Wind intermittency and supply-demand imbalance: Evidence from U.S. regional power markets*

Victoria Godwin Matthew E. Oliver

Georgia Institute of Technology
School of Economics

December 2023

Abstract

Wind is a prominent source of carbon-free electric power despite being highly variable due to random changes in wind speeds. This variability is problematic because total electricity supply must match demand at all times with little margin for error. Imbalances are costly, and system operators must respond to them instantaneously. We investigate the relationship between wind intermittency and supply-demand imbalances in electricity systems, using data from major regional power markets in the United States. Results show greater variation in wind generation leads to a robust, highly statistically significant increase in electricity system imbalance. We discuss the implications for system operators and renewable energy policies.

Keywords: wind generation, intermittency, area control error, renewable energy.

JEL Codes: L94, Q41, Q42, Q54

*We thank the following for many helpful comments and suggestions: Dylan Brewer, Daniel Dench, Gaurav Doshi, Ben Gilbert, Paul Turner, and Matt Woerman. This research was presented in a brown-bag seminar at the Georgia Tech School of Economics (Oct. 2023), at the Southeastern Workshop in Energy and Environmental Economics & Policy (Georgia Tech, Oct. 2023), and at the 93rd Annual Meeting of the Southern Economic Association (New Orleans, Nov. 2023). We are grateful to the participants of those events for their constructive feedback. Denis Schwirtz and Emily O'Day provided research assistance related to data collection. Finally, we thank the Georgia Tech Ivan Allen College of Liberal Arts for minor financial support.

1 Introduction

By 2030, wind turbines will overtake hydroelectric power stations as the largest source of renewable electricity generation globally (Energy Information Administration, 2021). In the United States, wind turbines supplied over 375 GWh of electrical energy in 2021, or roughly 9% of total generation, and by 2050 are projected to supply over 730 GWh annually, or 14% of total generation (Energy Information Administration, 2022).

Apart from being a carbon-free source of electricity, the two distinguishing characteristics of wind power are (i) that it can be supplied at zero (or near zero) marginal operating cost, and (ii) that it is intermittent due to natural variations in wind speeds and direction (Joskow, 2019). Short-term fluctuations in wind generation can be significant because the power produced by a wind turbine is proportional to the cube of the wind speed (Blume, 2017).[1] This intermittency has a variety of important implications for electricity markets and system operations, including increased variability of wholesale prices (Woo et al., 2011; Johnson and Oliver, 2019), episodes of negative prices (Genoese et al., 2010; Seel et al., 2021), increased need for costly stand-by capacity (Skea et al., 2008; Brouwer et al., 2014; Liebensteiner and Wrienz, 2020), and greater reliance on costly ancillary services such as balancing, load following, voltage control, and frequency control (Østergaard, 2006; Vandezande et al., 2010; Katzenstein and Apt, 2012; Kiviluoma et al., 2012; Singarao and Rao, 2016; Godoy-González et al., 2020). Recent empirical work has shown that greater intermittency negatively impacts net economic welfare by increasing operational costs, diminishing both wind generator profits and consumer benefits (Petersen et al., 2022).

The market effects of wind intermittency stem mainly from an inability to store electricity at a sufficiently large scale. Simply put, the absence of storage requires electricity generation (supply) to perfectly balance load (demand) at all times. To maintain system balance, system operators must respond to fluctuations in net load—i.e., load net of wind generation—in real-time.[2] Accurately forecasting wind generation is challenging, and while methods are improving, perfectly anticipating sudden changes in wind speeds is effectively impossible (DeCesaro et al., 2009; Notton et al., 2018; Yang et al., 2021). As a result, increasing penetration of large-scale wind generation is expected to result in larger and more frequent system imbalances.[3] Even minor imbalances can be potentially costly, in part because they

[1]To illustrate, doubling wind speed increases the power generated by a factor of eight.

[2]Because (i) both load and wind generation are inherently variable, and (ii) wind is non-dispatchable, system operators treat wind generation as 'negative load' (Durrwachter and Looney, 2012; Ela and Edelson, 2012).

[3]Multiple studies have proposed a variety of market mechanisms designed to mitigate wind energy imbalances at the generator level, defined as differences between contracted energy and produced energy (Bourry et al., 2008). A full review is not possible here. See, for example, Wan et al. (2007); King et al. (2011);

cause system frequency to deviate from its target (60 Hz in U.S. systems), which can damage sensitive electrical equipment connected to the grid. Large imbalances can lead to rolling blackouts when demand exceeds supply,[4] or excess energy being lost as waste heat (and potential damage to the transmission grid) when supply exceeds demand. Despite this, surprisingly few studies have documented the empirical relationship between wind energy intermittency and electricity system imbalances.

This paper fills this gap in the literature by empirically estimating the relationship between wind intermittency and a widely used measure of system imbalance known as 'area control error' (ACE), using data from four major electricity system operators in the United States. ACE is a measure of the total, system-wide energy imbalance at any given moment in time, taking into account the effects of such imbalances on system frequency.[5] As explained by Stoft (2002), "ACE is the main indicator of the supply-demand balance in every control area in the United States and when there is a market, it is the signal that determines whether the price will be increased or decreased by the system operator."

We estimate the effect of wind intermittency on ACE using an instrumental variables approach, due to the likely endogeneity of our primary measure of wind intermittency: the hourly variance in (system-wide) wind generation. Results of our preferred model indicate that a doubling of the hourly variance in wind generation leads to an increase in hourly average system imbalance—i.e., the hourly average absolute value of the difference between generation and load—of between 2% and 12%, depending on the system operator. In energy units, relative to sample means for the system operators to which each corresponds, these point estimates translate roughly to 1.2 MW and 9.8 MW increases in hourly average system imbalance, respectively. To put this into perspective, 1 MW of generating capacity is enough to power 400 to 900 U.S. homes for one year.[6] Although this may not seem like an especially large effect relative to system-wide levels of generation and load, it is important to keep in mind that from the system operator's perspective, *any* increase in ACE is undesirable. Customers value power quality, which depends crucially on system balance (Bollen, 2003; Heydt, 2005). The goal is therefore to minimize ACE, targeting an error of zero at all times.

We supplement our main findings by investigating the effects of directional changes in wind generation on the ACE.[7] First, using a logit model, we show that a sudden increase in wind generation increases the probability that ACE will be positive (that is, total supply

Zugno et al. (2013).

[4]See Borenstein et al. (2023) for a detailed review of the economic costs of power system outages.

[5]Additional details on how ACE is calculated are provided in Section 2.

[6]`https://www.nrc.gov/docs/ML1209/ML120960701.pdf`

[7]By "directional changes" we mean *increases* and *decreases* in wind generation, not changes in geographic wind direction (e.g., east versus southeast).

exceeds demand) over the same interval, whereas ACE is more likely to be negative (total supply falls short of demand) for a sudden drop in wind generation. Second, in terms of magnitudes, we find evidence that ACE responds slightly differently to sudden increases in wind generation than to sudden decreases, suggesting differences in ISO response strategies and/or technical constraints associated with ramping other generators up or down as needed.

The most closely related studies to our own are from the electrical engineering literature and employ numerical simulation methods to compute estimates of the effects of wind intermittency on ACE (and other system stability indicators) under differing assumptions about system characteristics. Sortomme et al. (2010) estimate the impact on ACE of integrating new wind generation capacity using a time-series modeling technique, with the goal of providing system operators with a method to forecast the impacts of increased wind generation on system imbalance and stability. Nguyen and Mitra (2015, 2016) model the effects of wind generation on system frequency deviations and ACE, providing guidance on how much wind can be added to a system while still maintaining frequency deviations below a given limit. The key distinction between these studies and ours is that each of these engineering analyses computes *ex ante* estimates of the effect of wind intermittency on ACE in a hypothetical power system. By contrast, we provide an *ex post* empirical estimate of this relationship by applying straightforward econometric methods to historical data on wind generation and ACE from major interconnections in the U.S. To our knowledge, ours is the first paper in the economics or policy literature on intermittent renewable generation to utilize ACE data to proxy for supply-demand imbalance in electricity markets.

The implications of our findings for electricity systems operations are highly relevant given the urgency of the climate crisis, the need to rapidly de-carbonize the electricity sector, and the prominent role that wind generation already plays in the clean energy transition. In many regions, wind has already achieved cost parity with traditional, thermal generation technologies and will continue to gain ground as costs are projected to decline yet further. As but one example of the current pace of growth, according to a 2022 report by the U.S. Department of Energy, the U.S. now has over 40,000 MW of offshore wind generation capacity in various stages of development, thanks in part to a recent expansion by the Biden-Harris Administration of the areas open to offshore wind development.[8] Yet, while offshore wind tends to be, on average, less variable than onshore, such a large infusion of intermittent generation could lead to increased stress on the electricity grid in terms of managing and responding to larger system imbalances, particularly given that the development of large-scale electricity storage technologies continues to lag considerably behind the pace of investment

[8] `https://www.energy.gov/sites/default/files/2022-09/offshore-wind-market-report-2022-v2.pdf`

in wind (and utility-scale solar).

A secondary but still important takeaway from our study is that ACE as a measure of system imbalance has thus far been almost entirely overlooked and under-utilized in economic and policy research on electricity markets. One might imagine any number of useful insights that could be gained from analyzing how the ACE responds to different policy interventions or other market perturbations. More research is also needed to understand how to translate ACE into a measure of the social cost of system imbalance. We hope our findings spur other researchers to think of creative ways that the ACE might be used to study the economics of electricity markets, the renewable energy transition, and related topics.

