NTNU – Trondheim
Norwegian University of
Science and Technology

Using Quantile Regression for Modeling of Electricity Price and Demand

Linh Phuong Catherine Do

Industrial Economics and Technology Management
Submission date: June 2015
Supervisor: Peter Molnar, IØT

Norwegian University of Science and Technology
Department of Industrial Economics and Technology Management

Objective

The German electricity market has undergone significant changes in the recent years. Increasing infeed from renewable sources had lead to higher risk in the German electricity market. This paper suggests that quantile regression model will give helpful insight of the electricity market and risk analysis.

First, we employ quantile regression model to investigate and compare electricity demand and residual demand model. Second, we analyze the electricity price by using quantile regression.

Preface

This mater thesis was written for the degree of Master of Technology at the Norwegian University of Science and Technology (NTNU), Department of Industrial Economics and Technology Management within the field of Investments, Finance and Management Accounting.

This thesis includes two articles. Article 1 “Demand and Residual Demand Modeling using Quantile Regression” and article 2 “Day Ahead Electricity Prices Modeling using Quantile Regression”.

The author would like to thank post doctor Peter Molnar for his valuable discussion and constructive feedback. Least, the author would also like to thank Statkraft for kindly providing data that was not publicly available.

Trondheim 6 June 2015

Link Phuong Catherine Do

Abstract

Article 1 “Demand and Residual Demand Modelling using Quantile Regression”.

Residual demand, the difference between demand and renewable production, is important variable in predicting the future price and the future need for energy storage for intermittent renewables production. The residual demand represents the load that can not be met by renewable production and must be served by conventional power plant, electricity imports or storage capacity. However, little is known about predicting the residual demand itself as well as its quantiles. We therefore model demand and residual demand using ordinary and linear quantile regression, and thereafter compare the results for the hourly electricity consumption in Germany. We find that that the residual demand is less predictable than demand. Our paper makes two contributions to the literature: (1) unlike other studies it analyses the residual demand by using quantile regression (2) it compares the results of demand and residual demand.

Article 2 “Day Ahead Electricity Price Modelling using Quantile Regression”.

This paper analysis the relation between several fundamental variables and German day-ahead electricity price for each hour. The study performed quantile regression on the electricity prices and reveals important effects that are missed by ordinary regression. Ordinary regression would assume that the relation to be the same for high and normal electricity prices on a specific hours. While the quantile regression measures the dependence of the extreme event. Examine these extreme event on the price is an important aspect of effective risk management. The results indicate that the effect from the factors on electricity price vary substantially across the quantiles, thus confirming the high complexity of the electricity price.

Sammendrag

Artikkel 1 “Etterspørsel og Residual etterspørsel modellering med Kvantilregresjon”.

Residual etterspørsel, differansen mellom elektrisitets forbruk og fornybar produksjon, er en viktig variabel for å predikere fremtidig kraftpriser og fremtidig behov for energilagring for fornybarproduksjon. Residual etterspørsel representerer forbruk som ikke kan dekkes av fornybar produksjon og må betjenes av termiske kraftverk, kraftimport eller energilagring. Lite er kjent om prediksjon av residual etterspørselen, så vel om dens kvantiler. Vi har derfor modellert og sammenlignet elektrisitets etterspørsel og residual etterspørsel ved hjelp av vanlig klassisk regresjon og lineær kvantilregresjon. Resultatene fra denne artikkelen tyder på at residual etterspørselen er mindre forutsigbart enn elektrisitets etterspørselen. Denne artikkelen gjør to bidrag til litteraturen: (1) I motsetning til andre studier analyserer denne artikkelen residual etterspørselen ved hjelp av kvantilregresjon (2) Den sammenligner resultatene av elektrisitets etterspørsel og residual etterspørsel.

Artikkel 2 ”Kraftpris modellering ved hjelp av Kvantilregresjon.”

Denne artikkelen analyserer forholdet mellom flere avhengig variabler og spot-priser i det tyske kraftmarkedet. Studiet utfører kvantilregresjon på strømpriser og avslører viktige effekter som er fraværende med klassisk regresjonsanalyse. Klassisk regresjon antar at forholdet er den samme for både høye og normale strømpriser for en bestemt time. Mens kvantilregresjon måler også avhengigheten av ekstreme kraftpriser. Å undersøke ekstrempriser er en viktig del av effektiv risikostyring. Resultatene fra denne artikkelen tyder på at effekten fra de avhengige variablene på kraftprisen varierer for de ulike kvantilene, og bekrefter at kraftprisen er kompleks.

DEMAND AND RESIDUAL DEMAND MODELLING USING QUANTILE REGRESSION

Linh Phuong Catherine Do[1], Peter Molnar[2]

Residual demand, the difference between demand and renewable production, is important variable in predicting the future price and the future need for energy storage for intermittent renewables production. The residual demand represents the load that can not be met by renewable production and must be served by conventional power plant, electricity imports or storage capacity. However, little is known about predicting the residual demand itself as well as its quantiles. We therefore model demand and residual demand using ordinary and linear quantile regression, and thereafter compare the results for the hourly electricity consumption in Germany. We find that that the residual demand is less predictable than demand. Our paper makes two contributions to the literature: (1) unlike other studies it analyses the residual demand by using quantile regression (2) it compares the results of demand and residual demand.

Keywords: demand modelling, residual demand, renewables, quantile regression

1. Introduction

In the recent years, Germany has established environmental policies to phase out nuclear power and promote progressive replacement of fossil fuels by renewables sources. From support schemes for renewables energy, Renewable Energy Act (EGG), the renewable got priority access to the grid and subsidies by fixed feed in tariffs. At the same time, the efficiency of the renewables technology is improving; the economies of scale lead to lower component cost. As a result, the renewable installed capacities have grown continuously. According to a report by Wirth H. (2015), the German installed renewables account for 31% of total production; the goal is to reach 35% of renewable energy by 2020, as well reduce CO_2 emissions and increase energy efficiency. The report also states that during 2014, the renewable sources have contributed 31% of net electricity consumption on a normal day and up to 50% on weekend.

[1] Norwegian University of Science and Technology, Norway, e-mail: linhphuo@stud.ntnu.no
[2] Norwegian University of Science and Technology, Norway, email: peter.molnar@iot.ntnu.no

The increasing amount of renewable sources and their volatility in production has introduced challenges for different market participants. The power producers need to consider the fluctuations from both load and renewable energy infeed when submitting daily price bids. A market with high infeed of renewable, like Germany, requires a more integrated demand model. As for the grid operators, increasing renewable infeed is challenging both from the perspective of stability of the grid and security of supply. They need to balance the demand and the supply. Since the production of renewable sources is price inelastic[3], it makes sense to look at the balancing problem as balancing the demand minus renewables with the supply of conventional power producer[4].

The penetration of renewable sources into the supply mix has introduced two extreme and challenging situations: high and low residual demand (Nicolosi M., 2012). Firstly, the maximum residual demand is the condition when the demand for electricity is high and at the same time the amount of renewable production is low. This situation requires flexible conventional power plants that can ramp up, electricity imports or storage systems. This has initiated discussions regarding different forms of capacity markets, potentially replacing the traditional energy market. Another solution to high residual demand is incorporating flexible demand, where the large industrial consumers are willing to reduce their consumption by selling the already purchased demand. The second situation is low residual demand; the demand for electricity is low and at the same time the amount of electricity produced by renewable is high. The transmission and distribution grid can develop into a bottleneck when the renewable energy sources generate sufficient electricity. This setting can happen in weekend or holiday with high renewable production. The situation with low residual demand requires enhance of transmission grid, flexible conventional power plants and the possibility to increase the export from Germany.

Residual demand is one of the main characteristics in German power market. It specifies the maximum market share left for the conventional power producer. We will therefore in this paper closely examine residual demand and its fundamental variables.

The word residual demand has not reached a common definition. In this paper we use the term residual demand as a demand minus wind and solar electricity production. This distinction is meaningful because wind and solar electricity producers supply electricity independently on the price. Hence, we can consider wind and solar electricity as a negative demand to the system.

[3] The renewable energy production is price inelastic because it does not react to price changes.
[4] Conventional power producer include power producer using fossil fuels and nuclear.

Earlier studies use the residual demand in strategic price bidding in day-ahead market (Baillo A. et al., 2004; Vazques S. et al., 2013), and forward market (Wagner A., 2014). Motamedi A. (2012) provides a residual demand model to forecast electricity prices. Schill W. P. (2014) uses residual demand for energy system analysis and analyzes flexibility options with storage technologies. The previously studies approach different aspects of residual demand. However, little is known about predicting the residual demand itself as well as its quantiles.

Most of the researchers have modeled electricity demand with a traditional ordinary least square method. This method is useful for finding the tendencies and the average relation between the demand and the explanatory variables. The alternative quantile regression method, introduced by Koenker R. & Basset Jr G. (1978), evaluates the dependence of the normal and the extreme event. The extreme event constitute a major source of risk to market participants in the electricity market. Hence, examine these extreme event on the electricity consumption is important part in risk management.

The quantile regression application has been widely applied in financial risk management and been recently used in energy market studies: household energy consumption (Kaza . 2010), oil prices (Lee C. C. and Zeng J. H., 2011), on electricity price (Hagfors L. I. et al. 2014), CO_2 emission allowance price (Hammoudeh S. et al. 2014). This paper aims to contribute to the quantile regression literature by applying this method on both the aggregated electricity demand and residual demand. This analysis is relevant because it provides a more comprehensive picture of the effects from the variables on the electricity demand/residual demand in normal time and periods with extreme demand/residual demand.

This paper is organized as follows. Section 2 describes the data used to model demand and residual demand. The results from ordinary and linear quantile regressions are presented and compared in sections 3 and 4, respectively. Finally, concluding remarks are given in section 5.

2. Data

In this section we will first describe the fundamental variables that we used to model demand. We will thereafter analyze the load and renewables data, and then combine these two variables in order to obtain the residual demand.

The electricity demand is influenced by at least the following variables: trend, weather and holidays (Genethliou D. et al. 2014). Instead of using some deterministic function as a trend, we use economic trend approximated by the industrial production. As we can see on figure 1 the electricity load pattern depends on the day of the week. We therefore introduce six dummies for the days of the week, where Wednesday is taken as base weekday

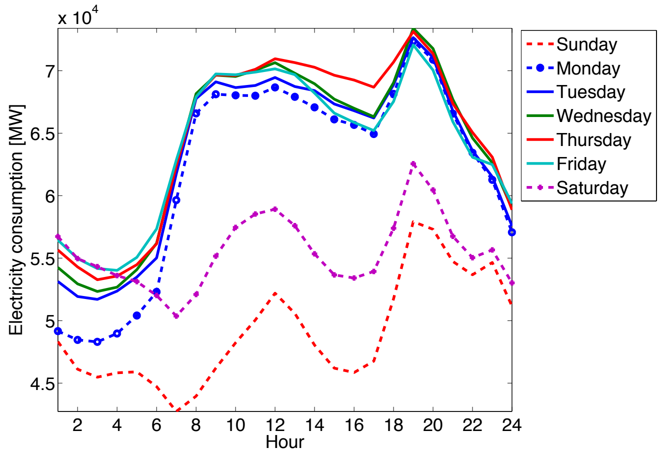

Figure 1 Typical load pattern

Incorporating religious and public holidays effects are important in creating load forecast, because the electricity consumption on a holiday is usually lower than normal day (Fezzi C., 2007). Similar to Pardo A. et al. (2002) we describe the holiday effect by incorporating binary dummy variables. We distinguish the different load reductions into two distinct groups, Minor and Major holiday, because the Minor holidays has lower load reduction than Major holidays. We also consider dummies for one day lagged Major holiday, because of the effect on adjacent days. Further, details on the composition of Minor and Major holidays variables are elaborated in Appendix A.

There are several weather variables that are likely to effect electric consumption. The average outside temperature is most commonly used among researchers, and we use this variable in our paper. The temperature data is taken from the cities with highest population densities and geographically dispersed. We choose to retrieve temperature data from Munich, Berlin, Dusseldorf and Stuttgart. The average daily temperature from these four cities is used in our models. Figure 2 depicts that the relationship between temperature and load is non linear; the temperature and the load has an increasing linear relationship when the temperature is above 20C^{o} and decreasing relationship when temperature is below

17C°. The break appears to be around 18C°. There are several options to model the temperature: One of them is quadratic function (Gupta, E., 2011), another method is Logistic Smooth Transition model (Cancelo, J. R. et al., 2008). The traditional approach is to divide the model into two linear parts by transforming the average temperature to Heating Degree Days (HDD) and Cooling Degree Days (CDD) (Pardo A. et al., 2002). We choose to omit CDD, because it is not relevant for Germany.

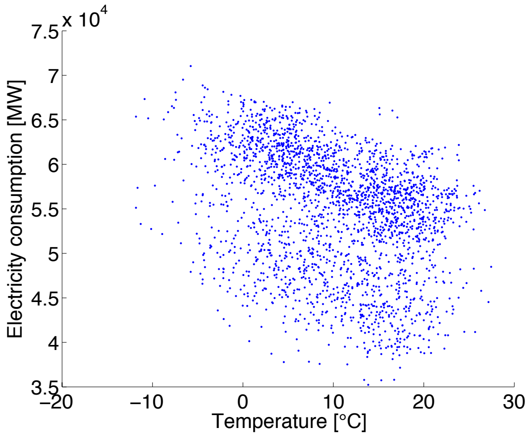

Figure 2 Scatterplot of the total load and the average outdoor temperature.

In additionally to the HDD, we use hours of daylight (DL) in order to reduce bias of electricity demand sensitivity to temperature variables. High DL will reduce energy usage for lighting and usage related to activities that are usually indoors (Molnar P., 2011). Furthermore, the DL can explain most of the calendar effect of the electricity consumption in Germany (Do L. & Molnar P., 2014).

Industrial Production (IP) captures economic conditions in the country. Electricity consumption depends on Industrial Production, and particularly in a country like Germany, where 43% of the industry belongs to energy intensive industries (IEA, 2014).

Table 1 presents the explanatory variables used in this paper to model demand and residual demand. Table 2 denotes whether the explanatory variables is daily or hourly granularity.

Table 1

Overview of fundamental variables used in the analysis

Variable	Description	Data source
Demand lag	The aggregated demand for the same hour of the previous day.	European Network of Transmission System Operators: www.entsoe.eu
Residual demand lag	Residual demand is actual renewable production subtracted from demand. Residual demand lag is residual demand for the same hour of the previous day	European Network of Transmission System Operators: www.entsoe.eu Transmission system operators: www.50Hertz.com , www.amprion.de www.transenbw.de , www.tennetso.de
Actual Solar electricity Infeed.	The actual aggregated solar electricity production in Germany.	Transmission system operators: www.50Hertz.com , www.amprion.de www.transenbw.de , www.tennetso.de
Expected Wind electricity Infeed	Forecasted aggregated wind infeed in Germany. German transmission system operators publish this data in the late afternoon the day before the delivery day.	Transmission system operators: www.50Hertz.com , www.amprion.de www.transenbw.de , www.tennetso.de
HDD	Heating degree days is an indication for the need of heating, $HDD = \max(T_{ref} - T, 0)$ where T_{ref} is the reference temperature equal 18 degrees, and T describes the weighted average outdoor temperature for the day. The temperature data is taken from the cities with highest population densities and are geographically spread: Munich, Berlin, Dusseldorf and Stuttgart.	The German Weather Service: www.dwd.de
IP lag	Three months moving average on the Industrial Production time-series (IP) is applied to smooth out jumps. IP lag is the moving average industrial production value on the previous day.	OECD Statistics: stats.oecd.org
Mon, Tue, Thu, Fri, Sat, Sun	Binary dummy variables, where Wednesday is taken as base weekday.	Calendar: www.timeanddate.com
Holiday	Binary dummy variable on major holiday and holidays with high load reduction. For more information about the composition of this variable, see appendix A.	Own data National holidays: www.bmi.bund.de School holiday: www.holidays-info.com
Holiday lag	Binary dummy variable on the day before holidays.	Own data National holidays: www.bmi.bund.de School holiday: www.holidays-info.com
Minor holiday	Binary dummy variable on minor holiday, local holidays and holidays with lower load reduction. For more information about the composition of this variable, see appendix A.	Own data Local holidays in Germany: www.timeanddate.com

DL	Hours of Daylight (DL) is determined by first calculating the sun's inclination angle λ_t where l_t is [1,365] and 1 represent January 1st etc. Thereafter calculate *DL*, where δ is the latitude in Germany, see Kamstra M. J. et al (2003). $\lambda_t = 0.4102 \sin\left(\frac{2\pi}{365}(l_t - 80.25)\right)$ $DL_t = 7.722 \arccos\left(-\tan\left(\frac{2\pi\delta}{360}\tan(\lambda_t)\right)\right)$	Own data

Table 2

Data granularity of the explanatory variables in our model

Variable	Resolution
Demand lag	Hourly
Residual Demand lag	Hourly
IP	Daily
HDD	Daily
DL	Daily
Expected Wind	Hourly
Actual solar Production	Hourly
Mon, Tue, Thu, Fri, Sat, Sun	Daily
Major holiday	Daily
Major holiday lag	Daily
Minor holiday	Daily

The German hourly electricity load data is retrieved from the European Network of Transmission System Operators for Electricity. Our dataset contains data from July 1, 2011 to July 1, 2013. The load data is the hourly average active power consumed by all installation connected to the central and the distribution network. This load data includes the production from conventional power plant and network feed-in from renewables.

