Figure 6.1: CUSUM test for models in column (3) in Table 6.2.



(a)



(b)

Figure 6.2: (a) Plot of the UEDT; (b) Plot of the CUSUM test.

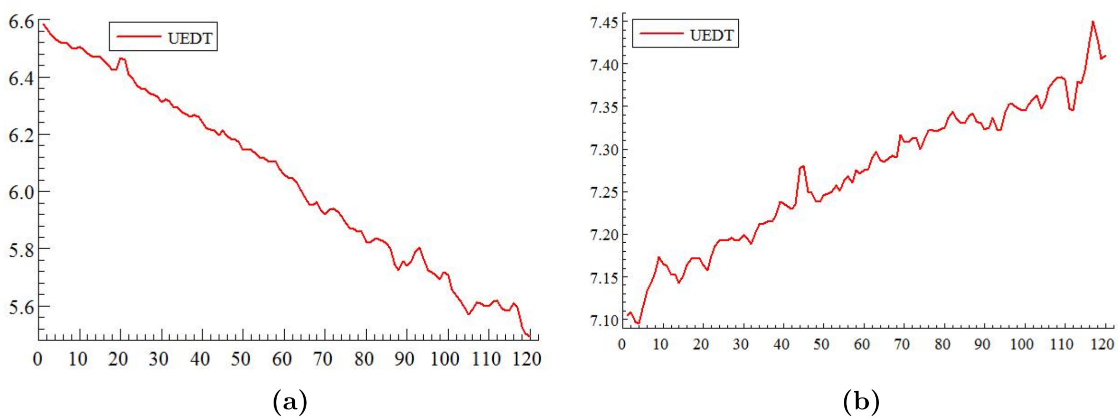


Figure 6.3: (a) Plot of the UEDT of model (6.2) without the temperature variable; (b) Plot of the UEDT of model (6.2) without the income variable.

Table 6.1: Estimates of model (6.1) in residential sector

	Dependent variable: $\log(y_t)$					Dependent variable: y_t			
	(1)		(2)			(3)		(4)	
$\log(p_t)$	0.1279	(0.1360)	−0.2500	(0.2443)	p_t	0.0453***	(0.0132)	0.0396	(0.0332)
$\log(p_{t-1})$	−0.3037***	(0.0975)	−0.2720**	(0.1383)	p_{t-1}	−0.0439*	(0.0106)	−0.0508***	(0.0167)
$\log(I_t)$	0.6335**	(0.2978)	1.1043**	(0.3045)	I_t	4.9292***	(2.8281)	5.3153**	(2.6911)
$\log(p_t^{gas})$	−0.4570**	(0.1850)	−0.3792**	(0.1804)	p_t^{gas}	−0.0118**	(0.0035)	−0.0100***	(0.0032)
$\log(CDD_t)$	−0.3835***	(0.0370)	−0.3194***	(0.0504)	CDD_t	−0.0981***	(0.0078)	−0.0944***	(0.0137)
$\log(HDD_t)$	−0.3753***	(0.0375)	−0.3168***	(0.0487)	HDD_t	−0.0961***	(0.0095)	−0.0923***	(0.0137)
$D_t^{Holiday}$	0.0443**	(0.0215)	0.0524**	(0.0228)	$D_t^{Holiday}$	0.5032	(0.2421)	0.6422***	(0.2391)
R^2	0.9577		0.9123			0.9288	0.9270		

Source: Author's estimations using data from 2003 to 2012.

Columns (1) and (2) are estimated without IVs by OLS; Columns (2) and (2) are estimated with IVs by GMM.

Instrumental variables: Yearly lagged price variables.

Standard errors are in parentheses.

Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

Table 6.2: Encompassing test

	Dependent variable: $\log(y_t)$				Dependent variable: y_t		
	(1)	(2)	(3)		(4)	(5)	(6)
$\log(p_t)$	-0.6314^{***}		-0.2500	p_t	-0.0432^{**}		0.0396
	(0.1556)		(0.2443)		(0.0192)		(0.0332)
$\log(p_{t-1})$		-0.3869^{***}	-0.2720^{**}	p_{t-1}		-0.0334^{***}	-0.0508^{***}
		(0.0806)	(0.1383)			(0.0086)	(0.0167)
R^2	0.8758	0.9058	0.8982		0.8988	0.9207	0.9270

Source: Author's estimations using data from 2003 to 2012.
Instrumental variables: Yearly lagged price variables.
Standard errors are in parentheses.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

Table 6.3: Diagnostic test statistics

(1)	(2)	(3)	(4)	(5)	(6)
Endogenity test					
12.327^{***}	4.433^{**}	8.261^{**}	8.935^{***}	3.416^{*}	6.149^{**}
Cumbly-Huizinga Autocorrelation Test statistics					
3.6206	6.4292^{*}	5.6429	5.3592	8.6526^{**}	5.3592
Pagan-Hall Heteroscedasticity general test statistics					
18.824	22.276	21.327	22.885	22.528	25.238

Standard errors are in parentheses.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

Table 6.4: Test statistics of the Hansen J test

(1)	(2)	(4)	(5)
2.893^{*}	0.993	5.756^{**}	1.253

Columns corresponding the model in the same column in Table 6.3.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

Table 6.5: Regression results with quadratic temperature variable

Dependent variable: $\log(y_t)$		
$\log(p_t)$	-0.1454	(0.2932)
$\log(p_{t-1})$	-0.3160^{**}	(0.1439)
$\log(I_t)$	0.6406^{**}	(0.2865)
$\log(p_t^{gas})$	-0.3615^{**}	(0.1581)
$\log(Temp_t)$	-0.0060^{***}	(0.0016)
$(\log(Temp_t))^2$	-0.0001	(0.0000)
$D_t^{Holiday}$	0.0522^{***}	(0.0158)
R^2	0.9290	

Source: Author's estimations using data from 2003 to 2012.
Instrumental variables: Yearly lagged price variables.
Standard errors are in parentheses.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

Table 6.6: Structural time-series model results

	Dependent variable: $\log(y_t)$			
	(1)	(2)	(3)	(4)
$\log(p_t)$	-0.2048^{*}	-0.1669	-0.1628	0.0763
	(0.1102)	(0.1217)	(0.1190)	(0.1175)
$\log(p_{t-1})$	-0.3454^{***}	-0.2926^{***}	-0.3503^{***}	-0.3839^{***}
	(0.0724)	(0.1016)	(0.0736)	(0.0950)
$\log(I_t)$	1.1264^{***}	1.0930^{***}	1.0930^{***}	1.0876^{***}
	(0.3133)	(0.3114)	(0.3114)	(0.2550)
$\log(p_t^{gas})$	-0.1974	-0.2089	-0.2089	-0.2125
	(0.2211)	(0.2194)	(0.2194)	(0.1346)
$\log(Temp_t)$	-0.3296^{***}	-0.3382^{***}	-0.3382^{***}	-0.3387^{***}
	(0.0342)	(0.0342)	(0.0342)	(0.03377)
$D_t^{Holiday}$	0.0599^{***}	0.0585^{***}	0.0585^{***}	0.0584^{***}
	(0.0181)	(0.0180)	(0.0180)	(0.0179)
Hyperparameters				
Irregular	0.0031	0.0031	0.0031	0.0003
Level	0.0000	0.0000	0.0000	0.0000
Slope	0.0000	0.0000	0.0000	N/A
Estimated component				
Level	7.3830	7.3296	7.3296	7.3458
Slope	-0.00001	-0.00002	-0.00002	N/A
R^2	0.8703	0.8725	0.8725	0.9278
Level[a]	F	S	S	S
Slope[a]	S	F	S	NO

Source: Author's estimations using data from 2003 to 2012.
Instrumental variables: Yearly lagged price variables.
Standard errors are in parentheses.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.
[a] F refers to fixed; and S refers to stochastic.

Table 6.7: Diagnostic tests

	(1)	(2)	(3)	(4)
Heteroscedasticity	1.2153	1.2185	1.2185	1.3025
BoxLjung	0.4431	0.5210	0.5211	0.5205
Skewness	1.4984	1.3518	1.3519	1.7459
Kurtosis	0.4257	0.5192	0.5192	0.2508
Bowman-Shenton	1.9241	1.871	1.871	1.9967
AIC	-5.6034	-5.6204	-5.6204	-5.6387
BIC	-5.1388	-5.1558	-5.1558	-5.1973
Auxiliary				
Irregular				
Skewness	0.1292	0.1206	0.1207	0.1274
Kurtosis	0.5733	0.8327	0.8326	0.8661
Bowman-Shenton	0.7025	0.9533	0.9533	0.9935
Level				
Skewness	0.2183	1.2172	0.5311	5.2073**
Kurtosis	0.0024	0.1421	0.2820	1.6569
Bowman-Shenton	0.2207	1.5963	0.8130	6.8642**
Slope				
Skewness	0.0657	0.3056	0.3897	N/A
Kurtosis	0.4323	0.0275	0.0340	N/A
Bowman-Shenton	0.4980	0.3331	0.4237	N/A
Level[a]	F	S	S	S
Slope[a]	S	F	S	NO

Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

[a] F refers to fixed; and S refers to stochastic.

Table 6.8: Structural time-series model results without the income variable

Dependent variable: $\log(y_t)$		
$\log(p_t)$	0.09349	(0.1289)
$\log(p_{t-1})$	-0.4432^{**}	(0.0825)
$\log(p_t^{gas})$	-0.1695	(0.2228)
$\log(Temp_t)$	-0.3485^{***}	(0.0424)
$D_t^{Holiday}$	0.03448^{*}	(0.0179)
Hyperparameters		
Irregular	0.0031	
Level	0.0002	
Slope	0.0000	
Estimated component		
Level	7.4101	
Slope	0.0026	
R^2	0.8600	

Source: Author's estimations using data from 2003 to 2012.
Instrumental variables: Yearly lagged price variables.
Standard errors are in parentheses.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.
[a] F refers to fixed; and S refers to stochastic.

6.2 Industrial and Commercial Sectors

In the industrial and commercial sectors, the tariff is time of use (TOU), involving seasonally varying rates. KEPCO offers several TOU pricing options for industrial and commercial customers.

In the TOU literature, electricity demand is generally modeled with respect to each pricing interval, since standard economic theory implies that TOU customers would respond to the corresponding price at specific times. In Section 6.2.1, I present a simple conceptual framework to model electricity demand under TOU in each pricing interval. The modeling procedure follows Patric and Wolak (2001), who provide a theoretical demand modeling framework under real time pricing (RTP). Empirically, however, this modeling method might suffer some limitations: (1) if the tariff has lack of variation in the pricing intervals, then the model would be invalid to identify the price effect; (2) some important features of electricity consumption, i.e., temporal inter-dependence of electricity consumption, might be obscured, since time-series consumption data are artificially split, according to pricing intervals, into a time-series cross section framework, breaking time series orders in the original data. Due to data limitations, I will not model demand following this line, since electricity consumption for each pricing interval is not available. Instead, I model demand as monthly aggregate consumption against the monthly average price for each sector in Section 6.2.2. With this modeling compromise, some pricing information regarding each pricing interval would be masked by this method. First, part of the cross-price effects among different pricing intervals would be absorbed by the price coefficients, but part of these effects would not. For instance, if firms shift the same amount of electricity usage from the peak interval to the off-peak interval due to peak price increases, then this price responsiveness cannot be captured in the model. Second, it is possible that firms are generally more responsive in peak than in off-peak pricing intervals, so this model would tend to underestimate price elasticity in peak times and overestimate it in off-peak times.

6.2.1 Theoretical Framework

To begin with, I present a heuristic conceptual framework to inspire model specification.

Assume there exists a sector level aggregate production function in a competitive market with the following technology assumptions.

Assumption 1: Total capital stock and technology are fixed.

Assumption 2: Electricity usage and the service-capital ratio is constant in each period. Labor and other materials are pre-determined but their provided services are variable and are complements to electricity usage. There is no possible substitution between labor service and machine service.

Assumption 3: No substitution between electricity and other fuels.

Assumption 3 can be justified by the fact that firms generally will not be able to substitute among fuels without changing the type of capital. Based on assumptions 1-3, it is equivalent to study a model in which there is only one factor of production, electricity.

I adopt the static deterministic production theory, and goods are distinguished by the two pricing intervals, peak, and off-peak periods, in use within a day.[10] Let $\{y_m\}_{i=1}^{M}$ denote M different outputs and e_i denote electricity input at time interval i which can be allocated to different outputs as,

$$e_i = \sum_{m=1}^{M} e_{i,m}$$

where $e_{i,m}$ is the amount of electricity used for output m in time interval i.

The corresponding production transformation function is given by,

$$F(Y, E, X) = 0 \tag{6.5}$$

where Y is the vector of outputs with dimension $M * 1$; E is the vector of electricity inputs

[10] Taylor and Schwarz (1990) treat the aggregate daily demand of different pricing intervals as different goods in the residential sector, and derive the demand for each good by Roy's identity based on the predetermined indirect utility function.

with dimension $2*1$; X is the composite of other quasi-fixed input quantities and other non-price exogenous variables; and F satisfies the regularity conditions in Lau (1972). Note that the transformation function F incorporates possible inter-dependence of production links among different outputs and the way in which each output is conducted in the corresponding periods.

Profit Maximization or Cost Minimization

Note that profit maximization and cost minimization have their own merits under different firm behaviors. In practice, if the level of output produced in real time is not affected by electricity prices, then a cost-minimization approach is a straight-forward application. This is because the interested price coefficients under this behavior is equivalent to estimation of the output-constant cross-price effect, $\frac{\partial u_i}{\partial p_l}$, u_i represents the conditional electricity demand. On the other hand, if the output level in real time is adjusted optimally with electricity prices, then a profit-maximization approach is a better choice. This is because the interested price coefficients in this case is equivalent to estimation of the gross cross-price effect, $\frac{\partial e_i}{\partial p_l}$, and,

$$\frac{\partial e_i}{\partial p_l} = \frac{\partial u_i}{\partial p_l} + \sum_m \frac{\partial u_i}{\partial y_m} \frac{\partial y_m}{\partial p_l}$$

where the first term on the left-hand side is the substitution effect along the isoquant; and the second term represents the output expansion effect; this is the change of the output that contributes additional variations to electricity demand.

Sakai (1974) shows that the expansion effect is non-positive, that is $\sum_m \frac{\partial u_i}{\partial y_m} \frac{\partial y_m}{\partial p_l} \leq 0$. Therefore, if actual output incorporates electricity price effects, using $\frac{u_i}{\partial p_l}$ in cost minimization to study price effects will either over-estimate the substitution effect or under-estimate the own-price effect and the absolute value of the complementarity effect, and vice versa.

