Dynamic Modeling of Price Responsive Demand in Real-time Electricity Market: Empirical Analysis

Jaeyong An, P. R. Kumar, and Le Xie

Abstract—In this paper, we study the price responsiveness of electricity consumption from empirical commercial and industrial load data obtained from Texas. Employing a dynamical system perspective, we show that price responsive demand can be modeled as a hybrid of a Hammerstein model with delay following a price surge, and a linear ARX model under moderate price changes. It is observed that electricity consumption therefore has unique characteristics including (1) qualitatively distinct response between moderate and extremely high prices; and (2) a time delay associated with the response to high prices. It is shown that these observed features may render traditional approaches to demand response and retail pricing based on classical economic theories ineffective. In particular, ultimate real-time retail pricing may be limitedly beneficial than as considered in classical economic theories.

Index Terms—Demand Response, Electricity Market, Dynamic System Modeling.

I. INTRODUCTION

In response to the emerging carbon emissions constrained world, the usage of renewable energy sources is increasing. The overall increase in penetration of renewable energy resources in the U.S. is depicted in Fig. 1. Such growth of renewable energy resource is not limited to the U.S. Globally, installed global renewable electricity capacity has continued to increase and represented 28.5% of total electricity capacity in 2014 [1] [2].

Renewable energy sources are generally characterized as variable energy resources (VER) due to their variability and uncertainty [3]. While there has been efforts for better control of resources such as wind farms as doubly fed induction generators [4], their limited controllability and lack of predictability pose new challenges for the operation of the power system.

Price-responsive demand, or demand response (DR) is a key mechanism to achieve system balancing. One avenue is by modifying consumption patterns through economically exposing customers to time-varying pricing that reflect supply-demand balancing status. A number of programs on DR have been implemented or proposed [5] [6]. While such price-responsive demand can potentially provide the key to system operability under high penetration of VERs, the presumed benefits of DR programs substantially depend on how responsive demand actually is to price: the price elasticity of demand [7].

Though price elasticity of demand is critical for the effectiveness of DR programs, previous empirical works based on data-driven static analysis of demand suggests that even load labeled as price-responsive is fairly inelastic [10]. On the other hand, viewing DR as a dynamical system, our previous work [11] analyzing industrial and commercial loads of which is directly exposed to Electric Reliability Council of Texas (ERCOT) real-time wholesale market makes two observations:

1) The consumer’s response to large prices (over the 95%-quantile: $144.42) can be modeled as a Hammerstein system, i.e., a static nonlinearity followed by a linear transfer function [24]. After accounting for this nonlinear transformation, which is typically concave since the response is sublinear, the response exhibits a reduction after a delay of about 0.75-2.5 hours, before subsequently reverting back to normal levels.

2) The response to moderate prices (up to $144.42) can be modeled as a linear stochastic system, specifically as an autoregressive exogenous (ARX) system, i.e., an autoregressive (AR) system with exogenous input and white noise.

This paper extends our previous work [11], by analyzing the economic effect of characteristics of consumer behavior that prevent real-time retail electricity pricing from optimal signaling and respn.

The rest of this paper is organized as follows. In Section II previous works analyzing the models and benefits of price responsive demand, mostly conducted in the economics literature, are reviewed. We introduce our observations on consumer behavior from empirical load data from ERCOT in Section III. In Section IV on the basis of our empirical observation we discuss the implication of our observations, presenting an alternative analysis of the potential benefits of DR in comparison to previous literature. Concluding remarks followed in Section V.

Fig. 1. The growth of renewable generation in the U.S. [1]
II. LITERATURE REVIEW

While the idea of DR is currently attracting wide interest as a solution for system operability under high penetration of VERs, the necessity of DR has been advocated for decades by economists from a market efficiency perspective. The volatility of load that has been challenge for system operators to cope with also entails abrupt and drastic changes in electricity price in the wholesale market. Though extreme price fluctuation is widely observed in today’s restructured electricity wholesale competitive markets, retail customers in most regions do not face frequent price change. While wholesale electricity prices vary from hour to hour, retail prices do not change for months in most electricity markets. Such discordance between rapid fluctuation in wholesale prices and near flat retail prices not only incurs economic inefficiency in terms of social welfare, but also creates price-inelastic wholesale demand that severely exacerbates the volatility of wholesale electricity prices. The combination of inelastic demand with the inherent real-time nature of the market makes electricity markets vulnerable to the exercise of market power [13].

