Forecasting wholesale electricity prices: a hybrid 2-stage perspective

LIN HUANG (MICROSOFT CORPORATION)
YONGLI ZHANG (UNIVERSITY OF OREGON)

ISF 2013
Introduction

Since 1998, USA wholesale electricity market has been deregulated and competitive
  ◦ Heavily dependent on location & time
  ◦ Highly volatile since electricity cannot be stored

Marginal cost of dispatching – Location Marginal Price (LMP)

Implied heat rate (IHR) is the ratio between LMP and natural gas price
  ◦ Higher when LMP is higher (e.g., dispatched from natural gas), and lower when LMP is lower (e.g., dispatched from hydro power)
Literature review & our new proposal

Reduced form models
- Ornstein-Uhlenbeck process (Schwartz 1997; Schwartz & Smith 2000)
- Jump diffusion process (Clewlow et al. 2001; Deng 2001)

Structural models
- Electricity price is modeled as function of exogenous variables, such as demand capacity, fuel cost and transmission constraint (Pirrong & Jermakyan 1999; German & Roncoroni 2006)

We propose a structural model consists of 2 stages
- First, model the temperature through an AR heteroscedasticity time series
- Second, model the relation between IHR and temperature through segmented regression
Data

Daily LMP of NYC Zone J (mean reversion, seasonality & jumps)
- And, temperature from EWR airport (2003-2008)

Figure 3: Historical Hourly Implied Heat Rate and Day Ahead LMP
Data (continued)

Spikes mostly occurred in extreme weather

- The relationship between temperatures and IHR varies across hours/weekdays/weekends

Figure 5: Relationship between Implied Heat Rate and Temperature
Model – time series for weather

Simplified GARCH model

\[ T_t = \text{Seasonal}_t + \sum_{l=1}^{L} \rho_{t-l} T_{t-l} + \sigma_m(t) \epsilon_t \]  \hspace{1cm} (2)

where

\[ \text{Seasonal}_t = \sum_{p=1}^{P} (\sigma_{c,p} \cos(2\pi p \frac{d(t)}{365}) + \sigma_{s,p} \sin(2\pi p \frac{d(t)}{365})), \]  \hspace{1cm} (3)

\[ \epsilon_t \sim iid \mathcal{N}(0,1). \]  \hspace{1cm} (4)
Model – segmented regression for the relationship between temperatures and IHR

Adopted instead of polynomial regression or nonparametric methods due to clear economic meaning, simplicity and manageability

- The segment are fitted by minimizing the distance between the dataset and the segmented line.

\[
\log(\text{IHR}_{w,h,t}) = f_{w,h}(T_t) + \sigma_{w,h}\delta_t
\]  \hspace{1cm} (5)

\[
f_{w,h}(T_t) = \sum_{i=1}^{I} (\alpha_{w,h,i} + \beta_{w,h,i}T_t)1_{T_t \in [T_i,T_{i+1})}
\]  \hspace{1cm} (6)
Model – the correlation between the power price and natural gas

The correlation coefficient is a time series and a function of temperatures

\[\text{Cov}(P^\text{power}, P^\text{NG}) = \text{Cov}(P^\text{NG}(IHR), P^\text{NG}) = E(IHR) \ast \text{Var}(P^\text{NG})\] (7)

Hence,

\[\text{Corr}(P^\text{power}, P^\text{NG}) = \frac{E(IHR)sd(P^\text{NG})}{sd(IHR)}\] (8)

\[= \frac{sd(P^\text{NG})}{\sqrt{E[\exp\{2f_{w,h}(T_t)\}]} \cdot e^{\sigma^2_{w,h} - 1}}\] (9)
Sample empirical results

Figure 8: Segmented Regression
Simulation

Nice prediction performance

Figure 15: Simulations
Concluding remarks and future research directions

Contribution of this method to the field

- Sound microeconomic foundation
- Parsimonious model fitting

Future research direction

- Model supply side factors (e.g., available capacity, and transmission constraints)