I argue that under a TOU tariff schedule, firms know the tariff schedule before-hand, and tariff changes generally remain stable for certain time periods, and thus it would be sensible to expect that production decisions made by firms incorporate the tariff changes; under an RTP tariff schedule, firms know the actual tariff one day before the production decision, so it is possible that firms' output production levels are not affected by tariff changes, as Wolak and Patric (2001), Taylor, Schwarz, and Cohell (2005), and Cochell, Schwarz, and Taylor (2012) suggested. Two aspects, however, should be raised. One is that, under RTP, real-time production decisions by rational firms possibly incorporate the information of the expected real-time electricity price. Thus, if the change in real-time price is reflected in the firm's expectation, it would in turn affect the production decision. The other is that, with the development of information technology, the short-run reaction of output under RTP is not implausible. For example, an advanced digital planning and scheduling solution system can help firms manage short-run inevitables and make faster production decisions with least cost. An example is Preactor, which facilitates firms continuously refining and adjusting their production plans. Thus, it is possible that firms' output planning might partially respond to electricity tariffs in a short period of time. Collectively, under RTP, it is not quite straightforward to argue that the predetermined output level is not affected by tariff changes without detailed investigation of the data or firm planning behaviors. In practice, if data permit, it will be better to test the relationship between output and electricity tariffs or apply a multi-level regression to take into account the output expansion effect.

Since the tariff structure in my study is TOU, I will adopt a profit-maximization approach to derive the factor demand equation.

The Theoretical Model

I start by assuming daily production is independent, and thus a sector's production decision under a certain decision horizon can be reduced to consider the daily-production decision only. I also assume there is no time discount within the day.

Given output and input prices, the optimization problem is,

$$\begin{aligned} & \max \textstyle\sum_{m=1}^{M} q_m y_m - \sum_{i=1}^{3} p_i e_i \\ \text{s.t.} \quad & F(Y, E, X) = 0 \\ & e_i \geq 0 \\ & y_m \geq 0 \end{aligned}$$

where q and p are normalized output and input prices by the price of the first output, and thus set $q_1 = 1$.

The solution to this problem can be represented as the vector $e(p, q, X)$ and $y(p, q, X)$. Thus, normalized profit is $\pi(p, q, X) = qy(p, q, X) - pe(p, q, X)$.

Instead of using the primal approach[11] by specifying the production function and solving for factor demand, I instead apply a dual approach[12] to specify the normalized profit function, given by a normalized quadratic (NQ)[13] form based on Lau (1978a),

$$\begin{aligned} \pi(q, p, X) = a + & \textstyle\sum_{m=2}^{M} \alpha_m q_m + \sum_{i=1}^{2} \beta_i p_i + \frac{1}{2} \sum_{m=2}^{M} \sum_{n=2}^{M} \alpha_{nm} q_m q_n + \frac{1}{2} \sum_{i=1}^{2} \sum_{l=1}^{2} \beta_{il} p_i p_l \\ & + \textstyle\frac{1}{2} \sum_{m=2}^{M} \sum_{i=1}^{2} \gamma_{mi} q_m p_i + \sum_{m=2}^{M} \eta_m q_m X + \sum_{i=1}^{2} \mu_i p_i X + \phi X^2 \end{aligned}$$

[11]The primal approach involves solving the demand equations based on an explicit production function or frontier, which can be arbitrarily pre-determined or estimated from cross-sectional or time-series data. In practice, however, the demand function may not be determined explicitly from the technology.

[12]The use of the duality theorem dates back to Shepard's 1953 work. The dual approach establishes that complete systems of output supply and input demand responses can be estimated from the underlying profit/cost functions in which cross-equation restrictions are imposed on parameters so that the system derives rigorously from a profit or cost function. In other words, duality implies that well-behaved cost and profit functions are equivalent to well-behaved production functions. Chambers (1988) points out that analysts' main interest does not pertain to the production function as such, but to reliable representation and prediction of economic behavior. Duality is indispensible for empirical work based on functional forms that are too complicated to be derived directly from the technology as explicit solutions of a problem of intertemporal optimization (Epstein 1981, 1982). The research question and the availability of data for this study support the use of the dual approach.

[13]Note that the parameters of the profit function contain adequate information from which to infer the properties of the underlying technology in which the production technology underlying this process can be described through production elasticities, input substitution possibilities, returns to scale, and the bias in technology (Thijssen, 1992).

In essence, the quadratic function is a second-order Taylor series expansion, which is appropriately within the class of a "locally-flexible" functional form (FFF) (Fuss, McFadden and Mundlak (1978); Appelbaum (1979)).[14] The advantage of FFF is that it imposes fewer main hypotheses and is thus less restrictive than many classical function forms, such as the Cobb-Douglas and CES forms. In practice, for consistency with the defined technology and competitive theory, the following conditions can be established on the restricted profit function,

(1) There is linear homogeneity in output and input prices, which is maintained through the price normalization.

(2) Convexity can be maintained by the Cholesky factorization based on Lau (1978b) or can be tested after the regression.

(3) There is symmetry which can be maintained by imposing restrictions $\alpha_{mn} = \alpha_{nm}$, $\beta_{il} = \beta_{li}$, and $\gamma_{mi} = \gamma_{im}$ or can be tested after the regression.

Since this section is intended to study electricity usage substitution possibilities, i.e., own- and cross- price effects on inputs, by Hotelling's lemma, I derive the electricity demand equation for each pricing interval,

$$\frac{\partial \pi}{\partial p_i} = e_i = \beta_i + \sum_{l=1}^{2} \beta_{il} p_l + \sum_{m=2}^{M} \gamma_{mi} q_m + \mu_i X \tag{6.6}$$

Some Explanations on the Choice of Functional Form Economic theory only provides generic characteristics for building production models with no explicit guidance for choosing the appropriate functional form, although it does aid in identifying relevant variables and homogeneity restrictions. Moreover, the researcher never knows the true functional form. Along with NQ, in practice, generally there are two other commonly used flexible functional forms: translog and Generalized Lenotief (GL). The following are some comparisons of these three forms.

[14] Applications involving flexible forms have been available since Diewert (1969).

First, a translog function is suitable when input substitution elasticity approaches one. It is better to adopt a GL function when input subsitution elasticity is low. NQ does not have these restrictions and it can be used to derive a simple expression for elasticities, which can be evaluated at any level of prices and quantities.

Second, a translog profit function, unlike GL and NQ, can be used to derive a share function rather than a direct factor demand function, which requires factor share information in estimation that is beyong the scope of this study.

Third, Lau (1978b) shows that imposition of global regularity reduces the translog model to a Cobb-Douglas form, which is not a flexible functional form anymore and has no estimable elasticities. GL, if imposing concavity restrictions, will only capture small price substitute effects. On the other hand, NQ can be easily imposed on homogeneous and global curvature restrictions, while remaining flexible.

Overall, adopting a translog or GL form is better when one has some prior information about input substitution effects. Without prior information, however, NQ is a better choice. Since all possible substitution and complementary possibilities are open, I take NQ as the functional form of my model.

A Digression: Analytic Properties

I make a digression by introducing some standard properties theoretically, as inputs demand at given input-output prices are governed by certain conditions. Sakai (1974) defines a technology as normal if the following conditions hold,

$$
\begin{aligned}
\frac{\partial^2 c}{\partial y_t \partial y_s} &\leq 0 \quad \forall s \neq t \\
\frac{\partial^2 c}{\partial p_i \partial y_s} &\geq 0 \quad \forall i, s \\
\frac{\partial^2 \pi}{\partial e_l \partial e_i} &\geq 0 \quad \forall i, l \\
\frac{\partial^2 \pi}{\partial q_s \partial e_i} &\geq 0 \quad \forall i, s
\end{aligned}
$$

where c and π are the cost and revenue functions respectively. The above conditions indicate that "the marginal cost of an output tends to increase when the quantities of other outputs decrease or when the prices of inputs increase; the marginal revenue of an input tends to go up when the quantities of other inputs go up or when the prices of outputs go up" (Sakai 1974, p. 272.)

Sakai establishes that, under normal technology, there is no gross substitution among inputs, that is $\frac{\partial e_i}{\partial p_l} \leq 0$[15], $\forall\, i, l$.

It is interesting to note that if the technology is normal, then there cannot be an electricity consumption shift in different pricing intervals. Further, its converse-negative proposition implies that if electricity consumption shifts, i.e., there are positive substitution effects across different pricing intervals, exist empirically, then the technology is not normal. Thus, the explanation for a demand shift, if detected, can be guided by the possible violation of the four normal technology conditions. I do not, however, in this study, go into this in further detail.

Some Criticisms

Firstly, assuming the independency of daily production seems quite unrealistic since production is generally correlated between days.[16] With a TOU tariff, however, prices at the same time interval in adjacent days are the same except for the day involving normalized price adjustments. Thus, the independency assumption is a practical simplification with respect to empirical identification, i.e., daily leading and lagging prices are the same as or highly correlated with current price, the effects of which are impossible to distinguish. Moreover, the price coefficients in the model should be interpreted with caution, as they incorporate

[15] This result is an extension of Rader (1968). Rader defines normal technology with a single output as $\frac{\partial^2 f}{\partial e_i e_j} \geq 0$, and shows that factor inputs are never gross substitutes under normal technology and they are in fact gross complements. Sakai (1974) points out that Rader's normal condition holds if and only if $\frac{\partial^2 r}{\partial e_j \partial e_i} \geq 0 \forall i \geq j$, and thus Rader (1968)'s result is a special case of Sakai (1974)'s.

[16] Taylor and Schwarz (2012) point out that there are possible inter-day price effects in RTP schedules and they estimate this effect by directly adding the next day's electricity price in their model without theoretical derivation.

inter-day price effects.

Secondly, this modeling strategy destroys the inherent time-series structure. I artificially manipulate the time-series structure into time-series-cross-section data. This transformation could lead to information loss, for instance, losing inherent dynamic adjustment structure and consumption dependency information carried in the time-series data, since, in reality, electricity demand is a flow process, highly correlated, cyclical, and persistent.

Thirdly, the profit function requires symmetry with price effects, which implies that the cross-price effect for the peak-hour to the off-peak hour is the same as the off-peak hour to the peak-hour. It is obvious, however, that if the off-peak price increases, it is still lower than the peak-hour price, so it would be too strict or dogmatic to expect the same consumption shift from the off-peak to peak period, and vice versa. There is a possibility of asymmetric price effects between peak and off-peak hours. Accordingly, in the empirical study, I should drop the symmetry restrictions and apply hypothesis testing for symmetric price effects.

Fourthly, electricity substitutes are not considered in the theoretical electricity demand model due to the assumption of a fixed capital stock.

6.2.2 Basic Regression Model

Economic theory suggests that, apart from the price of electricity, the price of substitutes and gross domestic product (GDP) also affect electricity usage. Consider the following estimation equations:

$$\log(y_t^{ind}) = \mu + \gamma \log(p_t^{ind}) + \alpha \log(GDP_t) + \log(p_t^s)'\beta + \log(Temp_t)'\theta + (D^{ind})_t'\delta + \epsilon_t^{ind} \quad (6.7)$$

$$\log(y_t^{com}) = \mu + \gamma \log(p_t^{com}) + \alpha \log(GDP_t) + \log(p_t^s)'\beta + \log(Temp_t)'\theta + (D^{com})_t'\delta + \epsilon_t^{com} \quad (6.8)$$

where y_t^{ind} and y_t^{com} are the average daily amount of electricity consumed in unit GWh at month t; the superscripts ind and com refer to the industrial and commercial sectors, respectively; p_t^{ind} and p_t^{com} are the average electricity price in million Korean won per GWh at month t; GDP_t is the real GDP at month t at a 2010 base year; p^s is the vector of the average prices of the electricity substitutes, high-sulfur fuel oil,[17] p_t^{oil}, and natural gas, p_t^{gas} for industrial customers in Korean won per tonne and per MWh GCV, respectively, at month t; $Temp_t$ is vector of cooling-degree days, CDD_t, and heating-degree days, HDD_t, in which CDD_t (HDD_t) equals the monthly average temperature in degrees F when it is greater (smaller) than 65 degrees F, and zero otherwise at month t. D_t^{ind} and D_t^{com} are the vectors of the dummy variables, and the common dummy variables include a yearly dummy and a holiday dummy, which indicates a specific month that has more than three days of national holidays.[18] D_t^{com} also includes winter and summer vacation dummies (January, February, December, and August) to control the possible monthly cyclical patterns; and ϵ_t^{ind} and ϵ_t^{com} are the unobservable random disturbances in the two equations. All nominal price and income data are deflated by the producer price index (PPI). 2010 is the base year for all real variables. GDP and the prices for high-sulfur fuel oil and natural gas are interpolated by a cubic spline to monthly frequency. Because the data of fuel oil and natural gas for commercial customers are not available, I apply their price for industrial customers as proxies in the commercial demand equation. Because there are no clear patterns of the relationships between average monthly temperature and sectors' consumption in Figures 4.10 (b) and (c), I thus set the breaking temperature point, 65 degrees F, to be the same as that in the wholesale market demand model.

Electricity demand is a derived demand, depending on firms' utilization behaviors and electrical equipment and appliance stocks. Changes in utilization behaviors and equipment and appliance stocks can be caused by two channels. One is induced by firms' overall products, presented as GDP and the price of electricity and its substitutes. The other is induced by other exogenous factors, such as firms' environmental protection liabilities. In this sense, models (6.7) and (6.8) could be viewed as reduced forms of electricity demand

[17] The correlation of the prices between low- and high-sulfur fuel oil for industry is around 0.99; the correlation of the prices between high-sulfur fuel oil and light fuel oil for industry is around 0.91.

[18] National holidays includes New Year's Day, Spring Festival, Moon Festival, Christmas Day, Independence Day, Children's Day, Memorial Day, Buddha's Birthday, Constitution Day, Liberation Day, and National Foundation Day.

models. The price and income coefficients not only capture firms' utilization behaviors but also absorb the effects of any adjustments to the equipment and appliance stocks induced by them on electricity consumption. Temperature, holiday, and other dummies are used to control the exogenous factors. In Section 6.2.4, as in the residential demand study in Section 6.1.2, I will introduce a Structural Time Series Model (STSM) to better control the changes in utilization behaviors and equipment and appliance stocks induced by exogenous factors.

The identification of the electricity price coefficients is substantiated by the nonlinear price variations in the system for two reasons. Firstly, either the pricing changes among different seasons of the same contract option or due to contract adjustments in different or the same year in the same season are not at the same proportion, for instance, for the contract classified as industrial (C) high voltage (A) option I in 2007, the price decreased 0%, 26.40%, and 41.49% from summer to spring/fall season and increased 13.84%, 9.16% and 8.83% in off-peak, mid, and peak load intervals respectively from 2007 to 2008 summer season. Secondly, the proportion of the amount of electricity consumed fluctuates, otherwise, I should expect to observe a constant electricity price under the same contract in the same season.