As a method to achieve price responsive demand, there has been a consensus on the potential benefits of real-time retail pricing (RTRP) among economists [14, 15, 16, 17, 20, 21]. The first potential benefit most discussed in the literature is the allocative efficiency improvement resulting from resolving the market inefficiency caused by (near) constant retail electricity prices [9, 17, 19, 21, 22, 23]. The second benefit studied is the increased robustness of the market with RTRP forestalling the exercise of market power [8, 12, 13, 18]. The last benefit considered is that the mitigation of demand volatility induced by real-time price signals will also relieve the need for excessive reserve requirement which incurs a large portion of the societal costs [7, 9, 14]. However, all the potential economic benefits of RTRP substantially depend on how responsive demand is to price, i.e., the price elasticity of demand [7, 9, 14, 15, 16, 17, 20, 21].

The efficiency improvement of RTRP is well analyzed in the literature [9, 19, 21, 23], as depicted in Figure 2. Since the demand curve is a time variant property, it is not likely to happen that the fixed rate meets $C$ or $C'$, which are the optimal market clearing prices in terms of social welfare maximization. Thus, the shaded triangles $\Delta ABC$ and $\Delta A'B'C'$ are the deadweight loss, the economic inefficiency caused by the fixed rate $P_0$. Due to the instantaneousness of electricity, it is reasonable to assume that the electricity at each time slot is a distinct commodity. Thus, RTRP advocates argue that the ultimate real-time retail price is the optimal pricing policy $[19]$ in terms of economic efficiency.

Although the analysis shown in Figure 2 seems reasonable, it requires a crucial assumption to be justified: Demand converges to $C$ or $C'$ almost immediately, in at most one time slot as determined by the market rules. However, this assumption is controversial when the market is fast-paced. Additionally, a fundamental limitation in the demand-supply curve model is that it is difficult to obtain any insight concerning dynamic behavior from the demand curve, which makes it difficult to estimate and predict demand from this static model. The goal of our work is to develop a model for demand response as a stochastic dynamical system where both past prices and past consumption influence future consumption probabilistically.

III. EMPIRICAL OBSERVATIONS ON CONSUMERS’ BEHAVIOR IN ERCOT

In this section, we introduce our prior work [11] that poses the problem of modeling price responsive demand at wholesale level. This work is based on an analysis of the data from an anonymous commercial/industrial (C/I) load in Houston, purchasing its power directly from ERCOT real-time wholesale market, gleaned over nine months (Jan. 1 - Sep. 30, 2008). Based on the empirical data, a dynamical model of consumer behavior is presented.

A. Preliminary Data Analysis

The C/I load and prices from Houston measured at intervals of 15 minutes from Jan. 1, 2008 to Sep. 30, 2008 is presented with respect to time in Fig. 3. The first notable point observed here is that the plot on price (Fig. 3(b)) shows many outliers while the plot on load rarely has any. This is called the “spiky” nature of electricity prices, an irregular sudden extreme price change for a very short duration of 15-30 minutes (Fig. 3(d)). This gives the price a highly non-normal heavy-tail distribution. The fundamental reason for the spiky nature of prices is explained in Section II. The other property we see in Fig. 3 is that the time-series of load shows a depressed demand in “peak hours” (afternoons), over time intervals that overlap with the time intervals exhibiting frequent large outliers in the price time series. Here, we surmise that the depressed demand is a manifestation of demand response, and that this demand response is highly connected to the outliers of price, because the depression is not likely to be explained by the plot of the median prices (Fig. 3(c)).

1 Anonymous even to us.
Fig. 3. Figs. 3(a) and 3(b) show the hourly plots of a C/I load and prices from ERCOT, based on 15-minute measurements from Jan. 1, 2008 to Sep. 30, 2008. Fig. 3(c) shows the median price over these nine months by time of day. Fig. 3(d) shows a particular sample of the price series on June 8, 2008.