The remainder of the paper is organized as follows. In Section 2 we provide a detailed description of ACE as a measure of system imbalance. Section 3 describes our data and presents descriptive evidence of the effect of wind intermittency on ACE. Our econometric strategy is described in Section 4. We present results and provide further discussion in Section 5. Section 6 concludes.

2 ACE: A Measure of System Imbalance

The economics of electricity markets cannot be decoupled from the physical laws of electricity generation and transmission or the engineering constraints of the electric power system itself. The most salient engineering constraint, when it comes to increasing penetration of intermittent renewable generation like wind (or solar), is that electricity cannot be stored at a large enough scale to smooth the variability of supply from these sources or to allow system operators to respond to fluctuations in demand by dispatching stored wind energy even when turbines are not producing it. As a counter-example, consider natural gas, which is similarly delivered on a vast transmission network, but is easily stored. Gas storage allows suppliers to smooth production over time, despite large seasonal variations in consumption. Gas supply (at the wellhead) and demand do not need to balance in aggregate at any given point of time, because storage inventories can be built up when supply exceeds demand and drawn down when demand exceeds supply. Storage helps stabilize market prices despite this temporal mismatch between production and consumption.[9] By contrast, in electricity systems the lack of large-scale storage requires supply and demand to balance at all times. Consequently, electricity prices can spike to astronomical heights during peak demand periods, because all power must be supplied instantaneously, whatever the marginal cost (or accept rolling blackouts if the marginal cost exceeds the value of lost load), rather than drawn from storage

[9]Similarly, gas storage mitigates the price effects of transmission capacity constraints by enabling intertemporal substitution of transmission services (Oliver et al., 2014).

inventories at relatively low cost.

Yet, even this is an oversimplification. Stoft (2002) makes the crucial distinction that electricity supply equals *consumption* at every point in time but may not equal *demand*, where electricity demand is defined as the amount "that would be consumed if system frequency and voltage were equal to their target values for all consumers." Holding voltage constant, any imbalance between supply and demand causes frequency to deviate from its target value.[10] Standard operating frequency in U.S. systems is 60 Hz, and generators are limited to a narrow bandwidth around this target (Blume, 2017). If supply exceeds demand, electrical devices connected to the grid must absorb the excess energy, and frequency increases to compensate. Similarly, if demand exceeds supply, electrical devices attempt to draw more power in aggregate than exists in the system, and frequency slows down in response. These frequency deviations occur *automatically* as a result of the physical laws governing electricity systems, rather than as a deliberate choice of the system operator. Moreover, because electrical devices are finely tuned to operate at precisely 60 Hz, even moderate deviations from this target can be damaging to sensitive loads (e.g., computer-based systems), resulting in potentially significant economic losses (Ahmed et al., 2020).

In addition to frequency deviations, a supply-demand imbalance affects net power flows into or out of a balancing authority's (i.e., system operator's) control region. Excess outflow—that is, relative to *scheduled* net outflow—is likely to occur when supply exceeds demand, whereas excess inflow is likely when demand exceeds supply.

These two signals—excess inflows/outflows and frequency deviations—are combined to form a single indicator of system imbalance, known as the *area control error*, or ACE. The North American Electric Reliability Corporation (NERC) defines ACE using the following formula:

$$ACE = (NI_a - NI_s) - 10B(F_a - F_s) - \epsilon, \tag{1}$$

where NI_a is net interchange (actual), NI_s is net interchange (scheduled), B is balancing authority frequency bias, F_a is frequency (actual), F_s is frequency (scheduled), and ϵ is interchange metering error (NERC, 2011). ACE is computed in MW and does not have a time dimension. It changes continuously and can be thought of as a snapshot of the system imbalance at any given moment.[11]

The first term, $(NI_a - NI_s)$, represents the error—i.e., system imbalance—associated with deviations from scheduled power deliveries, without consideration of any effects on

[10] Frequency, measured in Hertz (Hz), refers to the number of times per second the current switches direction in an alternating current (AC) circuit.

[11] ACE data are commonly reported at sub-hourly intervals—e.g., 1-minute, 5-minute, or 10-minute.

system frequency. To illustrate, say a balancing authority (BA) is scheduled to purchase (import) 300 MW of power from adjacent jurisdictions with which it shares interconnecting 'tie-lines'. By definition, flow into (out of) a BA's jurisdiction is a negative (positive) value, thus $NI_s = -300$. Now, suppose wind generation falls short of its forecast, causing actual flow into the BA to be $NI_a = -305$ MW; one might say that (part of) the missing power is 'sucked into' the system from the tie-lines to adjacent systems. The system imbalance due to deviation from scheduled power deliveries is therefore $[(-305) - (-300)] = -5$ MW.

The second term, $-10B(F_a - F_s)$, corresponds to the BA's obligation to maintain system frequency. NERC defines a BA's frequency bias (B) as a value that approximates the BA's response to frequency error. This value is negative and is defined in terms of MW/0.1Hz; thus, the scalar 10 converts frequency bias to MW/Hz. In our illustrative example, say $B = -50$ and the sudden loss of wind generation reduces system frequency to 59.99 Hz. The additional system imbalance due to the frequency deviation would therefore be $-10(-50)(59.99 - 60) = -5$ MW.

Assuming no interchange metering error ($\epsilon = 0$), the total system imbalance as measured by ACE is therefore -10 MW. The interpretation is that demand (load) unexpectedly exceeds supply (generation plus scheduled imports) by 10 MW due to the loss of wind generation. Half of this energy deficiency (5 MW) was compensated for by an increase in actual inflows relative to scheduled inflows, and the other half (5 MW) was compensated for by a 0.01 Hz reduction in system frequency. The BA would need to increase supply from other sources by 10 MW to alleviate the sudden loss of wind generation, bring actual inflows in line with scheduled inflows, and correct the frequency deviation.

System operators are required to keep ACE as close to zero as possible at all times. Although in practice ACE is rarely precisely equal to zero, as we will show in the next section, deviations of ACE away from zero are, on average, quite small relative to total generation levels. This is important because the costs of exceptionally large system-wide energy imbalances can be significant. As we have noted, episodes of either oversupply or unmet demand each have clear implications for economic and technical efficiency.

3 Data & Descriptive Evidence

To investigate the relationship between wind intermittency and system imbalance, we construct an unbalanced panel of hourly wind generation and ACE data from three regional interconnections in the United States: the Bonneville Power Administration (BPA), the New York Independent System Operator (NYISO), and the Southwest Power Pool (SPP). BPA, while not technically classified as an ISO, is a federal power marketing administration charged with marketing wholesale electricity from 31 federally owned dams in the Pacific Northwest and operating a major transmission system covering Idaho, Oregon, Washington, western Montana, and parts of California, Nevada, Utah, and Wyoming. NYISO is a not-for-profit corporation that operates a wholesale electricity market exclusively in the state of New York. SPP manages the transmission grid and wholesale power market across 15 states, serving all or parts of Kansas, Oklahoma, New Mexico, Texas, Arkansas, Louisiana, Missouri, South Dakota, North Dakota, Montana, Minnesota, Iowa, Wyoming, and Nebraska. In addition to these three major regional markets, we analyze twice-daily (12-hour) measures of wind generation and ACE from the PJM Interconnection, an independent wholesale power market operator serving over 65 million people, and covering all or parts of Delaware, Illinois, Indiana, Kentucky, Maryland, Michigan, New Jersey, North Carolina, Ohio, Pennsylvania, Tennessee, Virginia, West Virginia and the District of Columbia. All data used in our analysis are publicly available, and the final sample was determined based on data availability. A table of data sources is provided in Appendix 1.

For BPA, NYISO, and SPP the dependent variable in our main specification is the hourly average of the absolute value of ACE. ACE data for BPA and SPP are reported in one-minute intervals, whereas ACE data for NYISO are reported every six seconds. Thus, when formatting the NYISO data, we kept only the values reported at each minute to match the one-minute intervals in the BPA and SPP data. We then took the absolute value of ACE in each minute and computed the average of those values for each hour. The resulting dependent variable is denoted $\overline{ACE}_{i,s} \equiv mean(|ACE_{i,t \in s}|)$, where i indexes ISO, s indexes hour (of sample), and $t = 0, ..., 59$ indexes the one-minute intervals at which ACE is observed. The measure of wind intermittency we use for our independent variable is the hourly variance in wind generation in the concurrent hour. Wind generation data are reported in MW, every five minutes. We specify the dependent variable as $V_{i,s} \equiv var(W_{i,\tau \in s})$, where $\tau = 0, 5, ..., 55$ indexes the five-minute intervals at which wind generation is observed. Thus, $V_{i,s}$ is the variance in hour s for ISO i of wind generation observations $W_{i,\tau \in s}$.

For PJM, although ACE data are reported in one-minute intervals, generation is only reported every hour, making it impossible to compute an hourly variance. We therefore aggregate to 12-hour intervals instead. The steps taken for PJM were similar to those for BPA, NYISO, and SPP, except that the ACE data were formatted first into one-hour intervals[12] and then aggregated with the generation data into 12-hour intervals. Thus, for PJM our main dependent variable is $\overline{ACE}_{\delta} \equiv mean(|ACE_{h \in \delta}|)$, where $h = 0, ..., 23$ indexes

[12]By convention, the one-hour ACE is calculated as a simple average of the one-minute ACE values over each hour, just as a 10-minute ACE is the simple average of the one-minute ACE values over each ten-minute interval.