In practice, average price is correlated with the amount of electricity consumed, and thus the monthly average electricity prices are endogenously determined in both models (6.7) and (6.8). Instrumental variables (IV) are thus needed to correct the simultaneous bias asymptotically. Moreover, the price variable is computed as monthly average revenue based on the aggregate revenue and consumption data in each sector provided by KEPCO. I am concerned that this arrangement might introduce measurement error in the model for two reasons: (1) aggregate revenue data might include other miscellaneous payments, such as delivery charges, and thus average revenue would serve as a proxy of the true consumption-weighted average price; (2) these data might suffer from calculation errors; as KEPCO points out in its release, "the sales figures of KEPCO have been prepared based on interval estimates of KEPCO for your convenience only, and have been neither audited nor reviewed by

KEPCO's independent accountants, Samjung & Co. or any other accountants." Therefore, measurement error would be another source of an endogeneity problem in the model. In detail, the difference between average price and average revenue is absorbed in the residuals and this could possibly lead to bias in the electricity price coefficients. Standard econometric theory indicates that classical measurement error[19] leads to a bias towards zero (Greene, 2008, p. 325). Therefore, if the measurement error is caused by reason (2), then the price coefficients tend to underestimate the absolute value of price elasticity. If, however, the measurement error is caused by reason (1), the direction of the bias then would depend on the correlation between the measurement error and the average electricity price variables.

To solve the endogeneity problem, I propose four and five instruments for the industrial and commercial sectors respectively. The three common instruments are choosen from economic indicators, including the monthly average Korean exchange rate against US dollars, the monthly value of the construction order received on a daily average basis, and the monthly retail trade volume on a daily basis. The rationale for these choices are the following: (1) the amount of electricity consumed in the industrial and commercial sectors are positively correlated with the sectors' performance, and this performance in each sector tends to be positively correlated with economic indicators, (2) these IVs do not directly affect the dependent variable, aggregate electricity consumption for both sectors, and (3) once GDP is controlled, they are not correlated with the unobservable disturbances for both sectors. Before proposing other IVs in these two sectors, I assume that there is no autocorrelation in the error terms, ϵ_t^{ind} and ϵ_t^{com}, in models (6.7) and (6.8). This assumption allows me to apply a monthly lagged IV for the industrial average price and yearly and monthly lagged IVs for the commercial average price. The reason for these choices is that in the industrial sector, y_t^{ind} and y_{t-1}^{ind} are highly correlated, as shown in Figure 4.11 (b), and with a relatively stable TOU tariff in the short run, p_{t-1}^{ind} are thus correlated with p_t^{ind}, but not ϵ_t^{ind} due to

[19]Classical measurement error refers to the error that has a mean zero and is independent of the true regressor and the economic error disturbance in the regression.

the assumption of no autocorrelation assumption. In the commercial sector, however, there is evidence of a yearly cyclical pattern as well as a consumption correlation between the current and lagged month, as shown in Figures 4.11 (c) and 4.12 (c). Thus, both yearly and monthly lagged prices are applied as IVs.

To ensure the proposed instruments are valid, I apply a standard F-test in the first-stage regression regarding the endogenous electricity price variables and the Cragg-Donald Wald F test. Results are shown in Table 6.9. As a rule of thumb, an F-statistic over 10 is required to suggest that the IVs are sufficiently strong.[20] The F-statistics imply that, for the industrial electricity price, only the construction IV is relevant and has explanatory power, while, for the commercial electricity price, both the monthly and yearly lagged instruments are valid. The Cragg-Donald test statistics[21] convey the same results. Therefore, the construction instrument will be used for industrial demand estimation, and the monthly and yearly lagged instruments will be used for commercial demand estimation. In the estimations, I will also provide the results with the monthly IV in the industrial sector for comparison.

Empirical Results and Implications

Tables 6.10 and 6.11 present the estimation results for both the industrial and commercial sectors. In both tables, column (1) contains the results without IV by OLS, and columns (2) to (4) are the results with IVs by GMM. All models fit well, since the adjusted R^2 are all greater than 0.9. Before interpreting the results, I firstly look at some diagnostic tests, shown in Tables 6.12 and 6.13. The columns in these two tables correspond to the same columns in Tables 6.10 and 6.11. The Hansen J test in column (4) in both tables tests the exogeneity of the IVs and a rejection of the null hypothesis would cast doubt on the validity of the instruments. Both statistics imply that I cannot reject the null at

[20] Staiger and Stock (1997) show that the weak instruments problem can arise even when the first-stage t- and F-tests are significant at conventional levels in a large sample.

[21] The Stock-Yogo weak ID test critical values for one endogenous variable and one instrument are 10% maximal IV size 16.38; 15% maximal IV size 8.96; 20% maximal IV size 6.66; 25% maximal IV size 5.53. Source: Stock-Yogo (2005).

the 10% level, suggesting the validity of the IVs' for both models; that is, there is no statistical evidence against the orthogonal condition of the IVs. Secondly, for the endogeneity test, the industrial electricity price variable suffers from an endogeneity problem when the construction IV is applied, since the test statistic is significant at the 1% level, but not so with the monthly lagged price instrument, which is actually redundant. This result is expected since the F and Cragg-Donald Wald F tests suggest the monthly lag price instrument has limited explanatory power. While the commercial electricity price can be viewed as an exogenous variable when monthly or/and yearly lagged price instruments are applied, the test statistics are all insignificant at the 10% level. Thus, both instruments are actually redundant, even though they are relevant based on the F and Cragg-Donald Wald F tests. Thirdly, for the Cumbly-Huizinga autocorrelation test, in the industrial sector, although the model has an autocorrelation problem without the IVs at the 5% level, I could accept the hypothesis of no autocorrelation for the models with IVs at the 10% level. In the commercial sector, however, test statistics suggest there is autocorrelation in all models at the 5% level, casting doubt on the suitability of using lagged prices as instruments, although the Hansen J test suggests their exogeneity. Fourthly, for the heteroscedasticity test, the industrial demand models are generally well behaved, and do not suffer from a heteroscedasticity problem. Heteroscedasticity is not detected in the commercial demand models except the one in column (3). In all, with these diagnostic results, the results in column (2) in Table 6.10 with the construction instrument and column (1) in Table 6.11 without instruments are preferred in the industrial and commercial sectors respectively. Figures 6.4 and 6.9 provide CUSUM test plots for these two specifications, suggesting that the estimated coefficients are stable within a 95% confidence band.

Next, I will interpret the results based on the two preferred specifications regarding the industrial and commercial sectors with some interesting findings.

Firstly, there is statistical evidence that firms in the industrial sector respond to electricity price variations, but that firms in the commercial sector do not. The price elasticity in the industrial sector is -0.1854, which is inelastic and significant at the 1% level, and the sign is aligned with economic theory. In the commercial sector, although the sign of the estimated price coefficient is also negative, it is highly insignificant. This lack of responsiveness could explain why the endogenity tests suggest that the price variable is exogenous when monthly or/and yearly lagged price instruments are applied. Recall that in Figure 4.7 (d), the real electricity price in the commercial sector has a decreasing trend, which might cause commercial firms to be unresponsive to price.

Secondly, in both the industrial and commercial sectors, GDP has significantly positive effects on electricity consumption. The corresponding income elasticities are 1.4585 and 1.7326 respectively. The magnitude is larger in the commercial sector, and based on the Wald test, GDP elasticities are significantly different according to the test statistic, 4.01, which is significant at the 5% level. Therefore, economic growth is a driving force of electricity usage in both sectors, but has higher effects in commercial sectors. Moreover, the firms in both sectors are more responsive to GDP growth than to the electricity price.

Thirdly, in the industrial sector, the effects of natural gas, as an electricity substitute, on electricity usage are significant at the 5% level, and the corresponding price elasticity is 0.1345. The estimated coefficient of the high-sulfur fuel oil price is not, however, significant at the 10% level, suggesting it has no effects on electricity consumption. In the commercial sector, both estimated coefficients of natural gas and high-sulfur fuel oil price coefficients are insignificant at the 10% level, suggesting fuel oil and natural gas are, statistically, not electricity substitutes. Generally, changing the type of energy usage in the industrial and commercial sectors requires a change in the type of equipment and appliances or starting up on-site generators by firms. A lack of evidence of substitution effects of fuel oil in the industrial sector and together with natural gas in the commercial sector might indicate that the gain from using a relatively cheaper form of energy is less than the cost of adjusting the type of equipment and appliances or starting up on-site generators. Moreover, since I apply the substitute prices for industrial firms as proxies for commercial firms due to data limitations, bias would be introduced which might cause the estimates to become insignificant.

Fourthly, temperature affects electricity consumption in both the industrial and commercial sectors. The corresponding elasticities in cooling-degree days are -0.0375 and -0.0915, and in heating-degree days are -0.0442 and -0.1112 for the two sectors respectively. Therefore, as temperature increases, electricity usage decreases. The magnitudes of the temperature effects, both in cooling- and heating-degree days, are greater in the commercial sector in absolute value, and these differences are statistically significant at the 1% level based on the Wald test with the test statistics 11.40 and 14.18, respectively. Thus, firms are more responsive to temperature in the commercial sector than in the industrial sector. I also test whether temperature effects are heterogeneous in cooling- and heating-degree days for each sector by using Wald tests. The tests statistics are 12.58 and 19.05, and both are significant at the 1% level, suggesting that the null of homogenous effects can be rejected.

Fifthly, electricity consumption falls in those months that have at least three days of national holidays in the industrial sector, but not in the commercial sector. In the industrial sector, aggregate electricity usages falls by about 3% on a daily basis during these holiday months. I also test the effects with other numbers of holidays within a month: the months with at least one and two days of national holidays, for both sectors, respectively. That is, I replace the holiday dummy with these two dummies separately in each sector's demand equation. The results are shown in Table 6.16. For both sectors, the months with at least one or two days of national holidays all have significant negative effects on electricity consumption. Since the estimated holiday dummy coefficient of the month with at least three days of holidays is not significant in the commercial sector, I suspect the different number of holidays within a month might have heterogeneous effects. I test this hypothesis by adding the dummies indicating whether the month has one day or two days of holidays in the sector demand models (6.7) and (6.8), respectively. Table 6.17 presents the estimation results. In the industrial sector, all of the coefficients are negative and significant at the 1% level.

Based on the Wald test, the effects of the month with one day of holidays is statistically different from the effects of the months having holidays with two or at least three days; the test statistics are 15.42 and 11.33, respectively. There is no significance difference between the months having holidays with two or at least three days; the test statistic is 0.03, which is insignificant at the 10% level. The percentage reduction in electricity is about 1.6% in the month with one day of holidays and 4% in the month with two or at least three days of holidays. An explanation for this phenomenon is that firms generally shut down some operations during holidays, leading to a reduction of electricity consumption load. In the commercial sector, the effect of the month with one day of holidays is insignificant at the 10% level, indicating these kinds of months do not affect firms' operations. The effects of months with two or at least three days of holidays are significant at the 1% and 10% levels, with 7.3% and 2.2% reductions in electricity usage on a daily basis, respectively. Based on the Wald test, the months with two days of holidays have statistically different effects compared with the months with at least three days of holidays in absolute value; the test statistic is 12.07, which is significant at the 1% level. These results are quite interesting, and seem counter-intuitive. Upon further reflection, however, my explanation is as follows. The most important holidays generally have at least three days off, i.e., Lunar New Year and Moon Festival. As well as families getting together and celebrating, these holidays are also great times for shopping, like the Black Friday holiday in US, sightseeing, entertaining, dining, etc. Therefore, commercial firms, like restaurants, hotels, department stores, etc. might need to operate at a higher level compared to normal days to satisfy customers' demand, thus involving smaller reductions in electricity consumption compared with the month with two days of holidays.

I also present estimation results with a linear specification for both sectors, as shown in Tables 6.14 and 6.15. Previous results are robust to this linear specification.

Moreover, I introduce a variation of the model. I add quadratic temperature terms for cooling- and heating-degree days in the basic models (6.7) and (6.8) to test whether there exist quadric temperature effects on firms' electricity usage. I will test whether the coefficients for $(\log(CDD_t))^2$ and $(\log(HDD_t))^2$ are significantly different from zero or not. The results are presented in Tables 6.18 and 6.19. The estimated quadratic effects are detected for heating-degree days with 5% significance in the industrial sector. In the commercial sector, the estimated coefficients of the quadratic terms are insignificant at the 10% level, indicating no quadratic effects. The other estimates are of similar sizes and and signs to those in the basic models.

KPX offers a document of electricity forecasting using monthly data in the retail market. The demand specification is log-linear. The estimated price coefficient in the commercial sector is also insignificant, aligning with my results. The estimated price coefficient in the industrial sector is significant, but the document does not provide the unit of the price variable, so I cannot compare their estimates' magnitude to mine. The sign of the estimated temperature coefficients in their model is, however, positive, which is suspicious, since it implies that the electricity usage, ceteris paribus, would be the lowest in winter.

In all, the empirical model should be interpreted as capturing the equilibrium demand relationship, and thus the price elasticity is described as long-run elasticity, from a time-series viewpoint, with a couple of months' window.

6.2.3 SUR Model

In this section, I estimate the industrial and commercial demand model under the framework of Seeming Unrelated Models (SUR). The choice of this framework is due to the conjecture that electricity usage in these two sectors might be correlated, and this correlation information could be useful in estimation of the coefficients. Models (6.7), (6.8) and an additional assumption that ϵ_t^{ind} and ϵ_t^{com} are correlated represent the SUR systems.

This representation can be simplified by first stacking the observations and then the two equations together as the following:

$$Y = XB + \epsilon \tag{6.9}$$

where Y and ϵ are of dimension $(2T \times 1)$, and X is of dimension $(2T \times n)$, provided T is the number of observations and n is the total number of the regressors in the two equations.

The estimation method is 3SLS. Table 6.20 presents the estimation results. The correlation coefficient between the two equations is 0.4155. I apply Breusch and Pagan's (1980) LM test[22] to check whether the two equations in the SUR system are related based on the results in these two tables. The test statistic is over 60, suggesting a rejection of the null that the two equations are not correlated at any conventional significance level and supporting the application of the SUR model.

The estimated price and income elasticities in the industrial sector are -0.1071 and 1.4808, and both are significant at 1%. Still, the estimated price coefficient in the commercial sector is insignificant, indicating no evidence of price responsiveness. Estimated income elasticity in the commercial sector is 1.8427, which is significant at the 1% level. I apply the Hausman test to check whether all coefficients are systematically different with the results if the two models were estimated separately. The test statistics are 13.77 and 13.72, suggesting that the null hypothesis, that the coefficients are not systematically different, cannot be rejected.