Fig. 4. The statistics of Price (P) and C/I load (Q) on workdays (i.e., weekends removed) based on 15-minute measurements from Jan. 1, 2008 to Sep. 30, 2008.

In Fig. 4, the statistics of the prices (P) and C/I load (Q) on workdays are shown. In Fig. 4(b), the empirical probability plot of load versus the normal distribution is indicated by the diagonal dashed line, and we can check that this empirical distribution of the load can be assumed to be a normal distribution. For further validation, we can also check an estimate of the kurtosis, $\mu_4/\sigma^4$, where $\mu_n$ is the nth moment about the mean and $\sigma$ is the standard deviation. It is 2.77, which is close to the value 3.0 for the normal distribution. Also, its skewness, $\mu_3/\sigma^3$, is 0.11, which is close to the value 0 for the normal distribution. (Table I). Therefore, we can conclude that the distribution of the load is near normal. In Fig. 4(c), the plot of autocorrelation (ACF) of the load shows a high correlation between the current and the past load, while the partial autocorrelation (PACF) of the load shown in Fig. 4(d) decays rapidly in no more than five quarter hours (75 minutes). Taking these facts into account, a simple autoregressive (AR) model of order 3 or 5 is concluded to sufficiently well describe the load process.

On the other hand, the first feature we can see in Fig. 4(a) is that the distribution of prices is highly non-normal. The cumulative distribution matches the diagonal dashed line, suggesting closeness to the normal distribution at low to moderate prices. However, the top 5% of the prices deviate severely from the line, reflecting the spiky nature of electricity prices. Such a long-tail property yields huge kurtosis (149.0002) and skewness (10.9133) as shown in Table I.

From the above, it is obvious that it is not feasible to find a linear relationship between load and price over all values of $P$ and $Q$. Hence, we conclude that it is not possible to obtain one single all encompassing universal linear dynamic system model between price and demand. As an alternative, we continue the analysis by assuming that there are two transfer functions (TFs), one for moderate prices which is a linear model, and one for high prices which is nonlinear. The deviation from normality of the top 5% in Fig. 4(a) provides a reasonably good demarcation between moderate prices and high prices.

Estimation of Dynamic model on Load and Price

From the preliminary data analysis in Section III-A, we infer that there exist two qualitatively distinct regimes, a moderate price regime, and a high price regime. In the former, we consider a linear transfer function between price and load, with additional noise to account for uncertainty, i.e., an ARX model driven by white noise. In the high price regime, we consider a concave transformation of peak prices to account for the non-normality of the process. In this section, we further address the problem of the dynamic model identification of DR.

1) Methodology: We briefly discuss the estimation and validation methodology for the estimation of the dynamic model of DR. As a dynamic model of DR, we consider an ARX model driven by white noise, one of the simplest but most utilizable models for forecasting and control. For estimation, we use the least squares (LS) method for estimating the unknown parameters of a linear regression model [26], [27]. For the verification of the existence of DR and the significance of the results of estimated parameters, we consider the
analysis of variance (ANOVA) method [28]. For examining the minimum net contribution of price information to load estimation, we conduct a two-step estimation procedure. To achieve parsimony of the model, we cross-validate the model by a random division of each complete data set under the two separate conditions (i.e., moderate prices and high prices) into two sets of the same size, namely, a training set for estimation and a test set for validation.

2) Autoregressive Exogenous (ARX) model: Denote by \( \{P(t)\}_{t=1}^{N} \) and \( \{Q(t)\}_{t=1}^{N} \), the time series of prices and loads, each consisting of \( N \) observations. Denoting by \( z^{-1} \) the backward shift operator \( z^{-1}X(t) := X(t-1) \), the ARX model can be represented as follows:

\[
\alpha(z^{-1})Q(t) = \beta(z^{-1})P(t) + \epsilon_t,
\]

where vectors \( \alpha := [1 - \alpha_1 - \alpha_2 \ldots - \alpha_m]^\top \) and \( \beta := [\beta_1 \beta_2 \ldots \beta_n]^\top \) are unknown parameters to be estimated, \( \alpha(z^{-1}) := \alpha' \cdot [z^{-i}]_{i=1}^{m} \) and \( \beta(z^{-1}) := \beta' \cdot [z^{-i}]_{i=1}^{n} \) are the characteristic and numerator polynomials of the TF respectively, and \( \epsilon_t \) is an error which is an independent and identically distributed (i.i.d.) noise process with expectation \( E\epsilon_t = 0 \) and variance \( \text{VAR} \epsilon_t = \sigma^2 \).

