Wind generation and zonal-market price divergence: evidence from Texas

(Energy Policy, 39:7, 3928-3938)

C.K. Woo\textsuperscript{a,b}, J. Zarnikau\textsuperscript{c,d}, J. Moore\textsuperscript{a,*}, I. Horowitz\textsuperscript{e,f}

\textsuperscript{a} Energy and Environmental Economics, Inc., 101 Montgomery Street, Suite 1600 San Francisco, CA 94104, USA

\textsuperscript{b} Hong Kong Energy Studies Centre, Baptist University of Hong Kong, Hong Kong

\textsuperscript{c} Frontier Associates LLC, 1515 S. Capital of Texas Highway, Suite 110, Austin, TX 78746, USA

\textsuperscript{d} The University of Texas at Austin, LBJ School of Public Affairs and Division of Statistics, Austin, TX 78713, USA

\textsuperscript{e} Warrington College of Business, University of Florida, Gainesville, FL 32611, USA

\textsuperscript{f} Senior Research Associate, Public Utilities Research Center, University of Florida, Gainesville, FL 32611, USA

Abstract

The extant literature on wind generation and wholesale electricity spot prices says little about how wind generation may affect any price differences between two interconnected sub-markets. Using extensive data from the four ERCOT zones of Texas, this paper develops a two-stage model to attack the issue. The first stage is an ordered-logit regression to identify and quantify, for example, the impact of wind generation in the West zone on the estimated probability of a positive or negative price difference between the North and West zones. The second stage is a log-linear regression model that identifies and quantifies the estimated impact of wind generation on the sizes of those positive and negative price differences. It is shown that high wind generation and low load in the wind-rich ERCOT West zone tend to lead to congestion and zonal price differences, that those differences are time-dependent, and that other factors such as movements in nuclear generation and natural-gas prices, as well as fluctuating non-West zone loads, also play a role. The results have broad implications for energy policy makers that extend well beyond the borders of Texas and, indeed, those of the United States.

Keywords: Wind energy; zonal-market price difference; ERCOT

* Corresponding author. Phone: 415-391-5100, Fax: 415-391-6500.

Email address: jack@ethree.com (Jack Moore)
1. Introduction

The debate over climate change has helped fuel interest in alternative sources of energy that promise low carbon emissions, including nuclear, solar, and wind energy, the latter of which is the focus of this paper. Renewable-energy policies in North America and Europe, in particular, have led to large-scale wind-energy development. In this regard, Texas has been the state with the most wind generation in the United States (Haas et al., 2008; Stahl et al., 2009; Abbad, 2010; Jacobsen and Zvingilaite, 2010; Sioshansi and Hurlbut, 2010). Globally, “[2009] was a record-setting year for wind energy in the IEA Wind member countries, which installed more than 20 gigawatts (GW) of new wind capacity. This growth led to a total of 111 GW of wind generating capacity, with more than 2 GW operating offshore” (IEA, 2010, p.2).

Integrating this rapidly expanding and intermittent generation from remote locations into electricity grids is challenging for grid operators, resources planners, and policy makers (Sovacool, 2009). While the physical impacts of wind generation on grid operations are well documented, the impact of increasing wind power on energy prices in wholesale markets has received less attention. That said, the displacement of generation sources with higher operating costs by wind power in Denmark and Germany has resulted in lower electricity prices (EWEA, 2010) and magnified wholesale-price volatility in Denmark (Jacobson and Zvingilaite, 2010). In Texas, which is the data source for the empirical study that follows, wind generation has occasionally led to negative prices in the state’s West zonal market (Lively, 2009). The West zone is one of the four competitive Electricity Reliability Council of Texas (ERCOT) power sub-markets, the other three being the North, South, and Houston zones.
In addition to promoting wind-energy development, Australia, New Zealand, and parts of North America, South America, and Europe have undergone electricity-market restructuring to introduce competition into the wholesale market for electricity generation (Sioshansi and Pfaffenberger, 2006; Woo et al., 2006a; Zarnikau, 2005, 2008). Wholesale electricity spot-market prices are inherently volatile. The volatility, accompanied by occasional sharp spikes, has sparked extensive research into spot-price behavior and dynamics.

The rich research on wind generation and wholesale spot prices, however, says little about how wind generation may affect any price differences between two interconnected sub-markets within a larger power market. If those sub-markets are only geographical artifacts, then absent frictions such as transmission constraints, the law of one price implies that the same price will prevail in all sub-markets under least-cost dispatch (Hogan, 1992; ERCOT, 2004) or active competitive trading (Woo et al., 1997). While a grid operator can ably manage transmission congestion (Kumar et al., 2005; ERCOT, 2010b), frequent and large price divergences indicate an urgent need for transmission expansion to resolve severe congestion. Indeed, Texas has already identified the least-cost transmission plan to bring abundant wind resources from its West zone to load centers around Dallas and Houston (Sioshansi and Hurlbut, 2010; ERCOT, 2009).

Texas, which is the largest electricity-consuming state in the nation, is a prime candidate for studying divergences from the law of one price. When the state’s zonal market prices are paired over 15-minute time intervals, depending upon the sub-market pairings, the paired prices have over the past three-and-a-half years defied the law of one price between 15% and 17% of the time, yielding strictly positive or negative price

---

2 Volatility is defined here as the standard deviation of the log returns of electricity prices.

3 The term “dynamics” refers to the behavior and evolution of the system over time.
differences that can be very large in size, up to $3,500/MWh. This particular datum has evolved from our access to a unique and rich ERCOT data base of 15-minute high-frequency electricity price data observed over the 41-month period extending from January 2007 through May 2010. These data allow us to identify and quantify the impact of wind generation on those paired-price divergences. To the best of our knowledge, we are the first to do so.