Table 1: Summary statistics by ISO.

	# Obs.	Mean	SD	Min	Max
BPA (2013-2020)					
ACE (one-minute)	841,536	1.424	86.099	-3101.294	3788.373
ACE (hourly avg. abs. val.)	70,128	59.183	35.226	3.775	1209.809
ACE (abs. val.) / total gen. (hourly avg.)	70,128	0.005	0.003	0.0003	0.092
Wind generation (hourly avg.)	70,128	1082.171	1157.103	0	4570.542
Wind / total generation (hourly avg.)	70,128	0.086	0.088	0	0.441
Variance of wind gen. (hourly)	70,128	5460.299	14007.9	0	629382.9
NYISO (2019-2021)					
ACE (one-minute)	281,088	14.459	68.185	-5211.538	1286.588
ACE (hourly avg. abs. val.)	23,424	48.371	35.226	21.607	18.517
ACE (abs. val.) / total gen. (hourly avg.)	23,424	0.003	0.002	0.001	0.060
Wind generation (hourly avg.)	23,424	477.865	409.523	0	1888.083
Wind / total generation (hourly avg.)	23,424	0.033	0.029	0	0.154
Variance of wind gen. (hourly)	23,424	1142.78	3337.946	0	250050.3
SPP (2021-2022)					
ACE (one-minute)	185,748	4.305	110.477	-2255.164	1017.681
ACE (hourly avg. abs. val.)	15,479	82.007	34.886	0.262	506.969
ACE (abs. val.) / total gen. (hourly avg.)	15,479	0.005	0.003	0.000	0.193
Wind generation (hourly avg.)	15,479	11369.4	5191.237	436.058	22696.92
Wind / total generation (hourly avg.)	15,479	0.584	0.178	0.038	0.886
Variance of wind gen. (hourly)	15,479	90975.13	179158.2	95.745	4706453
PJM (2018-2021)					
ACE (one-hour)	35,040	84.497	3177.468	-570.365	594621
ACE (12-hr avg. abs. val.)	2,920	99.136	917.012	15.92	49606.73
ACE (abs. val.) / total gen. (12-hr avg.)	2,920	0.001	0.010	0.0002	0.561
Wind generation (12-hr avg.)	2,920	2837.977	1842.368	68.542	8354.667
Wind / total generation (12-hr avg.)	2,920	0.032	0.024	0.0002	0.109
Variance of wind gen. (12-hr)	2,920	520300	692941.1	965.789	7603996

All ACE and generation values are in MW.

hour of day and δ indexes 12-hour intervals.[13] Similarly, for our independent variable we compute the 12-hour variance in hourly wind generation, $V_\delta \equiv var(W_{h\in\delta})$.

Table 1 presents summary statistics on ACE and wind generation by ISO, from which a few key points emerge. First, the ACE (before taking the absolute value) is positive on average for all four ISOs, suggesting dispatchers generally err on the side of excess generation rather than excess load. Second, ACE is, on average, only a fraction of one percent of total generation. The extreme maximum value for PJM was a one-off event that lasted for three minutes in October of 2018 and is dropped as an outlier for the remainder of our analysis. Third, for BPA and NYISO the minimum value of zero for wind generation and, more

[13]Note that $h \in [0, 11]$ for the "a.m." interval (12:00am-11:00am) and $h \in [12, 23]$ for the "p.m." interval (12:00pm-11:00pm).

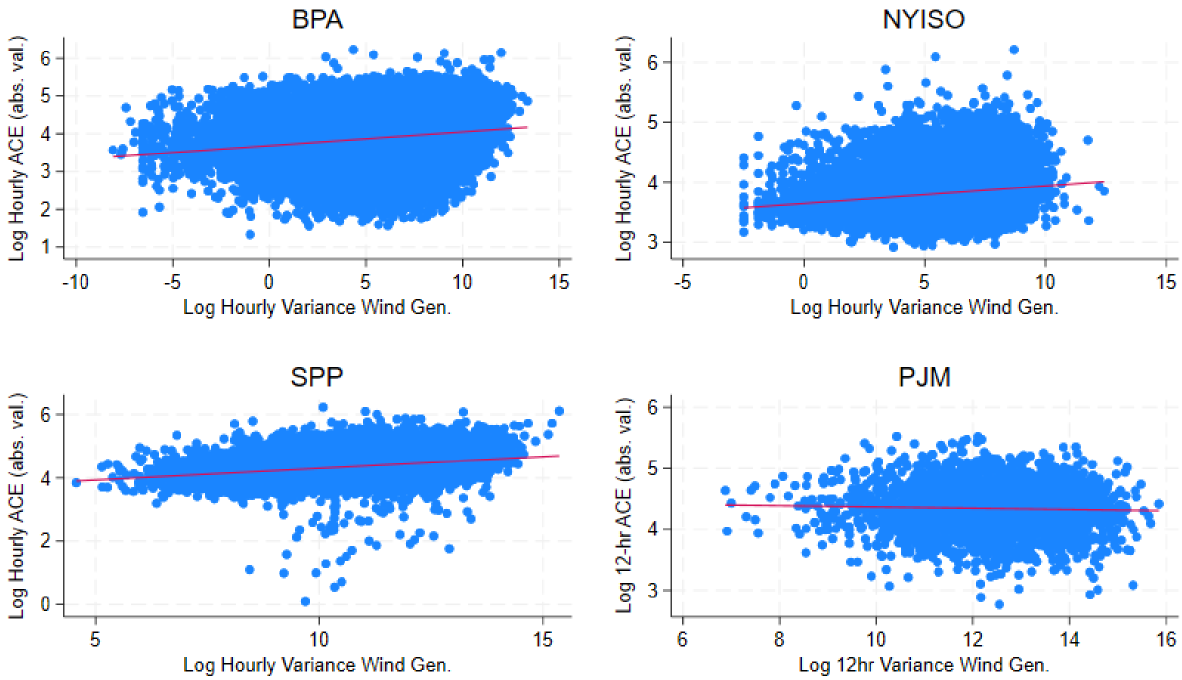

Figure 1: Scatter plots of log ACE (abs. val.) against log variance of wind generation.

importantly, variance of wind generation implies that some observations will be dropped from the main analysis as a result of taking natural logs. The number of observations dropped is sufficiently low compared to the total that we do not view this as cause for concern.[14] Fourth, SPP has by far the most wind generation of the four ISOs in our sample, both in absolute terms and as a percentage of total generation, due to its geographic location. The area served by SPP overlaps one of the most productive on-shore wind generation corridors in the world.

Figure 1 presents scatter plots by ISO of log ACE (absolute value) against log variance of wind generation. For BPA, NYISO, and SPP, we see in the unconditional data clear visual evidence of a positive relationship, which is consistent with our expectation that increased wind intermittency is associated with greater system imbalances. Conversely, for PJM the relationship is slightly negative. However, after controlling for regular temporal variation using a variety of time fixed effects (discussed below), what appears as a negative effect in the unconditional PJM data will prove to be statistically indistinguishable from zero.

Finally, in Figure 2 we plot selected intra-week time series of the two hourly variables for BPA, NYISO, and SPP. These graphs provide additional visual evidence that wind in-

[14]Moreover, the overwhelming majority of dropped observations were for BPA in the early years of the sample when very little wind capacity had been installed.