\( \alpha(z^{-1})Q(t) = \beta(z^{-1})P(t) + \epsilon_t \)

a) Two-step Estimation: Our primary objective is to determine the existence of DR, and understand it, if it exists, from a dynamic system perspective. We employ the following two-step estimation procedure to examine the net contribution of price to load.

1) First estimate the regression parameters \( \hat{\alpha} \), and obtain \( Q_{res}(t) := (1 - \sum_{i=1}^{m} \hat{\alpha}_i z^{-i})Q(t) \).

2) Estimate \( \hat{\beta} \) using the equation \( Q_{res}(t) = (\sum_{i=1}^{n} \hat{\beta}_i z^{-i})P(t) + \epsilon_t \).

Then, the overall estimated ARX model is the following:

\[
Q(t) = \sum_{i=1}^{m} \hat{\alpha}_i z^{-i}Q(t) + (\sum_{i=1}^{n} \hat{\beta}_i z^{-i})P(t) + \epsilon_t,
\]

where \( \hat{\alpha}_i \) and \( \hat{\beta}_i \) are the LS estimators of \( \alpha_i \) and \( \beta_i \).

C. Demand Response to Moderate Price

In this section, an ARX model for DR in the moderate price regime, the prices below the 95%-quantile, is presented. The overall estimation result of fitting an ARX model are shown in Tables IV-VI. The estimated TF of the model is:

\[
TF_{\text{Low}} = \frac{-0.8555 z^{-1} + 0.5273}{1 - 0.8127 z^{-1} - 0.0461 z^{-3} - 0.0366 z^{-5}}.
\]

Tables II and III present the results of the analysis for each of the two steps of estimation. The Estimate column shows the estimated coefficient value, \( SE \) refers to the standard error of the estimate, \( t\text{Stat} \) indicates the t-statistic for a hypothesis test that the coefficient is zero, and \( \text{pValue} \) is the p-value for the t-statistic. This model explains 77.6% of the variance that \( Q(t) \) initially possesses.

Though price has sufficient statistical significance due to its low p-value (0.0147), what we see here is that its innovative contribution to the load forecast is relatively small (less than 0.1%), and most of the change in \( Q(t) \) can be explained by the past of the load itself (AR(5) model). This suggests that a moderate price has very little impact in terms of eliciting demand response. This is also consistent with our observation in the preliminary analysis in Section III-A.

D. Demand Response to High Price

In this section, an ARX model for the high price regime, where the prices are over the 95%-quantile (144.4187 $/MWh), is presented. A sample load evolution time-series after a high price spike is shown in Fig. 5(a). What we can notice here is a huge drop of the load after a one and half hour lag. Fig. 5(b) indicates that such a load drop phenomenon is not an isolated event; we commonly see such a load drop and recovery pattern over two and half hours after price surges. The ANOVA result of Fig. 5(b) in Table V, showing its extremely low p-value (3.86 \times 10^{-3}), supports our observation that there exists a significant load drop 0.5-1.5 hours after a price surge. This significance level is sufficiently low to reject the null hypothesis of a price unresponsive model for the load.

In addition, we also see from Fig. 5(c) and 5(d) that the height of the price surges is correlated to the depth of load drop. Fig. 5(c) depicts the average change \( Q(k) \), at a certain level of price surge \( P \) at time \( t \), where \( Q(k) := \frac{1}{|P|} \sum_{P(t) \in P} |Q(t+k) - Q(t)| \) for all \( P \) in a subset of sample

**TABLE II**

**Estimated AR Model of \( Q(t) \)**

| Coeff. | Estimate | SE  | tStat | pValue |
|--------|----------|-----|-------|--------|
| \( \alpha_0 \) | 238.07 | 13.999 | 17.018 | 8.883 \times 10^{-6} |
| \( \alpha_1 \) | 0.81268 | 0.0085477 | 95.075 | 0 |
| \( \alpha_3 \) | 0.046086 | 0.010267 | 4.4886 | 7.2744 \times 10^{-6} |
| \( \alpha_5 \) | 0.036614 | 0.0085466 | 4.284 | 1.8579 \times 10^{-5} |