The biggest share of renewables production in Germany consists of wind and solar. Moreover, production of these two renewables is completely price inelastic. We therefore focus only on these two renewables in our paper and use the term renewables as interchangeably with wind and solar. The wind and solar production data have been converted from 15 min data to hourly data.

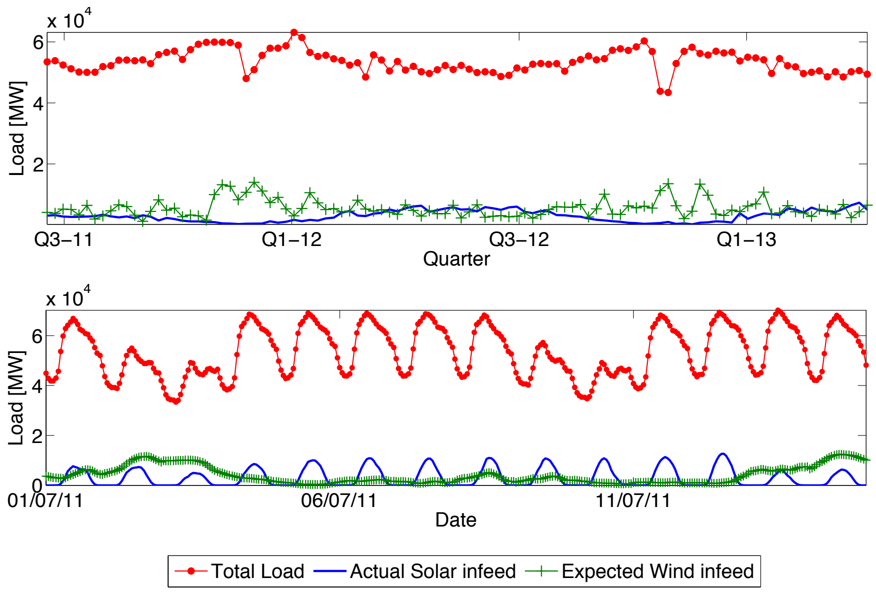

Figure 3 Total load, wind infeed and solar infeed structure in Germany. The top panel use weekly values and bottom panel use hourly values.

We further want to study how renewables sources are related to the electricity demand. The top panel in figure 3 shows that the demand and the wind production have slightly similar seasonally pattern. In general, the energy consumption and the average wind production are higher during winter than during summer. As oppose to the wind production, the average solar production is highest in summer and lowest in winter.

The relationship between the renewables and the load data is examined by plotting the wind and solar production against total load, see figure 4 and 5. The wind and solar production is weakly correlated with the hourly load. However, we can depict four extreme situations. Both the first and the second situations are not challenging situation, because the market can cope with low/high infeed when the demand is low/high. The third and the fourth situation illustrate the maximum and the minimum residual demand, respectively. We have in the introduction discussed that both situations are challenging for the market participants. As illustrated on figure 4 and 5, the maximum residual demand occurs statistically more often than the minimum residual demand.

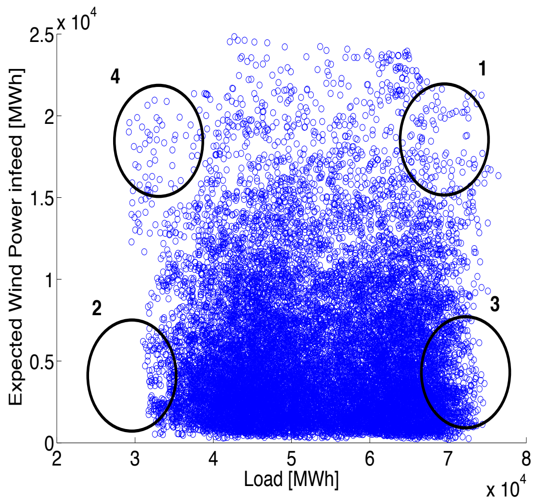

Figure 4 Scatterplot Wind infeed and total load in Germany 2011-2013.

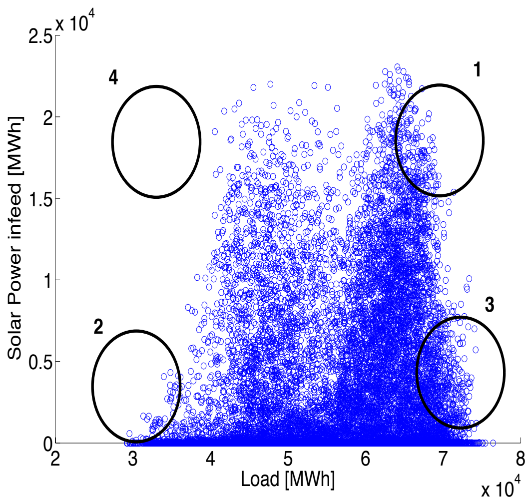

Figure 5 Scatterplot Solar infeed and total load in Germany 2011-2013.

Table 2 shows the descriptive statistics of the load, wind and solar time-series. The actual solar production is highest around the noon and zero during the night. In contrast to solar production, the wind production is high throughout the day. Both wind and solar production has an average production that is lower than the median, which indicate that there is a strong effect of outliers. In general, the wind and solar production are highly volatile.

The electricity load fluctuations are higher than the deviation of the wind and solar production. Further, the combination of demand and renewables, also called residual demand, has higher volatility than the deviation of demand.

Table 2

Descriptive Statistics Demand, Wind, Solar and Residual Demand for hour 8am, 12pm and 12am

Hour	Mean (MW)	Median (MW)	Max (MW)	Min (MW)	St.dev (MW)
Demand 8am	54700.1	58322.0	70205.0	29644.0	10229.9
Demand 12pm (noon)	62437.0	65542.0	74271.0	38327.0	7769.0
Demand 12am	47691.8	47476.0	63020.0	37101.0	4766.6
Expected Wind 8am	5291.3	3943.8	23911.0	372.8	4292.2
Expected Wind 12pm	5513.1	3936.3	23698.0	253.0	4688.5
Expected Wind 12am	5379.0	4200.0	24216.8	490.0	4229.6
Solar production 8am	1058.8	473.6	4834.8	0.0	1244.7
Solar production 12pm	8662.3	8178.0	21481.1	311.6	5108.6
Solar production 12am	0.0	0.0	0.0	0.0	0.0
Residual Demand 8am	49181.8	52042.7	70785.5	14941.9	11372.8
Residual Demand 12pm	48143.9	48752.3	74173.3	17697.6	10435.6
Residual Demand 12am	42818.0	43150.6	60959.0	18106.7	6377.3

3. Demand and residual demand modeling

Logarithmic transformation of the demand is sometimes used when the purpose is to overview the price elasticity of demand (Bianco V. et al, 2009). However, we found the electricity load data to have a linear relation to almost all variables and will therefore model the data directly.

We use a linear regression models for demand and residual demand, which is specified in Eq.1 and Eq.2, respectively. We estimate 24 separate linear regression model for each hour during the day. This approach is based on Do L. & Molnar P. (2014) earlier work. They find that 24 separate linear model performed overall better than single equation model for short-term prediction of electricity demand in Germany. The separate linear model assumes that each hour have different features (Ranaweera D. et al., 1997). Hence, each hour can not be explain by the coefficient in the same systematic way. This approach requires fewer variables than single equation model, because the insignificant variables are omitted.

$$Y_{i,t} = a_{i,1} + a_{i,2}HDD_t + a_{i,3}IP_{t-1} + \sum_{\substack{n=1 \\ n \neq 3}}^{7} a_{i,4n}W_{n,t} + a_{i,5}H_t + a_{i,6}H_{t-1} + a_{i,7}MH_t + a_{i,8}DL_t + a_{i,9}Y_{i,t-1} + \varepsilon_t \quad (1)$$

$$Z_{i,t} = a_{i,1} + a_{i,2}HDD_t + a_{i,3}IP_{t-1} + \sum_{\substack{n=1 \\ n \neq 3}}^{7} a_{i,4n}W_{n,t} + a_{i,5}H_t + a_{i,6}H_{t-1} + a_{i,7}MH_t + a_{i,8}DL_t + a_{i,9}Z_{i,t-1} + \varepsilon_t \quad (2)$$

where *Y* is demand, *Z* is residual demand, *HDD* is Heating Degree Days, W_n are dummy variables for days of the week, *H* is a major holiday variable, *MH* is a minor holiday variable, *DL* is Hours of Daylight and *i* represent the hour.

The coefficients from these regressions are presented in table 3 for three selected hours of the day (other hours are not reported in this paper due to space limitations). These results illustrate significant differences in modeling demand and residual demand.

First of all, R^2 shows that the models are able to explain much more of the variation of demand then residual demand. This is due to stochastic nature of wind and solar production. As previous discussed the wind and solar production exhibit different yearly seasonality.

We therefore propose a second residual demand model in Eq.3, which incorporates the dynamic nature of the wind and solar separately. Eq.3 is a modification of Eq.2 where the lag residual demand is replaced by three variables: lagged demand, forecasted wind production and lagged solar production.

$$Z_{i,t} = a_{i,1} + a_{i,2}HDD_t + a_{i,3}IP_{t-1} + \sum_{\substack{n=1 \\ n \neq 3}}^{7} a_{i,4n}W_{n,t} + a_{i,5}H_t + a_{i,6}H_{t-1} + a_{i,7}MH_t + a_{i,8}DL_t + a_{i,9}Wind_{i,t} + a_{i,10}PV_{i,t-1} + a_{i,11}Y_{i,t-1} + \varepsilon_t \quad (3)$$

where *Z* is residual demand, *Y* is demand, *Wind* is expected wind, *PV* is actual solar production and the other variables are already defined under Eq. 2.

The results from ordinary regression of Eq. 3 are shown in table 3. The result from R^2 shows that Eq. 3 describes the data better than Eq. 2. Hence, Eq. 3 is an improved residual demand model compared to the previous model, Eq. 2. Moreover, the three new variables in Eq.3, lagged demand, forecast wind and lagged solar production, are significant for almost all hours of the day. The estimated coefficients of wind are around -1, indicating the expected wind infeed to be almost the same as actual wind production. Unlike the wind, the coefficient of PV is above -1. This might be due to the inaccuracy of using lagged values. We will from now on consider Eq.3 when we use the term residual demand.

The demand model, Eq. 1, is compared to the residual demand model, Eq. 3 in the following paragraphs. The coefficient sign for the day type dummies (day of the week) is mostly negative for all hours. This implies that the electricity consumption is normally lower than Wednesday (base day). Further, we observe that the magnitude in demand reduction is higher for weekend than for weekday. This is also visible on the residual demand. Additionally, the level of demand and residual demand reduction is quite similar on Major and Minor holidays. Another similarity between demand and residual demand model is the impact of HDD variable (transformation of temperature). In both model, we observe the effect from HDD to be higher during early morning and night hours than noon.

The impact from DL is different on demand and residual demand model. The estimated coefficient of DL has higher effect on residual demand than on demand. This is because on average days with less daylight is also days with low solar production.

There are two main reasons why we investigate the model using linear quantile regression. Firstly, the results of ordinary regression show differences in demand and residual demand modeling, which indicate that the renewables production change the affect of the variables on the demand. We therefore employ econometric techniques to investigate in detail the relationship between the intermittent renewable resources and the demand. The investigations are based on modeling the demand and residual demand by using quantile regression.

Secondly, applying quantile regression brings new insight that can not be obtained with other estimators; the quantile regression approach analyzes the relationship at mean and at the different point on the demand/residual demand distribution. The variables that affect the demand/residual demand may have a weak relationship to the mean of the demand/residual demand, but stronger relationship with other parts of the demand/residual demand distribution. The quantile regression will give a more complete picture of the effect of the explanatory variables on the demand/residual demand.

Table 3

OLS estimates for demand and residual demand for hour 8am, 12pm, and 12am.
*****, ** and * indicates that the coefficient is significant at 10%, 5% and 1 % level, respectively**

	Demand Eq.1			Residual demand Eq.2			Residual demand Eq.3		
Hour	8am	12pm	12am	8pm	12pm	12am	8am	12pm	12am
Demand lag	0.5*	0.5*	0.6*				0.4*	0.6*	0.7*
Residual demand lag				0.5*	0.5*	0.5*			
Wind forecast							-1.1*	-1.0*	-1.1*
PV actual lag							-0.4	-0.6*	0.0*
HDD	136.5*	105.5*	181.8*	232.3*	234.8*	308.6*	148.9*	138.1*	153.5*
IP lag	539.0*	539.2*	324.9*	840.4*	932.7*	567.1*	581.8*	630.3*	141.9***
Sunday	-16416.1*	-12194.7*	-2236.9*	-17229.9*	-12860.3*	-3353.8*	-17201.2*	-11805.5*	-2341.7*
Monday	9393.2*	7681.4*	3014.7*	9141.8*	7158.8*	1840.9*	8416.3*	8366.2*	2816.6*
Tuesday	-147.3	-112.4	84.9	-1089.***	-1301.***	465.2	-181.5	-816.5	-5.2
Thursday	-994.5*	-845.5*	-395.3**	-2032.8*	-1823.7**	-672.8	-946.0**	-917.3	-542.4*
Friday	-1014.5*	-895.5*	-1319.1*	-1584.3*	-1382.***	-847.7	-1002.8**	-1020.5	-1446.7*
Saturday	-15798.8*	-11653.6*	-5765.2*	-16334.6*	-12608.2*	-6137.1*	-16005.8*	-12200.9*	-5969.5*
Holiday	-19040.5*	-14664.5*	-4466.0*	-19896.7*	-14989.1*	-5985.5*	-19036.6*	-14634.6*	-4147.9*
Holiday lag	4813.9*	4118.7*	1831.6*	3095.8*	2286.7***	298.6	4082.2*	5422.7*	2010.8*
Minor holiday	-3895.5*	-2697.4*	-1243.6*	-4311.4*	-3113.9*	-2341.8*	-3932.7*	-2456.3*	-884.7*
DL	-259.3*	-116.3*	-37.0	-68.9	-385.7*	242.2*	-469.8*	-598.8*	-112.5*
Constant	-21193.2**	-20423.8**	-16356.8*	-57533.5*	-64717.3*	-40440.**	-20513.***	-32167.**	3642.4
R^2	0.95	0.93	0.91	0.87	0.75	0.60	0.95	0.87	0.92

4. The Linear Quantile Regression

The quantile regression is an extension of ordinary regression method, where the optimization objective change from minimizing the residual sum of square to minimizing the residual sum with different q weights on residual above than below the mean value, see Eq. 4.

$$\min \sum_{t=1}^{T} \left(q - 1_{Y_t \le \alpha_i^q X_{i,t}}\right)\left(Y_t - (\alpha_i^q X_{i,t})\right) \tag{4}$$

$$1_{Y_{t \le \alpha_i^q X_{i,t}}} \begin{cases} 1 & if \quad Y_t \le \alpha_i^q X_{i,t} \\ 0 & \quad otherwise \end{cases}$$

where Y is the actual value, $\alpha_i^q X_{i,t}$ is the predicted quantile from the model, X is a vector with independent variables and q is specific quantile from 0 to 1.