We accomplish this through a two-stage regression analysis of a rich sample of over 115,000 widely varying observations. The first stage entails the estimation of three ordered-logit regressions, one for the North and West zonal pairing, one for the South and West zonal pairing, and one for the Houston and West zonal pairing. We obtain parameter estimates for three sets of binary monthly, daily, and hourly dummy variables and seven metric variables including zonal MWh loads, nuclear generation, natural-gas prices, and of greatest interest for present purposes, wind generation for the ERCOT system. These estimates allow us to determine the direction and impact of wind generation on the probability of a positive electricity-price difference between, say, the North and West zones.

The second stage entails the estimation of two log-linear regressions for each of the three zonal market pairings. In these regressions, the dependent variables are, on the one hand, the positive paired-price differences and on the other the absolute values of the negative paired-price differences. The independent exogenous variables are the same dummy and metric variables that are employed in the ordered-logit regressions. The parameter estimates attached to the metric variables are directly interpretable as price difference elasticities.
Our regression analysis yields the following key findings. First, on average, rising wind generation increases the likelihood and size of a strictly positive paired-price difference between a non-West zone and the West zone in Texas. So too do an increase in the price of natural gas and an increase in nuclear generation in the ERCOT system. Second, non-West zonal loads tend to move the likelihood and size of a zonal market pair's strictly positive price difference. Third, a rising West zonal load tends to reduce the West zone's wind-energy exports, thus reducing the likelihood and size of a zonal market pair's strictly positive price difference. Fourth, wind generation, the natural-gas price, nuclear generation, and zonal loads on average have directionally opposite effects on a market pair’s strictly negative price difference. Finally, price divergence is time-dependent, reflecting the effects of month of the year, day of the week, and hour of the day. From a policy-making standpoint, these findings confirm that wind generation tends to cause zonal prices to diverge in Texas, supporting Texas’ planned transmission expansion to facilitate the state’s projected rapid growth in wind generation.

2. Wind generation in the Electric Reliability Council of Texas

2.1 The ERCOT market

Roughly 85% of the electricity needs in Texas are satisfied through the ERCOT market. Figure 1 portrays ERCOT’s generation mix in 1997-2009, underscoring the state’s rising wind generation.
ERCOT has undergone gradual restructuring since the mid-1990s (Baldick and Niu, 2005; Zarnikau, 2005; Adib and Zarnikau, 2006). Legislation enacted in 1995 required the Public Utility Commission of Texas (PUCT) to establish rules to foster wholesale competition and create an Independent System Operator (ISO) to ensure non-discriminatory transmission access, leading in the summer of 1997 to ERCOT’s establishment as the first operating ISO in the nation. Sweeping reforms were introduced through Senate Bill 7 (SB 7) in 1999, allowing customers of the investor-owned utilities within ERCOT to choose among competitive retail electric providers (REPs) for retail supply of electricity beginning on January 1, 2002. SB 7 also enhanced the ISO’s centralized control over, and operation of, the wholesale market, replacing ten “control centers” formerly operated by various utilities. This led to the establishment of formal markets for ancillary services and balancing energy.

From January 2002 until December 2010, ERCOT maintained a zonal market structure, containing between four and five zones. The boundaries of each zone were re-
examined each year through ERCOT’s stakeholder process. Dynamic transmission limits among the zones could change between each 15-minute interval, depending upon system conditions. Figure 2 provides the average limits between the zones in effect in 2009.

For the five Commercially Significant Constraints (CSCs), the simple average of the constraint quantities for all 15-minute intervals of 2009 was:

1. South to North: 1,249 MW
2. North to South: 728 MW
3. North to Houston: 3,198 MW
4. West to North: 1,015 MW
5. North to West: 741 MW

Figure 2. Zones and commercially-significant transmission constraints in ERCOT

Figure 3 indicates that three of the four zones have sufficient summer-peak resources to meet their within-zone anticipated peak demands. The exception is the Houston zone. While geographically large, the West zone is sparsely populated and hence has very little demand for electricity.
Zonal-balancing energy market prices are determined through least-cost dispatch (ERCOT, 2004). Absent commercially significant transmission constraints (“CSCs”), balancing energy prices are equal across zones; when inter-zonal CSCs are binding, however, zonal market prices diverge (ERCOT, 2010b).

2.2 Wind generation in Texas

The growth in wind power in Texas has actually had little to do with concern over climate change (Caputo, 2007; Price, 2009). The promotion of wind energy has, however, proven attractive in a state with declining oil production, the largest non-hydro renewable-resource potential in the nation, an entrepreneurial business climate, and an interest in diversifying the state’s energy mix.

Texas has a wind-resource potential of 1,901,530 MW and 6,527,850 GWh (NREL, 2010), which is over 17% of the nation’s total resource potential, far in excess of the state’s power needs. SB 7 established initial goals for renewable-energy capacity of
2,000 MW by 2009 and a renewable-energy credit (REC) trading program, as well as a process to facilitate the construction of electrical transmission facilities to interconnect an expanding amount of wind power. In the 2005 legislative session, SB 20 increased the state’s goal for renewable energy to 5,880 MW by 2015 and 10,000 MW by 2025.

Figure 4 indicates that Texas has already met the 2015 goal and is on track to meet the 2025 goal well ahead of schedule. Since 1999, ERCOT has also invested over $5 billion in transmission upgrades, mirroring the growth in wind generation. Figure 5 shows that the vast majority of wind-energy development has occurred in the West zone.

![Figure 4. ERCOT installed wind generating capacity and transmission investment since 1999 (Source: ERCOT, 2006, 2008, 2010)](image-url)
Figure 5. ERCOT installed generation by zone. (Source: EIA, 2009; matched to 2010 ERCOT zone based on the county location of generator)

3. Descriptive analysis of the zonal paired-price differences

3.1 Example of ERCOT price movements

ERCOT has distinctive zonal-market price patterns, thanks to the concentration of wind power in the West zone, limited transmission transfer capability between zones, and federal production tax credits for wind power. To presage our regression results in Section 4, here we provide a descriptive analysis of the zonal-market paired-price differences between the North zone and its immediately adjacent West zone. The findings for the South-West and Houston-West market pairs are similar to those of the North-West pairing and therefore are not repeated.

