Market Inefficiency and Environmental Impacts of Renewable Energy Generation

Hyeongyul Roh

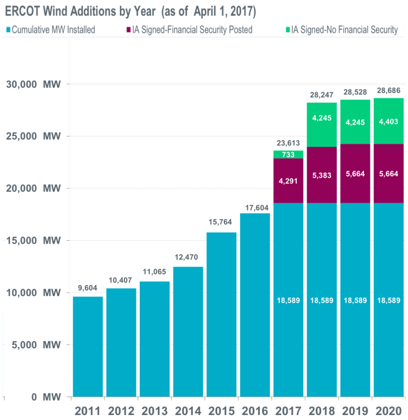
North Carolina State University

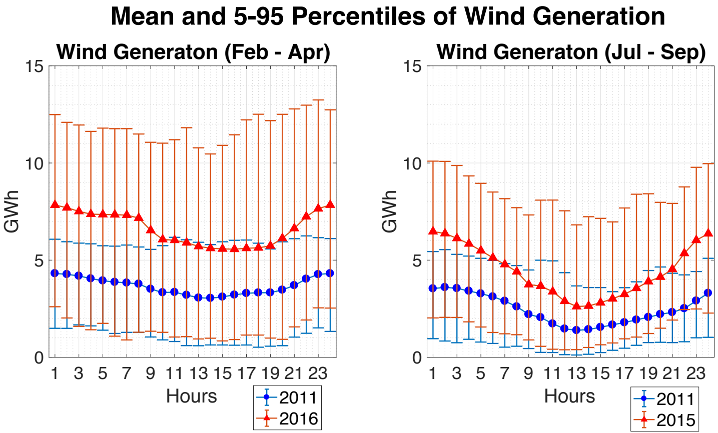
hroh@ncsu.edu

August 8, 2017

Recent Dramatic Changes to the U.S. Electricity Market

- **Increase in Renewable (wind) Capacity** involves more volatility with relatively less increase in mean. Inverse relationship between wind and load incurs another seasonal and diurnal volatility.





Motivation: Wind Intermittency & Market Inefficiency

- Channels that intermittency affects social costs: additional start up costs, more frequent ramping up, backup capacity, and operating reserve:

  $\Rightarrow$ Few papers empirically investigate the market inefficiency.

  - Intermittency accounts for $46.00 per MWh for 20 percent solar generation (Gorwrisankaran et al. (JPE, 2016))
  - Renewable 'curtailments' were at record levels in March 2017 in California, amounting to over 80 GWh, which is more than a typical day's worth of solar production that month. (Catherine Wolfram, Blog posted on April 24, 2017)

- I develop a method to quantify the market inefficiency with **increased forward premiums** (day-ahead minus spot market prices) and **inefficient dispatch** and compare them to **environmental benefits**.

Motivation: Surge of Interests in Machine Learning

- There has been **a surge of interest in Machine Learning methods (ML) in economics.**
    - Athey and Imbens (2015), Athey and Imbens (2017), Bajari et al. (2015), Belloni et al. (2017), Burlig et al. (2017), Chernozhukov et al. (2017), Kleinberg et al. (2015), Mullainathan and Spiess (2017), Varian (2014)
- ML flexibly control for a large number of covariates as part of an estimation strategy.
- ML typically rely on data-driven model selection.
    - *"When causal interpretations of estimates are more plausible, and inference about causality can reduce the reliance of these estimates on modeling assumptions (like those about functional form), the credibility of policy analysis is enhanced."*, Athey and Imbens (2017)
- I deploy dynamic neural networks in machine learning method.
    - Ultimately, the method returns a **forecasting model**, which is used to generate a **counterfactual data series**.

Research Question I

1. **Estimate Forward Price Premiums (forward price - spot price, FPP) caused by wind generation increase**
    - A lower premium on average has been associated to higher efficiency, sending accurate signals for generation planning
    - Previous studies address the rationales behind systematic FPP in electricity market by: **(a).** Market Power
        - Negative FPP with Monopsony power (Borenstein et al. (2008)).
        - Positive FPP with Monopoly power (Ito and Reguant (2016)).
    - **(b).** Transaction Costs and Limits to Arbitrage
        - California market: Jha and Wolak (2014)
        - Midwest electricity market (MISO): Mercadal (2016) & Birge et al. (2017)
    - **(c).** Risk Premium
        - Bessembinder and Lemmon (2002) & Longstaff and Wang (2004)
    - **My results**: **Increased wind generation increases FPP**, due to integration costs of intermittent wind.