The optimization objective estimates the parameters for the linear regression. The linear regression can be described as in Eq.5:

$$Q_q(Y_{i,t} \mid X_{i,t}) = \alpha_i^q X_{i,t} + \varepsilon_t \tag{5}$$

$$Q_q(Z_{i,t} \mid X_{i,t}) = \alpha_i^q X_{i,t} + \varepsilon_t \tag{6}$$

where $Q_q(Y_{i,t} \mid X_{i,t})$ is conditional quantile of the demand, $Q_q(Z_{i,t} \mid X_{i,t})$ is conditional quantile of the residual demand, X is independent variables, q is the quantile and ε_t is the error term.

Eq. 5 and 6 uses the same equation specification as Eq.1 and 3, respectively. However, we estimate these equations for different quantiles. The model is estimated for 5th, 25th, 50th, 75th and 95th quantile for each hour of the day. Based on previous results from ordinary regression we focus on the impact of lagged demand and lagged residual demand, HDD, DL, weekend dummy variables and holidays dummy variables.

One of quantile regression's most appealing features is that it enables to describe the relationship between the independent variable and the demand/residual demand not only on the mean but also on the tail of the conditional demand/residual demand. Furthermore, it also reveals the risk of immediate changes of the independent variable and the effect they will have on the demand/residual demand.

Additionally, the quantile regression model provides a set of different sensitivities for each quantiles compare to one. The distribution of independent variables gives information about asymmetric and non-linear effects on the demand/residual demand. This insight can be useful when making strategies to hedge against future loss and risk (Alexander C., 2009).

Another advantage to this approach is its reveals information about the tail, or how various risk factors affect the extreme demand/residual demand. The extreme demand/residual demand constitute a major source of risk to market participants in the electricity market. Hence, examining the tail can show the risk exposure that the conventional power producer have regarding to weather, renewables, among others.

The quantile regression is run in Stata 12.1. The standard error for the estimated coefficients for demand and residual demand model is obtained by using the pair bootstrapping procedure proposed by Buchinsky M. (1995). This bootstrapping method does not require the standard error to be identically distributed or homoscedastic.

The following sections describe the results of demand and residual from quantile regression. Each section begins with a description of the results from demand model, and is followed by a comparison of the demand and residual demand model. The estimated coefficients of the explanatory variables from quantile regression are displayed in figures 6-24.

Lagged demand: Figure 6 illustrates the estimated coefficients from quantile regressions for demand. The estimated coefficients of lagged demand are significant and positive for all hours, but greater in magnitude for the intermediate quantiles (median, 25^{th} and 75^{th} quantile) than the extreme quantiles (5^{th} and 95^{th}). This suggests that the current demand provides more information about the future mean electricity consumption than about the possible future extremely high or low consumption. Moreover, this difference is more pronounced during the day than during the night. This can be explained by the fact that the load variation is higher during the day period.

Figure 7 depicts the estimated coefficients from quantile regressions for residual demand. The most notable difference between figure 6 and 7 is the difference between the 5^{th} and 95^{th} quantile. Both the extreme quantile in figure 6 is lower than the intermediate quantiles. However, the 5th quantile in figure 7 is higher than the other quantile during the night, indicating that the previous low residual demand are more likely to be on the same level the next day.

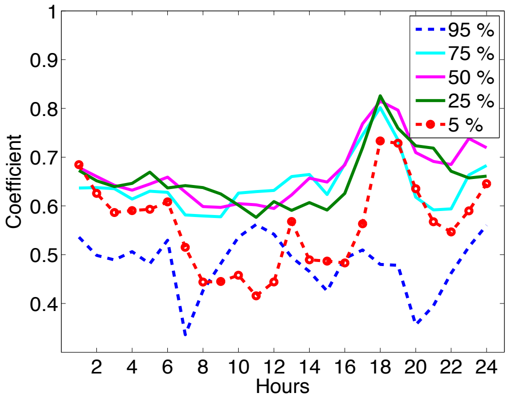

Figure 6 Coefficient Demand lag

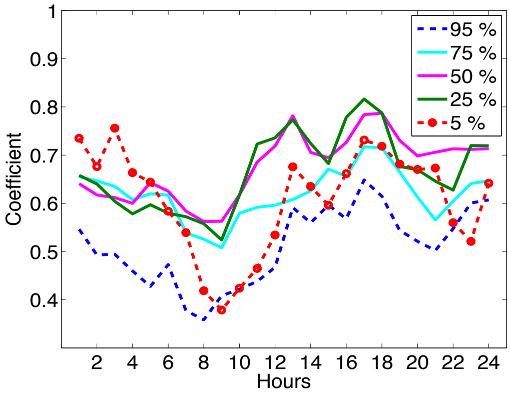

Figure 7 Coefficient Residual Demand lag

Friday dummy variable: Figure 8 presents the estimated Friday dummies on demand under different quantiles plotted against hours. The extent of negative effect of Friday on demand corresponds to the demand reduction on Friday compare to the previous day. Hence, our analysis shows that the electricity consumption is lower on Fridays compare to the previous day. Or more specific, the load reduction is larger during night, afternoon and evening period. These three observations are consistent with the typical load profile of Friday in Germany, see figure 1. As we observe on figure 1, the difference between the load profile on Thursday and Friday is the load reduction during the night, afternoon and evening.

The median coefficient of Friday has the same pattern on both demand and residual demand, figure 8 and 9 respectively. Unlike figure 8, the quantiles in figure 9 are more dispersed.

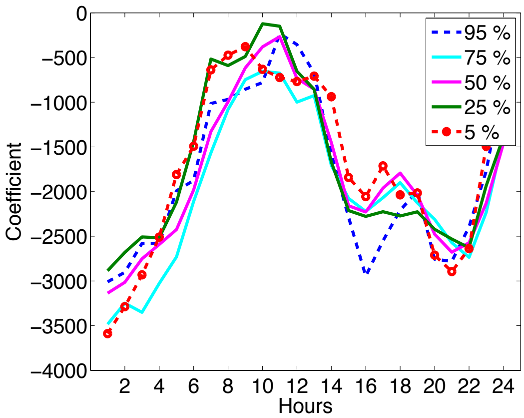

Figure 8 Coefficient of Friday for Demand

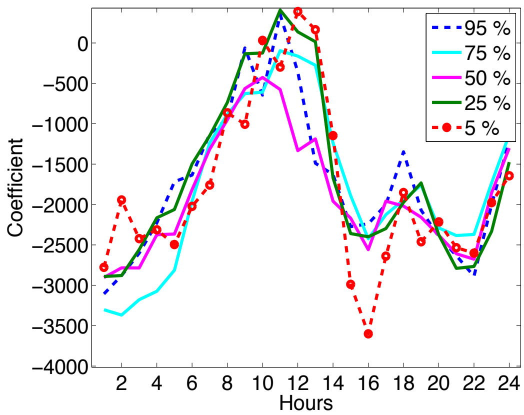

Figure 9 Coefficient of Friday for Residual Demand

Saturday dummy variable: The Saturday coefficient has same effect on demand and residual demand for all hours. The quantiles coincide for most of the period, illustrating that the quantile regression approach is not useful for this variable.

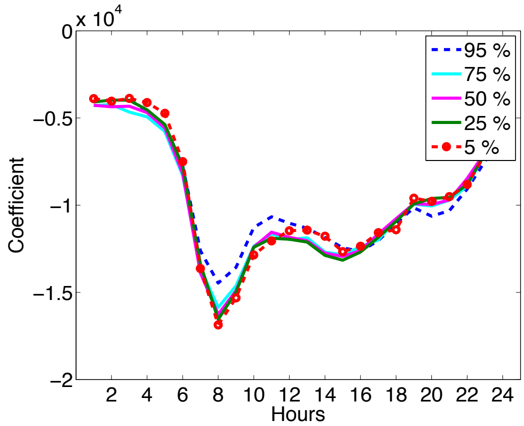

Figure 10 Coefficient of Saturday for Demand

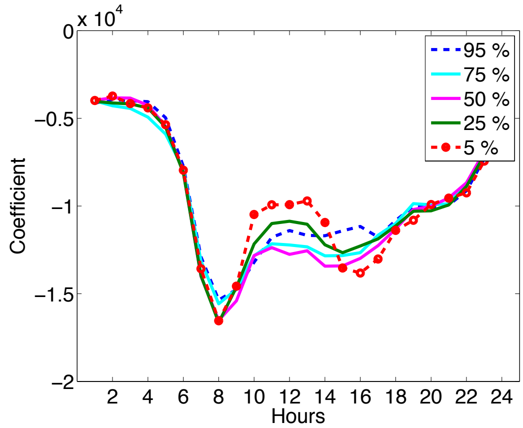

Figure 11 Coefficient of Saturday for Residual

Sunday dummy variable: As figure 12 depicts, the electricity consumption is lower on Sundays than the base day (Wednesday). The level of load reduction is higher for the day period, and is largest at the morning and afternoon hours. An explanation for this is that the business activities are lower on weekends than on weekdays. Furthermore, the impact on the demand is higher for the extreme quantiles (5^{th} and 95^{th}) than the intermediate quantiles (median, 25^{th} and 75^{th}). Hence, there are tail dependencies. Sunday dummy variable has quite similar effect on both demand and residual models.

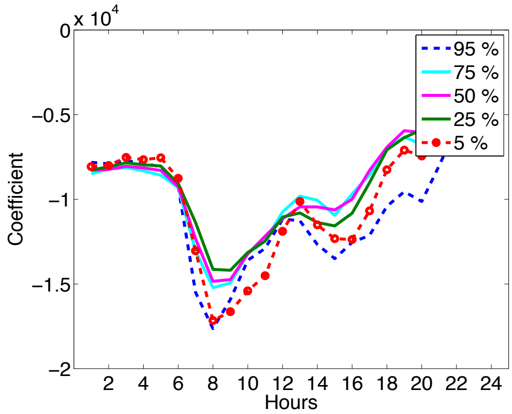

Figure 12 Coefficient of Sunday for Demand

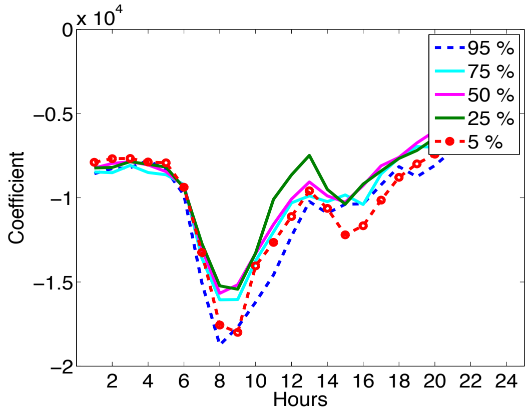

Figure 13 Coefficient of Sunday for Residual Demand

Major holiday dummy variable: Our quantile regression results suggest that the electricity consumption is lower on holidays. The disparity between electricity consumption on holiday and workday is about than 20 GW for the median, 5th, 25th and 75th, and represents the load reduction on a typical holiday. The 95th quantile is not statistically significant for all hours. We therefore conclude that the conditional high demand is not affected by Major Holiday dummy variable. However, the holidays with mean or low demand will be mostly explained by the Major holidays dummy variable.

Figure 14 and 15 exhibit that Major holiday has the same impact on the demand and residual demand during the night hours. When looking at the day period, the quantiles of residual demand are much more dispersed than the quantiles of demand.

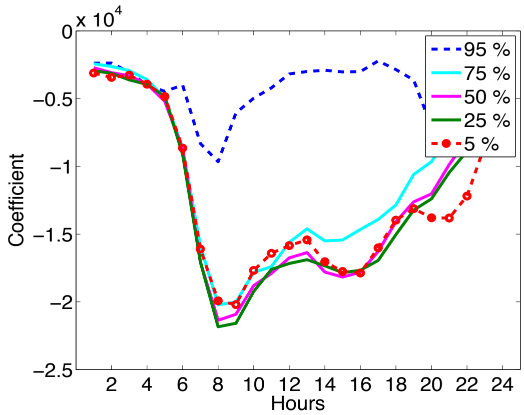

Figure 14 Coefficient of Holiday for Demand

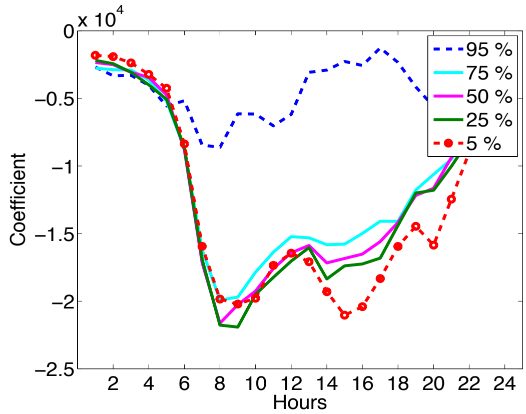

Figure 15 Coefficient of Holiday for Residual Demand

Major holiday lagged dummy variable: Figure 16 illustrates the estimated coefficient of Major Holiday from quantile regression. The results suggest that the average day after a holiday has higher electricity consumption than a holiday. In details, the results reveal that the consumption level during the night is lower than previous night. But the consumption level during the day is higher than previous day period. The spread between quantiles denotes the different load reduction that depends on the holiday.

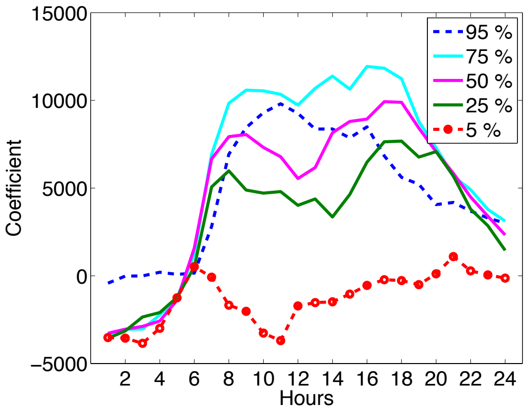

Figure 16 Coefficient of Holiday lagged for Demand

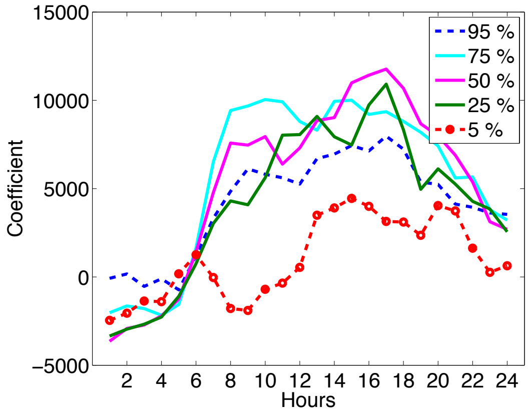

Figure 17 Coefficient of Holiday for lagged lagged Residual Demand

Minor holiday dummy variable: The Minor holiday component has high impact on the consumption, where the impact is higher during day than the night. As figure 18 depicts, the Minor holiday coefficients show different load reduction depending on the conditional quantile. The reason for this is that not all business and industry activities in Germnay are closed on minor holidays. Further we observe that the electricity consumption is lower on the 5th quantile than on 25th, 50th, 75th and 95th quantile. This means that the Minor Holiday variable is important when predicting lower levels of consumption (5th quantile).

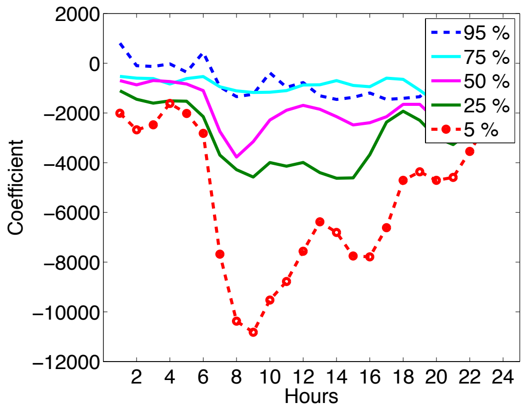

Figure 18 Coefficient of Minor Holiday for Demand

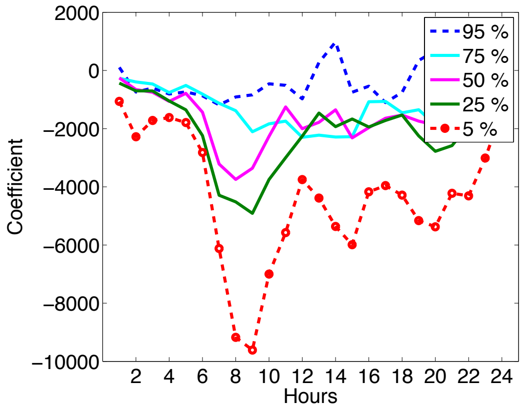

Figure 19 Coefficient of Minor Holiday for Residual Demand

Heating degree days: As figure 20 depicts, the estimated coefficient of HDD is positive for all hours, whereas the magnitude of the coefficiets is higher for the night hours. This finding can be explained by the fact that lower temperature during night leads to higher heating activities. Moreover, the 95th quantile is more sensitive to temperature compare to other quantiles during the night hours. This suggests that temperature has higher impact on extreme high electricity usage than on electricity usage in ordinary times.