The upper panel of Figure 6 shows the 15-minute electricity spot-price movements in the ERCOT West and North zones for one sample month, April 2009,
when prices are mostly identical across the zones and range from $15/MWh to 45/MWh. There are, however, a few hours when North-zone prices spike up for 15 minutes to one hour, before returning to the typical range of $15/MWh to $45/MWh. West-zone prices follow North prices during some of these spikes, but remain low in others, suggesting that transmission constraints prevent an export response from the West to the price spikes in the North. Still further, prices in the West diverge from those in the North and even dip into negative territory during a limited number of hours. Negative wholesale electricity prices in west Texas occur because federal production-based tax credits motivate wind-energy producers to make negative price bids in order to have their output dispatched into the grid (Lively, 2009).

The lower panel of Figure 6 compares ERCOT's wind generation and the West zone's load. Sustained dips in West-zone prices appear to occur more frequently when wind output is high, such as during the April 25-27 three-day span. High wind output alone, however, is not the only cause for the zonal price divergence (e.g., April 1-2), suggesting that there exist other influencing factors.
Figure 6. 15-minute zonal prices, West-zone load and wind generation in April 2009. (Source: http://www.ercot.com/mktinfo/services/bal/index)

Figure 7 plots the paired-price differences between the North and West zones for the entire 3.5 years of 15-minute data for our more than 115,000 observations, with the West-zone prices being subtracted from those in the North. This figure shows a noisy data pattern, with over 80% of the paired observations having a zero price difference. While to the naked eye wind output and the contemporary paired-price difference would seem to be positively related, this relationship would also appear to be quite weak, corroborating the intuitively plausible notion that wind generation is not the sole cause of zonal price divergences.
To explore how system conditions might influence any zonal-market price differences, Figures 8.A through 8.C contain descriptive statistics for four of our metric independent variables, notably: 15-minute wind generation; 15-minute nuclear generation; 15-minute zonal loads; and daily Henry Hub natural gas prices. As discussed below, these variables are considered exogenous and, subject to statistical verification, a priori useful for explaining any deviations from the law of one price (i.e., any non-zero paired-price differences). These figures suggest that wind generation tends to be higher when there are strictly positive paired-price differences between the North and West zones. The West-zone load appears to be slightly higher when there are strictly negative paired-price differences between the two zones. To untangle the less-than-clear influence of the remaining variables on the paired-price differences, however, it is necessary to use
the regression analysis detailed in Section 4.

Figure 8.A: Box plot of generation, zonal load, and gas price, conditional on a positive paired-price difference, when the West-zone price is less than that in the North

Figure 8.B: Box plot of generation, zonal load, and gas price, conditional on a zero price difference, when the West-zone price equals that in the North

Figure 8.C: Box plot of generation, zonal load, and gas price conditional on a negative price difference, when the West-zone price exceeds that in the North
Finally, Figure 9 reveals that the distribution of the zonal paired-price differences is rather skewed to the left with the mean price difference falling well above the median difference. The distribution of the natural logarithms of the differences is more symmetric, a property that serves us well in our log-linear regression model of the size of the price difference.

![Figure 9: Distribution of positive zonal paired-price differences (y) vs. distribution of the natural-logarithms of the paired-price differences (ln y)](image)

4. Regression analysis of the zonal paired-price differences

Consider the electricity spot-market paired-price difference $y_t$ in time interval $t$ between a non-West ERCOT zone (e.g., Houston, North, or South) and the West zone. For over 80% of the 15-minute observations for each zonal market pair, $y_t = 0$, rendering ordinary least squares (OLS) unsuitable for estimating the effect of such variables as wind generation on the pair's price difference (Maddala, 1983). In addition, up to 4% of the observations have $y_t < 0$, and 14% have $y_t > 0$. The presence of many observations
with \( y_t < 0 \) precludes using a censored regression (e.g., Tobit) to estimate the wind-generation effect on \( y_t \), even though the technique is well-suited to a data file with many zeroes and some strictly positive values for the dependent variable, as in an outage-cost analysis (e.g., Woo and Train, 1988; Hartman et al., 1991). Thus, we adapt the generalized econometric models with selectivity in Lee (1983) to formulate a two-stage approach that accounts for the peculiar features of the ERCOT zonal paired-price difference data. Notwithstanding the highly noisy data described in Section 3, this approach proves to be fruitful in delineating the effect of wind generation on the zonal-market paired-price differences.

In the first stage of this approach, we estimate an ordered-logit regression model in which the natural logarithms of the odds of a positive paired-price difference, a non-negative difference, and a negative difference comprise the dependent variable. In the second stage, the natural logarithm of the size of the price difference is the dependent variable in a log-linear regression. In either stage, the independent explanatory variables comprise a single set of binary monthly, weekly, and hourly dummy variables, as well as a single set of metric variables that we might plausibly expect to influence both the probability of a positive or negative paired-price difference, and the magnitudes of those differences.

4.1 The independent variables

The following variables and corresponding notation enter both regression formats.

*Binary variables*
The month-of-the-year, day-of-the-week, and hour-of-the-day binary indicators aim to capture that both the probabilities of a positive or negative paired-price difference and the magnitudes of those differences may be time-dependent for any time period, \( t \), after controlling for the influence of the exogenous metric variables listed below. After suppressing the subscript that would distinguish between zonal pairs, these binary variables are: (a) \( M_i = 1 \) for \( i = 1 \) (January), \( \ldots \), 11 (November), and is zero otherwise; (b) \( W_j = 1 \) for \( j = 1 \) (Sunday), \( \ldots \), 6 (Friday), and is zero otherwise; and (c) \( H_k = 1 \) for \( k = 1 \) (an hour ending at 1:00), \( \ldots \), 23 (an hour ending at 23:00), and is zero otherwise.