6.2.4 Structural Time-Series Model

In this section, I relax models (6.7) and (6.8) by allowing the intercept coefficient to evolve stochastically over time under the framework of the structural time-series model, as a way to better control electricity usage trend effects should they exist. On the other hand, Leamer (1983) advocates conducting sensitivity analysis to help diagnose misspecification, and the

[22] The test statistic is $\lambda = N\rho_{12}^2$, where N is the sample size and ρ_{12} is the correlation.

"fragility" of regression coefficient estimates is indicative of specification error. Therefore, the study in this section could also be viewed as a simple sensitivity analysis in line with Leamer to check the robustness of the estimates.

I revise model (6.7) and (6.8) as the following:

$$\log(y_t^{ind}) = \mu_t^{ind} + \gamma \log(p_t^{ind}) + \alpha \log(GDP_t) + \log(p_t^s)'\beta + \log(Temp_t)'\theta + (D^{ind})_t'\delta + \epsilon_t^{ind} \quad (6.10)$$

$$\log(y_t^{com}) = \mu_t^{com} + \gamma \log(p_t^{com}) + \alpha \log(GDP_t) + \log(p_t^s)'\beta + \log(Temp_t)'\theta + (D^{com})_t'\delta + \epsilon_t^{com} \quad (6.11)$$

where μ_t^{ind} and μ_t^{com} are now called the stochastic trend component.

μ_t^{ind} and μ_t^{com} are then expressed as

$$\mu_t^{ind} = \mu_{t-1}^{ind} + \beta_{t-1}^{ind} + \eta_t^{ind} \quad (6.12)$$

$$\beta_t^{ind} = \beta_{t-1}^{ind} + \zeta_t^{ind} \quad (6.13)$$

and

$$\mu_t^{com} = \mu_{t-1}^{com} + \beta_{t-1}^{com} + \eta_t^{com} \quad (6.14)$$

$$\beta_t^{com} = \beta_{t-1}^{com} + \zeta_t^{com} \quad (6.15)$$

where $\eta \sim N(0, \Sigma_\eta^2)$ and $\zeta_t \sim N(0, \Sigma_\zeta^2)$.[23] Equations 6.12 and 6.13, 6.14, and 6.15 represent the level and the slope of the trend, respectively. The exact form of the trend depends on the size of the variances (referred as the hyperparameters), σ_η^2 and/or σ_ζ^2. In other words, η_t shifts the level of the trend up and down and ζ_t allows the slope to change. Models (6.7) and (6.8) in Section 6.2.2 are special cases here as both σ_η^2 and σ_ζ^2 equal zero.

μ_t is generally interpreted as capturing the underlying structure of the residential sector or the proxy for the Underlying Energy Demand Trend (UEDT) as it is defined in the literature. Structural factors include technology progress, economic structure changes, energy usage

[23]The superscripts are dropped for presentation simplicity.

standards, environmental concerns, etc., which can be assumed to be exogenous factors that influence the amount of electricity consumed in both the industrial and commercial sectors. For instance, a more energy efficient technology or higher environmental standard would reduce electricity usage, while, on the other hand, some economic structure changes might be working in the other direction. Thus, compared with models (6.7) and (6.8), stochastic μ_t could better control the effects of the exogenous factors imposed on industrial and commercial sector electricity usage either through changes in equipment and appliance stocks or utilization behaviors.

Since electricity price variables are endogenous in the industrial demand model and IVs are redundant in the commercial demand model based on the endogeneity test, I apply LIML and LM estimation method to estimate industrial and commercial demand equations respectively in STAMP 8.3.

For the industrial demand equation, results are presented in Table 6.21 with four different structure specifications. The presented component variables are estimated at the end period. The estimated coefficients of electricity price are significantly at the 1% level and with expected signs for all specifications. The magnitudes of the estimated price elasticities are generally less responsive than the results of model (6.7) in column (2) in Table 6.10 in absolute value, but are not significantly different. Unlike the results in column (2) in Table 6.10, the effects of natural gas are insignificant at the 10% level for all specifications, indicating that some UEDT effects are confounded in the natural-gas coefficient when it is not specifically controlled. Therefore, the estimated significant substitution effect with natural gas without the UEDT has low credibility. The rest of the estimated coefficients all have similar sizes and signs. The estimated slopes at the end periods are also all positive, indicating an upward consumption trend. Table 6.22 provides statistics of the diagnostic tests. Based on the AIC and BIC criterion, models with a *stochastic level/fixed slope* specification are preferred to the other three. For this specification, autocorrelation and heteroscedasticity are not detected, and the normality standard can be accepted based on the Bowman-Shenton test at the 10% level, except for the level error in the auxiliary regression, which is significant at the 1% level. The corresponding price and income elasticities are -0.1678, and 1.3227. Figure 6.6 plots the UEDT from the *stochastic level/fixed slope* specification over time throughout the sample period. The plot is stochastic with an upward trend. In 2001 to 2003, when there were venture capital booms after the 1997 financial crisis, the corresponding UEDT had a faster growth rate but it returned to the original path afterwards. In the 2007-08 global financial crisis period, the trend slowed down and began to decrease. After the crisis, aggregate electricity usage quickly catches up and shows a steady growth. Figure 6.7 plots the CUSUM test based on the same model. The shape of the figure suggests that the model is stable.

In the commercial sector, the fittings are not good. The *stochastic level/no slope* specification is the most preferred in terms of an absence of heteroscedasticity, normality standard, and AIC/BIC. Autocorrelation problems still exist. I present the results of this specification in Table 6.23. Its diagnostic test is listed in Table 6.24. The estimate price coefficient is still insignificant. Income elasticity is 1.2483, lower than the results in Sections 6.2.2 and 6.2.3, and this might be explained by the upward UEDT trend in Figure 6.8. When the UEDT is not controlled, its effects are confounded in the GDP coefficient, yielding a higher estimate. The 2007-08 global financial crisis imposes downside effects on electricity usage, and the UEDT seems more volatile since then. The CUSUM plot is presented in Figure 6.9, suggesting that the model is generally stable, although it might cast some doubts around period 58.

6.2.5 Figures and Tables

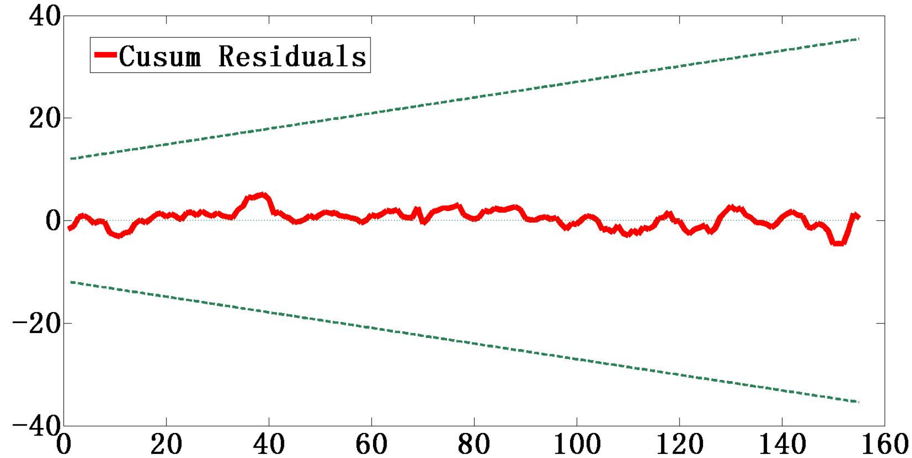


Figure 6.4: Plots of CUSUM test for industrial demand model (6.7) in column (2) in Table 6.10.

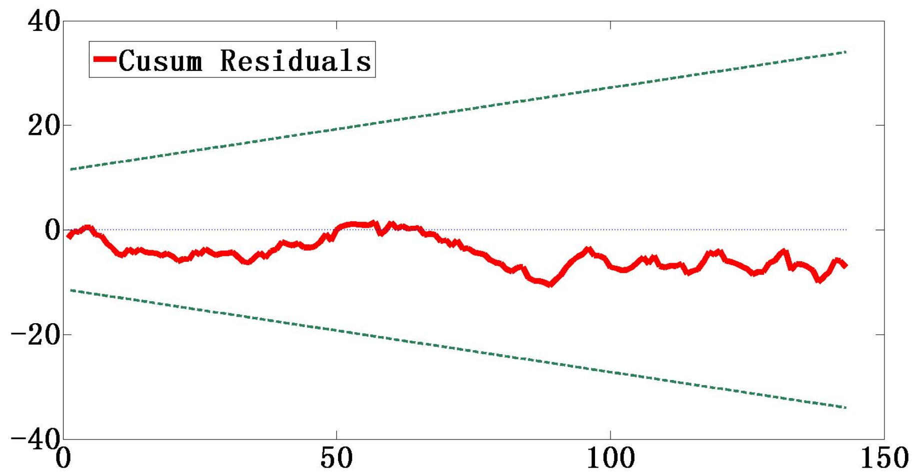


Figure 6.5: Plots of CUSUM test for commercial demand model (6.8) in column (1) in Table 6.11.

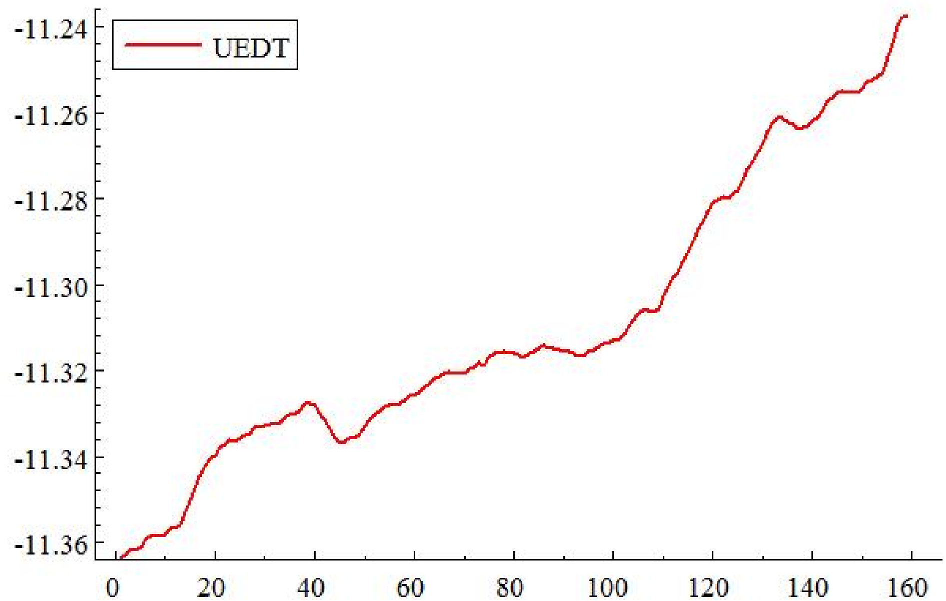


Figure 6.6: Plot of the UEDT for industrial demand model (6.10)

Figure 6.7: Plot of CUSUM test for industrial demand model (6.10)

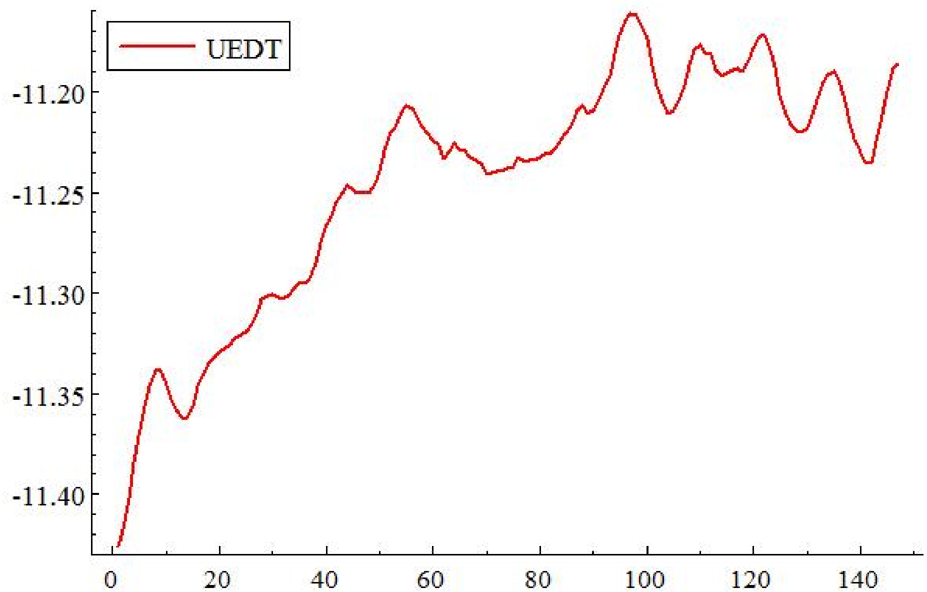


Figure 6.8: Plot of the UEDT for commercial demand model (6.11)

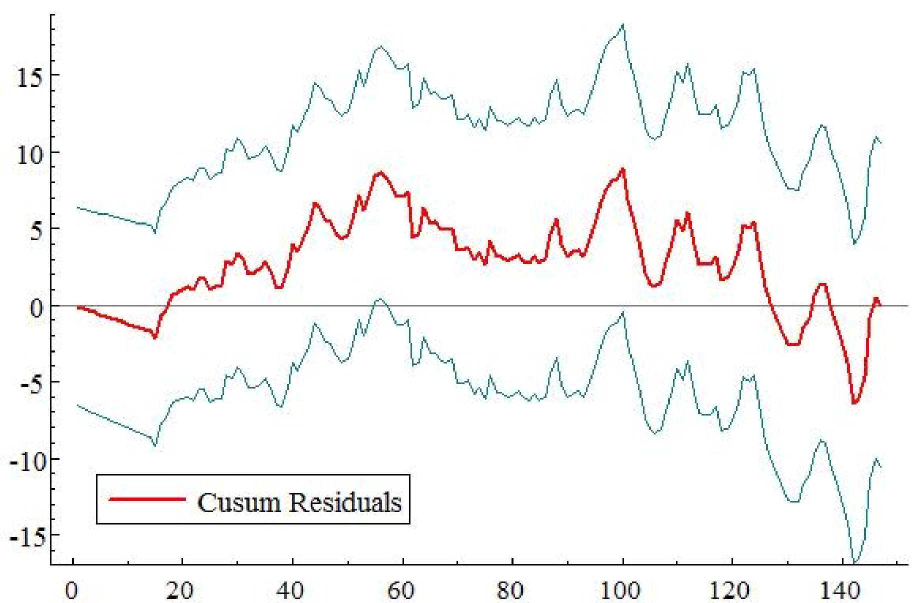


Figure 6.9: Plot of CUSUM test for commercial demand model (6.11)

Table 6.9: F and Cragg-Donald Wald F tests

	Endogenous variable: $\log(p_t^{ind})$		Endogenous variable: $\log(p_t^{com})$	
	F-Statistics	CD statistics[a]	F-Statistics	CD statistics[a]
Construction[b]	29.32	40.94	3.42	5.92
Exchange rate	0.01	0.01	0.37	0.33
Retail trade volume	1.32	0.65	2.72	1.84
Monthly lag	3.04	4.68	65.61	126
Yearly lag	-	-	80.43	168.54

Source: Author's estimations using data from 2000 to March 2013.