\( \sqrt{\text{MSE}} = 301 \)

F-statistic vs. constant model: 8.81 \times 10^{3} \ p-value = 0

**TABLE III**

**Estimated Linear Model of \( Q_{res}(t) \)**

| Coeff. | Estimate | SE  | tStat | pValue |
|--------|----------|-----|-------|--------|
| \( \beta_0 \) | 22.506 | 10.054 | 2.2385 | 0.025218 |
| \( \beta_1 \) | -0.8555 | 0.42677 | -2.0046 | 0.045043 |
| \( \beta_2 \) | 0.5273 | 0.43006 | 1.2261 | 0.2202 |

\( \sqrt{\text{MSE}} = 301 \)

F-statistic vs. constant model: 4.22 \ p-value = 0.0147

**TABLE IV**

**The ARX Model on \( Q(t) \)**

\[
(1 - \alpha_1 z^{-1} - \alpha_3 z^{-3} - \alpha_5 z^{-5})Q(t) = \hat{\beta}_1 z^{-1} + \hat{\beta}_2 z^{-2} P(t) + \epsilon_t + \epsilon_0
\]

| Coeff. | Estimate | Coeff. Estimate |
|--------|----------|-----------------|
| \( \alpha_1 \) | 0.81268 | \( \beta_1 \) |
| \( \alpha_3 \) | 0.046086 | \( \beta_2 \) |
| \( \alpha_5 \) | 0.036614 | \( \epsilon_0 \) |

\( \sqrt{\text{MSE}} = 301 \)

\( R^2 = 0.776 \)
prices \( P = \{P(t) : P_{\text{min}} \leq P(t) \leq P_{\text{max}}\}\) for given \( P_{\text{min}} \) and \( P_{\text{max}} \). We check that higher \( P_{\text{min}} \) and \( P_{\text{max}} \) result in a greater load drop. The negative correlation between the height of the price surge (\( \Delta P = P(t) - P(t-1) \)) and the load \( Q \) is plotted in Fig. 5(d), which is negatively significant after \( k = 5 \) quarter-hour periods (i.e., one hour and 15 minutes) following a price surge.

On the basis of the above observations, we establish a simple dynamic model between the magnitude of the price surge and the load, in the case of high price surges. Taking into account the long-tailed characteristic of prices, we consider a linear model for a concave transformation \( \log P(t) \), instead of \( P(t) \). In this paper, we present a TF for a specific time period, from 2:00pm to 2:30pm, due to the innate time-dependency on DR. The estimation results for the ARX model of DR at high price are presented in Tables VI, VII, and VIII. The estimated TF of the ARX model up to 51.2%, as presented in Table VIII, compared to the AR model for the moderate price regime (Table II). However, we observe that a relatively high portion (27%) of the variance of \( Q_{\text{peak}}(t) \) is explained by the estimated model of \( Q_{\text{peak}}(t) \) shown in Table VIII, from which we conclude that the innovation from the price information is significant to improve \( R^2 \) of the ARX model up to 51.2%, as presented in Table VIII.

In Fig. 5 The validity of our model is shown by sample load forecast. Figs. 6(a) and 6(b) delineate the errors in the load forecast at 3:15pm after a price surge at 2:15pm. We see that the forecasted \( \hat{Q}(t) \) and the actual \( Q(t) \) at \( t = 3:15pm \) are reasonably well correlated (correlation \( r_{Q\hat{Q}} = 0.7160 \) in Fig. 6(b) and that the errors exhibit normality (Kurtosis = 3.1809) in Fig. 6(a).