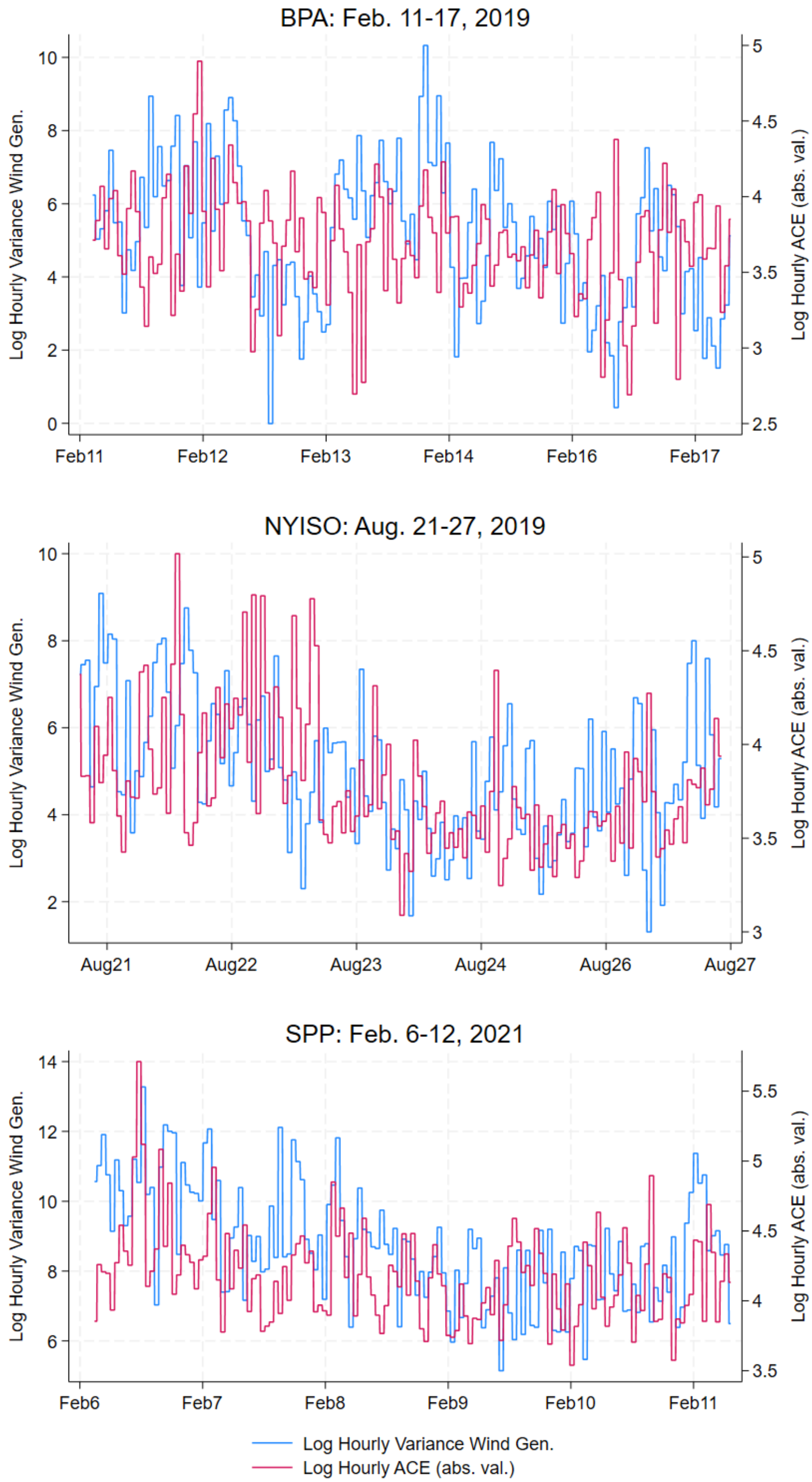

Figure 2: Selected intra-week time series plots of log hourly ACE (abs. val.) and log hourly variance of wind generation.

termittency and system imbalances generally move together over time. Despite considerable noise in the unconditional data, it is fairly clear that when the hourly variance of wind generation is relatively high, the hourly average ACE tends to be large as well, and *vice versa*. Having thus established sufficient descriptive evidence that greater wind intermittency leads to increased system imbalances, we now turn to our econometric exercise to estimate the magnitude of the effect.

4 Econometric Specifications

For reasons explained below, in our main specification we employ an instrumental variables (IV) identification strategy in which the variance in wind generation in a given hour (BPA, NYISO, SPP) or half day (PJM) is treated as endogenous. Because each ISO's raw data are reported differently and because each is unique both geographically and across a variety of characteristics such as generation mix, total load, transmission constraints, and other factors, we estimate the effect of wind intermittency on ACE separately for each ISO, rather than estimate an average effect across all three using a panel estimator and ISO-specific fixed effects. For BPA, NYISO, and SPP, our main regression equation is

$$\ln(\overline{ACE}_{i,s}) = \beta_i \ln(V_{i,s}) + \alpha_{i,h} + \theta_{i,d} + \gamma_{i,m} + \sigma_{i,q} + \varepsilon_{i,s}, \tag{2}$$

where $\alpha_{i,h}$ are hour-of-day fixed effects, $\theta_{i,d}$ are day-of-week fixed effects, $\gamma_{i,m}$ are month-of-year fixed effects, and $\sigma_{i,q}$ are quarter-of-sample fixed effects.[15] For PJM, we estimate

$$\ln(\overline{ACE}_{\delta}) = \beta \ln(V_{\delta}) + \alpha_h + \theta_d + \gamma_m + \sigma_q + \varepsilon_{\delta}, \tag{3}$$

where α_h is an indicator for the am/pm interval. In each equation, β represents a pure elasticity; it can be interpreted as the elasticity of the average hourly (or 12-hour, for PJM) system imbalance (in absolute value) with respect to a change in the hourly (12-hour) variance of system-wide wind generation.[16]

One might expect OLS to provide reasonably unbiased estimates of the effect of wind variance on ACE for each ISO. Any variation in wind generation is going to partly planned

[15] The first three, $\alpha_{i,h}$, $\theta_{i,d}$, and $\gamma_{i,m}$, are intended to pick up regular intra-day, intra-week, and seasonal variation, whereas $\sigma_{i,q}$ is included to control for longer-term, system-wide macroeconomic, climate, and technology shocks. One might be concerned that if daily and weekly patterns vary by month of year, some spurious correlation might be left over when using additive fixed effects. We also ran models using interacted hour-by-day-by-month fixed effects, which resulted in only negligible differences in the estimated effects of wind variance on ACE (not reported).

[16] Each model uses robust standard errors. Although Figure 1 mitigates concerns about heteroskedasticity, Figure 2 suggests serial correlation in the error structure is likely.

based on wind forecasts and partly unplanned due to forecast errors. If wind could be forecast perfectly, we should expect the effect to be zero because the system operator could perfectly match the wind generation profile. Conversely, if wind generation were perfectly random, we should expect the estimated effects to be large because the dispatcher would always be trying to catch up to unanticipated changes in wind generation. OLS estimates should therefore reflect an 'average' between these two extremes.

However, there may be other sources of endogeneity that could bias OLS estimates in either direction. Our primary motivation for utilizing an IV specification is that both ACE and wind generation variance are endogenous to unobserved planned and unplanned outages of wind generators, introducing omitted variable bias. This endogeniety is likely to bias OLS estimates of the effect of wind intermittency on ACE downward or upward, depending on whether planned or unplanned outages predominate. The underlying logic is as follows.

A planned outage of a wind turbine (or set of turbines) would result in a change in wind generation both at the time the turbine is shut down and at the time it is brought back online. This would increase the variance of wind generation in both the hour of shutdown and the hour of restart. However, if a wind generator outage is planned—that is, *scheduled in advance*—the dispatcher is aware of the impending change in wind generation and can respond in real time by ramping other generators up at the moment of shutdown and down at the moment of restart, thus minimizing any resulting imbalance. In other words, a planned outage increases the variance in wind generation, but has only a minimal impact on ACE. By contrast, pure intermittency cannot be perfectly anticipated, meaning there is a greater likelihood of a larger imbalance because the dispatcher cannot respond as quickly. Thus, the impact on ACE is larger. If planned outages are more common than unplanned outages, we would expect the OLS estimates to be biased downward.

Conversely, unplanned outages would have a similar effect on the observed ACE, except in the other direction, because they would exacerbate the unanticipated variation in wind generation. Thus, if unplanned outages are more frequent, we would therefore expect our OLS estimates to be biased upward.

Although we view planned/unplanned outages as the primary source of endogeneity, another potential source of endogeneity arises if unobserved weather shocks contemporaneously affect both wind variance and electricity demand. Because ACE is a measure of system imbalance, it is also affected by fluctuations in load. Thus, any weather shocks that affect both wind variance and load would introduce additional endogeneity, although the direction of the resulting bias in the OLS estimates is unclear.

Our preferred instruments are combinations of lags of the variance in wind generation. The IV strategy is empirically valid if and only if the instruments are independent of the main outcome variable and jointly uncorrelated with the error term. Intuitively, the first of these conditions is satisfied simply because the present cannot affect the past; ACE in hour s has no impact on wind generation variance in hour $s-1$, $s-2$, and so on. The second condition is satisfied if the variation in past wind generation variance is uncorrelated with the current period's error term. One might be concerned that unobserved weather shocks that persist through time could affect both prior periods' wind generation variance and the present period's ACE. However, weather events, including the very short-term changes in wind speeds from which the hourly variance in wind generation ultimately derives, are relatively localized compared to the immense geographic areas of the ISOs in our sample.[17] It is therefore plausible that variation in prior periods' wind generation variance, which changes due to variation in wind speeds over both time and space, is sufficiently orthogonal to the current period's error term, and the further back in time one moves, the more likely this is to be true. We support our choices of lagged wind generation variance as instruments using standard IV test statistics, discussed below. In fact, favorable IV test statistics were the decisive factor in choosing which lags were to be used as instruments in a given model, as some combinations of shorter lags were found to be jointly correlated with the error term in particular models, indicating misspecification.

4.1 Directional Changes

Our main specification tells us how wind intermittency, measured as variance, impacts the magnitude of system imbalance. It does not, however, distinguish situations in which electricity supply exceeds demand from situations in which demand exceeds supply. We now alter our model to capture the effects of *directional* changes in wind generation on the ACE, taking into account the *sign* of ACE. As illustrated in Section 2, ACE is negative when demand exceeds supply and positive when supply exceeds demand. Thus, we should generally expect ACE to be positive when wind generation is rising and negative when wind generation is falling. To test this hypothesis, we first set up a logit model for each ISO to estimate the effect of a change in wind generation over a given interval on the probability that the average ACE over that interval is positive or negative.[18] We restrict our focus here to only BPA, NYISO, and SPP due to the greater intertemporal granularity of the data.

As noted above, wind generation is observed every five minutes. Denote wind generation for ISO i at five-minute interval $\bar{\tau}$ as $W_{i,\bar{\tau}}$.[19] Our main independent variable in the logistic

[17] NYISO, the smallest in geographical area, covers over 140,000 km^2. PJM, BPA, and SPP cover areas of roughly 475,000 km^2, 775,000 km^2, and 1.43 million km^2, respectively.

[18] Importantly, in this exercise we do not take the absolute value of ACE before computing the average.

[19] We use $\bar{\tau}$ to index five-minute interval of sample, to distinguish it from $\tau = 0, 5, ..., 55$, which indexes

regression is simply $\Delta W_{i,\bar{\tau}} \equiv W_{i,\bar{\tau}} - W_{i,\bar{\tau}-5}$, or the change in wind generation over each five-minute interval.