[a] CD refers to Cragg-Donald Wald F test.

[b] Construction refers to the monthly value of the construction order received.

Table 6.10: Estimates of model (6.7) in industrial sector

	(1)		(2)		(3)		(4)	
$\log(p_t)$	-0.0693^{***}	(0.0247)	-0.1854^{***}	(0.0416)	-0.1016	(0.0799)	-0.1426^{***}	(0.0373)
$\log(GDP_t)$	1.3831^{***}	(0.1797)	1.4585^{***}	(0.1790)	1.4256^{***}	(0.1690)	1.4485^{***}	(0.1684)
$\log(p_t^{oil})$	-0.0398	(0.0757)	-0.0895	(0.0803)	-0.0582	(0.0753)	-0.0703	(0.0746)
$\log(p_t^{gas})$	0.0610	(0.0610)	0.1345^{**}	(0.0621)	0.1139	(0.0885)	0.1513^{*}	(0.0606)
$\log(HDD_t)$	-0.0285^{***}	(0.0064)	-0.0375^{***}	(0.0077)	-0.0316^{***}	(0.0090)	-0.0349^{***}	(0.0072)
$\log(CDD_t)$	-0.0324^{***}	(0.0075)	-0.0442^{***}	(0.0093)	-0.0363^{***}	(0.0110)	-0.0405^{***}	(0.0086)
$D_t^{Holiday}$	-0.0212^{***}	(0.0061)	-0.0244^{***}	(0.0067)	-0.0220^{***}	(0.0061)	-0.0229^{***}	(0.0064)
IV1[a]			√				√	
IV2[b]					√		√	
R^2	0.9883		0.9861		0.9881		0.9872	

Source: Author's estimations using data from 2000 to March 2013.

Standard errors are in parentheses.

Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

[a] IV1: Value of construction orders received.

[b] IV2: monthly lagged electricity price..

Table 6.11: Estimates of model (6.8) in commercial sector

	(1)		(2)		(3)		(4)	
$\log(p_t)$	-0.0776	(0.1364)	-0.0054	(0.1333)	0.0062	(0.1314)	0.0005	(0.1166)
$\log(GDP_t)$	1.7326***	(0.2572)	1.7601***	(0.1848)	1.7645***	(0.1768)	1.7652***	(0.1765)
$\log(p_t^{oil})$	-0.0818	(0.1696)	-0.0660	(0.1052)	-0.0635	(0.1099)	-0.0664	(0.1052)
$\log(p_t^{gas})$	0.0897	(0.1412)	0.0490	(0.0941)	0.0435	(0.0947)	0.0897	(0.1412)
$\log(HDD_t)$	-0.0915**	(0.0346)	-0.0923***	(0.0289)	-0.0924**	(0.0290)	-0.0916***	(0.0279)
$\log(CDD_t)$	-0.1112**	(0.0363)	-0.1127***	(0.0309)	-0.1130***	(0.0313)	-0.1122**	(0.0301)
$D_t^{Holiday}$	-0.0044	(0.0129)	-0.0040	(0.0172)	-0.0040	(0.0173)	-0.0042	(0.0171)
IV1[a]			√				√	
IV2[b]					√		√	
R^2	0.9578		0.9576		0.9575		0.9576	

Source: Author's estimations using data from 2001 to March 2013.

Standard errors are in parentheses, in which standard errors in column (1) are autocorrelation robust.

Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

[a] IV1: Yearly lagged electricity price in commercial sector.

[b] IV2: Monthly lagged electricity price in commercial sector.

Table 6.12: Diagnostic test for the model in columns (1) - (4) in Table 6.10

	(1)	(2)	(3)	(4)
Endogeneity test statistics	-	10.42***	0.239	7.360***
Hansen J statistics	-	-	-	0.316
Cumby-Huizinga test statistics	6.0873	1.6760	2.7036	1.7976
Heteroscedasticity test statistics[a]	30.620**	25.345	24.390	26.002

Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

[a] Heteroscedasticity test is White/Koenker nR^2 test in column (1) and Pagan-Hall test in columns (2)-(4).

Table 6.13: Diagnostic test for the model in columns (1) - (4) in Table 6.11

	(1)	(2)	(3)	(4)
Hausman test statistics	-	0.598	0.956	1.488
Hansen J statistic	-	-	-	0.009
Cumby-Huizinga test statistics	13.3661***	9.9303**	9.8443**	9.8563**
Heteroscedasticity test statistics[a]	37.128***	28.008**	22.322	28.307**

Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

[a] Heteroscedasticity test is White/Koenker nR^2 test in column (1) and Pagan-Hall test columns (2)-(4).

Table 6.14: Estimates of linear model specification of model (6.7) in industrial sector

			Elasticity
p_t^{ind}	-1.1186^{***}	(0.3503)	-0.1615
GDP_t	0.0008^{***}	(0.0001)	1.3744
p_t^{oil}	-0.0001	(0.0000)	-
p_t^{gas}	0.0013^{*}	(0.0007)	0.1215
CDD_t	-0.3470^{***}	(0.0670)	-0.0493
HDD_t	-0.6217^{***}	(0.1282)	-0.0523
$D_t^{Holiday}$	-11.4945^{***}	(3.2426)	-
R^2	0.9881		

Source: Author's estimations using data from 2000 to March 2013.
Dependent variable: Industrial monthly consumption, y_t^{ind} in level.
Instrumental variable: Value of construction orders received.
Standard errors are in parentheses.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

Table 6.15: Estimates of linear model specification of model (6.8) in commercial sector

			Elasticity
p_t^{com}	0.1275	(0.3392)	
GDP_t	0.0004^{***}	(0.0001)	1.6941
p_t^{oil}	-0.0000	(0.0000)	-
p_t^{gas}	0.0004	(0.0007)	-
CDD_t	-0.2273	(0.1614)	-0.0729
HDD_t	-0.4913^{***}	(0.1840)	-0.0994
$D_t^{Holiday}$	0.3390^{***}	(3.1519)	-
R^2	0.9456		

Source: Author's estimations using data from 2001 to March 2013.
Dependent variable: Commercial monthly consumption, y_t^{com} in level.
Standard errors in parentheses are autocorrelation robust.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

Table 6.16: Estimated coefficients with different holiday dummies based on models (6.7) and (6.8)

	Industrial sector		Commercial sector	
Month with national holidays	-0.0249^{***}	(0.0045)	-0.0276^{***}	(0.0087)
Month with at least two days national holidays	-0.0275^{***}	(0.0051)	-0.0381^{***}	(0.0119)

Source: Author's estimations using data from 2000 to March 2013 in industrial sector, and 2001 until Mar. 2013 in commercial sector.
Instrumental variable in industrial sector: Value of construction orders received.
Standard errors are in parentheses. Standard errors for the commercial sector are autocorrelation robust.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

Table 6.17: Estimated coefficients with different holiday dummies based on models (6.7) and (6.8)

	Industrial sector		Commercial sector	
Month with one day of national holidays	-0.0161^{***}	(0.0042)	-0.0140	(0.0109)
Month with two days of national holidays	-0.0393^{***}	(0.0057)	-0.0760^{***}	(0.0125)
Month with three days of national holidays	-0.0380^{***}	(0.0074)	-0.0219^{*}	(0.0125)

Source: Author's estimations using the data from 2000 until Mar. 2013 in industrial sector, and 2001 until Mar. 2013 in commercial sector.
Instrumental variable in industrial sector: Value of construction orders received.
Standard errors are in parentheses. Standard errors for the commercial sector are autocorrelation robust.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

Table 6.18: Estimates of model (6.7) with quadratic temperature term in industrial sector

$\log(p_t)$	-0.2275^{***}	(0.0861)
$\log(GDP_t)$	1.4121^{***}	(0.1732)
$\log(p_t^{oil})$	-0.7272	(0.1732)
$\log(p_t^{gas})$	0.1488^{*}	(0.0761)
$\log(CDD_t)$	0.2127	(0.2021)
$(\log(CDD_t))^2$	-0.0069	(0.0332)
$\log(HDD_t)$	0.4869^{**}	(0.2299)
$(\log(HDD_t))^2$	-0.0738^{**}	(0.0318)
$D_t^{Holiday}$	-0.0219^{***}	(0.0063)
R^2	0.9864	

Source: Author's estimations using data from 2000 to March 2013.
Instrumental variable: Value of construction orders received.
Standard errors are in parentheses.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

Table 6.19: Estimates of model (6.8) with quadratic temperature term in commercial sector

$\log(p)$	-0.0693	(0.1377)
$\log(GDP_t)$	1.7317^{***}	(0.2390)
p_t^{oil}	-0.0779	(0.1591)
$\log(p_t^{gas})$	0.0745	(0.1366)
$\log(CDD_t)$	0.3644	(0.5136)
$(\log(CDD_t)^2)$	0.1126	(0.0798)
$\log(HDD_t)$	0.3753	(0.5024)
$(\log(HDD_t))^2$	-0.0663	0.0653
$D^{Holiday}$	-0.0055	(0.0122)
R^2	0.9642	

Source: Author's estimations using data from 2001 to March 2013.
Instrumental variable: Value of construction orders received.
Standard errors in parentheses are autocorrelation robust.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

Table 6.20: Estimates of SUR model (6.9)

Dependent variable: $\log(y_t^{ind})$)			Dependent variable: $\log(y_t^{com})$		
$\log(p_t^{ind})$	-0.1071^{***}	(0.0228)	$\log(p_t^{com})$	0.0305	(0.0923)
$\log(GDP_t)$	1.4808^{***}	(0.1474)	$\log(GDP_t)$	1.8427^{***}	(0.1446)
$\log(p_t^{oil})$	-0.0129	(0.0616)	$\log(p_t^{oil})$	-0.1249	(0.1034)
$\log(p_t^{gas})$	0.0676	(0.0569)	$\log(p_t^{gas})$	0.1137	(0.0931)
$\log(CDD_t)$	-0.0349^{***}	(0.0061)	$\log(CDD_t)$	-0.1123	(0.0272)
$\log(HDD_t)$	0.0401^{**}	(0.0071)	$\log(HDD_t)$	0.1307	(0.0293)
$D_t^{Holiday}$	-0.0228^{***}	(0.0047)	$D_t^{Holiday}$	0.0174	(0.1436)
R^2	0.9895			0.9611	

Source: Author's estimations using data from 2000 to March 2013.
Instrumental variable in industrial sector: Value of construction orders received.
Standard errors are in parentheses.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

Table 6.21: Structural time-series model results in the industrial sector

$\log(p_t^{ind})$	-0.1603^{***}	-0.1678^{***}	-0.1739^{***}	-0.1635^{***}
	(0.0226)	(0.0225)	(0.0227)	(0.0228)
$\log(GDP_t)$	1.3221^{***}	1.3227^{***}	1.3618^{***}	1.4260^{***}
	(0.2315)	(0.2738)	(0.0847)	(0.1192)
$\log(p_t^{oil})$	-0.0599	-0.0572	-0.0590	-0.0713
	(0.0769)	(0.0865)	(0.0846)	(0.0848)
$\log(p_t^{gas})$	0.0734	0.0808	0.0847	0.1192
	(0.0872)	(0.0990)	(0.0973)	(0.0799)
$\log(CDD)$	-0.0234^{***}	-0.0238^{***}	-0.0241^{***}	-0.0239^{***}
	(0.0053)	(0.0054)	(0.0052)	(0.0052)
$\log(HDD)$	-0.0365^{***}	-0.0373^{***}	-0.0382^{***}	-0.0364^{***}
	(0.0068)	(0.0075)	(0.0069)	(0.0070)
$D^{Holiday}$	-0.0427^{***}	-0.0439^{***}	-0.0450^{***}	-0.0428^{***}
	(0.0079)	(0.0087)	(0.0080)	(0.0079)
Hyperparameters				
Irregular	0.0006	0.0005	0.0005	0.0005
Level	0.0000	0.0000	0.0000	0.0000
Slope	0.0000	0.0000	0.0000	N/A
Estimated component				
Level	-11.1347	-11.2377	-11.7925	-13.1806
Slope	0.0003	0.0003	-0.0002	N/A
R^2	0.6597	0.6091	0.5977	0.9882
Level[a]	F	S	S	S
Slope[a]	S	F	S	NO

Source: Author's estimations using data from 2000 to March 2013.
Instrumental variable in industrial sector: Value of construction orders received.
Standard errors are in parentheses.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.
[a] F refers to fixed; and S refers to stochastic.

Table 6.22: Diagnostic tests for STSM model (6.10)

Heteroscedasticity	1.7444	1.8555	1.8888	2.1580
BoxLjung	2.2246	2.7732	2.6531	2.1759
Skewness	0.9459	0.3803	0.7106	0.2031
Kurtosis	0.0557	0.0310	0.0169	0.0004
Bowman-Shenton	1.0016	0.4113	0.7275	0.2035
AIC	-7.1832	-7.2228	-7.1941	-7.2029
BIC	-6.7153	-6.7981	-6.7694	-6.7976
Auxiliary				
Irregular				
Skewness	2.6329	3.3313*	3.5196*	3.3395*
Kurtosis	0.3145	0.0758	0.0888	0.0001
Bowman-Shenton	2.9474	3.4061	3.1646	3.3396
Level				
Skewness	2.208	4.8828**	2.9301*	4.3302**
Kurtosis	0.0010	1.6807	0.5079	1.6972
Bowman-Shenton	2.209	6.5635**	3.435	6.0093**
Slope				
Skewness	0.0037	0.0859	0.1472	N/A
Kurtosis	5.0524**	3.6934	2.9854*	N/A
Bowman-Shenton	5.4961*	3.7792	2.9854	N/A
Level[a]	F	S	S	S
Slope[a]	S	F	S	NO

Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

[a] F refers to fixed; and S refers to stochastic.