### Summary

Our empirical study suggests that (1) the demand only responds to high price surges at peak hour and (2) there exists a demand response delay consequent on a high price surge. The second finding shows that there exists a certain “inertia” in consumption, resulting in a certain time delay to reduce power consumption after a peak price observation.

| Source | SS    | DF   | MS    | F     | p-value |
|--------|-------|------|-------|-------|---------|
| Groups | 1.21 \times 10^7 | 10   | 1.21 \times 10^6 | 3.21 | 3.86 \times 10^{-4} |
| Error  | 3.89 \times 10^9 | 10351 | 3.76 \times 10^5 | | |
| Total  | 3.90 \times 10^9 | 10361 | | | |

SS: Sum of squares; DF: Degree of freedom of error; MS: Mean square; F: F-statistic.

\[ F(t)_{\text{2:15pm}}^{\text{Peak}} = \frac{-220.1 z^{-4}}{1 - 0.4015 z^{-1} + 0.2383 z^{-2} - 0.2512 z^{-4}}, \]

which accounts for 51.2% of the variance of \( Q(t) \). The notable feature we find here is that the accuracy of the AR model for \( Q(t) \) is severely degraded (\( R^2 = 33.2\% \)) in Table VI compared to the AR model for the moderate price regime (Table II). However, we observe that a relatively high portion (27%) of the variance of \( Q_{\text{peak}}(t) \) is explained by the estimated model of \( Q_{\text{peak}}(t) \) shown in Table VIII, from which we conclude that the innovation from the price information is significant to improve \( R^2 \) of the ARX model up to 51.2%, as presented in Table VIII.

### Table VI

**Estimated AR Model for \( Q(t) \)**

| Coeff. | Estimate | SE  | tStat | pValue |
|--------|----------|-----|-------|--------|
| \( \alpha_0 \) | 748.26 | 233.72 | 3.2013 | 0.0025097 |
| \( \alpha_1 \) | 0.40153 | 0.11763 | 3.4133 | 0.0013678 |
| \( \alpha_2 \) | -0.23826 | 0.14611 | -1.6308 | 0.10992 |
| \( \alpha_4 \) | 0.25124 | 0.11516 | 2.1816 | 0.0344 |

\( \sqrt{\text{MSE}} = 336 \)

\( R^2 = 0.332 \)

F-statistic vs. constant model: 7.44 \ p-value = 0.000377

### Table VII

**Estimated Linear Model for \( Q_{\text{peak}}(t) \)**

\[ Q_{\text{peak}}(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + \epsilon_t \]

| Coeff. | Estimate | SE  | tStat | pValue |
|--------|----------|-----|-------|--------|
| \( \beta_0 \) | 1213.4 | 293.68 | 4.1316 | 0.00014688 |
| \( \beta_1 \) | -220.1 | 52.774 | -4.1707 | 0.00012965 |

\( \sqrt{\text{MSE}} = 281 \)

\( R^2 = 0.27 \)

F-statistic vs. constant model: 17.4 \ p-value = 0.00013

### Table VIII

**The ARX Model for \( Q(t) \)**

\[ (1 - \alpha_1 z^{-1} - \alpha_2 z^{-2} - \alpha_4 z^{-4}) Q(t) = \beta_1 z^{-4} \log P(t) + \epsilon_t + \epsilon_0 \]

| Coeff. | Estimate | Coeff. | Estimate |
|--------|----------|--------|----------|
| \( \alpha_1 \) | 0.40153 | \( \beta_1 \) | -220.1 |
| \( \alpha_2 \) | -0.23826 | \( \epsilon_0 \) | 1961.66 |
| \( \alpha_4 \) | 0.25124 | | |

\( \sqrt{\text{MSE}} = 281 \)

\( R^2 = 0.5124 \)
IV. THE EVALUATION OF THE BENEFIT OF DEMAND EXPOSURE TO REAL-TIME PRICING

In this section, we shall further analyze the results from the data. We will first examine the rationality of consumer behavior. Then we will show that the observed delay in demand response changes classical arguments about the role of prices and the equilibrium process, as well as classical efficiency results of markets.

A. Is the Observed Consumer’s Behavior Rational?

One of our observations in Section III is that price responsive demand exhibits delayed response to price shock at peak hours. It may seem to be irrational to decrease one’s demand after a price shock has already occurred. However, if we consider that consumer behavior is based on prediction of price, rather than the current price itself, then we can explain the delayed response based on the inertia of demand. In this sense, decreasing one’s demand after a price spike, specifically, during or after the price plummets after price surge, can be well explained as a rational behavior if there is a high chance of a price increase after price spike. The chance of such a price increase relapse after a price spike is presented in Figure 7.