Research Question II & III

2. **Estimate changes in structure of economic dispatches**
    - For the grid stability, system operator relies almost entirely on controlling supply (exogenous and inelastic demand).
      $\Rightarrow$ Increased wind potentially requires more expensive alternative supply options.
    - **My results**: Wind generation **increases market share of expensive and less efficient units** (such as Simple Cycle), while **reducing the share of the most efficient units** (such as Combined Cycle) to satisfy decreased but more volatile net load

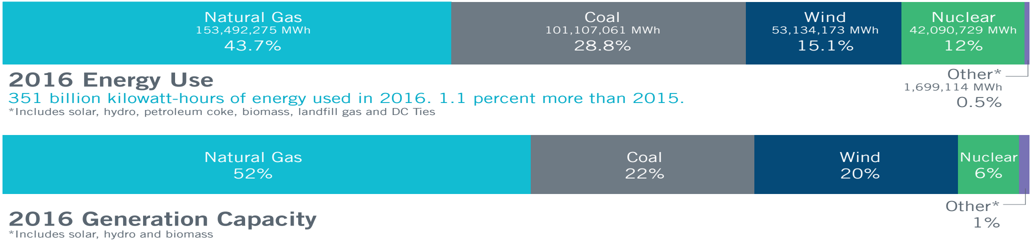
3. **Estimate external benefits (emission reduction) provided by wind generation to quantify the tradeoff**
    - Cullen (2013), Fell and Kaffine (2017), and Novan (2015)
    - **My results**: 15% wind market share from 1% induces **10∼14% emission reductions** for CO2, NOx, and SO2, depending on natural gas prices.

Identification Strategy: Alternative way to derive a causal inference with ML

1. Employ dynamic neural networks (recurrent time series model) to find prediction functions for hourly nodal prices and quantities.
2. Construct counterfactual paths for the prices and quantities in forward and spot markets based on the prediction functions.
3. Take difference between counterfactual paths for the **ATE of wind generation increase on the market prices and quantities**.
4. Classify transmission nodes for analysis of heterogenous impacts over (a) firms (market power), (b) generator types (limits to arbitrage), (c) locations (transmission constraints).

Electricity Market in Texas: ERCOT

- 570+ generating units owned by ∼70 firms
  ⇒ Conventional generators are owned by 40 firms
- Market share of four biggest firms with fossil fuel fleets is approximately **45%**
  ⇒ **65%** among dispatchable fossil fueled fleets
- Relatively large share of natural gas power plants and wind farms
  ⇒ U.S. average is 33.8% for natural gas and 5.5% for wind as of 2016

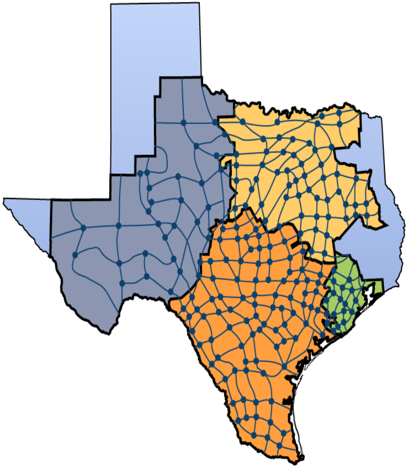


ERCOT: Sequential Markets

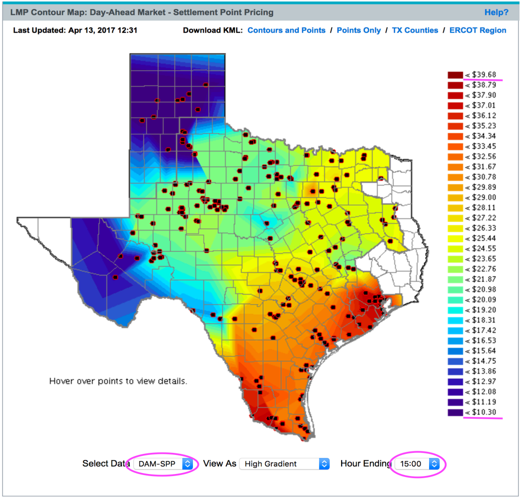
- **Day-Ahead Market (DAM)** operations run the day before the operating day to ensure reliability of transmission grid.
  - 10AM-1:30PM: DAM is cleared for next day (00:00-23:00).
  - The volume of day-ahead purchases was approximately **51% of real-time load** as of 2015.
- **Real-Time Market (RTM)** runs a market clearing process at least every five minutes in operating period (immediately before operation) to match generation to load.