**Metric variables**

The following numeric metric variables also enter all regressions: (a) \( x_1 = 15 \)-minute wind generation of the ERCOT system, which is largely at the mercy of random wind conditions; (b) \( x_2 = \) Daily Henry Hub natural-gas price; (c) \( x_3 = 15 \)-minute MWh nuclear generation in the ERCOT system; (d) \( x_4 = 15 \)-minute MWh load in ERCOT’s Houston zone; (e) \( x_5 = 15 \)-minute MWh load in ERCOT’s North zone; (f) \( x_6 = 15 \)-minute MWh load in ERCOT’s South zone; and (g) \( x_7 = 15 \)-minute MWh load in ERCOT’s West zone.

4.2 Stage 1: Ordered-logit model

As each paired-price difference can only belong to one of the three mutually-exclusive ordered categories (i.e., strictly positive, zero, or strictly negative), our first-stage estimation uses an ordered-logit regression model (Greene, 2003, pp. 736-739).

Specifically, and still suppressing a paired-zone subscript, as well as subscripts to delineate month, day, and hour, let \( p_{1t} \) denote the probability of a positive paired-price
difference in time period $t$, let $p_{2t}$ denote the probability of a zero price difference in that time period, and let $p_{3t} = 1 - p_{1t} - p_{2t}$ denote the probability of a negative difference. Then, the *odds* of a positive paired-price difference in $t$ are given by:

$$O_{1t} = \frac{p_{1t}}{1 - p_{1t}}.$$

Similarly, the odds of a non-negative paired-price difference are given by:

$$O_{2t} = \frac{(p_{1t} + p_{2t})}{p_{3t}}.$$

The odds of a negative paired-price difference are defined accordingly. The logits that define the dependent variable in the regression are given by the natural logarithms of the odds, or $L_{1t} = \ln(O_{1t})$ and $L_{2t} = \ln(O_{2t})$. Since the probabilities sum to unity, it is not necessary to define a third logit, which can be inferred from the first two.

The logit regression’s parametric specification that incorporates the above-listed variables may be written as

$$L_{pt} = \alpha_p + \sum_i \mu_i M_{it} + \sum_j \alpha_j W_{jt} + \sum_k \eta_k H_{kt} + \sum_r \beta_r x_{rt}$$

(1)

for $p = 1, 2; i = 1, ..., 11; j = 1, ..., 6; k = 1, ..., 23; r = 1, ..., 7$. Thus, a single set of slope parameters is estimated, and two intercepts are simultaneously estimated such that $\alpha_2 \geq \alpha_1$, since the odds of a non-negative outcome must be at least as great as the odds of a positive outcome. Our hypotheses as to the impact of the exogenous variables on the odds and consequent probabilities directly translate into testable hypotheses as to the signs of the regression coefficients.

4.2.1 Hypotheses

We have no *a priori* basis for speculating as to directions and magnitudes of, or indeed as to whether there are any, time factors that impact the probabilities of positive,
zero, and negative paired-price differences. Rather, we include the time-related dummy variables to allow for any impact that month, day, or hour might have on those probabilities. The metric variables, however, are another matter.

Based on the underlying hypothesis that wind generation congests the ERCOT grid, we expect $x_{1t}$ to have a positive marginal effect on the probability of a positive paired-price difference, implying a coefficient of $\beta_1 > 0$.

The exogenous Henry Hub natural-gas price, $x_{2t}$, is highly correlated ($R = 0.99$) with the Texas natural-gas price at Houston Ship Channel. Since most of the thermal generation is outside the West zone, we would expect a rising Henry Hub price to magnify the marginal generation cost and hence the probability of a positive paired-price difference, thus implying a coefficient of $\beta_2 > 0$.

Nuclear generation in the ERCOT system, $x_{3t}$, is exogenous, reflecting the fact that variations in nuclear generation are unrelated to wind generation. Since nuclear generation is outside the West zone, rising nuclear generation limits the West zone's wind exports, suppresses prices in the West zone, and therefore raises the probability of a positive paired-price difference, which implies a coefficient of $\beta_3 > 0$.

We contend that the MWh load in ERCOT’s Houston zone, $x_{4t}$, is exogenous, and that it is largely determined by the weather and time of use (e.g., the hour, the day, and any holiday and seasonal factors). Our contention is corroborated by ERCOT’s zonal-market price determination that assumes a price elasticity of zero (ERCOT, 2004). Because the non-West zones in ERCOT are inter-connected, we do not know a priori the direction in which a rising load in Houston might move the probability of a positive price, and hence we do not speculate as to the sign of $\beta_4$. 
We expect $x_{5t}$ and $x_{6t}$, the 15-minute MWh loads in ERCOT’s North and South zone, respectively, to have marginal effects on the probability of a positive price, but hesitate to speculate on their directions and hence the signs of $\beta_5$ and $\beta_6$.

This is not, however, the case with the MWh load in ERCOT’s West zone in which a rising load reduces wind-generation exports. Hence, we hypothesize $x_{7t}$ to have a negative marginal effect on the probability of a positive price, which implies a coefficient of $\beta_7 < 0$.

4.3 Stage 2: Log-linear regressions for the size of the paired-price difference

The second stage begins with an OLS log-linear model for the size of the paired-price differences, either strictly positive or negative, to be explained by the now familiar sets of dummy and metric variables included in the ordered-logit model as independent variables. The parameter estimates attached to the seven metric variables are then directly interpretable as paired-price-difference elasticities. In the following discussion we focus on the positive paired-price-difference regression because it analogously applies to its negative paired-price counterpart.