Figure 7 shows a comparison of the estimated conditional probability of a high price in different situations, based on the obtained data from Houston. We can easily discern from Figure 7 that the conditional probability of a second price spike following the occurrence of a price spike quickly reduces in off-peak hours. However, we also observe that the conditional probability of a price spike after the occurrence of a price spike during peak-hours remains at a significantly higher level than the probability of price spike without any conditioning. This rationalizes our observation that if we assume a consumer has limited ability for immediate load reduction, then a rational consumer adjusts its load in response to price spike in spite of its inertia, because the relative chance of repeated price surge after a price spike is significantly high, and the demand is still not able to respond quickly to that following price surge. This also explains why demand is responsive to price only when it is during peak hours. The rigorous analysis of the consumer rationality behind such behavior will be presented in subsequent work.
While the analysis on the allocative efficiency of RTRP compared to fixed price under volatile demand analyzed in previous literature, is depicted in Figure 2, this analysis can be extended to the allocative efficiency of RTRP compared to fixed price under supply fluctuation by VERs as depicted in Figure 8. Analogously, the shaded triangles $\Delta ABC$ and $\Delta AB'C'$ are the deadweight loss, representing the economic inefficiency caused by the fixed rate $P_0$, or, conversely, the expected allocative efficiency benefit from RTRP compared to a fixed tariff. According to the classical arguments on RTRP, more frequent price change would be more beneficial, because they would more accurately the balance supply and demand in real time, so that it is more informative for consumers to make an optimal decision.

However, our observation suggests that such allocative efficiency is not likely to be achievable because of the distortion in demand curve caused by demand inertia. The inability of customers to respond instantly may distort the demand curve to a vertical line in quantity-price plot. Noting this distortion, the demand behavior in practice may be realized as if it is exposed to a fixed price. The impact on allocative efficiency from the demand curve distortion is presented in Figure 9.

Figure 9(a) depicts the situation where VER drops due to events such as sudden diminution of wind or cloud cover blocking the sun. Suppose that the market equilibrium point before the VER drop is $A$. Then, the VER drop shifts the supply curve to the left. After the supply curve shift, the optimal market clearing point maximizing social welfare is $C$ on the nominal demand curve. However, the actual market clearing point is realized at $B$ due to demand curve distortion. Hence, the shaded triangle $\Delta ABC$ is the deadweight loss, exhibiting the economic inefficiency following from demand inertia.

In Figure 9(b), the analysis of the situation where VER is restored by incidents e.g. the increase in wind generation caused by a gust of wind, or that of solar generation following a cloud gap is provided. Suppose the market equilibrium point before VER restoration is $A$. The increase in generation followed by VER restoration event results in the newly formed supply curve to lie on the right hand side of the previous one. While $C$ on the nominal demand curve is the optimal market clearing point after the supply curve shift, the demand curve distortion caused by the hedging of demand against the risk of a repeated VER drop-off may result in the actual market clearing point being realized at $B$. Again, the shaded triangles $\Delta ABC$ indicates the deadweight loss, the economic inefficiency from demand inertia.

The allocative (in)efficiency analysis of RTRP under demand inertia suggests that there is a fundamental limitation to achieving market efficiency that can be expected from traditional market efficiency analysis without consideration of demand inertia. This necessitates a redefinition of market efficiency from an optimal control theoretical perspective. In addition, the demand behavior under RTRP is as if it is exposed to a fixed price, leading to another crucial implication that RTRP may not significantly resolve the vulnerability of markets from the exercise of market power. The differences in the ability of various market participants to control their behavior endows different market powers to each market participant; the more instantaneously responsive market participant has an advantage over market participants with larger delay. Such a combination of differentially endowed market powers makes market more vulnerable to the exercise of market power or market manipulation. A similar argument is found in financial markets with high-frequency trading (HFT) practices, in terms of the robustness with respect to market manipulation and market fairness [25]. Moreover, the inability of demand response to instantaneously respond also suggests that RTRP may contribute negatively to demand volatility mitigation, so that the savings in the cost of maintaining reserve capacity may be less than expected under previous literature.