As previous discussed in section 3, the HDD variable has quite the same predicting power on demand and residual demand. Further examination on figure 20 and 21 shows that quantiles of HDD are more dispersed for the residual demand compared to demand model. The larger spread between the quantiles can be explained by the uncertainty regarding the renewable production.

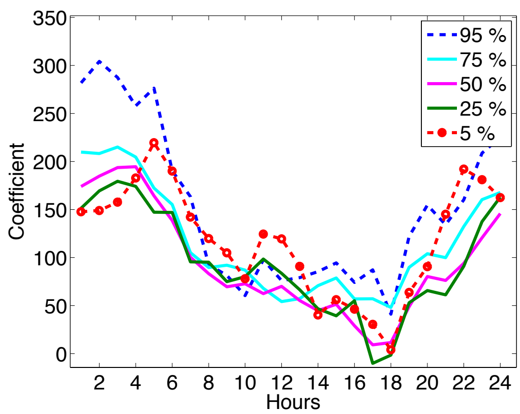

Figure 20 Coefficient of HDD for Demand

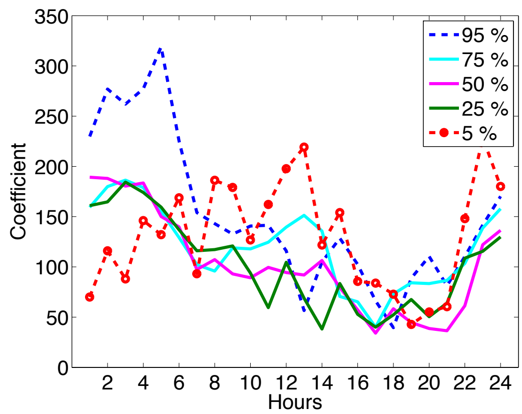

Figure 21 Coefficient of HDD for Residual Demand

Hours of Daylight: The DL has almost no effect on demand during night hours and describes the nonexistence of daylight during night, see figure 22. The figure also shows that DL is negative for all hours, and drops around 8am and 6pm. This finding means that DL has higher impact on the early morning and afternoon hours. Moreover, the quantiles are more dispersed during the early morning and afternoon hours than other time periods of the day. The distributions of the quantiles reflect the temporal differences in sunrises and sunsets throughout the year. During the year, the reduction of number of hour of daylight happens during those early morning and afternoon hours; for instant number of hour of daylight in December is lower than in June. Further, on the early morning and afternoon hours, the impact of the DL on the electricity demand is higher for the extreme quantiles (5^{th} and 95^{th}) than the intermediate quantiles (median, 25^{th} and 75^{th}). Hence, we conclude that hours of daylight have higher influence on the extreme than the ordinary electricity consumption values.

A comparison of the demand and residual demand model suggests that DL has higher effect on residual demand than on demand, since the coefficient in figure 23 exhibits higher magnitude than the coefficient in figure 22. This suggests that there is a strong relation between renewable production and hours of daylight. We observe two observations that suggest the combination of both wind and solar has some kind of strong cyclic yearly pattern. Firstly, the estimated coefficient of DL in figure 23 exhibit clear pattern during the night. Secondly, the peaks are shifted from 8am and 6pm towards 10am and 4pm.

Another notable difference between the demand and residual demand is the position of the peaks of the extreme quantiles (5^{th} and 95^{th}). While figure 22 shows that both extreme quantiles have nearly the same effect on demand, figure 23 shows that DL has lower influence on high residual demand values (95^{th} quantile) than on low residual demand values (5^{th} quantile).

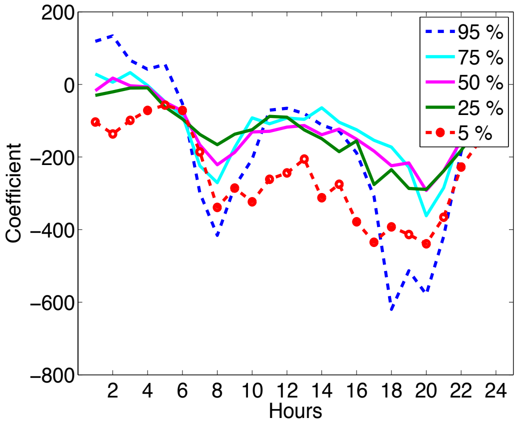

Figure 22 Coefficient of DL for Demand

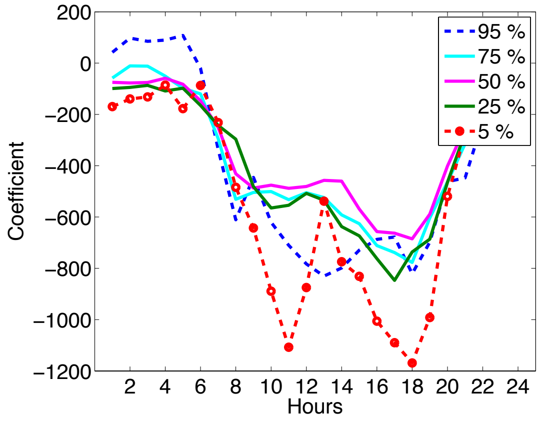

Figure 23 Coefficient of DL for Residual Demand

Wind forecast: Figure 24 draws the estimated coefficients from quantile regression, showing the relation between residual demand and wind production. The estimated coefficients of expected wind production are significant and negative for all hours, whereas the magnitude of coefficients is highest during the night hours. The spread between the quantiles is also larger during the night.

PV production lag: Similar to the wind, the coefficient of PV has negative effect on the residual demand across different quantiles. The coefficient of solar production is zero during the night and around -0.5 during the day.

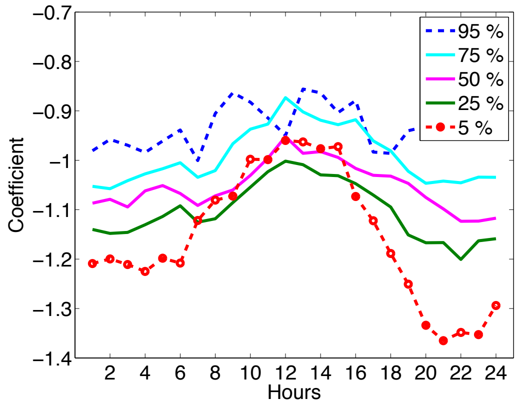

Figure 24 Coefficient Expected Wind

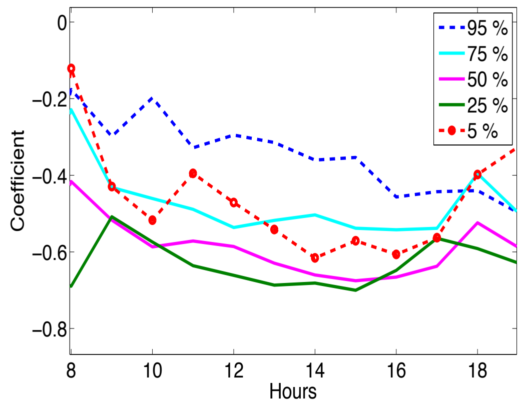

Figure 25 Coefficient PV Productions

6. Conclusions

In this empirical analysis we studied the differences and similarity in modelling demand and residual demand (the difference between demand and price inelastic solar and wind production) using both ordinary regressions and quantile regressions.

Overall, the estimated results from quantile regression indicated that fundamental variables have different effect on electricity demand/residual demand across the quantiles. This means that the quantile regression aproach clearly provided a more comprehensive picture of the underlying range of disparities in the fundamental variables of demand/residual demand than the ordinary regression. However, the quantile regression approach was not necessary when the quantiles of the variable coincided. This situation did happen for the Saturday dummy variable.

We found that residual demand was less predictable than demand. The quantile regression model showed that the conditional quantiles of the residual demand are more widely spread than conditional quantiles of demand. The effect was visible for all hours. Our results did not only confirm that the renewables lead to more challenge in predicting the load, but also illustrate how this challenge can be addressed. Our findings have implications for future research on the demand modeling, particularly in countries with increasing renewables infeed.

Appendix A. Major and Minor holiday

The electricity load reduction is different on local holidays and public holidays. We choose to distinguish the different load reduction by transforming the type of the day to a percentage weight. The weighting is based on dividing the load at the day with Wednesday at the same week. The calculated weights are compared to Wednesday at the same. If Wednesday is a holiday, use the Wednesday from the previous week. The result is grouped in two categories, higher and lower load reduction. These two groups is in table 4 consider as category A and B, or major- and minor holiday. We introduced binary dummy variable H_t and MH_t, which respectively represent major- and minor holiday.

Table 4

Categorize Major Holiday and Minor-Holiday in respectively Category A and B, * denotes National Public holiday except for 24th Dec and 31th Jan. ** denotes local holidays in the biggest cities as Munich, Dusseldorf, Berlin and Stuttgart

Type of day	**Date**	**Monday**	**Tuesday**	**Wednesday**	**Thursday**	**Friday**	**Saturday**	**Sunday**
Non holiday		0.95-1	0.97-1	1	0.96-1	0.98-0.99	0.82-0.9	0.72-0.83

(Category A Major holiday)

Current Special Day *								
New Years	1st January		0.78		0.70	0.80	0.79	0.64
Good Friday	Change each year					0.74-0.82		
Easter Monday	Change each year	0.70-0.72						
Labor Day	1st May		0.74		0.75	0.70	0.76	0.68
Ascension Day	Change each year				0.75-0.78			
Whit Monday	Change each year	0.68-0.77						
German Unity Day	3rd October	0.74		0.77		0.76	0.83	0.76
Christmas Day	24th December	0.68		0.74	0.75	0.82	0.73	
Christmas 25	25th December		0.62		0.68	0.73	0.77	0.68
Christmas 26	26th December	0.68		0.62		0.68	0.74	0.77
New Years Eve	31th December	0.84		0.76	0.74	0.80	0.73	

(Category B Minor holiday)

Local holiday**								
Epiphany	6th January	0.87	0.91	0.95	0.93		0.95	
Whit Sunday	Change each year							0.67-0.73
Corpus Christ	Change each year				0.82-0.87			
Peace Festival	8th August	0.98			0.95	0.97	0.87	0.78
Assumption of Mary	15th August	0.93			0.98	0.95	0.82	0.77
Reformation Day	31th October	0.88		0.94		0.94	0.83	0.75
All Saints Day	1st November	0.82	0.85		0.85		0.79	0.79
Repentance day	Change each year			0.98-1				
Bridging proximity days that follow a special day that occur in weekday								
Day after Ascension Day	Change each year					0.86-0.89		
Day before Christmas	23th December		0.84		0.95			0.71
Bridging proximity days that follow a special day that occur in weekend								
Day after New	2nd January	0.85						0.77

Year								
School holiday/ Non Bridging proximity days that follow a special day and occur on a weekday								
Easter Holiday	Change each year	0.98-1	0.95-1	0.98-1	0.92-1	0.84-0.97	0.79-.90	0.68-0.74
Christmas Holiday	Change each year	0.80-0.89	0.80-0.9	0.79-0.91	0.74-0.82	0.76-0.95	0.70-0.88	0.65-0.92

REFERENCES

Alexander C., (2009) Market Risk Analysis: Practical Financial Econometrics, **vol. 2,** John Wiley & Sons.

Baillo A., Ventosa M., Rivier M., Ramos A., (2004) Optimal Offering Strategies for Generation Companies Operating in Electricity Spot Market, IEEE transactions of power systems, **vol. 19**, no. 2.

Bianco V., Manca O., Nardini S., (2009) Electricity consumption forecasting in Italy using linear regression models, Energy, **vol. 34,** pp.1413-1421.

Buchinsky M., (1995) Estimating the asymptotic covariance matrix for quantile regression models a Monte Carlo study, Journal of Econometrics, **vol. 68**, pp. 303–338.

Cancelo, J. R., Espasa, A., Grafe, R. (2008) Forecasting the electricity load from one day to one week ahead for the Spanish system operator, International Journal of Forecasting, **vol. 24.**

Do L. & Molnar P. (2014) Demand modeling for Germany, NTNU, Project Thesis.

Fezzi C., (2007) Econometric models for the analysis of electricity markets, Alma Mater Studiorum Universita' Di bologna, Doctor of Philosophy Thesis.

Genethliou D. & Feinberg A. E., (2004) Load forecasting, Chapter 12, pp. 269-285.

Gupta, E. (2011) Climate Change and the Demand for electricity: A Non-linear Time Varying approach, Working paper, Indian Statistical institute.

Hagfors L. I., Kristoffersen E., Staver T., Westgaard S., (2014) Modelling the UK electricity market using quantile regression- Scenario analysis of non linear sensitivities to fundamental variables, NTNU, master thesis.

Hammoudeh S., Nguyen D., Sousa R., (2014) Energy prices and CO2 emission allowance price: A quantile regression approach, Energy Policy, **vol. 70**, pp. 201-206.

IEA., (2014) Key World Energy Statistic. Downloaded on 10/09/2014 from 〈http://www.iea.org/publications/freepublications/publication/KeyWorld2013.pdfl〉.

Kamstra M.J., Kramer L.A., Levi, M.D (2003) Winter blues: A SAD} stock market cycle, American Economic, **vol. 93**

Kaza N., (2010) Understanding the spectrum of residential energy consumption: A quantile regression approach, Energy Policy, **vol. 38**, pp 6574-6585.

Koenker R., Basset Jr G., (1978) Regression Quantiles, Econometrica, **vol. 46**, pp. 33-50.

Lee C. C. & Zeng J. H., (2011) The impact of oil price shacks on stock market activities: Asymmetric effect with quantile regression, Mathematics and computers in Simulations, **vol. 81**, pp. 1910-1920.

Molnar P., (2011) Daylight and Electricity Consumption, NTNU.

Motamedi A., (2012) Electricity Price Forecasting considering residual demand, IEEE Innovative Smart Grid Technologies, pp. 1-8.

Nicolosi M., (2012) Necessity of Capacity Mechanisms, German Association of Energy and Water Industries –BDEW.

Pardo A., Meneu V., Valor E., (2002) Temperature and seasonality influences on Spanish electricity load, Energy Economics, **vol. 24**, pp. 55-70.

Ranaweera D. K., Karady G. G., Farmer F. G. (1997) Economic impact analysis of load forecasting, IEEE Trans Power Systems, **vol. 12,** pp. 1388–1392.

Schill W. P., Oct. (2014) Residual load, renewable surplus generation and storage requirements in Germany, Energy policy, **vol. 73**, pp. 65-79.

Vazques S., Rodilla P., Battle C., (2013) Residual demand models for strategic bidding in European power exchanges: Revisiting the methodology in the presence of a large penetration of renewables, Electric Power systems Research, **vol. 108**, pp. 178-184.

Wagner A., (Mar. 2012) Residual demand modeling and application to Electricity Pricing, Fraunhofer ITWM.

Wirth H., (2015) Recent facts about Photovoltaics in Germany, Fraunhofer ISE.

DAY AHEAD ELECTRICITY PRICE MODELLING USING QUANTILE REGRESSION

Linh Phuong Catherine Do[1], Peter Molnar[2]

This paper analysis the relation between several fundamental variables and German day-ahead electricity price for each hour. The study performed quantile regression on the electricity prices and reveals important effects that are missed by ordinary regression. Ordinary regression would assume that the relation to be the same for high and normal electricity prices on a specific hours. While the quantile regression measures the dependence of the extreme event. Examine these extreme event on the price is an important aspect of effective risk management. The results indicate that the effect from the factors on electricity price vary substantially across the quantiles, thus confirming the high complexity of the electricity price.