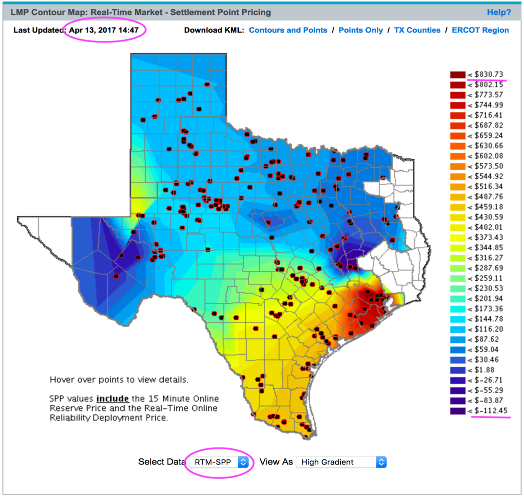
ERCOT: Nodal Markets

- ERCOT market uses **a network operation model.**
    - 4,000 points of interconnection with system may be an **energy source (injection point)**, sink (withdrawal point), or switching station (transformation point).
- **Locational Marginal Pricing (LMP)** establishes a price per MW at a given network node.



ERCOT: Heterogenous LMP Examples



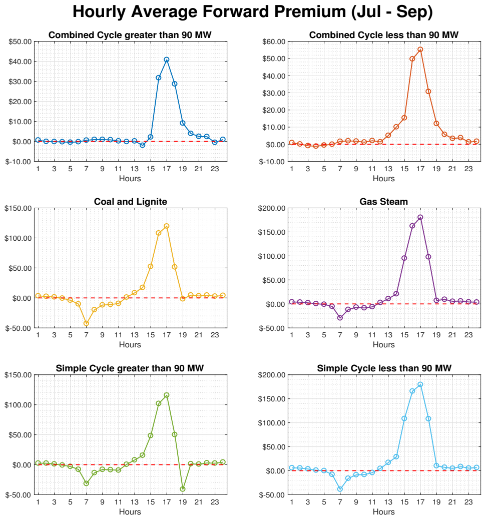


Summary of Data

Hourly data from June, 2011 - June, 2016 (43176 observations) obtained via three sources

1. Publicly available data sets from ERCOT
   - Hourly settlement point prices and bid awards
   - Backcasted hourly demands at sink nodes
   - Hourly zonal demand
2. Data obtained by the energy data service company Ventyx
   - System-wise hourly generations: Wind, Nuclear, and Others ($\leq 1\%$)
   - Coal and natural gas prices sold to power plants in ERCOT
3. Data obtained by NOAA
   - Hourly dummies for extreme weather events
   - Regional hourly temperatures in Texas
4. Other covariates
   - 59 dummy variables for time effects
   - To capture the dynamic response, the lagged dependent and independent variables are also added $\Rightarrow$ total # of covariates $\sim$ 10,000.

Forward Price Premium from Data by Hours

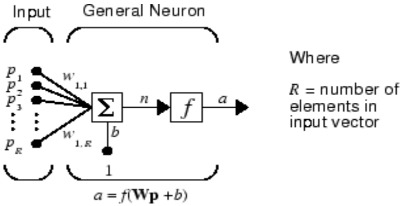


Machine Learning Method: Artificial Neural Networks

- A single neuron for input vector P:

$$a_{w,b}(P) = f(W^T P) = f(\sum_{i=1}^{R} w_i p_i + b)$$

where w_i is weight for i^{th} input,
b denotes network bias term,
and f can be any differentiable transfer function.