V. CONCLUSION

A market is a dynamical system that is designed to proceed toward an optimal state as its equilibrium. However, such a process necessarily requires a certain amount of time to reach its equilibrium. While dynamic modeling and control on the generation side in power systems has been well understood, the understanding of dynamic behavior on the demand side in response to price has been unclear. In this paper, we consider a consumer’s dynamic behavior in response to real time price change, by studying empirical data on a price-responsive load in the ERCOT area.

Our empirical study suggests the following: (1) the price responsiveness of demand may exhibit qualitatively different behavior in response to “normal price” and “high price”;
and (2) there exists a demand response delay consequent on a high price surge at peak hours.

Such behavioral features imply that frequent price changes do not necessarily bring economic efficiency in the sense of social welfare maximization.

This idea provides important guidance for designing two fundamental factors in time-varying retail electricity prices, frequency and timeliness. Here “frequency of price” is the frequency at which retail prices change, and “timeliness of price” is the time lag between when a price is set and when it is effective [19]. It is generally assumed among economists that RTRP with high frequency and just-in-time timeliness would be ideal in terms of economic efficiency in the electricity market, as RTRP is an attempt to provide more accurate signals closely reflecting the actual supply/demand status in the market. However, the inference based on our work is that neither argument is necessarily right. The inherent delay in the responsiveness of loads to high price volatility exacerbates the predictability of price, thereby making demand less responsive to RTRP, which in fact worsens economic efficiency. Consumers which are more exposed to market volatility stiffen their demand to be more inelastic and tend to be more conservative due to the inertial nature of demand. This suggests that there exists a trade-off between controllability of demand and observability of markets, so that there may exist an optimal frequency and timeliness which should be carefully considered for optimal pricing design. This also supports the importance of relatively long-term contract markets such as day-ahead electricity markets. Market efficiency should be reanalyzed taking into consideration the trade-off between the controllability of demand and the observability of the market. In subsequent work, we aim to provide a rigorous analysis of consumer rationality and develop a quantitative prediction model for demand response.

REFERENCES

[1] P. Beiter, “2014 Renewable Energy Data Book,” National Renewable Energy Laboratory (NREL) of U.S. Department of Energy (DOE), 2015.
[2] “Electric Power Annual 2014,” U.S. Energy Information Administration (EIA) of U.S. Department of Energy (DOE), Feb. 2016.
[3] “The Future of the Electric Grid,” Research Report, Massachusetts Institute of Technology Energy Initiative (MITEI), 2011.
[4] I. A. Hiskens, “Dynamics of Type-3 Wind Turbine Generator Models,” Power Systems, IEEE Transactions on, vol. 27, no. 1, pp. 465-474, Feb. 2012.
[5] M.H. Albadi, E.F. El-Saadany, “A summary of demand response in electricity markets,” Electric Power Systems Research, vol. 78, no. 11, pp. 1989-1996, Nov. 2008.
[6] P. Siano, “Demand response and smart grids-A survey,” Renewable and Sustainable Energy Reviews, vol. 30, pp. 461-478, Feb. 2014.
[7] H. Allcott, “Rethinking real-time electricity pricing,” Resource and Energy Economics, vol. 33, pp. 820-842, 2011.
[8] H. Allcott, “The Smart Grid, Imperfect Competition, and Entry in Electricity Supply,” NBER Working Paper, no. 18071, May 2012.
[9] H. Allcott, “Real-Time Pricing and Electricity Market Design,” Working paper, 2013.
[10] L. Xie, S. Puller, M. Ilie, and S. Oren, “Quantifying Benefits of Demand Response and Look-ahead Dispatch in Systems with Variable Resources,” PSERC Final Report M26, Aug. 2013.
[11] J. An, P. R. Kumar, and L. Xie, “On Transfer Function Modeling of Price Responsive Demand: An Empirical Study,” in Proc. IEEE Power & Energy Society General Meeting 2015, pp. 1-5, 2015.
[12] S. Borenstein and J. Bushnell, “An Empirical Analysis of The Potential for Market Power in California’s Electricity Market,” Journal of Industrial Economics, vol. 47, pp. 285-323, Sep. 1999.