Keywords: price modelling, renewables, ordinary regression, quantile regression

1. Introduction

The liberalization of power market in German has introduced competition and increased market participants' exposure to risk. In the new market structure, the extreme electricity price volatility can be higher than for financial instruments and commodities. This has forced market participants to consider not only volume risk but also price risk. Coupling to the price risk, the market participants face risk associated to unexpected outage, fluctuation in demand, fuel price and emission allowances. The expansion of renewable energy in Germany and its volatility in production has increased the day-ahead electricity price variance even further (Jacobsen, H. K. and Zvingilaite E 2010, Green, R. and Vasilakos, N, 2010). Hence, being able to understand how the fundamental drivers of electricity price affect the electricity price is necessary in order to manage the risks involved in the market. This has turn major interest in modeling and forecasting electricity price.

The day-ahead electricity price exhibits a number of intrinsic features, which are unique in comparison to commodity prices, gas and oil. The electricity price is more volatile than any commodity price, because it is not storable at reasonable

[1] Norwegian University of Science and Technology, Norway, e-mail: linhphuo@stud.ntnu.no
[2] Norwegian University of Science and Technology, Norway, email: peter.molnar@iot.ntnu.no

economic cost and it is limit by transmission constraints. Furthermore, the electricity price is characterized to have volatility clustering and large spikes. The possibility for extreme price movements increases the risk for the market participants. Hence, modeling the probability of simultaneous extreme price observations, usually called tail dependence, can be more important than the central expectations (Bunn D. et al, 2013). We will in this paper analyze how electricity price react to fundamental variables, when the price is abnormally low or high, for instant negative price and spikes.

Prior studies estimate the dependency of the extreme electricity price by Markow regime switching (Lindstrom E. & Regland F., 2012; Eichler M, Turk D., 2013). This model involves multiple equations that characterize the time series behavior in different regime: the drop, the base and the spike process. This approach separates the extreme price dependence from the normal dependence. Cherubini et al. (2004) propose an alternative framework; they use copula function to define the degree of dependence and the structure of dependence between electricity price and its fundamental variables. Most of the researchers used Markow regime switching or copula function in order to evaluate the dependency of the electricity price, however a strand of the literature decomposes the dependence in linear quantile regressions (Bunn D. et al 2013; Hagfors L. I. et al 2014). Compare to the other empirical methods, the quantile regression model is relatively easy to use and interpret. In additionally, the conditional variables are estimated directly for each quantile of the distribution.

The quantile regression application has been widely applied in financial risk management and been recently adopted in energy market studies: household energy consumption (Kaza N. 2010), oil prices (Lee C. C. and Zeng J. H., 2011),), CO_2 emission allowance price (Hammoudeh S. et al, 2014) and UK electricity price (Bunn D. et al 2013; Hagfors L. I. et al 2014). This paper contributes to the existing quantile regression literature by study the impact of the fundamental variables on the German day-ahead electricity price across different quantiles. This analysis is relevant because it provides a more comprehensive picture of the effects from the variables on the electricity price in normal time and periods with extreme price.

The first feature of our analysis is to examine the determinants of the day-ahead electricity price from the point of view of ordinary regression, using hourly instead of daily electricity prices. However, the ordinary regression model will give an incomplete picture of the relationship of fundamental variables and the price when the electricity price distribution is not normal. We will therefore in our second analysis use the quantile regression method, introduced by Koeker R. & G. Basset Jr. G. (1978).

This paper is organized as follow. Section 2 provides a short overview of the German electricity market. Section 3 discusses the relationship between renewables and extreme prices. Section 4 describes the fundamental variable used in this study. Since there does not exist public data for forecasted demand, we create our own demand forecast model in section 5. The results from ordinary and linear quantile regression of the electricity price are presented in sections 6 and 7, respectively. Finally, conclusions are drawn in section 8.

2. The German electricity market

The German electricity market was fully liberalized in 1998. In the liberalization process eight utilities merged to four utilities: RWE, E.ON, Vattenfall and EnBw Energie. These four vertically integrated utilities were responsible for the supply transmission and balancing of electricity. As the European directive considered that the liberalization was slow, they establish an unbundling policy. For this reason the four utilities sold a majority stake of their transmission share to third parties. Today, there are still four larger electricity generator and four transmission companies, but they act independently. The market is liberalized for both supply and retail electricity market. The German market is considered as competitive environment although there exist some degree of market power[3] (Janssen M. & Wobben M, 2008).

The Merit Order curve (the supply curve) and the demand curve are important components in understanding the electricity market. Figure 1 illustrates the Merit Order curve as a sorted short-term marginal cost curve of electricity production; the renewables has the lowest marginal cost, followed by nuclear energy, lignite, hard coal, natural gas and oil power plant. As depict in figure 1 the short-term marginal cost consists mainly of fuel and CO_2 cost. This suggests that increasing the marginal cost of the input variables leads to an increase in the electricity price. The demand curve is inelastic, meaning that the demand remains almost unchanged with change in electricity price (Sensfu F. et al, 2008). The electricity consumption is therefore predictable.

The intersection between the supply and the demand curve determinates the clearing price and the given demand for electricity. Every day, a day-ahead auction

[3] Market power is the market participants' ability to set to the price above short-term marginal cost, or withholding generation to create prices above marginal cost.

for each of the 24 hours takes place at 12pm. Each hour is dominated by different type of power plant (Murray B., 2009); where the conventional power plants remain to be the price setting utilities in the German market. Normally, the nuclear energy, lignite and the coal power plants cover the base load[4], while the gas power plants cover the peak load[5] (Sensfu F. et al, 2008).

The renewables production got priority access to the grid and has nearly zero marginal cost. As a result, the renewable production enters at the base of the Merit Order curve and shifts the curve to the right, so that cheaper conventional power plant set the price (Zachman G., 2013). This means that additional renewable infeed to the grid will reduce the electricity prices.

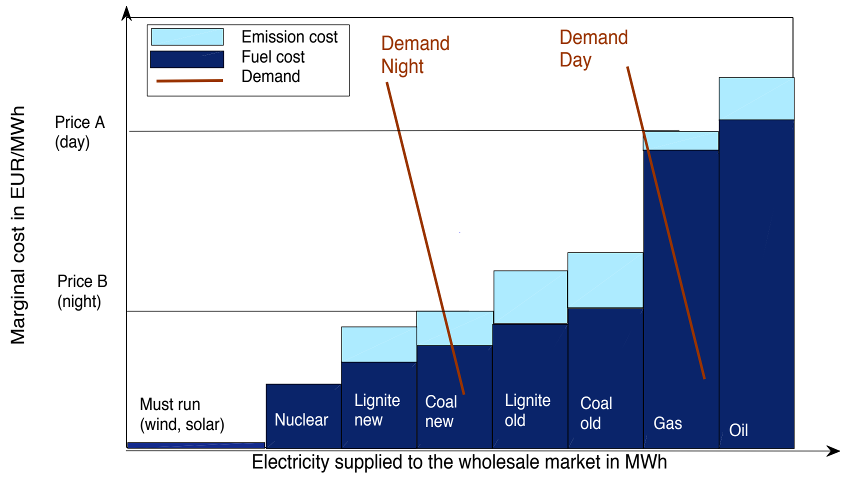

Figure 1 Stylized example of the stepwise marginal cost function and demand function for day and night.

3. Renewables and extreme prices

Lindstrom E. & Regland F. (2012) study the electricity prices for six European electricity markets and find that frequency of extreme event is positively correlated with amount of installed renewables sources to the grid. The extreme events are in this paper denotes as drops (negative prices) and spikes (price that is

[4] The base load is supply that has constant capacity during the delivery period and generally operates 24 hours.
[5] The peak load is the load higher than average supply.

three standard deviation above the mean value). In the German market, the frequency of the electricity price drops is higher than the spikes, see figure 2 and 3. The amount of negative prices increased with the penetration of renewables sources.

Further examination of our data shows that the negative prices occur more often during the night than day, while the price spikes appear during the day, see figure 4 and 5.

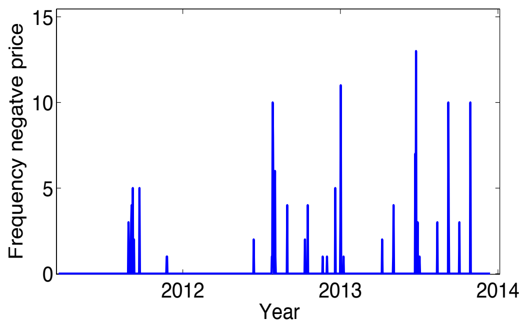

Figure 2 Amount negative price during a day.

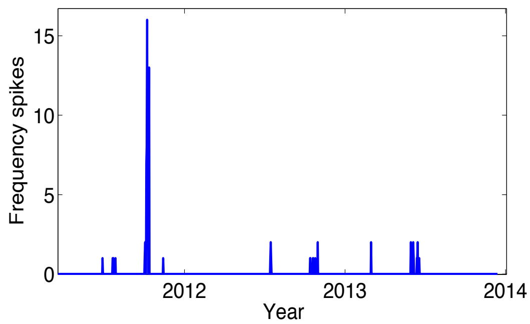

Figure 3 Amount price spikes during a day.

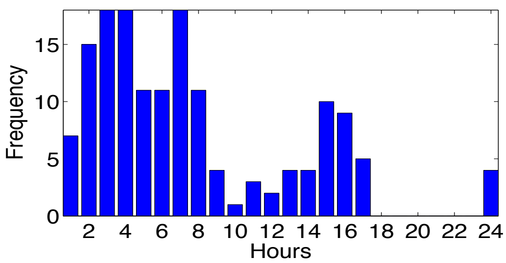

Figure 4 Distribution of negative prices.

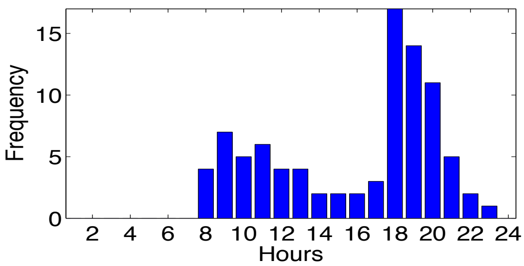

Figure 5 Distribution of price spikes.

The reason behind negative electricity prices is that must-run inflexible utilities, like nuclear power plants, are willing to pay the consumer, because the cost of shutting down excess the loss of accepting negative price (Keles et al. 2011). Additionally, high amount of solar and wind generation with essentially zero marginal cost, coupled with lower demand leads to negative electricity prices, see figure 6.

The electricity price spikes can occur for many different reasons, for instance unpredicted generation outage or transmission failures. Another reason is high demand coupled with low renewable production, which results in additional firing of power plants higher on the merit order curve and push the prices up, see figure 7.

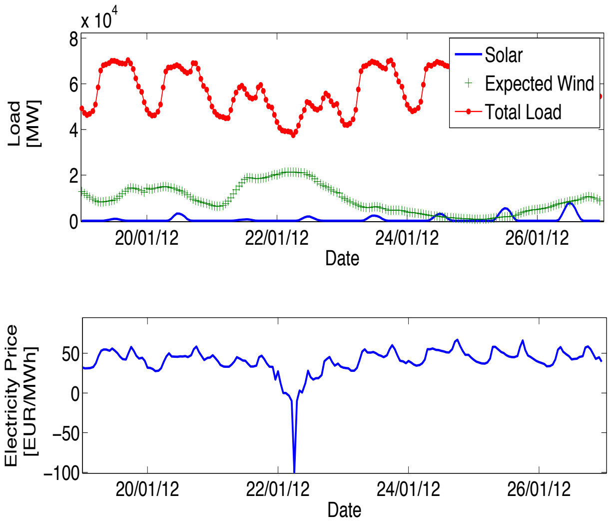

Figure 6 Snapshot view of the EEX Market negative electricity price

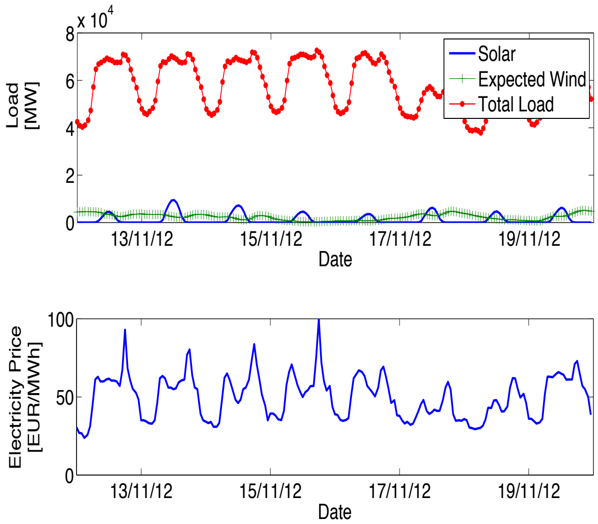

Figure 7 Snapshot view of the EEX Market electricity price spikes.

4. Data

This paper uses hourly day-ahead German electricity prices (Physical Electricity Index) provided by the European Energy Exchange market (EEX). We choose the day-ahead prices over intra day prices because they represent a larger share of the trading volume. The electricity price dataset cover the period from July 1st 2011 to July 1st 2014.

Earlier studies applied logarithmic transformation to the electricity price series in order attains variance stabilization (Conejo A. J. et al, 2005; Bunn D. et al 2013; Hagfors L. I. et al 2014). As opposed to this, Karakatsani N. V. & Bunn D. W. (2010) argue that the variance stabilization is not relevance in the electricity market, because the method conceals detailed statistical properties and gives error effects. Further, the electricity price data consists of some periods with negative prices, which can not be logarithmised. We choose to use the prices directly rather than the log prices.

As previous discussed demand and supply are important in electricity price formation process, and their components should be included in our model. The supply side is determined by several factors; the underlying fuel of the power plants, emission allowances and production of renewables energy (Paraschiv F. et al, 2014). The electricity consumption represents the demand side. As the demand is almost inelastic, the day-ahead electricity prices are strongly affected by unscheduled plant outages (Bunn et al, 2013). A clear understanding of the underlying factors is important in developing insight about the electricity price. In the following paragraphs, we will describe the choice of the fundamental drivers of the electricity price.

According to Sensfu F. et al. (2008) the different electricity generation has distinct fuel price dependencies, like the coal power plants is depending on the coal price and the gas power plants is depending on the gas price. Mjelde J. W. & Bessler D. A. (2009) study two US electricity markets, and include uranium prices along with other fuel prices. Ferkingstad et al (2011) use cointegrated model on the Northern Europe and find the electricity prices to have strong connection with the gas prices, while coal and oil prices are less important. However, Parashiv F. et al (2014) studies the German electricity price and finds that coal, gas and oil are important fundamental variables driving the electricity prices, whereas coal price is more notable during off peak hours[6] and the gas and oil prices are important for the peak hours[7]. This finding is inline with Murray B. (2009) theory; the relationship between fuel prices and the electricity price is depended on the marginal electricity price setting technology used at the specific hour. We will now describe the fuels price drivers of the German electricity price.

Normally, the nuclear power plants run at almost constant power due to economic reasons, even when the load is lower (International Atomic Energy Agency, 1999). The nuclear power plants have low marginal cost on the Merit Order curve. They are therefore less important in modeling the electricity price and we omit this variable.

The coal and lignite are the primary fuels (45%, in 2013) in the electricity market used to cover the base load (AG Energiebilanzen, 2015). Unlike hard coal, the lignite is based on the local distribution and there is currently no market price formation for the lignite. The coal price is represented by future contract on the price of coal imported to northwestern Europe via Amsterdam, Rotterdam and Antwerp.

As oppose to the coal power plants, the gas power plants are mainly used to cover the peak load due to its flexibility to ramp up and down. Both contracts from

[6] The off peak hours refer to periods with low electricity demand. The following periods is specified as off peak hours: (12am-7am), (12pm-5pm) and (9pm-12am).

[7] The peak hours refer to periods with high electricity demand. The following periods are specified as peak hours: (7am-12pm) and (5pm-9pm).

Gaspool and NetConnect Germany (NCG) are traded in Germany. We choose to use NCG contracts because of the higher liquidity in this market.

The production from oil power plants serve as a small fraction, (1%), of the total electricity production (AG Energiebilanzen, 2015). The oil price has therefore a low impact on the Merit Order curve (Sensfu F. et al., 2008). Further, the oil consumption is dominated by the transport sector and the industry (IEA, 2012). Hence, the oil price might serves as a proxy for the economic activity and the transport fuel for the coal fuel. In this paper, the European Brent spot price is used to represents the oil fuel cost.