To construct our dependent variable, we first compute the average ACE over each five-minute interval, denoted $\bar{A}_{i,\bar{\tau}} \equiv mean(ACE_{t\in\bar{\tau}})$, which can be positive or negative (or zero). Our binary outcome variable is therefore

$$A^{logit}_{i,\bar{\tau}} = \begin{cases} 0 & \text{if } \bar{A}_{i,\bar{\tau}} \leq 0 \\ 1 & \text{if } \bar{A}_{i,\bar{\tau}} > 0. \end{cases}$$

The logit model is

$$Pr(A^{logit}_{i,\bar{\tau}} = 1|\Delta W_{i,\bar{\tau}}) = f(\beta_{i,0} + \beta^{logit}_i \Delta W_{i,\bar{\tau}}), \tag{4}$$

where $f(\cdot)$ is the logit function. The logistic regression estimate of β^{logit}_i for each ISO is expressed as the factor change in the odds of $\bar{A}_{i,\bar{\tau}}$ being positive resulting from a one-MW increase in $\Delta W_{i,\bar{\tau}}$. The constant term, $\beta_{i,0}$, represents the baseline odds ratio, or $\Pr(\bar{A}_{i,\bar{\tau}} > 0)/\Pr(\bar{A}_{i,\bar{\tau}} \leq 0)$, for each ISO.[20]

However, we are also interested to know if the magnitude of the effect on ACE differs depending on whether the change in wind is positive or negative. Dispatchers may respond differently to positive versus negative changes in wind generation. Alternatively, if other generation assets in the system can be ramped up or down at different rates due to differing technical constraints, we might expect the magnitudes of the effect of changes in wind generation on ACE to differ depending on sign. To investigate this, we split the sample for each ISO into subsamples containing only positive and only negative changes in wind generation. We then estimate the following linear regressions using OLS:

$$\bar{A}_{i,\bar{\tau}} = \beta^{\pm}_i \Delta W^{\pm}_{i,\bar{\tau}} + \alpha_{i,h} + \theta_{i,d} + \gamma_{i,m} + \sigma_{i,q} + \mu_{i,\bar{\tau}}, \tag{5}$$

where the $\pm$ superscript indicates positive versus negative changes in wind generation. In other words, for each ISO we estimate two regression models, where coefficients β^+_i and β^-_i correspond to the positive and negative wind generation change subsamples, $\Delta W^+_{i,\bar{\tau}}$ and $\Delta W^-_{i,\bar{\tau}}$, respectively. Note that, in order to maintain distinct signs across subsamples, we do not take logs here (as in our main specification), implying that our estimates of β^+_i and β^-_i should be thought of as linear slope terms and not elasticities.

within-hour five-minute intervals.

[20] We do not include time fixed effects in the logistic regression for two reasons. First, the computational burden is excessive relative to the gain in accuracy of the estimates. Second, it is not clear that the estimates are more accurate with time fixed effects added, due to the well known 'incidental parameters' problem associated with including fixed effects as dummy variables in discrete choice models (Heckman, 1981).

5 Results

Results of our main specifications (Equations 2 and 3) are presented in Table 2. For each ISO, we present the results of both OLS and 2SLS specifications. Standard IV diagnostic statistics are described and reported for each 2SLS model in Appendix 2. The only IV diagnostic tests that fail to yield the expected result are the tests of endogeneity for $\ln(V_\delta)$ in the PJM regressions, indicating this variable can be treated as exogenous and that the OLS estimates are more efficient. For all others, the IV diagnostic tests indicate that 2SLS is preferred to OLS and that the instruments used are jointly uncorrelated with the error term. We discuss the results for each ISO in turn.

First, consider the estimates for BPA, in columns (1) and (2). In each model, the effect of wind variance on ACE is highly statistically significant, with estimates ranging between 0.2 and 0.4. As the coefficient estimates can be interpreted as pure elasticities, our preferred model—the 2SLS model in column (2)—indicates that a doubling of the hourly variance in wind generation in the BPA control area leads to a 2% increase in the average hourly ACE. In energy terms, relative to the sample mean hourly average ACE for BPA of 59.18 MW (see Table 1), a 2% increase translates to roughly 1.2 MW. Interestingly, BPA is the only ISO (of the three for which we have hourly observations) where the endogeneity of hourly wind variance biases the OLS estimate upward, suggesting unplanned wind turbine outages were more frequent than unplanned outages during the BPA sample period.[21] This may be explained by the much longer sample period for BPA compared to NYISO and SPP. In the earlier years of our BPA sample, large-scale wind turbine technology was not nearly as mature, and unplanned outages likely were more frequent and problematic.

Next, the estimates for NYISO are presented in columns (3) and (4). Again, we have a highly statistically significant effect, ranging between 0.31 and 0.41. The point estimate in our preferred model therefore suggests a doubling of hourly wind variance in NYISO leads to a 4.1% increase in the average ACE. In energy terms, relative to the sample mean ACE of 48.37 MW, this translates to roughly 2 MW. In contrast to the BPA results, endogeneity biases the OLS estimates downward, suggesting unplanned turbine outages were more common in NYISO during the sample period.

For SPP, in columns (5) and (6), we see that the effect of wind intermittency on ACE is again highly statistically significant but much larger in magnitude (in both the OLS and 2SLS models) compared to BPA and NYISO, ranging from 0.67 to 0.125. The point estimate in our preferred model indicates a doubling of hourly wind variance in the SPP control region

[21]Note that although the reported coefficient estimates in column (1) appear to be identical due to rounding, they are not. The 2SLS estimate is slightly lower.

Table 2: OLS and 2SLS estimates: Main specification.

OLS estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ISO	**BPA**	**BPA**	**NYISO**	**NYISO**	**SPP**	**SPP**	**PJM**	**PJM**
Coeff. (β)	0.040***	0.033***	0.031***	0.031***	0.068***	0.067***	0.003	0.001
Std. Err.	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.005)	(0.005)
h,d,m FE's	Y	Y	Y	Y	Y	Y	Y	Y
Quart. FE's	N	Y	N	Y	N	Y	N	Y
Sample period	2013-2020	2013-2020	2019-2021	2019-2021	2021-2022	2021-2022	2018-2021	2018-2021
Time interval	Hourly	Hourly	Hourly	Hourly	Hourly	Hourly	12-hour	12-hour
N	69621	69621	23403	23403	15477	15477	2919	2919

2SLS estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ISO	**BPA**	**BPA**	**NYISO**	**NYISO**	**SPP**	**SPP**	**PJM**	**PJM**
Coefficient	0.040***	0.020***	0.041***	0.041***	0.125***	0.121***	0.007	-0.007
Std. Err.	(0.001)	(0.001)	(0.002)	(0.002)	(0.005)	(0.005)	(0.025)	(0.029)
h,d,m FE's	Y	Y	Y	Y	Y	Y	Y	Y
Quart. FE's	N	Y	N	Y	N	Y	N	Y
Sample period	2013-2020	2013-2020	2019-2021	2019-2021	2021-2022	2021-2022	2018-2021	2018-2021
Time interval	Hourly	Hourly	Hourly	Hourly	Hourly	Hourly	12-hour	12-hour
N	69159	68982	23377	23377	15477	15477	2916	2916

All models use robust SE's. For PJM, hourly FE's indicate am/pm. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

leads to a 12.1% increase in the average hourly ACE, which, relative to the SPP sample mean ACE of 82 MW, translates to an increase of around 9.9 MW. Again, because SPP has the most wind generation capacity of any ISO in our sample, it is not surprising that the estimated effect of intermittency on the ACE would be largest. Like NYISO, the endogeneity of wind variance biases the OLS estimates downward for SPP, suggesting unplanned turbine outages were more common in the SPP control region during the sample period.

Finally, the results for PJM are presented in columns (7) and (8). In contrast to the other ISOs in our sample, we find no statistically significant effects for PJM. We attribute this primarily to the lack of granularity in the PJM wind generation data that required aggregating at the 12-hour level instead of hourly. This level of aggregation may simply introduce too much noise, which, in combination with a severely limited number of observations, undermines any ability to detect a statistically significant effect. On the other hand, the standard errors are of roughly the same magnitude as those of the other ISOs' coefficient estimates, suggesting the PJM estimates do not suffer from a high degree of inaccuracy. The implication may therefore be that the effect of intermittency on system imbalance exists at the sub-hourly level but not at the 12-hour level. If sub-hourly data for PJM were available that enabled a more granular level of aggregation, we might reasonably expect to find similar results to those of BPA, NYISO, and SPP. However, for the remainder of the analysis we drop PJM from the sample and focus only on the three ISOs for which we have sub-hourly generation data.

5.1 Results: Directional Changes

We begin with the estimates of our logistic regression model, based on Equation 4, which are presented in Table 3. Estimates of the effect of a five-minute change in wind generation are presented as the factor change in the odds ratio $\Pr(\bar{A}_{i,\bar{\tau}} > 0)/\Pr(\bar{A}_{i,\bar{\tau}} \leq 0)$. In each regression, the constant term represents the baseline odds ratio.[22] Note that for each ISO, the estimate is highly statistically significant and greater than unity. Thus, a 1 MW increase in $\Delta W_{i,\bar{\tau}}$ increases the odds ratios for BPA, NYISO, and SPP by factors of 1.015, 1.013, and 1.002, respectively. Each model has a highly significant likelihood ratio χ^2 value, indicating each fits significantly better than an empty model (i.e., a model with no regressors).