Table 6.23: Structural time-series model results in the commercial sector with *stochastic level/no slope* specification

$\log(p_t^{com})$	0.0188	(0.1320)
$\log(GDP_t)$	1.2483***	(0.3508)
$\log(p_t^{oil})$	-0.1541	(0.1420)
$\log(p_t^{gas})$	0.1396	(0.1442)
$\log(CDD_t)$	-0.1075***	(0.0257)
$\log(HDD_t)$	-0.1324	(0.0273)
$D_t^{Holiday}$	0.0040	(0.0119)
Hyperparameters		
Irregular	0.0015	
Level	0.0002	
Estimated component		
Level	-11.1864	
R^2	0.9704	

Source: Author's estimations using data from 2001 to Mar. 2013.
Standard errors are in parentheses.
Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

Table 6.24: Diagnostic tests with *stochastic level/no slope* specification

Heteroscedasticity	2.3929
BoxLjung	12.537***
Skewness	0.0352
Kurtosis	3.1383*
Bowman-Shenton	3.1735
AIC	-6.0187
BIC	-5.6526
Auxiliary	
Irregular	
Skewness	0.3720
Kurtosis	3.7626*
Bowman-Shenton	4.1346
Level	
Skewness	0.1777
Kurtosis	1.0039
Bowman-Shenton	1.1816

Significance levels: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

Chapter 7

Conclusion

I conclude this dissertation by answering the three questions in the introductory chapter regarding the demand-side restructure in the Korean electricity market.

Question 1: What are the estimates of the price elasticity of electricity demand in the wholesale and retail markets, including the residential, industrial, and commercial sectors?

In the wholesale market, based on hourly aggregate data, the estimated real-time price elasticity is -0.0034, and the corresponding long-run price elasticity is -0.0640 with the window of a couple of days. The highest price elasticities, in absolute value, during a day are around noon, and the lowest in the early hours of the morning. The price elasticities, in absolute value, are higher in heating-degree days and when district electricity businesses joined the market. Moreover, no statistical evidence shows that the linear demand curve kinks inward when the electricity price is high during the sampling periods.

In the retail market, based on monthly aggregate data, the estimated price elasticities are -0.2923 and -0.1854 in the residential and industrial sectors. Price elasticity in the commercial sector is insignificant, indicating commercial firms are not responsive to the electricity price. Since the empirical models I specified capture the equilibrium demand relationship, the price elasticities are better interpreted as long-run effects, from the viewpoint of time series, with a couple of months' window. Moreover, I show empirically that households in the residential sector respond to the previous monthly average price in aggregate.

The estimated price elasticities are all inelastic. The low price elasticity in the wholesale market might be due to the small size of the responsive customers. Although KEPCO purchases about 90% of the total load, and can respond to electricity price variations with direct load control, its operations are still mostly restrained by the unresponsive customers in the retail market. Suppose the actual responsive trade volume is approximately 14% of the total load, as in Lijsen (2007). A rough idea of the order of the price elasticities if all trade volume is under RTP can be obtained by dividing my estimated price elasticity by 0.14, yielding -0.0243 in real time, and -0.4571 in the long run. These figures are still low especially if I consider the current responsive customers are generally more responsive than the rest of the unresponsive customers. However, the size effects, in aggregate, are consequential. With real-time price elasticity -0.0243, 1% of price increase could relieve a large nuclear power station, or a coal, LNG, or oil power plant. Moreover, low price elasticities generally imply potential market power on the supply side. Most major power plants are, however, state-controlled, so market power would not be a big concern in the Korean electricity market. The low price elasticity in the retail market might be due to the fact that Korean policy makers advocate a low retail tariff in the retail market with certain political and economical considerations. The decreasing real monthly average price in the commercial market might explain why the commercial firms, in aggregate, are not responsive to the electricity price. Moreover, since the tariff structure in the industrial and commercial sectors are Time Of Use (TOU), my results would mask some own- and cross-price effects in different pricing intervals. Presumably, the absolute value of the estimated price elasticities might underestimate price elasticities in peak pricing intervals, and overestimate them in the off-peak pricing intervals. Further investigation is required with better interval consumption data.

Question 2: How do inter-temporal price changes affect consumption patterns, and what are the estimates of the inter-temporal electricity cross-price elasticities in the wholesale market?

In the wholesale market, within the previous 22 hours, the cross-price elasticities are negative, while for the 23 hour lag, the cross-price elasticity is positive but with a very small magnitude. Overall, I conclude that inter-temporal lag price changes impose major complementarity effects on current electricity load. Moreover, as the lag periods are further away, the magnitude of the cross-price elasticities weakens. The highest complementarity effects are in adjacent hours, and the cross-price elasticities with respect to the previous five hours are higher than the absolute value of the real-time price elasticity, indicating customers in the market are more responsive to the previous adjacent hours' prices than the current price, probably because customers need some response time and tend to adjust usage in block hours. I conjecture the reason for a lack of statistical evidence of a consumption shift might be that (1) SMP is not high enough in the sample to induce a consumption shift, and therefore the reduced electricity consumption might only involve some infra-marginal usage of the electricity, not the real operational work, i.e., lighting and air-conditioning, and these reductions will not shift electricity usage to later hours, and probably will get back to its original level some hours later; and (2) it is possible that the electricity usage of some individual consumers do exhibit a demand shift, but aggregation masks this behavior.

Question 3: Except for the electricity price, what other factors affect electricity consumption in the wholesale and retail markets, including the residential, industrial, and commercial sectors?

In the wholesale market, temperature, holidays, and certain types of days also affect electricity demand. Temperature elasticities are 0.0153 and -0.0109 in the cooling and heating degree days. The breaking point of cooling- and heating-degree days is 65 degrees F. Electrical cooling- and heating-usage in summer and winter cause the increasing electricity load in the market. Holidays lead to 0.22% and 0.34% hourly electricity usage reductions on normal and important holidays, respectively, and these reductions are not statistically different. The electricity usage reduction on Saturday is insignificant, but the reduction on Sunday is significant, which is about 0.29% hourly usage reduction. There is also about 0.20% reduction in electricity usage on the day before important holidays, while there is a

0.65% increase in hourly usage on the day after an important holiday and a 0.31% increase in hourly usage on Mondays. Generally, many firms reduce or stop operations on Sundays and holidays and also tend to reduce operations on the day before important holidays, leading to the electricity usage reductions. Operations need to resume and catch up after Sundays and holidays, leading an increase in electricity usage. GDP and electricity substitutes effects are not estimated in the model because mismatched frequencies and their effects would probably be nil given the short response time in the model.

In the retail market, firstly, households' income and GDP have positive and significant effects on residential, industrial, and commercial electricity usage, respectively. The estimated income elasticity in the residential sector is 1.0388, and in the industrial and commercial sectors the elasticities are 1.4585 and 1.6939, respectively. All estimates are elastic, indicating households and firms, in aggregate, are quite sensitive to economic growth. When the UEDT is controlled in the STSM specification, however, the magnitude of the estimated income elasticity, in the commercial sector, becomes smaller in absolute value, indicating that part of the UEDT effects are confounded in the income coefficient in the commercial sector when the UEDT is not specifically controlled. The estimated figure of 1.2483 is still elastic, however. These results are aligned with the literature when appliance stocks are not specifically controlled. Resis and Wright (2005) estimate residential electricity demand based on micro-data with specific appliance stocks information in California. The estimated income elasticity is 0.00 for all households. I thus conjecture that income effects on electricity usage in the residential sector might occur mostly through changes in appliance stocks.

Secondly, in the residential sector, natural gas is applied as a electricity substitute, and in the industrial and commercial sectors, both natural-gas and high-sulfur fuel oil are applied as electricity substitutes. No statistical substitution effects are detected. There is some evidence of the substitution effects of natural-gas in the industrial sector. The estimated elasticity of substitution is 0.1345. This result, however, is not robust to the STSM specification; the elasticity of substitution of natural-gas becomes insignificant once the UEDT is controlled in the STSM, indicating that the significance of the estimated natural-gas effect is probably due to the confounded UEDT. Overall, there is no strong evidence of substitution effects in all three sectors. Generally, changing the type of energy usage requires changes in the type of equipment and appliances for households and firms or starting up the on-site generators for larger firms. A lack of evidence of substitution effects might indicate that the gain associated with using a relatively cheaper form of energy is less than the cost of adjusting the type of equipment and appliances or starting up the on-site generators.

Thirdly, temperature elasticities are around -0.32 in the residential sector in both cooling- and heating-degree days, and are -0.0915 and -0.0375 in cooling-degree days, and -0.1112 and -0.0442 in heating-degree days in the commercial and industrial sectors, respectively. Electricity usage decreases as the temperature increases in all three sectors. An effect in which electrical cooling usage leads to increased electricity consumption is not detected. The reason for this might be due to the fact that monthly average data could mask the high temperature effects as detected in the wholesale market, and that the data points of cooling-degree days are limited.

Fourthly, in the residential sector, for the month with at least three days of holidays, electricity usage increases by 6%, or around 0.65 KWh, on a daily basis, but does not so in the month with less than three days of holidays. Generally, the most important holidays, such as Lunar New Year and Moon Festival, all involve people taking three days off. Families tend to get together and celebrate, which might lead to an increase in electricity usage. In the industrial and commercial sectors, holidays generally have negative effects on electricity usage. In the industrial sector, the reduction in electricity usage in the month with one day of holidays is about 1.6%, and with two days or at least three days of holidays are about 4%, on a daily basis. In the commercial sector, the months with two days of holidays have statistically the largest effects on electricity reduction, with about a 7.3% usage reduction on a daily basis. For the months with one or at least three days of holidays, the effects are not significant at the 5% level. On important holidays, not only do families get together and celebrate, as with the Black Friday holiday in the US, people also tend to use these days for things such as shopping, sight-seeing, entertaining, dining, etc. Therefore, commercial firms, including restaurants, hotels, and department stores, might need to operate at higher levels compared with normal days in order to satisfy customer demand, thus imposing smaller reductions in electricity consumption compared with the month with two days of holidays.

Fifthly, under the STSM specification, the UEDT shows a steady decreasing usage trend in the residential sector, while it exhibits an increasing trend for both the industrial and commercial sectors. I conclude that improved energy efficiency is the driving force of the UEDT in the residential sector, but not so in the other sectors.

Finally, some main policy implications I propose are that (1) More responsive customers in the wholesale market would help increase overall electricity price elasticities, and thus retail market reform by introducing more utility competition and more flexible tariff structures, for instance, RTP or CTP, to better reflect actual production costs, would provide sound choices. Such reforms, however, would need to be assessed using careful cost-benefit analysis and possible experiments before applying them to all customers to ensure the transition is smooth and will not create excessive social burdens; (2) Encouraging and subsidizing the development and application of energy-efficient appliances and equipment would help to counter the electricity increases caused by the cold and hot weather during winter and summer, economic growth, and the upward UEDT trend in the industrial and commercial sectors.

Appendices

A Data Source and Descriptive Statistics

Tables A.1 and A.2 list the source and descriptive statistics of relevant variables, respectively.

Table A.1: Analysis variables

Variable	Frequency	Data source
Aggregate electricity consumption	Hourly	KPX, 2006-2010
Wholesale System Marginal Price	Hourly	KPX, 2006-2010
Seoul temperature 1	Hourly	wunderground.com, 2006-2010
Electricity generator capacity	Hourly	KPX, 2007
Electricity generator cost	Hourly	KPX, 2007
Seoul temperature 2	Monthly	wundergroud.com 2000-2013.3
Time of Use (TOU) tariff	Daily	KEPCO 2005-2010
Consumer Price Index (CPI)	Monthly	KOSIS, 1999.12-2012
Producer Price Index (PPI)	Monthly	KOSIS, 2005-2010
Residential sector electricity consumption	Monthly	KEPCO, 2000-2012
Residential sector electricity average price	Monthly	KEPCO, 1999.12-2012
Industrial sector electricity consumption	Monthly	KEPCO, 1999.12-2013.6
Industrial sector electricity average price	Monthly	KEPCO, 1999.12-2013.3
Commercial sector electricity consumption	Monthly	KEPCO, 1999.12-2013.6
Commercial sector electricity average price	Monthly	KEPCO, 1999.12-2013.6
Korean dollar exchange rate	Monthly	Issues of Monthly Statistics of international trade, 2000-2013.5
Unemployment rate	Monthly	KOSIS, 2006-2010
Value of exported goods	Monthly	KOSIS, 2006-2010
Value of construction order received	Monthly	KOSIS, 1999.12-2013.5
Average income per household	Quarterly	KOSIS, 2003-2012
GDP	Quarterly	OECD data set and Bank of Korea, 1999.Q4-2013.Q2
Average natural-gas price for household	Quarterly	Issues of Energy Prices and Taxes, 2000-2012
Average natural-gas price for industry	Quarterly	Issues of Energy Prices and Taxes, 1999.Q4-2013.Q1
Average high-sulfur fuel oil price for industry	Quarterly	Issues of Energy Prices and Taxes, 1999.Q4-2013.Q1

Table A.2: Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Aggregate electricity consumption (MWh)	43824	47934.00	7089.50	71308	25050
Wholesale System Marginal Price (Won/KWh)	43824	99.37	33.11	335.17	3.72
Seoul temperature 1 (F)	43824	53.71	19.97	95	-5.8
Electricity base generator capacity (MW)	8760	42603.84	2298.80	46188	37831
Seoul temperature 2 (F)	162	53.19	18.62	80.84	16.45
Deflated Time of Use (TOU) tariff (peak-mid) (Won/2010 KWh)	1826	147.09	29.19	240.18	103.23
Deflated Time of Use tariff (Off) (Won/2010 KWh)	1826	34.28	3.53	43.50	26.86
Residential sector electricity consumption (KWh per household per day)	156	11.66	2.40	18.18	7.55
Deflated residential electricity average price (Won/2010 KWh)	157	109.50	14.12	162.22	76.63
Industrial sector electricity consumption (GWh per day)	163	524.43	111.65	736.24	357.13
Commercial sector electricity consumption (GWh per day)	163	211.77	55.12	340.77	111.48
Deflated industrial electricity average price (Won/2010 KWh)	163	75.45	8.97	104.60	53.36
Deflated commercial electricity average price (Won/2010 KWh)	163	113.02	15.37	159.77	89.11
Average income per household (2010 Kwon)	40	3522.74	185.40	3892.72	3168.12
Real GDP (2010 billion won)	55	910719.94	139479.51	1131864.01	668982.01
Average natural-gas price for household (2010 won/MWh GCV)	52	53454.62	12160.49	78062	35614
Average natural-gas price for industry (2010 won/MWh GCV)	54	48425.29	11372.23	71300.83	30523.02
Average high-sulfur fuel oil price for industry (2010 won/tonne)	54	768037.18	177254.39	1090882	768037.18

Source: Author's own calculation.

B Demand Identification with No Exclusion Condition

I firstly provide a heuristic proof to enlighten the mechanism of demand identification with nonlinear supply. This simple proof is revised from Maddala (1977, p.229).