Based on the arguments above, we choose to include coal, gas and oil prices in our model. The fuels are converted to the same currency, by using time series of spot change ($/Euro) from Skandinaviska Enskilda Bank (SEB).

The CO_2 markets are a national and international attempt to increase investment in cleaner technology, by fuel switching or reducing usage of carbon intensive power plants to less carbon intensive power plants. Both Fell H. (2010) and Parashiv F. et al (2014) find the short-term influence of CO_2 price on the electricity price to be higher in off peak hours than in peak hours. This is because the coal emits twice as much CO_2 as natural gas. We will therefore expect the CO_2 price to have higher effect on the electricity for periods where the coal power plant is the price setting technology.

When the CO_2 price is also considered to the coal and the gas price, the price difference between these two commodities can be reduced or even reversed. This phenomenon where the marginal cost of the gas power plants is lower than coal power plants is known as fuel switching (Zachmann, G. 2013).

The installed capacity and production from renewables sources have increased in the recent years. We therefore incorporate a RES variable to denote the long-term trend of renewable production. The RES variable is represented by the ratio of renewable production to total electricity production.

The biggest share of renewables production in Germany consists of wind and solar. Moreover, production of these two renewables is completely price inelastic. We therefore focus only on these two renewables in our paper and use the term renewables as interchangeably with wind and solar. The wind and solar production data have been converted from 15 min data to hourly data. Woo C. K. et al. (2011) and Keles D. et al. (2013), employ econometric techniques to investigate the impact of wind on the electricity price level in the respectively Netherlands and German markets. Both papers find that the wind productions have reduced the electricity prices. We therefore expect the renewables to have negative impact on the electricity price in Germany.

The demand variable will be represented by the expected aggregated electricity consumption in Germany. We create our own demand forecast model since there is currently no publicly available forecast demand data for Germany. The forecast demand data is further elaborated in section 5.

The reserve margin, the share of the total supply that is available, is negative correlated to the electricity price and represents the level of scarcity in the market (Boogert et al. 2008). Electricity prices can rise above marginal operating cost to include a scarcity premium. We therefore include ex ante available power plant capacity. This data is reported voluntary and do not reflect the total available capacity in Germany.

Table 1 denotes whether the fundamental variables of the electricity price is daily or hourly granularity. Table 2 is an overview of the fundamental variables used to model electricity price.

Table 1

Data granularity of the explanatory variables in our model

Variable	Resolution
Price lag	Hourly
Coal price	Daily
Gas price	Daily
Oil price	Daily
CO2 price	Daily
Expected Wind	Hourly
Actual solar Production	Hourly
Available power plant capacity	Daily
Forecast demand	Hourly
Demand lag	Hourly
RES	Daily

Table 2

Overview of fundamental variables used to model electricity price

Variable	Units	Description	Data source
Price lag	EUR/MWh	The electricity wholesale price for the same hour of the previous day.	European Energy Exchange: www.exx.com
Coal price	EUR/metric tonnes	The front month contract on API2 coal index. This index based upon the price of coal imported to northwestern Europe via Amsterdam, Rotterdam and Antwerp.	Intercontinental Exchange: www.theice.com/index
Gas price	EUR/MWh	The NetConnect Germany (NCG) day-ahead prices.	European Energy Exchange: www.eex.com
Oil price	EUR/barrel	The European brent crude spot price.	U.S. Energy information Administration: http://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=pet&s=rbrte&f=d
CO2 price	EUR/1000t CO2	The front month contract on European Union Emission allowance (EUA).	Intercontinental Exchange: www.theice.com/index
Expected Wind	MWh	Forecasted aggregated wind infeed in Germany. German transmission system operators publish this data in the late afternoon the day before the delivery day.	Transmission system operators: www.50Hertz.com , www.amprion.de www.transenbw.de , www.tennetso.de
Actual solar Production	MWh	The actual aggregated solar electricity production in Germany.	Transmission system operators: www.50Hertz.com , www.amprion.de www.transenbw.de , www.tennetso.de
Available power plant capacity	MWh	Ex ante expected power plant availability, reported voluntary by utilities. EEX publish this data at 10am the day before the delivery day.	European Energy Exchange: www.eex.com
Forecast demand	MWh	Expected aggregated demand in Germany. Details regarding this variable will be further considered in section 5.	Own data German Weather Service: www.dwd.de OECD Statistics: stats.oecd.org European Network of Transmission System Operators: www.entsoe.eu
Demand lag	MWh	The aggregated demand for the same hour of the previous day.	European Network of Transmission System Operators: www.entsoe.eu
Share of renewables production (RES)	%	The share of electricity production from the renewables sources. Three months moving average on the monthly electricity production time-series is applied.	European Network of Transmission System Operators: www.entsoe.eu

The descriptive statistics of the hourly variables: electricity price, load, wind and solar time-series are shown in table 3. The electricity price fluctuations are higher for the peak hour; also the extreme spikes appear during peak hours. The extreme negative prices occur during off peak and in the morning peak hours.

The actual solar production is highest around the noon and zero during the night. In contrast to solar production, the wind production is high throughout the day. Both wind and solar production have an average production that is lower than the median, which indicate that there is a strong effect of outliers. In general, the wind and solar production are highly volatile.

The electricity consumption has higher volatility during the day than the night. The average demand is quite near the median value. This suggests that there are few outliers.

Table 3

Descriptive Statistics electricity price, wind, solar and demand for hour 3am, 8am, 12pm, 19and 12am

	Mean (MW)	Median (MW)	Max (MW)	Min (MW)	St.dev (MW)
Price 8am	45.31	47.45	183.49	-156.92	19.20
Price 12pm (noon)	46.52	46.36	130.27	-8.30	15.43
Price 19pm	53.46	52.82	210.00	11.01	18.51
Price 12am	35.56	35.73	57.94	-90.98	9.92
Wind 8am	5473.4	3943.8	23911.0	372.8	4292.2
Wind 12pm	5666.1	3936.3	23698.0	253.0	4688.5
Wind 19pm	5761.5	4208.0	23708.3	362.5	4489.8
Wind 12am	5644.7	4200.0	24216.8	490.0	4229.6
PV 8am	1171.4	528.5	5524.9	0.0	1372.2
PV 12pm	9601.4	9096.6	22417.1	311.6	5535.4
PV 19pm	1966.0	779.3	8673.3	0.0	2346.7
PV 12am	0.0	0.0	0.0	0.0	0.0
Demand 8am	55573.7	58484.0	72982.0	29644.0	10293.2
Demand 12pm	63166.1	65651.0	76324.0	38327.0	7841.4
Demand 19pm	60060.0	60179.0	76860.0	38192.0	7916.3
Demand 12am	48400.6	47911.0	63020.0	37101.0	4955.5

5. Demand Forecasting

There is currently no publicly available data on forecast demand for Germany. We will therefore begin our analysis by constructing our own demand model.

The electricity consumption in Germany can be influenced by the following fundamental variables: trend, holidays and weather (Genethliou D. et al., 2014). Based on the work by Do L. & Molnar P. (2014a) we choose following explanatory variables for electricity demand: industrial production as the trend, holiday dummies and weather factors, like hours of daylight and heating degree days. Another important feature of the electricity load is seasonality, because the electricity load pattern depends on the day of the week. We therefore incorporate dummies variables for six out of seven days of the week. An overview of the fundamental variables used to model electricity demand and its source are depicted in Appendix A.

We use 24 separate linear regressions to estimate demand forecast values for each hour of the day. The model is performed with a rolling window of one year. Eq. 1 shows the specification of the demand in hour *i*:

$$Y_{i,t} = a_{i,1} + a_{i,2}HDD_{t-1} + a_{i,3}IP_{t-1} + \sum_{\substack{n=1 \\ n\neq 3}}^{7} a_{i,4n}W_{n,t} + a_{i,5}H_t + a_{i,6}H_{t-1} + a_{i,7}MH_t + a_{i,8}DL_t + a_{i,9}Y_{i,t-1} + \varepsilon_t \tag{1}$$

where Y is demand, HDD is Heating Degree Days, W_n are dummy variables for days of the week, H is a major holiday variable, MH is a minor holiday variable, DL is Hours of Daylight and i represent the hour.

Table 4 and 5 show performance of the demand model in terms of mean absolute (MAPE), R^2 and mean absolute error (MAE) for in-sample and out-sample, respectively. The in-sample period is from July 1, 2010 to July 1, 2011. The out-sample periods is from July 1, 2011 to July 1, 2014. The overall forecast performance of the demand model is good. The model performs better during the night than the day.

Table 4

In-sample results for estimated demand Eq. 1
The hourly MAPE (%) and R^2

	1am	2am	3am	4am	5am	6am	7am	8am	9am	10am	11am	12pm
MAPE	2.06	2.21	2.15	2.07	2.01	2.27	3.05	3.21	2.78	2.51	2.50	2.69
R^2	0.93	0.93	0.93	0.94	0.94	0.94	0.93	0.93	0.93	0.93	0.92	0.92

	1pm	2pm	3pm	4pm	5pm	6pm	7pm	8pm	9pm	10pm	11pm	12am
MAPE	2.50	2.69	2.81	2.89	2.81	2.59	2.20	2.04	2.00	1.86	1.78	1.91
R^2	0.92	0.92	0.92	0.92	0.93	0.94	0.95	0.94	0.94	0.93	0.94	0.94

Table 5

Out of-sample results for estimated demand Eq. 1,
The hourly MAPE (%) and MAE (GW)

	1am	2am	3am	4am	5am	6am	7am	8am	9am	10am	11am	12pm
MAPE	2.23	2.34	2.38	2.32	2.18	2.20	2.78	2.99	2.75	2.48	2.38	2.32
MAE	0.99	0.99	0.98	0.96	0.92	0.96	1.32	1.54	1.50	1.41	1.40	1.41

	1pm	2pm	3pm	4pm	5pm	6pm	7pm	8pm	9pm	10pm	11pm	12am
MAPE	2.33	2.45	2.53	2.59	2.60	2.49	2.28	2.16	2.11	2.06	2.07	2.20
MAE	1.40	1.44	1.46	1.46	1.47	1.43	1.32	1.24	1.17	1.10	1.07	1.04

6. Price modeling

Similar to the demand model, we generate a multiple regression model for the price. The reason we use separate equation is because each hour displays a rather distinct price profile, reflecting the daily variation of demand, fuel costs and operational constraints (Chen D. & Bunn D.W., 2010). Furthermore, the extensive research on price forecasting has generally favored the multi-model specification for short-term predictions (Chen D. & Bunn D.W., 2010; Florentina E. et al., 2014).

Based on the description of the electricity market in Germany given in section 2 and the availability of the data, we specify 24 separate linear regression models to estimate electricity prices in Germany. Eq.2 is estimated for each hour *i*:

$$
\begin{aligned}
Price_{i,t} &= a_{i,1} + a_{i,2} Price_{t-1} + a_{i,3} Coal_{t-1} + a_{i,4} Gas_{t-1} + \\
& a_{i,5} Oil_{t-1} + a_{i,6} CO2_{t-1} + a_{i,7} Wind_{i,t} + a_{i,8} Solar_{i,t-1} + a_{i,9} AC_t + \\
& a_{i,10} FD_{i,t} + a_{i,11} D_{i,t-1} + a_{i,12} RES_t + \varepsilon_{i,t}
\end{aligned} \tag{2}
$$

where *Y* is electricity day-ahead price, *Coal* is front month coal contract, *Gas* is day-ahead gas price, *oil* is spot oil price, *CO2* is the price on emission allowance, *Wind* is expected wind production , *PV* is actual solar production, *AC* is available power plant capacity, *FD* is forecast demand and *D* is demand and *RES* is share of renewables production.

The coefficients from Eq.2 are presented in table 6 for four selected hours of the day (other hours are not reported in this paper due to space limitations). In order to give a more comprehensive picture of the results, we provide a graphic representation on the estimated coefficient from ordinary regression in figure 8.

Table 6

OLS estimates for electricity price for hour 8am, 12pm, 19pm and 12am.
*** and ** indicates that the coefficient is significant at 5% and 1 % level, respectively**

	8am	12pm	19pm	12am
Price lag	0.2777**	0.2555**	0.4165**	0.1789**
Coal price	0.0634	0.1273**	0.1170**	0.0705**
Gas price	0.9106**	0.9925**	0.9644**	0.3302**
Oil price	-0.0145	-0.0610	-0.0743*	-0.0255
CO2 price	0.1746	0.5009*	0.5037	0.5552**
Expected wind	-0.0013**	-0.0010**	-0.0012**	-0.0012**
Actual lag solar	-0.0017**	-0.0005**	-0.0009**	0.0000**
Available capacity	-0.0005**	-0.0005**	0.0000	-0.0005**
Expected demand	0.0015**	0.0011**	0.0013**	0.0006**
Demand lag	-0.0005**	-0.0003**	-0.0005**	-0.0001**
Share of renewables	-31.8660*	-13.5641	5.2588	-42.4102**
Constant	-7.7822	-9.4167	-42.9155**	32.7119**
R^2	0.77	0.68	0.74	0.70

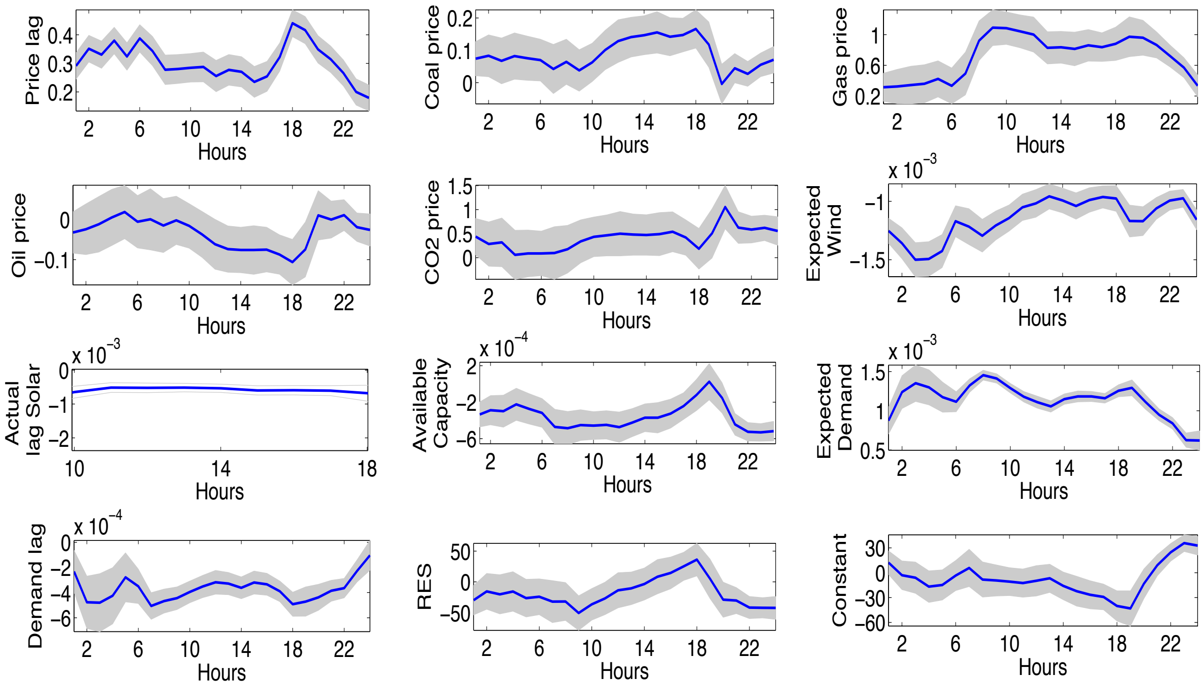

Figure 8 OLS estimates for day-ahead electricity price.
The shaded area represents 95% confident band for ordinary regression estimates.

We choose to standardize the estimated regression coefficient of wind, coal and gas because it enables us to compare the relative effects of the independent variables that have different units of measurement. The coefficients of the independent variables represent the number of standard deviation Y changes with an increase of on standard deviation in X. These coefficients can be interpreted as the measures of importance of the explanatory variables. The coefficient is standardize by following formula (Allen M. A, 1997):

$$b_{yx}{}^{*} = b_{yx}\left(\frac{ss_x}{ss_y}\right) \qquad (3)$$

where $b_{yx}{}^{*}$ is the standardize coefficient, b_{yx} is the unstandardized coefficient, ss_x is the standard deviation of the dependent variable and ss_y is the standard deviation of the independent variable. The standardize coefficients for wind, coal and gas are presented in table 7 for four selected hours of the day.