Following Lee (1983, p.511), the log-linear form of the regression may be written as follows:

\[
\ln( y_t | y_t > 0) = \theta_0 + \sum_i \delta_i M_{it} + \sum_j \lambda_j W_{jt} + \sum_k \varphi_k H_{kt} + \sum_r \theta_r x_{rt} + \gamma c_t + \epsilon_t. \tag{2}
\]

In equation (2), $c_t$ is a term to correct any OLS bias due to sample selection, and $\epsilon_t$ is a time-independent, normally-distributed, heteroskedastic random error with zero mean and finite variance (Lee, 1983, p.509). We use OLS to consistently estimate the intercept $\theta_0$ and coefficients attached to the dummy and metric variables.
4.3.1 Hypotheses

The coefficient $\theta_r$ for $r = 1, \ldots, 7$, is the elasticity of $y_t > 0$ with respect to variable $x_{rt}$. In accordance with the discussion of the stage-1 estimation and our hypotheses, we now hypothesize that rising wind generation magnifies the size of the positive paired-price difference, which implies $\theta_1 > 0$. The same is true of the Henry Hub price, which translates into $\theta_2 > 0$ and nuclear generation, which translates into $\theta_3 > 0$. How rising non-West zonal loads may affect the size and signs of their related coefficients of $\theta_4$, $\theta_5$, and $\theta_6$ is a question to be answered empirically. Nonetheless, we do expect a rising West load to reduce a positive paired-price difference due to declining wind exports to other zones, which translates into $\theta_7 < 0$.

The coefficient $\gamma$ helps determine if a sample selection of observations with $y_t > 0$ matters when using OLS to estimate equation (2). If $\gamma = 0$, then ignoring sample selection by excluding $c_t$ as an explanatory variable does not bias the OLS coefficient estimates. If, however, $\gamma < 0$, an unobserved factor that increases the likelihood of a positive paired-price difference enlarges the size of that difference.$^{10}$

4.4 Results

4.4.1 The probability of a strictly positive paired-price difference

Table 1 reports the ordered-logit regression results for the three zonal market pairs of interest: North-West, South-West, and Houston-West. Considering the large size of our sample, our regressions have a relatively good fit, with a pseudo-$R^2$ of 0.15 to 0.26
indicating a log-likelihood increase of 15% to 26% attributable to the explanatory variables other than the intercepts.\textsuperscript{11}

While too numerous to be included in Table 1, the coefficient estimates for the binary indicators confirm that the odds of a positive paired-price difference are time-dependent ($p < 0.01$), in regard to month, day, and hour, even though not all coefficients are statistically significant. The estimates of $\beta_1$, $\beta_2$, and $\beta_3$, which for the North-West pairing are $b_1 = 0.0030$, $b_2 = 0.0295$, and $b_3 = 0.0006$, suggest that in the first case rising wind generation increases the estimated odds of a paired-price difference, as do increases in nuclear generation and the natural-gas price in the second two cases, all of which accord with our hypotheses. The analogously-interpreted estimates of $b_4 = 0.0005$, $b_5 = 0.0009$, and $b_6 = -0.00019$ indicate that higher zonal loads in the Houston and North zones increase the estimated odds of a positive paired-price difference between the North and West zones, whereas higher zonal loads in the South zone reduce those odds. Finally, the coefficient estimate of $b_7 = -0.0007$, along with its associated $p$-values, suggests that rising West load has a statistically-significant negative impact on the estimated odds of a positive paired-price difference between the North and the West. For the other two price pairings, however, $b_7 = 0.00001$ and is statistically insignificant. Thus, our hypothesis on the effect of West load is supported for the North-West pairing, but the data are inconclusive regarding the other two pairings.

Finally, although the proportional-odds assumption of ordered logit, which forces fixed slope coefficients across the ordered strata is rejected ($p = 0.001$), from a practical standpoint this is not necessarily a problem (Kim, 2003). In point of fact, when we estimated a generalized logit model that allows for different slope-parameter estimates
across the different strata (Williams, 2006), the resulting estimates, even when they differ in value, nevertheless lead to the same inferences with regard to the impacts of the variables on the probability of a positive or a negative paired-price difference.

Table 1: Ordered-logit regressions estimates for a strictly positive price difference ($y_t > 0$) and a positive price difference ($y_t \geq 0$); $p$-values in ( ); odds ratios in [ ]. This table does not report the numerous estimates for the binary indicators, which are available from the corresponding author.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Zonal market pair</th>
<th>North-West</th>
<th>South-West</th>
<th>Houston-West</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of observations with $y_t &gt; 0$</td>
<td></td>
<td>12,394</td>
<td>15,636</td>
<td>16,087</td>
</tr>
<tr>
<td>Number of observations with $y_t = 0$</td>
<td></td>
<td>99,127</td>
<td>99,106</td>
<td>99,109</td>
</tr>
<tr>
<td>Number of observations with $y_t &lt; 0$</td>
<td></td>
<td>4,823</td>
<td>1,602</td>
<td>1,148</td>
</tr>
<tr>
<td>Pseudo $R^2$ (= percent increase in the log likelihood due to the covariates)</td>
<td></td>
<td>0.154</td>
<td>0.259</td>
<td>0.257</td>
</tr>
<tr>
<td>$\alpha_1$: Intercept for $y_t &gt; 0$</td>
<td></td>
<td>-4.761 ( &lt;0.001)</td>
<td>-10.138 ( &lt;0.001)</td>
<td>-10.258 ( &lt;0.001)</td>
</tr>
<tr>
<td>$\alpha_2$: Intercept for $y_t \geq 0$</td>
<td></td>
<td>1.539 ( &lt;0.001)</td>
<td>-2.244 ( &lt;0.001)</td>
<td>-2.105 ( &lt;0.001)</td>
</tr>
<tr>
<td>$\beta_1$: Coefficient for $x_{1t} = 15$-minute wind generation of the ERCOT system.</td>
<td></td>
<td>0.0030 ( &lt;0.001) [1.003]</td>
<td>0.0041 ( &lt;0.001) [1.004]</td>
<td>0.0039 ( &lt;0.001) [1.004]</td>
</tr>
<tr>
<td>$\beta_2$: Coefficient for $x_{2t}$ = the daily Henry Hub natural-gas price.</td>
<td></td>
<td>0.0295 ( &lt;0.001) [1.030]</td>
<td>0.2961 ( &lt;0.001) [1.345]</td>
<td>0.2710 ( &lt;0.001) [1.311]</td>
</tr>
<tr>
<td>$\beta_3$: Coefficient for $x_{3t} = 15$-minute nuclear generation of the ERCOT system.</td>
<td></td>
<td>0.0006 ( &lt;0.001) [1.001]</td>
<td>0.0010 ( &lt;0.001) [1.001]</td>
<td>0.0010 ( &lt;0.001) [1.001]</td>
</tr>
<tr>
<td>$\beta_4$: Coefficient for $x_{4t} = 15$-minute MWh loads in ERCOT’s Houston zone.</td>
<td></td>
<td>0.0005 ( &lt;0.001) [1.001]</td>
<td>0.0010 ( &lt;0.001) [1.001]</td>
<td>0.0017 ( &lt;0.001) [1.002]</td>
</tr>
<tr>
<td>$\beta_5$: Coefficient for $x_{5t} = 15$-minute MWh load in ERCOT’s North zone.</td>
<td></td>
<td>0.0009 ( &lt;0.001) [1.001]</td>
<td>-0.0011 ( &lt;0.001) [0.999]</td>
<td>-0.0009 ( &lt;0.001) [0.999]</td>
</tr>
<tr>
<td>$\beta_6$: Coefficient for $x_{6t} = 15$-minute MWh load in ERCOT’s South zone.</td>
<td></td>
<td>-0.0019 ( &lt;0.001) [0.998]</td>
<td>0.0015 ( &lt;0.001) [1.001]</td>
<td>0.0007 ( &lt;0.001) [1.001]</td>
</tr>
</tbody>
</table>
4.4.2 Size of the strictly positive price differences