To aid in the interpretation of these estimates, consider Figure 3. For each ISO, the histograms show that the five-minute changes in wind generation are evenly distributed around a mean of (effectively) zero.[23] More importantly, however, the graphs of predicted

[22]In other words, let ρ_i be the fraction of sample observations for ISO i for which $\bar{A}_{i,\bar{\tau}} > 0$. The baseline odds ratio is $\rho_i/(1-\rho_i)$.

[23]The exact sample means are, respectively, 0.0004 (BPA), 0.0011 (NYISO), and 0.005 (SPP). Note also

Table 3: Logistic regression estimates by ISO.

ISO	(1) **BPA**	(2) **NYISO**	(3) **SPP**
Wind Generation (5-min. change)	1.015***	1.013***	1.002***
Std. Err.	(0.000)	(0.000)	(0.000)
Constant	1.077***	1.676***	1.089***
Std. Err.	(0.002)	(0.006)	(0.005)
Log likelihood	-565467.61	-184552.38	-127187.11
Likelihood ratio χ^2	34693.38	2791.90	2799.38
LR p-value	0.000	0.000	0.000
N	841,536	281,087	185,748

Estimates presented as factor changes in odds ratios. Constant terms represent baseline odds ratios. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

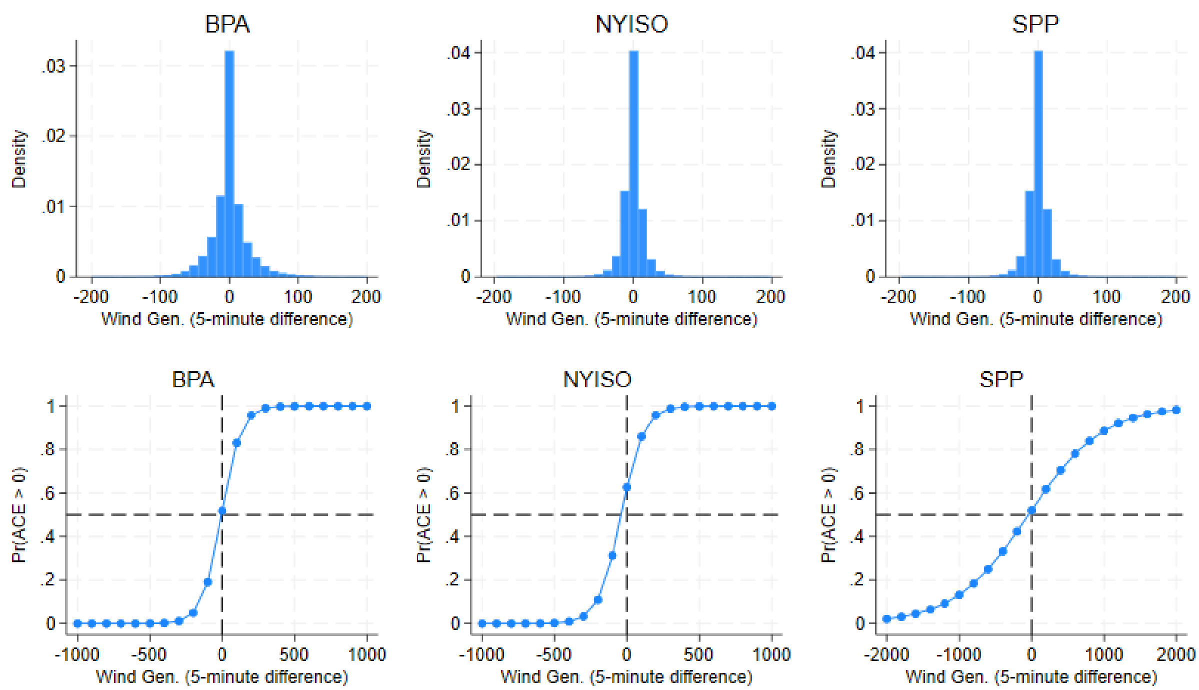

Figure 3: Histograms of 5-minute change in wind generation (upper) and predicted probabilities that $\bar{A}_{i,\bar{\tau}} > 0$ over selected range of $\Delta W_{i,\bar{\tau}}$ (lower), by ISO.

probabilities, which are computed from the logistic regression estimates of the marginal effect of a 1 MW change in $\Delta W_{i,\bar{\tau}}$, trace out what are effectively predicted (conditional) cumulative distribution functions (CDFs) of the probability that $\bar{A}_{i,\bar{\tau}} > 0$ over the range of $\Delta W_{i,\bar{\tau}}$. A clear pattern emerges, with only slight variations specific to each ISO. Consistent with our hypothesis in Section 4.1, a negative ACE is more probable for negative changes in wind generation,[24] whereas a positive ACE is more likely for positive changes in wind generation. However, we see that the CDFs are much steeper for BPA and NYISO than for SPP. To illustrate, consider a positive change in wind generation of $\Delta W_{i,\bar{\tau}} = 200MW$. For both BPA and NYISO, the predicted probability that $\bar{A}_{i,\bar{\tau}} > 0$ is over 95%, whereas for SPP it is around 70%. This qualitative difference in the shapes of the predicted CDFs implies system imbalances are much more tightly distributed and more closely linked to system-wide changes in wind generation for BPA and NYISO than for SPP. Wider variation in the ACE conditional on aggregate changes in wind generation is to be expected for SPP, given that it manages a much larger scale and proportion of wind generation in its system, which implies greater variability not only in aggregate, but also across the system's sub-regions.

Finally, estimates of the magnitudes of the effects on ACE of positive/negative changes in wind generation are presented in Table 4. As noted above, these estimates should be interpreted as linear slope coefficients. Point estimates suggest positive changes in wind generation are treated differently than negative changes, but the difference in responses appears unique to each ISO. These differences may reflect dispatchers' preferences over positive versus negative system imbalances, or they may reflect technical constraints specific to each ISO that affect the ability to ramp other generators up or down as needed.

To illustrate using the BPA point estimates, recall that our logit model showed that the ACE is more likely to be positive (supply exceeds demand) for a positive five-minute change in wind generation. Thus, a 1 MW increase in the magnitude of a positive five-minute change in wind generation increases the (likely positive) five-minute average ACE by 0.505 MW. Conversely, a change of -1 MW in the magnitude of a negative five-minute change in wind generation—for example, a change from -10 MW to -11 MW—leads to a change in the (likely negative) five-minute average ACE by -0.597 MW. By contrast, the relative magnitudes of positive versus negative effects are flipped; for NYISO the effect on ACE of a negative change in wind generation is smaller in magnitude than the effect of a positive change. For SPP, the positive/negative changes are both smaller in absolute terms and much

that the histograms are truncated at -200 MW and 200 MW for the purpose of visual presentation, although for each ISO a non-trivial number of observations fall beyond these limits.

[24] To be precise, because the baseline odds ratio for each ISO is greater than unity, there is a small range of negative values of $\Delta W_{i,\bar{\tau}}$ close to zero for which the predicted probability of $\bar{A}_{i,\bar{\tau}} > 0$ is greater than 50%. This range is largest for NYISO.

Table 4: OLS estimates of positive/negative changes in wind generation on ACE, by ISO.

	(1)	(2)	(3)	(4)	(5)	(6)
ISO	**BPA**	**BPA**	**NYISO**	**NYISO**	**SPP**	**SPP**
Coefficient ($\beta_i^{\pm}$)	0.505***	0.597***	0.449***	0.353***	0.107***	0.128***
Std. Err.	(0.009)	(0.010)	(0.039)	(0.031)	(0.010)	(0.006)
Sub-sample ($\pm$)	(+)	(−)	(+)	(−)	(+)	(−)
N	385,087	412,303	185,748	132,588	92,943	92,730

Dependent variable is the five-minute average ACE. All models contain hour, month, day, and quarter FE's and use robust SE's. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

closer to each other in relative magnitude.

5.2 Discussion

The results of the preceding analysis have shown that electricity system imbalances, as measured by ACE, are highly sensitive to short-term—i.e., sub-hourly—fluctuations in wind generation. The basic intuition is that, because intermittent variability in wind generation cannot be perfectly forecast, system operators cannot respond to sudden changes in wind generation with perfect accuracy, leading to greater system imbalances. We found qualitatively similar results for BPA, NYISO, and SPP but are careful to emphasize that certain features of the effect of wind intermittency on ACE are unique to each ISO. Although we did not find statistically significant results for PJM in our main specification, we do not view this as especially problematic for our core hypothesis, given the much less granular 12-hour aggregation required by the PJM data.

Our results have important implications for the clean energy transition. As the levelized cost of energy for wind power continues to decline, and as the public and private response to the climate crisis continues to gain momentum, wind will only increase in prominence in the electricity sector, both domestically and globally. Our results suggest increased penetration of intermittent generation will likely lead to increased system imbalances, imposing additional costs that should be accounted for in benefit-cost analyses of policies designed to promote investment in these technologies. They also emphasize the urgency of R&D in large-scale energy storage technologies to complement wind and solar by mitigating the deleterious effects of intermittency on the transmission grid. Gowrisankaran et al. (2016) find that the ability to perfectly offset solar intermittency through storage reduces the social costs of providing 20% of total generation via solar by roughly one-third.