Table 7

Standardize coefficient for wind, coal and gas for hour 8am, 12pm, 19pm and 12am.

	8am	12pm	19pm	12am
Coal price	0.01	0.02	0.02	0.02
Gas price	1.05	1.43	1.16	0.74
Exp Wind	-0.30	-0.32	-0.29	-0.51

Price lag: The estimated coefficients from the ordinary regression of the lagged price are positive and denote the tendency of the electricity price to continue to move in its present direction. The electricity price that was recently high is more likely to continue to remain high and vice versa. This is in accordance with the study by Bunn D. et al (2013), which also found the lagged electricity price to have positive effect on the electricity price.

Forecast demand and lag demand: Figure 8 shows that a marginal increase in demand will increase the electricity price by a larger amount during the day than the night.

The variable of lag demand is significant and negative for all hours. This can be an indication that our load forecasts are not perfect and yesterday's load still provides useful information.

Available power plant capacity: The linear regression estimates indicate that the available power plant capacity has negative impact on the electricity price. A marginal 1MWh decrease in available power plant capacity, for instance unscheduled maintenance, will increase the electricity price.

Fuel price: The coefficients of coal price are significant for the hours before and after peak load. This can be explained by the fact that the lignite power plants are also a price setting technology during off peak hours.

The gas price is significant and positively correlated to the electricity prices. This degree of influence of the gas price on the electricity price is higher during the peak hours. Further as table 7 shows, the standardize coefficients of gas price are larger in comparison to the standardize coefficients of the coal price during peak hours, and reflect the greater degree of important of gas price. However, the gas price is also significant during the off peak hours. This finding is different from the recent study by Paraschiv F. et al. (2014), which reported that gas price is only important for the peak hours.

The coefficient of oil price is not significant. This suggests that the oil price has no on the electricity prices.

Emission allowance CO_2 price: The positive coefficient of CO_2 price indicates that there is co movement between the electricity price and the CO_2 price; an increase in CO_2 price will increase the electricity price. This effect is visible for most of the time throughout the day, because the CO_2 price is affected by the usage of fossil fuel; the conventional power plant is usually price setting technology for most of the hours thought the day, and the CO_2 price will increase the marginal cost of generating electricity. Our results shows that the effect of CO_2 price is not largely different on the off peak or peak hours, which appears out of line with our hypothesis. As previous discussed in section 4, we believed that the influence of CO_2 price on the electricity price to be higher in off peak hours than in peak hours, because the coal emits twice the CO_2 content of natural gas.

Renewables sources: The effects of renewables on German electricity prices are estimated by using exogenous terms for total expected wind production, actual solar production, and share of electricity production from the renewables sources (RES). The share of total renewable capacity has a positive impact on the electricity prices during the peak hours, but on average a decreasing effect on the electricity price. The positive coefficients of RES during peak can be explained by displace high cost natural gas electricity that average out the low cost renewable production. For example the situations when the demand for electricity is high and at the same time the renewable production is low, requires flexible gas power plants that have high marginal cost to cover the demand. As a result, turning additional high cost power plant will average out the price lowering effect from renewables.

The negative coefficients of RES during off peak hours support the theory that the renewables is driving down wholesale prices. Moreover, this is confirmed by negative coefficients of the wind and solar production.

The impact from the wind on the prices varies throughout the day. A marginal increase in produced wind will decrease the electricity price with a higher amount during the night than the day. This result is inline with the previous discussion in section 3; the negative prices usually happen during the night hours, because of low demand and excess wind production.

The standardize coefficient of the wind is higher than the standardize coefficient of the coal. This suggests that the electricity price is more directly connected to the wind production and less to the marginal cost of coal power plants.

The infeed from solar production has lower impact on the electricity price compare to the wind production, because the total installed capacity of the wind account twice the solar sources (AG Energiebilanzen, 2015). We further observe the coefficient of solar production to be high during day and zero during night. This finding is accordance with the solar production level.

Our findings from ordinary regression shows that the price lag, fuel prices, emission allowances and demand have a positive effect on the electricity prices. The results also show that available power plant capacities and renewables have a negative impact on the electricity prices.

Range of R^2 is between 0.58 and 0.79 suggest a quite credible fit to the data. The lowest R^2 value is during night period, which means that the parameters describe better the fundamentals during day period, than the night period.

We will further provide additional empirical evidence on how the fundamental variables influence the electricity prices, by using linear quantile regression.

6. Quantile Regression of the Electricity Price

The quantile regression is a further extension of ordinary regression method, where the optimization objective change from minimizing the residual sum of square to minimizing the residual sum with different q weights on residual above than below the mean value, see Eq. 4. (Koenker R., 2005).

$$\min \sum_{t=1}^{T} \left(q - 1_{Y_t \le \alpha_i^q X_{i,t}} \right) \left(Y_t - (\alpha_i^q X_{i,t}) \right) \tag{4}$$

$$1_{Y_t \le \alpha_i^q X_{i,t}} \begin{cases} 1 & if \quad Y_t \le \alpha_i^q X_{i,t} \\ 0 & \quad otherwise \end{cases}$$

where Y is the actual value, $\alpha_i^q X_{i,t}$ is the predicted quantile from the model, X is a vector with independent variables and q is specific quantile from 0 to 1.

The optimization objective estimates the parameters for the linear regression. The linear regression can be described as in Eq.5:

$$Q_q(Y_{i,t} \mid X_{i,t}) = \alpha_i^q X_{i,t} + \varepsilon_t \qquad (5)$$

where $Q_q(Y_{i,t} \mid X_{i,t})$ is conditional quantile of the electricity price, X is independent variables, q is the quantile and ε_t is the error term.

Eq.5 uses the same equation specification as Eq. 2, but estimates these equations for different quantiles. The model is estimated for 5th, 25th, 50th, 75th and 95th quantile for each hour of the day.

The advantage of quantile regression, relative to the ordinary method, is that it estimates the effect of the explanatory variables not only on the conditional mean (OLS), but also the effect on the conditional quantiles. In order words, the quantile regression estimate a set of regression lines compare to just one line. As a result, the distribution of the electricity price can be fully captured by using several quantiles. The quantile can reveal the risk of immediate changes of the independent variable and the effect they will have on the electricity price.

Additionally, the quantile regression gives a set of different sensitivities for each quantiles distribution of the independent variables. The distribution of independent variables gives information about asymmetric and non-linear effects on the electricity price. This can be useful when making hedging strategies, buying/selling weather derivatives to hedge against future loss and risk (Alexander C., 2009).

The quantile regression method reveals information about the tail, or how various risk factors affect the extreme prices. Extreme prices can have a devastating impact on the returns. Hence, the tail can show the risk exposure that the market participants have regarding to fuel price, renewable production, among others. Further, computing the tail risk (VaR or expected shortfall) can give valuable information for the market participants in developing optimal strategies for risk management.

The quantile regression is run in Stata 12.1. The standard error for the estimated coefficients for demand and residual demand model is obtained by using the pair bootstrapping procedure proposed by Buchinsky M. (1995). This bootstrapping method does not require the residuals to be homoscedastic, because it derives clustering adjusted robust standard error of the quantile regression estimates.

Figures 9-18 provide graphic representations on the estimated coefficient of the explanatory variables from quantile regression.

Price lag: Figure 9 illustrates the estimated coefficient from quantile regression of the lag price. The quantiles of the price are more dispersed during the night than the day. The 5th quantile is prominently higher than the other quantiles during the night, meaning that the low off peak prices[8] are likely to be on the same level the next day. The 95th quantile is higher than the other quantiles for the peak hours. This suggests that the current price provides more information to the future higher peak prices[9] than the lower peak prices.

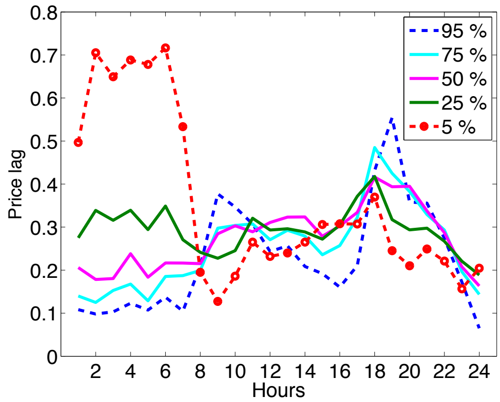

Figure 9 Coefficient of Price lag

Expected Demand and lag demand: The distribution of the estimated coefficient of the expected demand is different during the day and night. The variable has a quite uniform effect over all quantiles for the morning peak prices, while dispersed effect for the off peak prices. Further, the positive correlation between demand and electricity prices is greater in magnitude at the lower tail (5th and 25th quantiles) of the distribution than the upper tail (75th and 95th quantiles) during the night. This suggests that the base load have diverse effect on the price.

[8] The electricity prices during off peak hours. The off peak hours are specified as: (12am-7am), (12pm-5pm) and (9pm-12am).

[9] The electricity prices during peak hours. The peak hours are specified as: (7am-12pm) and (5pm-9pm).

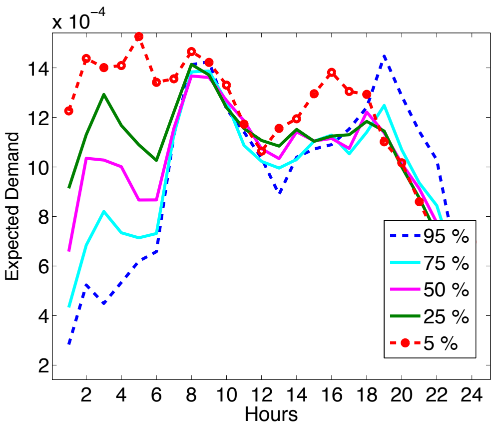

Figure 10 Coefficient of Expected Demand

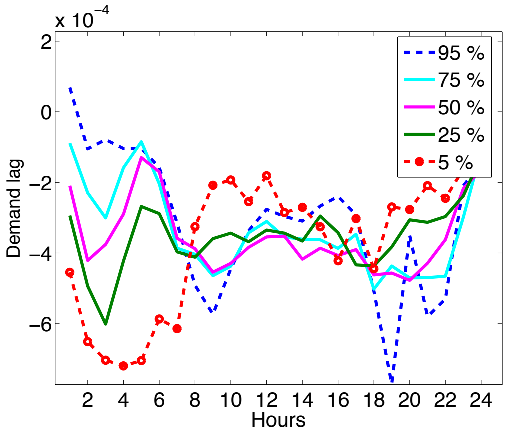

Figure 11 Coefficient of Demand lag

Available power plant capacity: As previous discussed, a reduction of available capacity will increase the electricity prices because of level of scarcity. We therefore observe negative relation between the available capacity and electricity price.

The quantiles of the power plant availability coincide for the night hours, and diverge for the day hours particularly the peak hours. This can be explain by the fact that residual demand, the difference between demand and solar and wind production, is more volatile during the peak hours due to the intermittent renewables production (Do L. & Molnar P., 2014b); a quickly change in residual demand can cause additional firing of power plants higher on the merit order curve pushing the prices up. As figure 12 depicts, this effect is significantly enhanced in the lower quantiles and weakened on the 95th quantile for peak prices. This suggests that the scarcity premium is lower for the already extreme high electricity prices than for the low and normal electricity prices. The low scarcity premium for the already extreme electricity peak prices can be explained by electricity import from cross border countries with lower electricity prices than in Germany. This is also known as market coupling; the cross border countries with lower electricity price will meet demand at higher price in Germany and the result is a reduction of price in Germany.

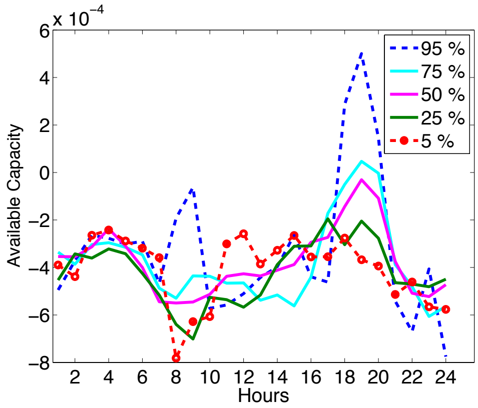

Figure 12 Coefficient of Available Capacity

Fuels: The estimated coefficients of the coal price are highest before and after peak hours. Further, the impact of the coal prices on the electricity prices is higher for the extreme quantiles, 5th and 95th, than the median, 25th and 75th quantile. Hence, there is tail dependence, meaning that the extreme coal prices movements have higher impact on the electricity prices.

As figure 14 depicts, the effect of the gas price on the electricity price is highest during peaks hours. The positive relationship between gas price and the electricity prices is greater in magnitude at the upper tail (75th and 95th quantiles) of the distribution than the lower tail (5th and 25th quantiles) during the peak hours. Gas prices at the 5th quantile of the electricity prices is related with an increase in the electricity prices around 0.8 EUR/MWh while the 95th quantile is related with 1.1 EUR/MWh during the peak hours. This can be explained by market power; when the demand is high the utilities have more room to exercise their market power by setting prices significantly above marginal costs.

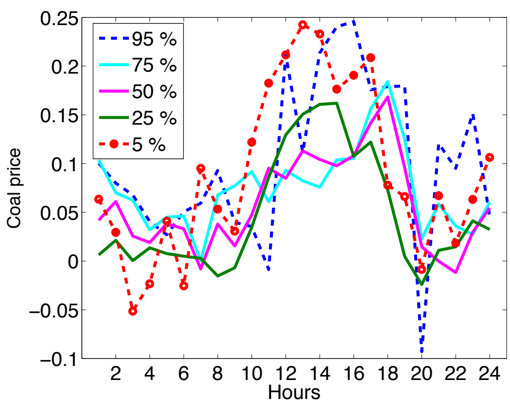

Figure 13 Coefficient of Coal price

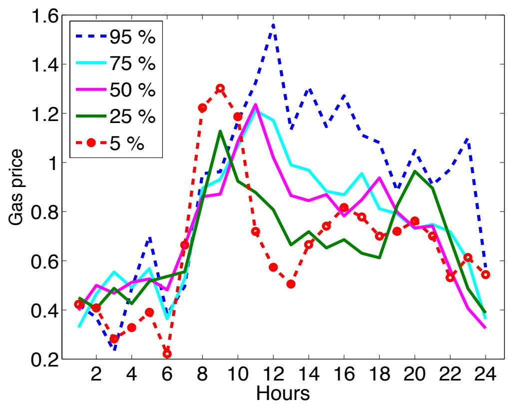

Figure 14 Coefficient of Gas price

Emission allowances CO_2 price: The confidence bands indicate how significant the coefficient is. They are not plotted in figure 15, but will be discussed in following paragraph.

The CO_2 price has positive impact on the electricity price. The extreme quantiles, 5th and 95th, of CO_2 price are not statistically significant for all hours. The 5th quantile is significant at the same time as the 95th quantile is insignificant, and vice versa. This implies the structure of dependence is asymmetric. The off peaks hours have upper tail (75th and 95th quantiles) dependency and lower tail (5th and 25th quantiles) independency. Hence, the lower electricity off peak price is not affected by increase in CO_2 price, because it is usually set by the renewable sources and nuclear power. The 95th quantile is not significant during peak hours. In other words, the CO_2 price has negligible effect on the extreme high electricity prices during peak hours. This can be explained by the lower additional CO_2 cost on the natural gas and oil compare to coal fuel.

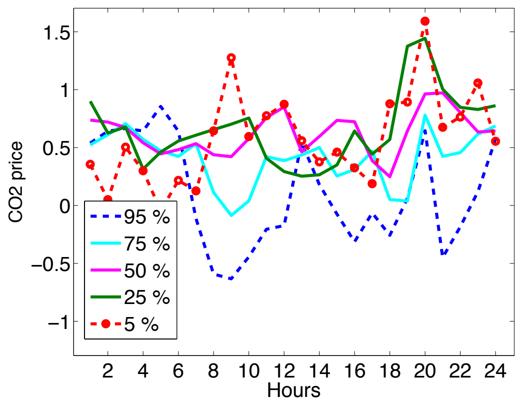

Figure 15 Coefficient of CO2 price

Renewables: As previously discussed in section 4, the RES variable has low impact on the electricity price during the evening peak hours due to high cost gas power plant even out the low cost renewables production. The estimated coefficient of RES is negative and significant for 0-12pm and 20-24pm, but only for the intermediate and upper quantiles (75th and 95th quantile), whereas for the lower quantiles (5th and 25th quantile) have no significant effects. This result indicates that the lowest electricity price is less affected by the trend of increasing share of renewables. We further see on figure 17 that the electricity price is more affected by the short-term production from renewables.