Part A of Table 2 reports the results for the OLS log-linear regressions for strictly positive paired-price differences for each of three zonal-market pairs. The $F$-statistics are all statistically significant ($p \leq 0.001$), and the regressions have eminently satisfactory adjusted $R^2$ values that range between 0.15 and 0.35 for a large sample of 12,000 to 16,000 15-minute observations.

Since the error term is heteroskedastic, the $p$-values for the coefficient estimates in Table 2 are based on the regression's heteroskedasticity-consistent covariance matrix (White, 1980). The coefficient estimates for the binary indicators, which in the interests of parsimony are omitted from the table, confirm the generally, although not uniformly, statistically-significant ($p \leq 0.001$) time-dependence of the paired-price differences for all zonal pairings. This time-dependence, however, does not have a clear and systematic pattern across the three market pairs.

The coefficient estimates in Table 2 indicate that rising wind generation tends to increase the positive paired-price differences, with estimated coefficients ranging from 0.210 to 0.907. In particular, the 0.907 elasticity estimate for the North-West market pair suggests that a 1% increase in wind generation can lead to an almost 1% increase in the pair’s positive price difference when the North-West transmission congests.

The estimated coefficients for nuclear generation and the natural-gas price are positive and statistically-significant estimates ($p < 0.001$). They support our hypotheses
that rising nuclear generation and a rising natural-gas price tend to magnify a market-pair’s positive price difference, albeit with relatively inelastic effects.

A market pair’s positive difference depends on non-West loads. Moreover, a rising West load tends to reduce the difference for the South-West and Houston-West pair, but it is statistically insignificant for the North-West pair.

Finally, the estimate for $\gamma$ is negative and statistically significant ($p < 0.001$) for all market pairings, indicating that if an unobserved factor increases the likelihood of a strictly positive paired-price difference, it also enlarges the size of that difference.

Table 2: OLS log-linear regressions for the size of a paired-price difference ($y_t \neq 0$); $p$-values in ( ). This table does not report the numerous estimates for the binary indicators, which are available from the corresponding author.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Part A</th>
<th>Part B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of observations</td>
<td>12,391</td>
<td>15,636</td>
</tr>
<tr>
<td>Mean of $\ln$ size of $y_t$</td>
<td>3.278</td>
<td>3.473</td>
</tr>
<tr>
<td>Root-mean-squared-error</td>
<td>1.381</td>
<td>1.138</td>
</tr>
<tr>
<td>$F$-statistic</td>
<td>138.91 (&lt;0.001)</td>
<td>70.10 (&lt;0.001)</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.348</td>
<td>0.175</td>
</tr>
<tr>
<td>$\theta_0$: Intercept</td>
<td>-14.182 (&lt;0.001)</td>
<td>-13.061 (&lt;0.001)</td>
</tr>
<tr>
<td>$\theta_1$: Coefficient for $\ln x_{1t} = $ natural log of 15-minute wind generation of the ERCOT system.</td>
<td>0.907 (&lt;0.001)</td>
<td>0.230 (&lt;0.001)</td>
</tr>
<tr>
<td>$\theta_2$: Coefficient for $\ln x_{3t} = $ natural log of the daily Henry Hub natural-gas price.</td>
<td>0.976 (&lt;0.001)</td>
<td>0.651 (&lt;0.001)</td>
</tr>
<tr>
<td>$\theta_3$: Coefficient for $\ln x_{4t} = $ natural log of 15-minute nuclear generation of the ERCOT system.</td>
<td>0.661 (&lt;0.001)</td>
<td>0.316 (&lt;0.001)</td>
</tr>
<tr>
<td>$\theta_4$: Coefficient for $\ln x_{5t} = $ natural log of</td>
<td>-0.727 (0.012)</td>
<td>1.005 (&lt;0.001)</td>
</tr>
</tbody>
</table>
15-minute MWh loads in ERCOT’s Houston zone.