There is, however, a 'silver lining' to this story. Figure 4 presents scatter plots by ISO

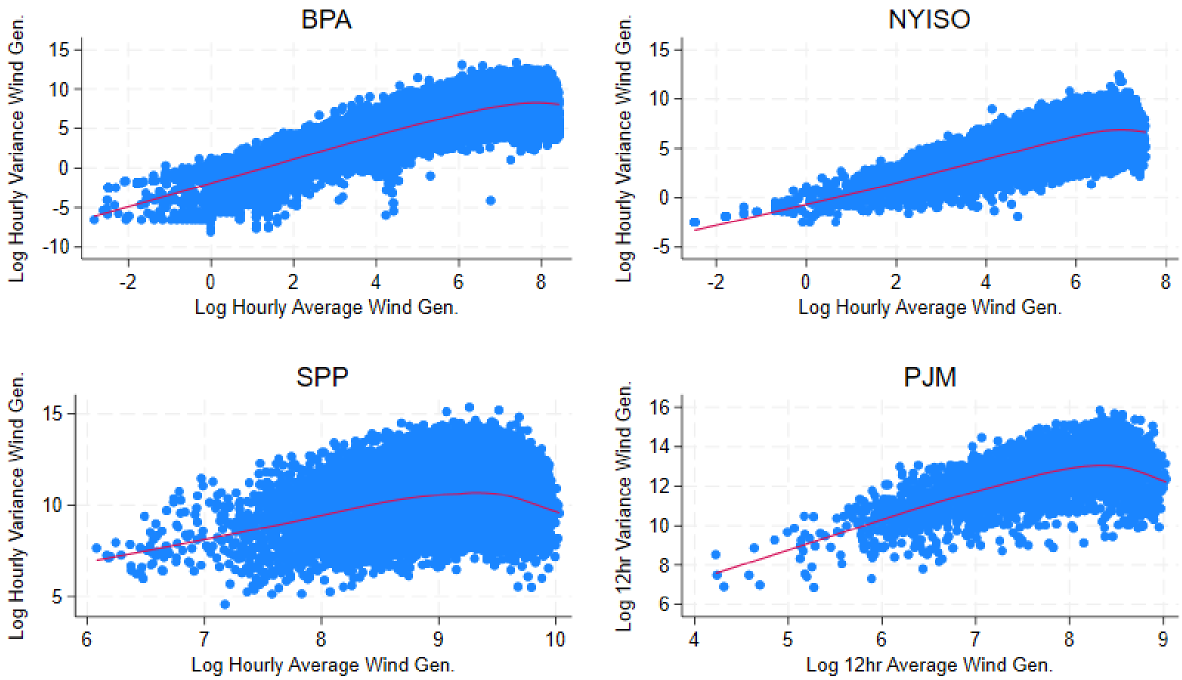

Figure 4: Scatter plots by ISO of log hourly variance of wind generation against log hourly average wind generation, with fitted lines using lowess smoothing.

of log hourly variance of wind generation against log hourly average wind generation, with fitted lines using lowess smoothing.[25] A consistent pattern emerges. As the level of wind generation increases, the variance of wind generation tends to increase also. However, for each ISO the fitted lowess line reaches a peak at a relatively high level of wind generation and then begins to fall. In other words, there appears to be a point at which increasing wind generation in the system begins to have a *lower* variance.

The implication of this finding is particularly intriguing. As more wind capacity is added to the system, not only does total wind generation increase, but the geographic distribution of wind generation assets also becomes more dispersed. This increases the likelihood at any given moment that somewhere in the system wind will be generating power, such that a local reduction in wind generation in one location is offset by an increase in another location, and *vice versa*. This would have the effect of reducing the system-wide variance in wind generation, which suggests there may be a 'critical mass' of wind generation beyond which the effects of localized intermittency on system-wide imbalances begin to fade.

[25]Lowess smoothing refers to **Lo**cally **We**ighted **S**catter plot **S**moothing. The fitted line is created such that a separate, locally-weighted regression is performed for every data point.

6 Conclusion

Wind has a large role to play in the clean energy transition, but expansion of wind in the electricity generation mix is not without significant challenges. This paper has investigated the effect of wind intermittency on electricity system imbalance, using publicly available data from major regional power system operators in the United States. Our results have shown that greater variability in wind generation increases system imbalances in systematic and predictable ways, as measured by a standard indicator used by system operators known as the *area control error*, or ACE.

We first show that greater hourly variance of wind generation leads to a larger hourly average ACE in absolute value—i.e., not taking into account whether total electricity supply exceeds demand or *vice versa*. This effect is consistent across all three system operators for which we have sub-hourly generation data: Bonneville Power Administration (BPA), New York Independent System Operator (NYISO), and the Southwest Power Pool (SPP). The point estimates of our preferred specification indicate a doubling of hourly variance of wind generation leads to increases in average hourly system imbalances of 2% (BPA), 4.1% (NYISO), and 12.1% (SPP). In energy units, relative to the sample mean ACE for each system operator, these elasticities translate to 1.2 MW, 2 MW, and 9.9 MW increases in average hourly system imbalances, respectively. Aside from being strongly statistically significant, these increases in system-wide imbalance are economically significant, as they imply potentially costly deviations in system frequency, in addition to requiring system operators to respond immediately through a variety of costly measures to bring the system back into balance. We were not able to find a similar effect for the PJM Interconnection, which we attribute to the less granular level of aggregation required because PJM wind generation data are reported only hourly.

We then estimated directional effects—that is, we examined whether sudden increases in wind generation affect the ACE differently than sudden decreases. Restricting our analysis to only BPA, NYISO, and SPP, we first used a simple logit model to show that ACE is more likely to be positive (supply exceeds demand) when wind generation is increasing and more likely to be negative (demand exceeds supply) when wind generation is falling. While this may seem intuitively obvious, to our knowledge ours is the first paper to confirm it empirically. We then show that the magnitudes of directional changes in the ACE depending on positive versus negative changes in wind generation differ for each ISO. Whether these positive or negative effects are larger in magnitude may depend on dispatcher preferences or on technical constraints associated with ramping other generation up or down as needed. Future research should look into how ramping costs are linked to system imbalances and whether investments in fast-ramping generation technologies mitigate the system imbalances resulting from intermittent wind (and solar) generation.

One could envision any number of other empirical applications in electricity market research for which ACE data might prove insightful. It is therefore surprising and unfortunate that ACE as a measure of system imbalance has been thus far largely overlooked and underutilized by energy economists and policy researchers, despite being a standard measure that is well known to electricity system operators and engineers. For example, future research might estimate the effects of ACE on outages, gaps between day-ahead and real-time prices, power plant heat rates, or any number of other economically relevant outcomes that could be used to estimate the welfare effects of generation intermittency. Cicala (2022) estimated the causal impact of U.S. electricity market liberalization on electricity demand and generation costs, showing significant gains to trade. If one also had the historical ACE data, it would be interesting to show whether there were also effects on system imbalances, which, as we have discussed, have implicit costs that may have offset or enhanced the gains from market liberalization.

Finally, we also showed evidence that, as wind continues to expand, the effects of localized intermittency on system-wide variation in wind generation, and thus energy imbalances, may begin to fade. More research is needed on this question, and a full investigation is beyond the scope of this paper. Yet, the implications of this possibility are profound. Given the lack of large-scale electricity storage solutions, the intermittency problem has thus far been one of the largest impediments to the expansion of wind and solar generation technologies. If, however, a sufficiently geographically dispersed portfolio of wind and utility-scale solar assets can reach a point at which system-wide variability from these generation sources begins to decline, the lack of storage may ultimately prove less problematic that it has been thus far. This would certainly be a boon to the clean energy transition, as it would encourage, rather than discourage, further penetration of intermittent generation technologies in the electric power system.

Appendix 1: Data Sources

ISO	Variable(s)	Source
BPA	ACE	https://transmission.bpa.gov/Business/Operations/ACE_FERC784/
	Generation	https://transmission.bpa.gov/Business/Operations/Wind/default.aspx
NYISO	ACE & Generation	http://mis.nyiso.com/public/P-38list.htm
PJM	ACE	https://www.pjm.com/markets-and-operations/etools/oasis/system-information/historical-area-control-error-data
	Generation	https://dataminer2.pjm.com/feed/gen_by_fuel
SPP	ACE	https://portal.spp.org/pages/integrated-marketplace-ace-chart
	Generation	https://portal.spp.org/pages/generation-mix-historical

Appendix 2: IV Test Statistics

To validate our choice of instrumentation in the 2SLS specifications, in Table 5 we report several IV test statistics for each 2SLS regression reported in Table 2.[26]

We first test whether the endogenous regressors—$\ln(V_{i,s})$ for BPA, NYISO, SPP and $\ln(V_\delta)$ for PJM—are in fact exogenous. Because robust standard errors are used, we report the p-values of the robust score test and robust regression test (Wooldridge, 1995). For each test, if the statistic is significant ($p < 0.1$), the regressor is considered endogenous.

Next, we report a first-stage summary, in which we indicate which lags were used as instruments and report the p-values of an F-test for joint significance in the first stage. Note that for each model, lags were chosen on the basis of the test for overidentifying restrictions, described below.

We then test for weak instruments using the Cragg-Donald minimum eigenvalue statistic. As explained in the Table 5 notes, these values are compared to the Stock-Yogo critical value corresponding to a 10% maximum rejection rate for a nominal 5% Wald test in 2SLS models. If the Cragg-Donald minimum eigenvalue statistic exceeds the Stock-Yogo critical value, the instruments are not considered weak (Cragg and Donald, 1993; Stock and Yogo, 2005).