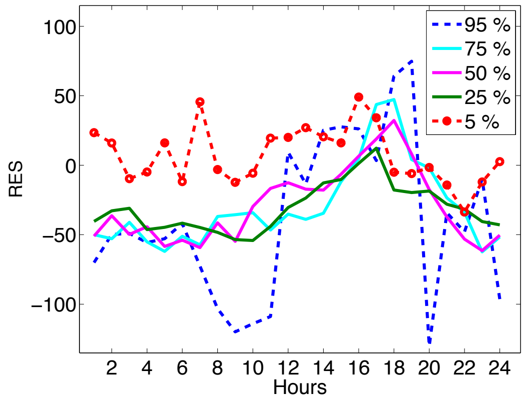

Figure 16 Coefficient of RES

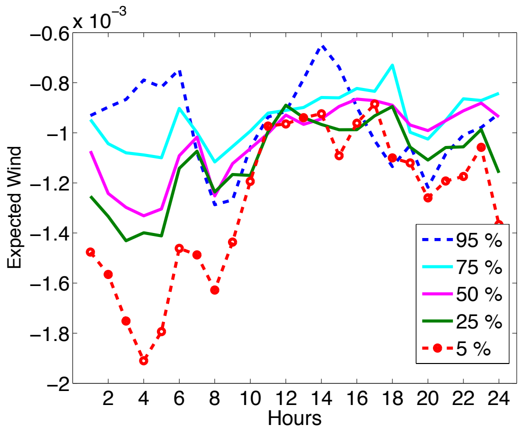

Figure 17 Coefficient of wind production.

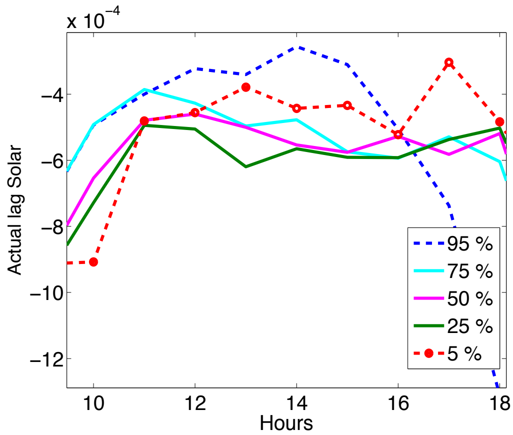

Figure 18 Coefficient of solar production.

The estimated coefficient of expected wind production is significant and negative for all hours, whereas the magnitude of coefficients is higher for the night hours. Further study on the night hours shows that the negative impact of wind production with electricity price is greater in magnitude at the 5th quantile than at the 95th quantile. During the morning and the afternoon peak hours, the impact of the wind production on the electricity prices is higher for the extreme quantiles, 5th and 95th, than the median, 25th and 75th quantile.

The infeed from solar production has quite uniform effect on the electricity price during the day, except for the morning and afternoon peak prices. The estimated coefficients from quantile regression of solar production exhibit tail dependency.

In overall, the quantile regression coefficients are found to give a more comprehensive picture of the relationship between fundamental variables and the electricity price. For most quantiles the estimates are outside the ordinary least square confident band, suggesting that the ordinary regression estimates is not sufficient. Moreover, this suggests that the quantile regression is important to identify the shortcoming of ordinary regression.

8. Conclusions

The study in this paper identified the main explanatory variables of day-ahead electricity price in Germany, and analyzes the impact of the variables on electricity price by ordinary regression and quantile regression models.

The estimation results from ordinary regression methods indicate that the price lag, fuel prices, emission allowances and demand have a positive effect on the electricity prices. The results show that available power plant capacities and renewables have a negative impact on the electricity prices. We find that the wind and solar production has a negative impact on the electricity price in short term. However, in long term the impact of renewables on electricity will be negated by the increase in more expensive flexible generation. We therefore see that the renewables have positive impact on the electricity price in the peak hours.

Our empirical finding from quantile regression shows high complexity of the electricity price, which makes it challenging to summaries the findings. We conclude that our findings show a more comprehensive picture of the electricity price in Germany. The finding can have important implications for the market participants who want to manage the risks involved in the electricity market. Particularly, examine the extreme event on price and computing the risk (VaR or expected shortfall) is an important aspect of effective risk management.

This model can be use as a reference for further work on the German energy market and it is also transferable to other electricity market with high penetration of renewables. However, the model needs to be adapted to the local conditional, but these changes do not affected the fundamentals of the model.

Appendix A. Variables used to forecast demand

Table 8

Overview of variable used to model demand

Variable	Description	Data source
Demand lag	The aggregated load for the same hour of the previous day. The load data include production from thermal energy and network feed-in from renewable energy.	European Network of Transmission System Operators: www.entsoe.eu
HDD	Heating degree days is an indication for the need of heating, $\text{HDD} = \max(\text{T}_{\text{ref}} - \text{T}, 0)$ where T_{ref} is the reference temperature equal 18 degrees, and T describes the weighted average outdoor temperature for the day. The temperature data is taken from the cities with highest population densities and are geographically spread: Munich, Berlin, Dusseldorf and Stuttgart.	The German Weather Service: www.dwd.de
IP lag	Three months moving average on the Industrial Production time-series (IP) is applied to smooth out jumps. IP lag is the moving average industrial production value on the previous day.	OECD Statistics: stats.oecd.org
Mon, Tue, Thu, Fri, Sat, Sun	Binary dummy variables, where Wednesday is taken as base weekday.	Calendar: www.timeanddate.com
Holiday	Binary dummy variable on major holiday and holidays with high load reduction. For more information about the composition of this variable, see appendix B.	Own data National holidays: www.bmi.bund.de School holiday: www.holidays-info.com
Holiday lag	Binary dummy variable on the day before holidays.	Own data National holidays: www.bmi.bund.de School holiday: www.holidays-info.com
Minor holiday	Binary dummy variable on minor holiday, local holidays and holidays with lower load reduction. For more information about the composition of this variable, see appendix B.	Own data Local holidays in Germany: www.timeanddate.com
DL	Hours of Daylight, (DL), is determined by first calculating the sun's inclination angle λ_t where l_t is [1,365] and 1 represent January 1st etc. Thereafter calculate *DL*, where δ is the latitude in Germany, see Kamstra M. J. et al (2003). $\lambda_t = 0.4102 \sin\left(\frac{2\pi}{365}(\text{l}_t - 80.25)\right)$ $\text{DL}_t = 7.722 \arccos\left(-\tan\left(\frac{2\pi\delta}{360}\tan(\lambda_t)\right)\right)$	Own data

Appendix B. Major and Minor holiday

The electricity load reduction is different on local holidays and public holidays. We choose to distinguish the different load reduction by transforming the type of the day to a percentage weight. The weighting is based on dividing the load at the day with Wednesday at the same week. The calculated weights are compared to Wednesday at the same. If Wednesday is a holiday, use the Wednesday from the previous week. The result is grouped in two categories, higher and lower load reduction. These two groups is in table 9 consider as category A and B, or major- and minor holiday. We introduced binary dummy variable H_t and MH_t, which respectively represent major- and minor holiday.

Table 9

Categorize Major Holiday and Minor-Holiday in respectively Category A and B, * denotes National Public holiday except for 24 Dec and 31 Jan. ** denotes local holidays in the biggest cities as Munich, Dusseldorf, Berlin and Stuttgart

Type of day	**Date**	**Monday**	**Tuesday**	**Wednesday**	**Thursday**	**Friday**	**Saturday**	**Sunday**
Non holiday		0.95-1	0.97-1	1	0.96-1	0.98-0.99	0.82-0.9	0.72-0.83
(Category A Major holiday)								
Current Special Day *								
New Years	1st January		0.78		0.70	0.80	0.79	0.64
Good Friday	Change each year					0.74-0.82		
Easter Monday	Change each year	0.70-0.72						
Labor Day	1st May		0.74		0.75	0.70	0.76	0.68
Ascension Day	Change each year				0.75-0.78			
Whit Monday	Change each year	0.68-0.77						
German Unity Day	3rd October	0.74		0.77		0.76	0.83	0.76
Christmas Day	24th December	0.68		0.74	0.75	0.82	0.73	
Christmas 25	25th December		0.62		0.68	0.73	0.77	0.68
Christmas 26	26th December	0.68		0.62		0.68	0.74	0.77
New Years Eve	31th December	0.84		0.76	0.74	0.80	0.73	
(Category B Minor holiday)								
Local holiday**								
Epiphany	6th January	0.87	0.91	0.95	0.93		0.95	
Whit Sunday	Change each year							0.67-0.73
Corpus Christ	Change each year				0.82-0.87			
Peace Festival	8th August	0.98			0.95	0.97	0.87	0.78
Assumption of Mary	15th August	0.93			0.98	0.95	0.82	0.77

Reformation Day	31th October	0.88		0.94		0.94	0.83	0.75
All Saints Day	1st November	0.82	0.85		0.85		0.79	0.79
Repentance day	Change each year			0.98-1				
Bridging proximity days that follow a special day that occur in weekday								
Day after Ascension Day	Change each year					0.86-0.89		
Day before Christmas	23th December		0.84		0.95			0.71
Bridging proximity days that follow a special day that occur in weekend								
Day after New Year	2nd January	0.85						0.77
School holiday/ Non Bridging proximity days that follow a special day and occur on a weekday								
Easter Holiday	Change each year	0.98-1	0.95-1	0.98-1	0.92-1	0.84-0.97	0.79-.90	0.68-0.74
Christmas Holiday	Change each year	0.80-0.89	0.80-0.9	0.79-0.91	0.74-0.82	0.76-0.95	0.70-0.88	0.65-0.92

REFERENCES

AG Energiebilanzen, (2015) Bruttostromerzeugung in Deutschland von 1990 bis 2014 nach Energieträgern. Downloaded on 20/04/2015 from ⟨http://www. ag-energiebilanzen.de/⟩.

Alexander C., (2009) Market Risk Analysis: Practical Financial Econometrics, **vol. 2,** John Wiley & Sons.

Allen M. P (1997) Understanding regression: regression analysis with standardized variable, Chapter 10, pp. 46-50.

Buchinsky M., (1995) Estimating the asymptotic covariance matrix for quantile regression models a Monte Carlo study, Journal of Econometrics, **vol. 68**, pp. 303–338.

Bunn D., Andresen A., Chen D. Westgaard S. (2013) Analysis and forecasting of electricity price risks with quantile factor models, Fortcoming Energy Journal.

Conejo A.J., Contreras J., Espinola R., Plazas M.A., (2005) Forecasting electricity prices for a day-ahead pool-based electric energy market, International Journal of Forecasting, **vol. 21**, pp. 435–462.

Chen D. & Bunn D. W. (2010) Analysis of the Nonlinear Response of Electricity Prices to Fundamental and Strategic Factors, IEEE Transaction on Power Systems, **vol. 25**, pp. 595-606.

Cherubini U., Luciano E., Vecchiato W., (2004) Copula Methods in Finance, John Wiley, Chichester.

Do L. & Molnar P., (2014a) Demand modeling for Germany, project thesis, NTNU.

Do L. & Molnar P., (2014b) Demand and Residual Demand modeling using Quantile Regression, NTNU.

Eichler M, Turk D., (2013) Fitting Semi parametric Markov regime-switching models to electricity spot price, Energy Economics, **vol. 36,** pp. 614–624.

Fell H., (2010) EU-ETS and Nordic Electricity: A CVAR Analysis, IAEE Resources for the Future, **vol. 31**, pp. 1-26.

Ferkingstad E., Løland A. Wilhelmsen M., (2011) Causal modeling and inference for electricity markets, Energy Economics, **vol. 33**, pp. 404-412.

Genethliou D., Feinberg E. A., (2014) Load forecasting, Chapter 12, pp. 269-285.

Green R. & Vasilakos N. (2010) Market behavior with large amounts of intermittent generation. Energy Policy, **vol. 38,** pp. 3211-3220.

Hagfors L. I., Kristoffersen, E., Staver T. and Westgaard, S. (2014) Modelling the UK electricity market using quantile regression- Scenario analysis of non linear sensitivities to fundamental variables, master thesis, NTNU.

Hammoudeh S., Nguyen D. & Sousa R. (2014) Energy prices and CO2 emission allowance price: A quantile regression approach, Energy Policy, **vol. 70**, pp. 201-206.

IEA, (2012) Oil & Gas security: Emergency response of IEA countries, Germany

International Atomic Energy Agency (1999), Modern Instrumentation and Control for Nuclear Power Plants: A Guidebook, Technical Report Series N°387.

Jacobsen, H. K. & Zvingilaite E. (2010) Reducing the market impact of large shares of intermittent energy in Denmark, Energy Policy, **vol. 38**, pp. 3403-3413.

Janssen M. & Wobben M., (2009) Electricity pricing and market power: evidence from Germany, European Transaction on Electrical Power, **vol. 8**, pp. 591-611.

Kamstra M.J, Kramer L.A., Levi M. D., (2003) Winter blues: A SAD stock market cycle, American Economic Review, **vol. 93**, pp. 324-343.

Karakatsani N. V. & Bunn. D. W., (2010) Fundamental and behavioral drivers of electricity price volatility, Studies in Nonlinear Dynamics & Econometrics, **vol. 14**, pp. 1-42.

Kaza. N. (Nov. 2010) Understanding the spectrum of residential energy consumption: A quantile regression approach, Energy Policy, **vol. 38**, pp. 6574-6585.

Keles D., Genoese M., Möst D., Fichtner W., (2011) Comparison of extended mean reversion and time series model for electricity spot price simulation consider negative prices, Energy Econ, **vol. 34**, pp. 1012-1032.

Keles D., Genoese M., Möst D., Ortlieb S., Fichtner W., (2013) A combined modeling approach for wind power feed-in and electricity spot prices, Energy Policy, **vol. 59**, pp. 213–225.

Koenker R. & Basset Jr G. (1978), Regression Quantiles, Econometrica, **vol.46**, pp. 33-50.

Koenker R. (2005), Quantiles Regression, Cambridge, UK, pp. 9-14.

Lee, C. C. & Zeng, J. H. (May 2011) The impact of oil price shacks on stock market activities: Asymmetric effect with quantile regression, Mathematics and computers in Simulations, **vol. 81**.

Lindström E. & Regland F., (2012) Modeling extreme dependence between European electricity markets, Energy Economics, **vol. 34**, pp. 899-904.

Mjelde J. W, & Bessler D. A, (2009) Market integration among electricity markets and their major fuel source markets, Energy Economics, **vol. 31**, pp. 482-491.

Murray, B. (2009) Power markets and economics: energy costs, trading, emissions. Chichester, U.K., Wiley.

Florentina P., David E., Ralf P. (2014) The impact of renewable energies on EEX day-ahead electricity prices, Energy Policy, **vol. 73**, pp.196-210.

SEB, (2015) Spot exchange rate Euro/US Dollar, Skandinaviska Enskilda Banken, Downloaded on 20/02/2015 from ⟨http://www.seb.se/⟩.

Sensfu F., Ragwitch M., Genoese M., (2008), The merit-order effect: a detailed analysis of the price effect of renewable electricity generation on spot market prices in Germany, Energy Policy, **vol. 36**, pp. 3086-3094.

Schill W. P., (2014) Residual load, renewable surplus generation and storage requirements in Germany, Energy Policy, **vol. 73**, pp. 65-79.

Wirth H., (2015) Recent facts about Photovoltaics in Germany, Fraunhofer ISE.

Woo, C. K., Horowitz, I., Moore, J., Pacheco, A. (2011) The impact of wind generation on the electricity spot-market price level and variance: The Texas experience, Energy Policy, **vol. 39**, pp. 3939-3944.

Zachman G., (2013) A stochastic fuel-switching model for electricity prices, Energy Econ, **vol. 35**, pp. 5–13.