- General function approximation method may employ a number of neurons and layers.
- Complementary tools and additional settings
  1. k-fold cross validations with random sample splitting and early stopping
     $\Rightarrow$ Trade off bias and variance, and penalize overfitting
  2. Normalization of inputs/outputs to fall in a certain range (ex) [-5, 5]

ML Estimation Model: Step I

- DAM is cleared at 13:30 for operation hours next day (00:00 - 23:00)

$$[P_{Hour0,Day1,i}, \dots, P_{Hour23,Day1,i}] = F(P_{Hour0,Day0,i}, \dots, P_{Hour23,Day0,i},\ X_{Hour13,Day0}, \dots, X_{Hour12,Day-1})$$

$$\Rightarrow \boldsymbol{P_{HourN,DayT,i}} : \textbf{Nodal Price} \text{ at N:00 of Day T on node i}$$

$$\Rightarrow \boldsymbol{X_{HourN,DayT}} : \text{A set of 389 covariates at N:00 of Day T}$$

- RTM is cleared every hour

$$P_{HourN,i} = F(P_{HourN-1,i}, \dots, P_{HourN-6,i},\ X_{HourN}, \dots, X_{HourN-23})$$

$$\Rightarrow \boldsymbol{P_{HourN,i}} : \textbf{Nodal Price} \text{ at N:00 on node i}$$

$$\Rightarrow \boldsymbol{X_{HourN}} : \text{A set of 389 covariates at N:00}$$

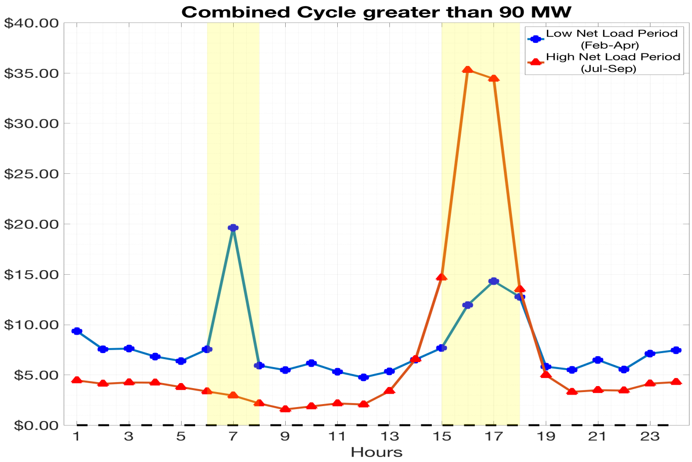
- Employ dynamic neural network (recurrent time series model) to find data-driven prediction function F (non-linear)
- **Multistep closed-loop predictions** to construct counterfactual paths from initial conditions after training the model F with **open-loop form**.

ML Estimation Model: Step II

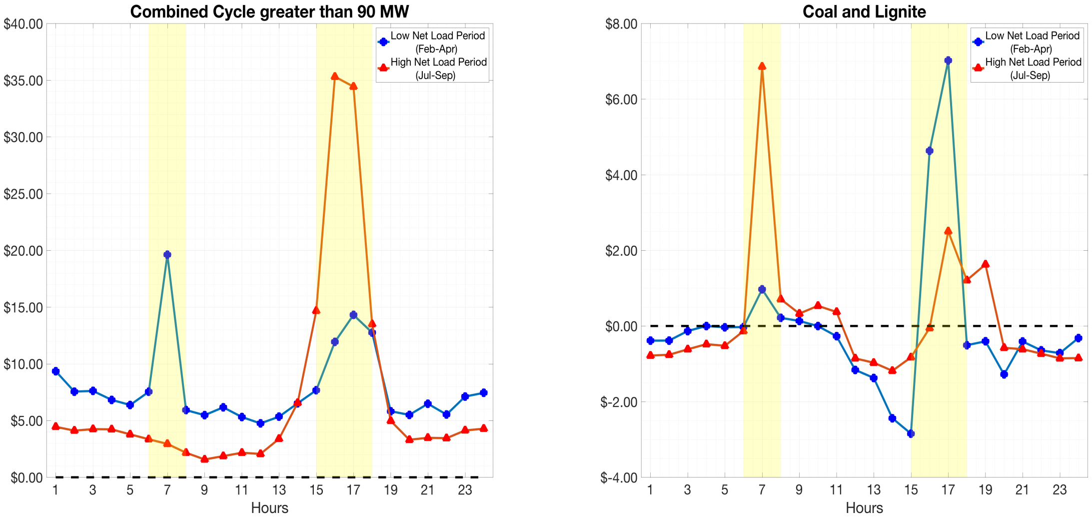
- The counterfactual time series describe what would have happened to electricity prices and supplies if there had been a variety of wind market shares.
- Calculate FPP time paths based on predicted time series above
  - FPP with low wind at node i : $F\hat{P}P_i^L = \hat{P}_i^{DAM_L} - \hat{P}_i^{RTM_L}$
  - FPP with high wind at node i : $F\hat{P}P_i^H = \hat{P}_i^{DAM_H} - \hat{P}_i^{RTM_H}$
- **ATE of wind generation increase on FPP at node i**:
  - $\hat{\beta}_i = E[F\hat{P}P_i^H - F\hat{P}P_i^L]$
- Classify transmission nodes for analysis of heterogenous impacts over (a) firms (market power), (b) generator types (limits to arbitrage), (c) locations (transmission constraints).