Finally, we report the p-values of the Wooldridge robust score test of overidentifying restrictions. In this test, failure to reject the null hypothesis is the desired outcome, as it indicates the instruments are jointly uncorrelated with the error term (Wooldridge, 1995).

[26]See: `https://www.stata.com/manuals/rivregresspostestimation.pdf`

Table 5: IV diagnostic statistics.

Model (see Table 2)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ISO	**BPA**	**BPA**	**NYISO**	**NYISO**	**SPP**	**SPP**	**PJM**	**PJM**
Endogeneity								
Robust score test (p-value)	0.05	0.00	0.00	0.00	0.00	0.00	0.87	0.78
Robust regression test (p-value)	0.05	0.00	0.00	0.00	0.00	0.00	0.87	0.78
First-stage Summary								
Instruments used (lags)	1,2	7,8	1,2	1,2	2,3	2,3	2,3	2,3
Robust F-test (p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Weak Instruments								
Cragg-Donald min. eigenvalue stat.	60313.5	11770.9	8463.73	8434.67	377.66	355.59	61.12	46.19
Overidentifying Restrictions								
Wooldridge robust score test (p-value)	0.424	0.558	0.22	0.22	0.69	0.77	0.58	0.65

(i) In the weak instrument test, the Stock-Yogo critical value corresponding to a 10% maximum rejection rate for a nominal 5% Wald test in 2SLS models is 19.93. If the Cragg-Donald minimum eigenvalue statistic is greater than this critical value, the instruments are not considered weak. (ii) Failure to reject the null hypothesis is the desired outcome for the Wooldridge robust score test of overidentifying restrictions.

References

Ahmed, H. M., Awad, A. S., Ahmed, M. H., and Salama, M. (2020). Mitigating voltage-sag and voltage-deviation problems in distribution networks using battery energy storage systems. *Electric Power Systems Research*, 184:106294.

Blume, S. W. (2017). *Electric Power System Basics for the Nonelectrical Professional*. IEEE Press, New Jersey, 2nd edition.

Bollen, M. H. (2003). What is power quality? *Electric Power Systems Research*, 66(1):5–14.

Borenstein, S., Bushnell, J., and Mansur, E. (2023). The economics of electricity reliability. *The Journal of Economic Perspectives*, 37(4):181–206.

Bourry, F., Juban, J., Costa, L., and Kariniotakis, G. (2008). Advanced strategies for wind power trading in short-term electricity markets. In *European Wind Energy Conference & Exhibition EWEC 2008*. EWEC.

Brouwer, A. S., van den Broek, M., Seebregts, A., and Faaij, A. (2014). Impacts of large-scale intermittent renewable energy sources on electricity systems, and how these can be modeled. *Renewable and Sustainable Energy Reviews*, 33:443–466.

Cicala, S. (2022). Imperfect markets versus imperfect regulation in us electricity generation. *American Economic Review*, 112(2):409–441.

Cragg, J. G. and Donald, S. G. (1993). Testing identifiability and specification in instrumental variable models. *Econometric Theory*, 9(2):222–240.

DeCesaro, J., Porter, K., and Milligan, M. (2009). Wind energy and power system operations: A review of wind integration studies to date. *The Electricity Journal*, 22(10):34–43.

Durrwachter, H. L. and Looney, S. K. (2012). Integration of wind generation into the ERCOT market. *IEEE Transactions on Sustainable Energy*, 3(4):862–867.

Ela, E. and Edelson, D. (2012). Participation of wind power in LMP-based energy markets. *IEEE Transactions on Sustainable Energy*, 3(4):777–783.

Energy Information Administration (2021). International Energy Outlook 2021 with projections to 2050. Technical report, U.S. Energy Information Administration.

Energy Information Administration (2022). Annual Energy Outlook 2022 with projections to 2050. Technical report, U.S. Energy Information Administration.

Genoese, F., Genoese, M., and Wietschel, M. (2010). Occurrence of negative prices on the German spot market for electricity and their influence on balancing power markets. In *2010 7th International Conference on the European Energy Market*, pages 1–6. IEEE.

Godoy-González, D., Gil, E., and Gutiérrez-Alcaraz, G. (2020). Ramping ancillary service for cost-based electricity markets with high penetration of variable renewable energy. *Energy Economics*, 85:104556.

Gowrisankaran, G., Reynolds, S. S., and Samano, M. (2016). Intermittency and the value of renewable energy. *Journal of Political Economy*, 124(4):1187–1234.

Heckman, J. J. (1981). An incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process. In Manski, C. and McFadden, D., editors, *The Structural Analysis of Discrete Data*, pages 179–195. Mit Press.

Heydt, G. T. (2005). Electric power quality. In *The Electrical Engineering Handbook*, pages 805–810. Elsevier Inc.

Johnson, E. P. and Oliver, M. E. (2019). Renewable generation capacity and wholesale electricity price variance. *The Energy Journal*, 40(5):143–168.

Joskow, P. L. (2019). Challenges for wholesale electricity markets with intermittent renewable generation at scale: The US experience. *Oxford Review of Economic Policy*, 35(2):291–331.

Katzenstein, W. and Apt, J. (2012). The cost of wind power variability. *Energy Policy*, 51:233–243.

King, J., Kirby, B., Milligan, M., and Beuning, S. (2011). Flexibility reserve reductions from an energy imbalance market with high levels of wind energy in the western interconnection. Technical report, National Renewable Energy Lab (NREL), Golden, CO (United States).

Kiviluoma, J., Meibom, P., Tuohy, A., Troy, N., Milligan, M., Lange, B., Gibescu, M., and O'Malley, M. (2012). Short-term energy balancing with increasing levels of wind energy. *IEEE Transactions on Sustainable Energy*, 3(4):769–776.

Liebensteiner, M. and Wrienz, M. (2020). Do intermittent renewables threaten the electricity supply security? *Energy Economics*, 87:104499.

NERC (2011). Balancing and frequency control: A technical document prepared by the NERC resources subcommittee. Technical report, North American Electric Reliability Corporation.

Nguyen, N. and Mitra, J. (2015). An analysis of the effects and dependency of wind power penetration on system frequency regulation. *IEEE Transactions on Sustainable Energy*, 7(1):354–363.

Nguyen, N. and Mitra, J. (2016). Effect of wind power on load frequency control. In *2016 IEEE Power and Energy Society General Meeting (PESGM)*, pages 1–5. IEEE.

Notton, G., Nivet, M.-L., Voyant, C., Paoli, C., Darras, C., Motte, F., and Fouilloy, A. (2018). Intermittent and stochastic character of renewable energy sources: Consequences, cost of intermittence and benefit of forecasting. *Renewable and Sustainable Energy Reviews*, 87:96–105.

Oliver, M. E., Mason, C. F., and Finnoff, D. (2014). Pipeline congestion and basis differentials. *Journal of Regulatory Economics*, 46(3):261–291.

Østergaard, P. A. (2006). Ancillary services and the integration of substantial quantities of wind power. *Applied Energy*, 83(5):451–463.

Petersen, C., Reguant, M., and Segura, L. (2022). Measuring the impact of wind power and intermittency. *SSRN 4291672*.

Seel, J., Millstein, D., Mills, A., Bolinger, M., and Wiser, R. (2021). Plentiful electricity turns wholesale prices negative. *Advances in Applied Energy*, 4:100073.

Singarao, V. Y. and Rao, V. S. (2016). Frequency responsive services by wind generation resources in United States. *Renewable and Sustainable Energy Reviews*, 55:1097–1108.

Skea, J., Anderson, D., Green, T., Gross, R., Heptonstall, P., and Leach, M. (2008). Intermittent renewable generation and the cost of maintaining power system reliability. *IET generation, transmission & distribution*, 2(1):82–89.

Sortomme, E., Irani, R., Nguyen, H., and El-Sharkawi, M. (2010). Estimating energy imbalance due to integration of wind power. In *IEEE PES General Meeting*, pages 1–6. IEEE.

Stock, J. H. and Yogo, M. (2005). Testing for weak instruments in linear iv regression. *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, page 80.

Stoft, S. (2002). *Power System Economics: Designing Markets for Electricity*. IEEE Press, New Jersey.

Vandezande, L., Meeus, L., Belmans, R., Saguan, M., and Glachant, J.-M. (2010). Well-functioning balancing markets: A prerequisite for wind power integration. *Energy Policy*, 38(7):3146–3154.

Wan, Y.-h., Milligan, M., and Kirby, B. (2007). Impact of energy imbalance tariff on wind energy. Technical report, National Renewable Energy Lab (NREL), Golden, CO (United States).

Woo, C.-K., Horowitz, I., Moore, J., and Pacheco, A. (2011). The impact of wind generation on the electricity spot-market price level and variance: The Texas experience. *Energy Policy*, 39(7):3939–3944.

Wooldridge, J. M. (1995). Score diagnostics for linear models estimated by two stage least squares. *Advances in econometrics and quantitative economics: Essays in honor of Professor CR Rao*, pages 66–87.

Yang, B., Zhong, L., Wang, J., Shu, H., Zhang, X., Yu, T., and Sun, L. (2021). State-of-the-art one-stop handbook on wind forecasting technologies: An overview of classifications, methodologies, and analysis. *Journal of Cleaner Production*, 283:124628.

Zugno, M., Jónsson, T., and Pinson, P. (2013). Trading wind energy on the basis of probabilistic forecasts both of wind generation and of market quantities. *Wind Energy*, 16(6):909–926.