Consider a demand and supply system,

$$\begin{aligned} y &= \gamma p + \delta x_t + \varepsilon_1 && \text{Demand Equation} \\ p &= m(y) + \varepsilon_2 && \text{Inverse Supply Equation} \end{aligned}$$

where x is the observable exogenous demand shifter; $\delta \neq 0$; and ε_1 and ε_2 are unobservable demand and supply disturbances. In our case, $\sigma(\varepsilon_1) \gg \sigma(\varepsilon_2)$; note that this condition does not affect identification.

Assume $m(y)$ is a nonlinear and non-decreasing function in y with a known form.

Proposition 1 *The demand parameters, γ_1 and δ, in the above system is identifiable.*

Proof. Based on Maddala (1977), I consider the system as consisting of four variables, y, p, x, and m(y), rather than three variables, y, p, and x. The coefficient matrix is thus given by,

$$\begin{bmatrix} 1 & -\gamma & -\delta & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

The corresponding restriction matrix ϕ for the demand equation is,

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Hence, $A\phi = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$ with $\rho(A\phi) = 1$, where $\rho(M)$ refers to the rank of matrix M.

Since I have two endogenous variables in the system, $\rho(A\phi) = 1$ satisfies the identification condition derived by Fisher. Nonlinear supply curve provides more restrictions of the system, so the demand function can be identified. ■

Note that the above proposition can be relaxed with an unknown functional form m up to the knowledge that it is not linear.

Next, I provide a general proof based on the non-parametric identification conditions on simultaneous equations. The non-parametric identification conditions are adopted from Matzkin (2008, 2012), introduced as the following.

Consider a simultaneous system. Let Y be the vector of G observable endogenous variables, X be a vector of K observable exogenous variables, ϵ be a vector of G unobservable variables that is distributed independently of X, and r^*: $R^{G+K} \to R^G$ specify the relationship among these variables,

$$\varepsilon = r^*(Y, X) \tag{1}$$

Define a set S consisting of vectors of function r: $R^{G+K} \to R^G$ and of density function $f_\varepsilon : R^G \to R$, satisfying the following conditions (Matzkin, 2008):

- all r are twice continuously differentiable;
- all f_ε are continuously differentiable with support R^G[1];
- for each $x \in R^K$, $r(\cdot, x)$: $R^G \to R^G$ is one-to-one and onto R^G; and
- for each $(y, x) \in R^{G+K}$, the Jacobian determinant $|\frac{\partial r(y,x)}{\partial y}|$ is strictly positive.[2]

The identification question here is whether I can uniquely recover the function r^* and the density f_ε within the set S from the conditional density $f_{Y|X=x}$ and unconditional density f_X.

[1] The differentiability of r and f_ε facilitate the condition expressions by derivative presentation.
[2] This condition is a normalization.

Theorem 1 *(Matzkin 2005) Let $M \times \Gamma$ denote the set of pairs $(r, f_\varepsilon) \in S$. The function r^* is identified in M if $r^* \in M$ and for all $f_\varepsilon \in \Gamma$ and all $\widetilde{r}$, $r \in M$ such that $\widetilde{r} \neq r$, there exist y, x such that the rank of the matrix*

$$\begin{bmatrix} (\frac{\partial \widetilde{r}(y,x)}{\partial y})' & \Delta_y(y,x;\partial r,\partial^2 r,\partial\widetilde{r},\partial^2\widetilde{r}) + (\frac{\partial r(y,x)}{\partial y})' \frac{\partial \log(f_\varepsilon(r(y,x)))}{\partial \varepsilon} \\ (\frac{\partial \widetilde{r}(y,x)}{\partial x})' & \Delta_x(y,x;\partial r,\partial^2 r,\partial\widetilde{r},\partial^2\widetilde{r}) + (\frac{\partial r(y,x)}{\partial x})' \frac{\partial \log(f_\varepsilon(r(y,x)))}{\partial \varepsilon} \end{bmatrix}$$

is strictly larger than G, where

$$\Delta_y(y,x;\partial r,\partial^2 r,\partial\widetilde{r},\partial^2\widetilde{r}) = \frac{\partial}{\partial y}\log\left|\frac{\partial r(y,x)}{\partial y}\right| - \frac{\partial}{\partial y}\log\left|\frac{\partial \widetilde{r}(y,x)}{\partial y}\right| \tag{2}$$

$$\Delta_x(y,x;\partial r,\partial^2 r,\partial\widetilde{r},\partial^2\widetilde{r}) = \frac{\partial}{\partial x}\log\left|\frac{\partial r(y,x)}{\partial y}\right| - \frac{\partial}{\partial x}\log\left|\frac{\partial \widetilde{r}(y,x)}{\partial y}\right| \tag{3}$$

Proposition 2 *Consider the same demand and supply system in Proposition 1. Assume that (1) $(\varepsilon_1, \varepsilon_2)$ has an everywhere positive and differentiable density f_ε^*; (2) x is distributed independently of $(\varepsilon_1, \varepsilon_2)$; and (3) $m(y)$ is a nonlinear, strictly increasing, and twice differentiable unknown function. Then, the demand parameters, γ and δ, are identifiable.*

Proof.

First rearrange the system as

$$\varepsilon_1 = r_1^*(y,p,x) = \quad y - \gamma^* p - \delta^* x \tag{4}$$

$$\varepsilon_2 = r_2^*(y,p,x) = \quad -m^*(y) + p \tag{5}$$

where r^* represents the true model.

The Jacobian determinant of r^*, $\left|\begin{bmatrix}\partial r^*/\partial y & \partial r^*/\partial p\end{bmatrix}\right|$, is given by,

$$D = \left|\begin{bmatrix} 1 & -\gamma^* \\ -\frac{\partial m^*(y)}{\partial y} & 1 \end{bmatrix}\right| = 1 - \frac{\partial m^*(y)}{\partial y}\gamma^*$$

D is positive if $1 > \frac{\partial m^*(y)}{\partial y}\gamma^*$. Two justifications can establish this inequality.

1. In theory, a downward-sloping demand curve and an upward-sloping supply curve indicate $\gamma^* \leq 0$ and $\frac{\partial m^*(y)}{\partial y} \geq 0$, and thus $1 > \frac{\partial m^*(y)}{\partial y}\gamma^*$ is satisfied automatically.

2. Based on Gale and Nikaido (1965), since the first diagonal element is positive, the function r^* is globally invertible if the condition $1 > \frac{\partial m^*(y)}{\partial y}\gamma^*$ holds for every y.

Let $\hat{r}$ and $\widetilde{r}$ be any two differentiable functions satisfying this invertible condition and the other properties assumed about r^*, and at some value of y, $\partial\widetilde{m}(y)/\partial y \neq \partial\hat{m}(y)/\partial y$.

Let $f_{\varepsilon_1,\varepsilon_2}$ denote any density satisfying the same properties that $f^*_{\varepsilon_1,\varepsilon_2}$ is assumed to satisfy.

Based on Theorem 1, define

$$\begin{aligned}
a_1(y,p,x) &= \tfrac{\widetilde{\gamma}(\partial^2\widetilde{m}(y)/\partial y^2)}{1-\widetilde{\gamma}(\partial\widetilde{m}(y)/\partial y)} - \tfrac{\hat{\gamma}(\partial^2\hat{m}(y)/\partial y^2)}{1-\hat{\gamma}(\partial\hat{m}(y)/\partial y)} + \tfrac{\partial}{\partial\varepsilon_1}\log f_\varepsilon(y-\hat{\gamma}p-\hat{\delta}x, p-\hat{m}(y)) \\
&\quad -\tfrac{\partial\hat{m}(y)}{\partial y}\tfrac{\partial}{\partial\varepsilon_2}\log f_\varepsilon(y-\hat{\gamma}p-\hat{\delta}x, p-\hat{m}(y)) \\
a_2(y,p,x) &= -\hat{\gamma}\tfrac{\partial}{\partial\varepsilon_1}\log f_\varepsilon(y-\hat{\gamma}p-\hat{\delta}x, p-\hat{m}(y)) + \tfrac{\partial}{\partial\varepsilon_2}\log f_\varepsilon(y-\hat{\gamma}p-\hat{\delta}x, p-\hat{m}(y)) \\
a_3(y,p,x) &= -\hat{\delta}\tfrac{\partial}{\partial\varepsilon_1}\log f_\varepsilon(y-\hat{\gamma}p-\hat{\delta}x, p-\hat{m}(y))
\end{aligned}$$

r^* is identifiable if there exists (y,p,x) such that the rank of the matrix

$$A = \begin{bmatrix} 1 & -\frac{\partial\hat{m}(y)}{\partial y} & a_1(y,p,x) \\ -\widetilde{\gamma} & 1 & a_2(y,p,x) \\ -\widetilde{\delta} & 0 & a_3(y,p,x) \end{bmatrix}$$

is 3.

I consider the identification of the demand equation. Suppose that there exist some values of (y,p,x) such that $\partial\log f_\varepsilon(\hat{r}(y,p,x))/\partial\varepsilon_1 = 1$ and $\partial\log f_\varepsilon(\hat{r}(y,p,x))/\partial\varepsilon_2 = 0$. Thus, the matrix A reduces to,

$$A' = \begin{bmatrix} 1 & -\frac{\partial\hat{m}(y)}{\partial y} & a_1' \\ -\widetilde{\gamma} & 1 & -\hat{\gamma} \\ -\widetilde{\delta} & 0 & -\hat{\delta} \end{bmatrix}$$

where $a_1' = \frac{\widetilde{\gamma}(\partial^2\widetilde{m}(y)/\partial y^2)}{1-\widetilde{\gamma}(\partial\widetilde{m}(y)/\partial y)} - \frac{\hat{\gamma}(\partial^2\hat{m}(y)/\partial y^2)}{1-\hat{\gamma}(\partial\hat{m}(y)/\partial y)} + 1$.

Suppose there exist constant λ_1 and λ_2 not all equal to zero such that the third column is a linear combination of the first two columns in A', thus,

$$\lambda_1 + \lambda_2(-\frac{\partial \hat{m}(y)}{\partial y}) = a_1' \tag{6}$$

$$\lambda_1(-\widetilde{\gamma}) + \lambda_2 = -\hat{\gamma} \tag{7}$$

$$\lambda_1(-\widetilde{\delta}) = -\hat{\delta} \tag{8}$$

The non-linearity of $m(y)$ implies the following:

1. In equation (6), $\partial \hat{m}(y)/\partial y$ cannot be a constant. Thus, in order to hold the equality in equation (7), I need to impose $\lambda_2 = 0$
2. $\partial^2 m(y)/\partial y^2 \neq 0$. Thus, $a_1' - 1$ is not a constant.[3] Thus, there cannot exist a constant λ_1 to hold the equality in equation (6) for all possible (y,p,x), which violates the equality in equations (7) and (8).

Thus, there exists some value (y,p,x) such that $\rho(A') = 3$, and thereby the coefficients in the demand equation are identifiable.

Note that if $m(y)$ is linear, then equation (6) will not depend on y. The linear combination among columns can be established by rearranging the values of the parameters to obtain $\rho(A') = 2$, which breaks the identification of the demand equation.

■

Matzkin (2012, p.31) shows that the first-order derivative of $m(y)$ is also identifiable.

Proposition 3 *Consider the same demand and supply system in Proposition 1 except y in demand equation is nonlinear in p, as* $y = g(p) + \delta x + \varepsilon_1$. *Assume that (1)* $(\varepsilon_1, \varepsilon_2)$ *has an everywhere positive and differentiable density* f_ε^*; *(2) x is distributed independently of* $(\varepsilon_1, \varepsilon_2)$; *and (3) both* $g(p)$ *and* $m(y)$ *are nonlinear and twice differentiable unknown functions,*

[3] It can be shown by simply algebra that $a_1' - 1$ is a constant only when $\partial^2 \widetilde{m}(y)/\partial y^2 = \partial^2 \hat{m}(y)/\partial y^2 = 0$.

in which g(p) is strictly decreasing in p, while m(y) is strictly increasing in y. Then, the demand parameters, δ, and the first-order derivative of $g(p)$ are identifiable.

Proof. The proof follows the same pattern as Proposition 2 so I omit it here. ■

Finally, I could also relax the restrictions on m(y) by allowing it to be a non-decreasing, non everywhere differentiable function, such as an increasing-step function,

$$P = \begin{cases} p_1 & y \in (0, k_1] \\ \vdots & \\ p_n & y \in (k_{n-1}, k_n] \\ \vdots & \\ p_N & y \in (k_{N-1}, \infty) \end{cases} \tag{9}$$

where $0 < k_1 < \cdots < k_n < \cdots < k_{N-1} < \infty$.

References

[1] Energy Power Exchange. http://www.kpx.kr/english/.

[2] Energy Vortex. http://www.energyvortex.com.

[3] Korea Electric Power Corporation. http://www.kepco.co.kr/eng/.

[4] Korean Statistical Information Service. http://kosis.kr/eng/.

[5] Weather Underground. http://www.wunderground.com/.

[6] Dennis Aigner. *The Residential Electricity Time-of-use Pricing Experiments: What Have We Learned?* Social Experimentation. University of Chicago Press, 1982.

[7] Dennis Aigner and J. Dale Poirier. *Electricity Demand and Consumption by Time-of-use: A Survey*. Palo Alto, CA, 1979.

[8] Dennis J. Aigner and Jerry A. Hausman. Correcting for truncation bias in the analysis of experiments in time-of-day pricing of electricity. *Bell Journal of Economics*, 11(1):131–142, 1980.

[9] Dennis J. Aigner and Edward E. Leamer. Estimation of time-of-use pricing response in the absence of experimental-data – an application of the methodology of data transferability. *Journal of Econometrics*, 26(1-2):205–227, 1984.

[10] M. H. Albadi and E. F. El-Saadany. A summary of demand response in electricity markets. *Electric Power Systems Research*, 78(11):1989–1996, 2008.

[11] Anna Alberini and Massimo Filippini. Response of residential electricity demand to price: The effect of measurement error. *Energy Economics*, 33(5):889 – 895, 2011.

[12] Hunt Allcott. Rethinking real-time electricity pricing. *Resource & Energy Economics*, 33(4):820–842, 2011.

[13] Joshua D. Angrist and Jorn-Steffen Pischke. The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives*, 24(2):3–30, 2010.

[14] Elie Appelbaum. On the choice of flexible functional forms. *International Economic Review*, 20(2):449–458, 1979.

[15] N. R. Badurally Adam, M. K. Elahee, and M. Z. Dauhoo. Forecasting of peak electricity demand in Mauritius using the non-homogeneous gompertz diffusion process. *Energy*, 36(12):6763–6769, 2011.