<table>
<thead>
<tr>
<th>( \theta_1 ): Coefficient for ( \ln x_1 )</th>
<th>( x_1 = \text{natural log of 15-minute MWh load in ERCOT’s North zone.} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.578 (0.012)</td>
<td>2.175 (&lt;0.001)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \theta_2 ): Coefficient for ( \ln x_2 )</th>
<th>( x_2 = \text{natural log of 15-minute MWh load in ERCOT’s North zone.} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.884 (0.008)</td>
<td>-1.187 (&lt;0.001)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \theta_3 ): Coefficient for ( \ln x_3 )</th>
<th>( x_3 = \text{natural log of 15-minute MWh load in ERCOT’s South zone.} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.061 (0.791)</td>
<td>-0.605 (&lt;0.001)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \gamma ): Coefficient for ( c_t )</th>
<th>( c_t = \text{term to correct OLS bias due to sample selection} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.411 (&lt;0.001)</td>
<td>-0.220 (&lt;0.001)</td>
</tr>
</tbody>
</table>

### 4.4.2 Size of the strictly negative paired-price differences

Part B of Table 2 reports the OLS results for the natural-log of the size of a strictly negative paired-price difference for each of three zonal market pairs, and the statistically-significant \( p < 0.001 \) results have adjusted \( R^2 \) values in a similar range of 0.20 to 0.37.

Based on the regression's heteroskedasticity-consistent covariance matrix, the \( p \)-values in Table 3 for the binary indicators' coefficient estimates, once again omitted from the table, indicate that the negative paired-price difference's time-dependence is generally statistically significant \( p < 0.01 \). This time-dependence, however, does not have a clear pattern of the month-of-year, day-of-week and hour-of-day effects.

Consistent with our earlier hypothesis, if an increase in a metric variable like wind generation tends to magnify the size of a positive price difference, it likely shrinks the
size of a negative price difference. Based on this line of reasoning, we expect the coefficient estimates for \((x_{1t}, \ldots, x_{7t})\) in Table 3 to have opposite signs as those in Table 2.

The coefficient estimates in Table 3 indicate that rising wind generation reduces the size of a strictly negative paired-price difference for the North-West pair. Its effect for the other pairs is positive, but statistically insignificant. The table also shows the statistically-significant result \((p < 0.01)\) that a rising natural-gas price tends to reduce the sizes of strictly negative paired-price differences for the North-West pair; but its negative effect on the other price pairs is statistically insignificant. Rising nuclear generation tends to reduce size of a negative price difference for the North-West but not the other two market pairs, even though these nuclear generation’s effects are not statistically significant. Moreover, the size of the negative price difference depends on non-West loads. A rising load in the West zone, however, tends to reduce the size of a negative price difference, because it reduces wind-generation exports from the West. Again, these results lend general support for our hypotheses. Finally, the negative coefficient estimate for \(\gamma\) indicates that if an unobserved factor increases the likelihood of a negative paired-price difference, it also tends to enlarge the size of that difference. This effect, however, is statistically significant only for the North-West and the Houston-West pairs, but not the South-West pair.

4.4.3 Remarks

We would be remiss if we failed to remark that the variance inflation factors (VIF) for the binary indicators in the log-linear regressions are mostly around 2, reflecting their orthogonal nature. The VIFs for wind generation, the Henry Hub natural-gas price, nuclear generation, and the OLS bias-correction term are between 2 to 5, which
are also satisfactory. Those for the zonal loads, however, are between 8 and 27, due largely to the correlated weather-dependence of zonal loads in Texas. While these VIF levels suggest variance-inflating multicollinearity, the coefficient estimates for the load variables in Tables 2 and 3 are mostly significant ($p < 0.01$) and pass the test of plausibility, which diminishes the likelihood that we have failed to detect an important and statistically-significant relationship, or incorrectly ascribed statistical significance to a relationship when one is unmerited. Hence, we decided not to modify the log-linear regression's specification by deleting one or more of the zonal-load variables.

Based on Figure 9, we recognized that the empirical distribution of the natural logarithms of the sizes of the paired-price differences is skewed and can have outliers. Hence, we re-estimated the log-linear equations using a robust-regression approach (Huber, 1973). The resulting coefficient estimates for the metric variables do not qualitatively alter our findings regarding the dependence of the paired-price differences on wind generation, nuclear generation, the Henry Hub natural-gas price, and zonal loads. Hence, our results are quite robust across alternative estimation approaches.

5. **Conclusions**

There is a vast and growing literature on the physical impacts of increasing wind generation upon the reliability and operation of power grids. Environmental concerns have further stimulated interest in wind generation as an environmentally-friendly alternative energy source. Despite these related and mutually-supportive interests, far too little attention has been paid to the potential price effects of this important but intermittent resource in a competitive electricity market. The research that we have
conducted on a huge and unique database accumulated in the largest energy-consuming state in the United States - Texas - and the results of that research described herein, represent our contribution to this literature and our effort to focus attention on a critical aspect of the renewable-energy debate.

To be sure, our most salient finding - namely, that high wind generation and low load in the wind-rich West ERCOT zone tend to lead to congestion and zonal price differences during any given time period - is based strictly on what has transpired within a single geographic region. Nonetheless, our results provide traction for our belief that they can be readily generalized, inasmuch as the results are not simply plausible, but also support our hypotheses and prior expectations, where such have been framed and set out.

In addition to wind generation, as one would expect, other factors contribute to divergences in zonal prices at any given time, even if the directions of their impacts may not be “clear” a priori. These factors include movements in nuclear generation and natural-gas prices, as well as fluctuating non-West loads, and the “timing effects” of month, day, and hour. When positive paired-price differences occur up to 14% of the time and can be as great as $3,500/MWh, they clearly signal the need for further investment in network infrastructure. Happily, this is what ERCOT is doing now with its ambitious transmission expansion plans.