Results: (a) Forward Premium Changes from 1 to 15 % Wind Shares: CC

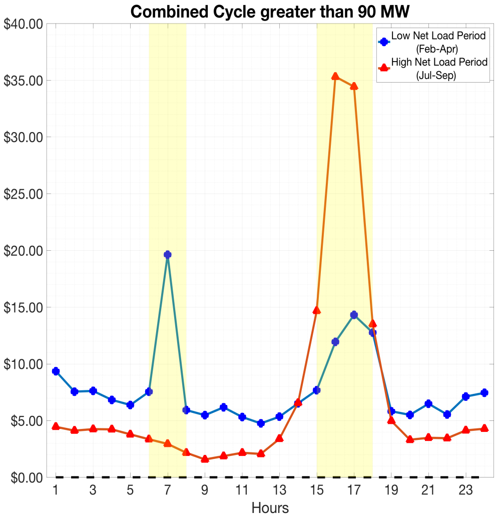
- The forward premium increases especially during certain hours of a day

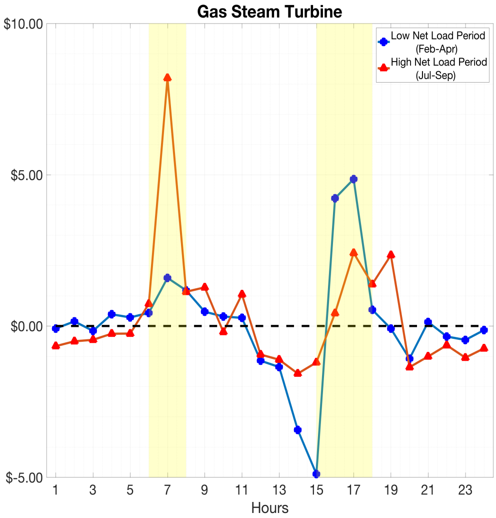


Results: (a) Forward Premium Changes from 1 to 15 % Wind Shares: CC vs. Coal

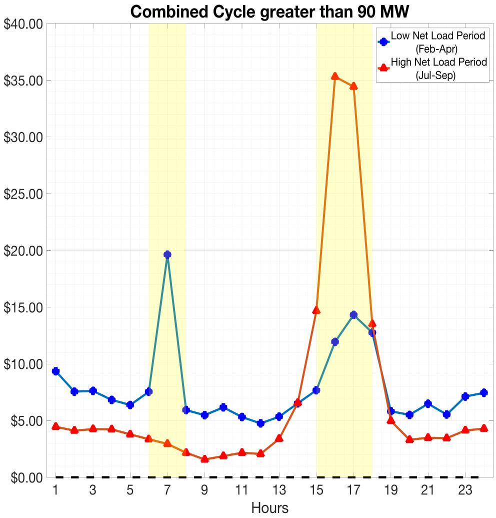


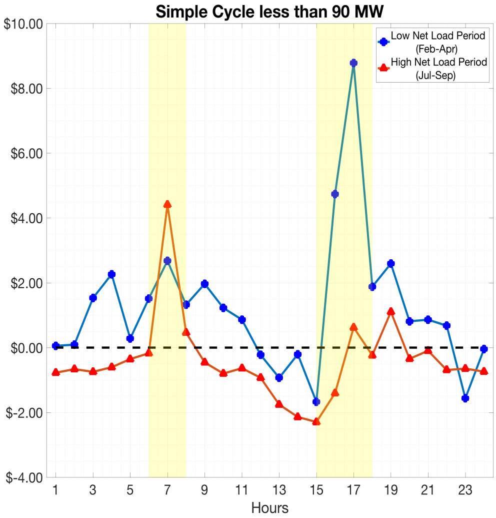
Results: (a) Forward Premium Changes from 1 to 15 % Wind Shares: CC vs. GS