[16] Ross Baldick, Ryan Grant, and Edward Kahn. Theory and application of linear supply function equilibrium in electricity markets. *Journal of Regulatory Economics*, 25(2):143–167, 2004.

[17] Cajsa Bartusch, Fredrik Wallin, Monica Odlare, Iana Vassileva, and Lars Wester. Introducing a demand-based electricity distribution tariff in the residential sector: Demand response and customer perception. *Energy Policy*, 39(9):5008–5025, 2011.

[18] David A. Belsley and Roy E. Welsch. Modeling energy consumption – Using and abusing regression diagnostics: Comment. *Journal of Business and Economic Statistics*, 6(4):442–447, 1988.

[19] Stephen R. Blough. The relationship between power and level for generic unit root tests in finite sample. *Journal of Applied Econometrics*, 7(3):295–308, 1992.

[20] Doglas Bohi. *Analyzing Demand Behavior: A Study of Energy Elasticities*. Baltimore: Published for Resources for the Future by the Johns Hopkins University Press, 1981.

[21] Richard Boisvert, Peter Cappers, Bernie Neenan, and Scott Bryan. Industrial and commercial customer response to real time electricity prices. *Neenan Associates*, 2006.

[22] Richard N. Boisvert, Peter Cappers, Charles Goldman, Bernie Neenan, and Nicole Hopper. Customer response to RTP in competitive markets: A study of Niagara Mohawk's standard offer tariff. *Energy Journal*, 28(1):53–74, 2007.

[23] Ettore Bompard, Ma Yuchao, Roberto Napoli, and Graziano Abrate. The demand elasticity impacts on the strategic bidding behavior of the electricity producers. *IEEE Transactions on Power Systems*, 22(1):188–197, 2007.

[24] Severin Borenstein. Equity effects of increasing-block electricity pricing. *CSEM Working Paper 180*, 2008.

[25] Severin Borenstein, James B. Bushnell, and Frank A. Wolak. Measuring market inefficiencies in California's restructured wholesale electricity market. *American Economic Review*, 92(5):1376–1405, 2002.

[26] Peter Cappers, Charles Goldman, and David Kathan. Demand response in U.S. electricity markets: Empirical evidence. *Energy*, 35(4):1526–1535, 2010.

[27] Robert G. Chambers. *Applied Production Analysis: A Dual Approach*. Cambridge University press, Cambridge, 1988.

[28] In-Koo Cho and Hyunsook Kim. Market power and network constraint in a deregulated electricity market. *Energy Journal*, 28(2):1–34, 2007.

[29] Wai Choi, Anindya Sen, and Adam White. Response of industrial customers to hourly pricing in Ontario's deregulated electricity market. *Journal of Regulatory Economics*, 40(3):303–323, 2011.

[30] Carl F Christ. *Econometric Models and Methods*. New York: Wiley, NW, 1966.

[31] Su Chua-Liang and Daniel Kirschen. Quantifying the effect of demand response on electricity markets. *IEEE Transactions on Power Systems*, 24(3):1199–1207, 2009.

[32] James Cochell, Peter Schwarz, and Thomas Taylor. Using real-time electricity data to estimate response to time-of-use and flat rates: an application to emissions. *Journal of Regulatory Economics*, 42(2):135–158, 2012.

[33] Dora L. Costa and Matthew E. Khan. Energy conservation "nudges" and environmentalist ideology: Evidence from a randomized residential electricity field experiment. *NBER working papers*, 2010.

[34] Theologos Dergiades and Lefteris Tsoulfidis. Estimating residential demand for electricity in the United States, 19652006. *Energy Economics*, 30(5):2722–2730, 2008.

[35] Walter E. Diewert. *Functional Form in the Theory of Production and Consumer Demand*. PhD dissertation, University of California at Berkeley.

[36] Walter E. Diewert and Terence J. Wales. Flexible functional forms and global curvature conditions. *Econometrica*, 55(1):43–68, 1987.

[37] V. Dordonnat, S. J. Koopman, M. Ooms, A. Dessertaine, and J. Collet. An hourly periodic state space model for modelling French national electricity load. *International Journal of Forecasting*, 24(4):566–587, 2008.

[38] Larry G. Epstein. Duality theory and functional forms for dynamic factor demands. *Review of Economic Studies*, 41(1):81–95, 1981.

[39] Larry G. Epstein. Comparative dynamics in the adjustment-cost model of the firm. *Journal of Economic Theory*, 27(1):77–100, 1982.

[40] Larry G. Epstein. Integrability of incomplete systems of demand functions. *Review of economic studies*, 49(3):411–425, 1982.

[41] Ahmad Faruqui and Sanem Sergici. Dynamic pricing of electricity in the mid-atlantic region: econometric results from the Baltimore gas and electric company experiment. *Journal of Regulatory Economics*, 40(1):82–109, 2011.

[42] Carlo Fezzi and Derek Bunn. Structural analysis of electricity demand and supply interactions structural analysis of electricity demand and supply interactions electricity demand and supply interactions. *Oxford Bulletin of Economics & Statistics*, 72(6):827–856, 2010.

[43] Denzil G. Fiebig, Robert Bartels, and Dennis J. Aigner. A random coefficient approach to the estimation of residential end-use load profiles. *Journal of Econometrics*, 50(3):297–327, 1991.

[44] Massimo Filippini. Short- and long-run time-of-use price elasticities in Swiss residential electricity demand. *Energy Policy*, 39(10):5811–5817, 2011.

[45] Lee S. Friedman. The long and the short of it: California's electricity crisis. *International Journal of Public Policy*, 4(1):4–31, 2009.

[46] Melvyn Fuss, Daniel Mcfadden, and Yair Mundlak. A survey of functional forms in the economic analysis of production. In Melvyn Fuss and Daniel McFadden, editors, *Production economics: A dual approach to theory and applications*, volume 1, pages 219–268. North-Holland, Amsterdam, 1978.

[47] Arthur S. Goldberger. Structural equation methods in the social sciences. *Econometrica*, 40(6):979–1001, 1972.

[48] Clive W. J. Granger. On modelling the long run in applied economics. *The Economic Journal*, 107(440):169–177, 1997.

[49] Richard Green. Increasing competition in the British electricity spot market. *Journal of Industrial Economics*, 44(2):205, 1996.

[50] Richard J. Green and David M. Newbery. Competition in the British electricity spot market. *Journal of Political Economy*, 100(5):929, 1992.

[51] William H Greene. *Econometric Analysis*. Prentice Hall, Upper Saddle River, NJ, 2008.

[52] Zvi Griliches. Data and econometricians-the uneasy alliance. *American Economic Review*, 75(2):196–200, 1985.

[53] Niels Haldrup and Morten Nielsen. A regime switching long memory model for electricity prices. *Journal of Econometrics*, 135(1/2):349–376, 2006.

[54] Karen Herter. Residential implementation of critical-peak pricing of electricity. *Energy Policy*, 35(4):2121–2130, 2007.

[55] Karen Herter, Patrick McAuliffe, and Arthur Rosenfeld. An exploratory analysis of California residential customer response to critical peak pricing of electricity. *Energy*, 32(1):25–34, 2007.

[56] Ali Hortacsu and Steven L. Puller. Understanding strategic bidding in multi-unit auctions: a case study of the Texas electricity spot market. *RAND Journal of Economics (Blackwell Publishing Limited)*, 39(1):86–114, 2008.

[57] Ronald Huisman, Christian Huurman, and Ronald Mahieu. Hourly electricity prices in day-ahead markets. *Energy Economics*, 29(2):240–248, 2007.

[58] R. Inglesi-Lotz. The evolution of price elasticity of electricity demand in South Africa: A Kalman filter application. *Energy Policy*, 39(6):3690–3696, 2011.

[59] Koichiro Ito. Do consumers respond to marginal or average price? Evidence from non-linear electricity pricing. *American Economic Review*, forthcoming.

[60] Faisal Jamil and Eatzaz Ahmad. Income and price elasticities of electricity demand: Aggregate and sector-wise analyses. *Energy Policy*, 39(9):5519–5527, 2011.

[61] John Johnston. *Econometric Methods*. New York: McGraw-Hill, NW, 1984.

[62] Paul L. Joskow. The difficult transition to competitive electricity markets in the U.S. *Cambridge Working Papers in Economics*.

[63] Paul L. Joskow. California's electricity crisis. *Oxford Review of Economic Policy*, 17(3):365–388, 2001.

[64] Paul L. Joskow and Edward Kahn. A quantitative analysis of pricing behavior in California's wholesale electricity market during summer 2000. *The Energy Journal*, 23(4):1–35, 2002.

[65] David R. Kamerschen and David V. Porter. The demand for residential, industrial and total electricity, 19731998. *Energy Economics*, 26(1):87, 2004.

[66] Nektaria V. Karakatsani and Derek W. Bunn. Intra-day and regime-switching dynamics in electricity price formation. *Energy Economics*, 30(4):1776–1797, 2008.

[67] Peter Kennedy. *A Guide to Econometrics*. Wiley-Blackwell, six edition, 2008.

[68] Christopher R. Knittel and Michael R. Roberts. An empirical examination of restructured electricity prices. *Energy Economics*, 27(5):791–817, 2005.

[69] Lawrence J. Lau. Profit functions of technologies with multiple inputs and outputs. *Review of Economics and Statistics*, 54(3):281–289, 1972.

[70] Lawrence J. Lau. Applications of profit functions. In Melvyn Fuss and Daniel McFadden, editors, *Production economics: A dual approach to theory and applications*, volume 1, pages 133–216. North-Holland, Amsterdam, 1978a.

[71] Lawrence J. Lau. Testing and imposing monotonicity, convexity and quasiconvexity constraints. In Melvyn Fuss and Daniel McFadden, editors, *Production economics: A dual approach to theory and applications*, volume 1, pages 269–286. North-Holland, Amsterdam, 1978b.

[72] Chien-Chiang Lee and Yi-Bin Chiu. Electricity demand elasticities and temperature: Evidence from panel smooth transition regression with instrumental variable approach. *Energy Economics*, 33(5):896–902, 2011.

[73] Mark G. Lijesen. The real-time price elasticity of electricity. *Energy Economics*, 29(2):249–258, 2007.

[74] C. A. Knox Lovell. Production economics: A dual approach to theory and applications. vol. 1: The theory of production. vol. 2: Applications of the theory of production. *Journal of Political Economy*, 88(6):1265–1269, 1980.

[75] G. S. Maddala. *Econometrics*. New York: McGraw-Hill, NW, 1977.

[76] G. S. Maddala. *Introduction to Econometrics*. New York: Macmillian, NW, 2nd edition, 1992.

[77] Rosa L. Matzkin. Identification in nonparametric simultaneous equations models. *Econometrica*, 76(5):945–978, 2008.

[78] Hassan Molana. The time-series consumption function: Error correction, random walk and the steady-state. *Economic Journal*, 101(406):382–403, 1991.

[79] Paresh Kumar Narayan and Russell Smyth. The residential demand for electricity in Australia: an application of the bounds testing approach to cointegration. *Energy Policy*, 33(4):467–474, 2005.

[80] Francisco J. Nogales, Javier Contreras, Antonio J. Conejo, and Rosario Espinola. Forecasting next-day electricity prices by time series models. *IEEE Transactions on Power Systems*, 17(2):342, 2002.

[81] Michael Parti and Cynthia Parti. The total and appliance-specific conditional demand for electricity in the household sector. *Bell Journal of Economics*, 11(1):309–321, 1980.

[82] Robert Patrick and Wolak Frank. Estimating the customer-level demand for electricity under real-time market prices. *NBER Working Paper*, 2001.

[83] Steven L. Puller. Pricing and firm conduct in California's deregulated electricity market. *Review of Economics & Statistics*, 89(1):75–87, 2007.

[84] Trout Rader. Normally, inputs are never gross substitutes. *Journal of Political Economy*, 76(1):38–43, 1968.

[85] B. Bhaskara Rao. Estimating short and long-run relationships: a guide for the applied economist. *Applied Economics*, 39(13):1613–1625, 2007.

[86] Peter C. Reiss and Matthew W. White. Household electricity demand revisited. *Review of Economic Studies*, 72(252):853–883, 2005.

[87] Suleiman Sa'ad. Electricity demand for the South Korean residential sector. *Energy Policy*, 37(12):5469–5474, 2009.

[88] Yasuhiro Sakai. Substitution and expansion effects in production theory: The case of joint production. *Journal of Economic Theory*, 9(3):255–274, 1974.

[89] Borenstein Severin. To what electricity price do consumers respond? Residential demand elasticity under increasing-block pricing. 2009.

[90] Lacir Jorge Soares and Leonardo Rocha Souza. Forecasting electricity demand using generalized long memory. *International Journal of Forecasting*, 22(1):17–28, 2006.

[91] Douglas Staiger and James H. Stock. Instrumental variables regression with weak instruments. *Econometria*, 65(3):557–586, 1997.

[92] Lester D. Taylor. The demand for electricity: a survey. *Bell Journal of Economics*, 6(1):74–110, 1975.

[93] Thomas N. Taylor and Peter M. Schwarz. The long-run effects of a time-of-use demand charge. *The RAND Journal of Economics*, 21(3):431–445, 1990.

[94] Thomas N. Taylor, Peter M. Schwarz, and James E. Cochell. 24/7 hourly response to electricity real-time pricing with up to eight summers of experience. *Journal of Regulatory Economics*, 27(3):235–262, 2005.

[95] Demand Forecasting Team. *Short-term Electricity Forecasting.*

[96] Mariano Ventosa, álvaro Baíllo, André Ramos, and Michel Rivier. Electricity market modeling trends. *Energy Policy*, 33(7):897–913, 2005.

[97] Frank Wolak. Residential customer response to real-time pricing: The Anaheim critical-peak pricing experiment. 2006.

[98] Frank A. Wolak, Robert Nordhaus, and Carl Shapiro. The competitiveness of the California energy and ancilliary services markets.

[99] Catherine D. Wolfram. Measuring duopoly power in the British electricity spot market. *American Economic Review*, 89(4):805–826, 1999.

[100] Ni Xiao, Jay Zarnikau, and Paul Damien. Testing functional forms in energy modeling: An application of the Bayesian approach to U.S. electricity demand. *Energy Economics*, 29(2):158–166, 2007.

[101] Jay Zarnikau. Functional forms in energy demand modeling. *Energy Economics*, 25(6):603–613, 2003.

[102] P. Zhou, B. W. Ang, and K. L. Poh. A trigonometric grey prediction approach to forecasting electricity demand. *Energy*, 31(14):2503–2511, 2006.

[103] Suling Zhu, Jianzhou Wang, Weigang Zhao, and Jujie Wang. A seasonal hybrid procedure for electricity demand forecasting in China. *Applied Energy*, 88(11):3807–3815, 2011.