The Texas experience has broader implications that extend throughout the United States and well beyond its borders. Given the emergence of China and India as the leading sources of growth in energy consumption with unabated appetites in the near future, as well as the specter of climate change and unwelcome carbon emissions looming large and ominously, it is incumbent upon policy makers the world over to direct
attention to alternative and climate-neutral sources of energy, such as wind generation, in particular. In doing so, it is also incumbent upon them to be aware of the wide array of consequences, both good and bad, that taking advantage of those alluring alternative generation options can have for competitive electricity markets. This paper is intended to point the way, and indeed to lead the way in that regard.
References


ERCOT, 2010b. ERCOT Protocols - Section 7: Congestion Management. Electric Reliability Council of Texas, Austin, Texas. Available at: http://www.ercot.com/content/mkrules/protocols/current/07-070110.doc


Kirby, B., Milligan, M., 2008. Facilitating Wind Development: The Importance of Electricity Industry Structure. Technical Report, NREL. Available at:


Millborrow, D., 2009. Managing Variability: A report to WWF-UK, RSPB. Greenpeace UK and Friends of the Earth EJNI. Available at:


Footnotes

1 On-going research in connection to wind generation includes: (a) renewable-energy benefits (Mosey and Vimmerstedt, 2009); (b) renewable-energy policy (Resch et al., 2005; Carlioz and Naseem, 2007; Haas et al., 2008; Schmalensee, 2009; Pollitt, 2010); (c) wind-energy integration (Buckley et al., 2005; CEC, 2007; Millborrow, 2009; Parsons et al, 2009; PNNL, 2010); (d) market design (Mitchell et al., 2006; IPA, 2008; Kirby and Milligan, 2008; Sioshansi and Hurlbut, 2010; Newberry, 2010; Vandezande et al., 2010); (e) transmission planning and cost recovery (Hiroux, 2005; Orans et al., 2007; Barroso et al., 2007; Scott, 2007; Decker, 2008; Fulli et al., 2009; Mills et al., 2009; Olson et al., 2009; Fraias et al., 2010; Mills et al., 2010; OPA, 2010); (f) marginal cost of renewable energy (Mahone et al., 2009); and (g) renewable-energy contract design (Johnston et al., 2008).

2 Well-known factors contributing to price volatility include: (a) daily fuel-cost variations, especially for the natural gas that is now widely used in combined-cycle gas turbines; (b) weather-dependent demands with intra-day and inter-day fluctuations that must be met in real time by generation and transmission already in place; (c) changes in available capacity caused by planned and forced outages of electrical facilities; and (d) lumpy capacity additions that can only occur with a long lead time (Li and Flynn, 2006; Tishler et al., 2008). Exacerbating the spot-price volatility are poor market designs and market-power abuse by generators (Borenstein, 2002; Borenstein et al., 2002; Woo et al., 2003; Trebilcock and Hrab, 2005).

3 Some examples of this research are Johnsen (2001), Bessembinder and Lemmon (2002), Goto and Karolyi (2004), Longstaff and Wang (2004), Knittel and Roberts (2005),
Haldrup and Nielsen (2006), Mount et al. (2006), Park et al. (2006), Guthrie and Videbeck (2007), Woo et al. (2007), Benth and Koekebakker (2008), Karakatsani and Bunn (2008), Marckhoff and Wimschulte (2009) and Redl et al. (2009). Useful applications of the research results include: risk management (Woo et al., 2004a, 2004b, 2006b; Kleindorfer and Li, 2005; Deng and Oren, 2006; Huisman et al., 2009); pricing of electricity options, futures, forwards, and generation assets (Deng et al., 2001; Woo et al., 2001; Kamat and Oren, 2002; Lucia and Schwartz, 2002; Eydeland and Wolyniec, 2003; Fleten and Lemming, 2003; Keppo and Lu, 2003); forward contracting of locational price spreads (Woo et al., 1998; Marckhoff and Wimschulte, 2009; ERCOT, 2010b); detection of market-power abuse and price manipulation (Borenstein et al., 2002; Joskow and Khan, 2002; Helman, 2006); investigation of generation-investment behavior (Neuhoff and De Vries, 2004); assessment of wholesale-market integration (De Vany and Walls, 1996; Woo et al., 1997; Park et al., 2006); and assessment of how retail competition may affect forward-contract pricing (Green, 2003).

4 Price and load data are available on ERCOT’s public website. We obtained 15-minute generation data by generator type directly from ERCOT.

5 In some years, a Northeast zone was carved out of the North zone. Beginning December 1, 2010, ERCOT will have a nodal wholesale-market structure.

6 We do not use the 15-minute data on hydro, coal, and natural-gas generation to explain the probability of a positive paired-price difference, because as a result of ERCOT’s least-cost dispatch they are endogenous (ERCOT, 2004).
The average own-price elasticity for the aggregated block of all energy consumers in ERCOT with interval data recorders is very small, about -0.000008 (Zarnikau and Hallett, 2008).

In addition to these seven variables, we included the 15-minute dynamic ratings of inter-zonal transmission in our initial set of explanatory variables. Our preliminary regression results indicated, however, that a rising load in the West zone has a statistically significant ($p \leq 0.01$) and positive impact on the probability of a strictly positive paired-price difference, a counter-intuitive result. Moreover, these ratings are endogenous (HEC, 2009, p.iii). Hence, we exclude the dynamic ratings from our final set of explanatory variables.

Based on Lee (1983, p.508) and ignoring the subscript $t$ for notational simplicity, $c = \phi(z) / p_1$, where $\phi(z)$ is a density function for a standard normal variate $z = \Phi^{-1}(p_1) =$ transformed value of $p_1$, and $\Phi(z)$ is a normal density function (e.g., $z = -1.65$ for $p_1 = 0.05$).

Equation (3.7) of Lee (1983, p.511) shows $\gamma = -\sigma \rho$, where $\sigma$ is the standard deviation of the error for the size regression, and $\rho$ is the correlation between the errors of the probability and size regressions. As $\sigma > 0$, $\gamma < 0$ implies $\rho > 0$.

This pseudo-$R^2$ measure is the most conservative of such measures of the goodness of fit for logit regressions (Hu et al., 2006).