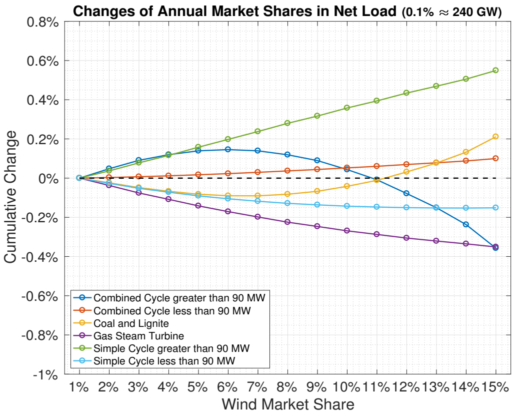
Results: (a) Forward Premium Changes from 1 to 15 % Wind Shares: CC vs. SC (small)



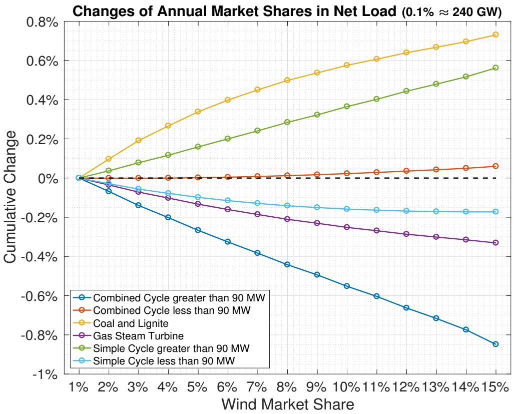


Results: (b) Economic Dispatch with Variable Wind Market Shares

- More deployment of SC units while less use of CC to satisfy decreased, but more volatile net load



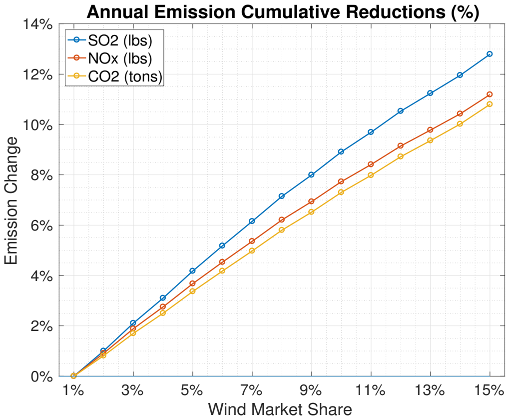
With Low NG Prices



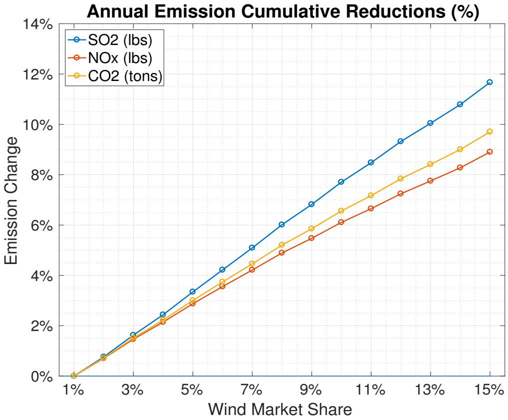
With High NG Prices

Results: (c) Emissions Avoided by Variable Wind Market Shares

- With 15% wind market shares, emission has been mitigated by 11∼13% (9∼12%)



With Low NG Prices



With High NG Prices

Conclusion

- Discussion
  1. Increased wind generation increases the forward premium especially during certain hours of a day, likely due to increased fluctuation of net load.
     $\Rightarrow$ Either Market power or Limits to arbitrage during certain hours of a day?
  2. A generation technology, rather than firm sizes, incurs heterogeneous implication of wind generation increase because it has different production cost convexity and faces different residual demand volatility.
     $\Rightarrow$ Finding detailed evidence is in progress.
  3. Inefficient dispatches with the increased market volatility.

- Future Research
  1. Find the rationale behind the results with hourly supply responses (market power or limits to arbitrage).