Course notes for EE394V
Restructured Electricity Markets:
Market Power

Ross Baldick

3

Empirical studies of market power

- This material is based on:
  - Catherine D. Wolfram, "Measuring duopoly power in the British electricity spot market," *The American Economic Review*, 89(4):805–826, September 1999.
  - Paul L. Joskow and Edward Kahn, "A quantitative analysis of pricing behavior in California's wholesale electricity market during summer 2000," *The Energy Journal*, 23(4):1–35, 2002.

– Scott M. Harvey and William W. Hogan, “Market power and market simulations,” manuscript, available from www.whogan.com, July 16, 2002.
– Severin Borenstein, James Bushnell, and Christopher R. Knittel, “Market Power in Electricity Markets: Beyond Concentration Measures,” *Energy Journal*, 20(4):65–88, 1999.
– Ali Hortaçsu and Steven L. Puller, “Understanding Strategic Bidding in Multi-Unit Auctions: A Case Study of the Texas Electricity Spot Market,” *RAND Journal of Economics*, 39(1):86–114, Spring 2008.
– Frank A. Wolak, “Identification and Estimation of Multi-Output Cost Functions Using Bid Data from Electricity Markets,” March 2004.
– Potomac Economics, Ltd., “2005 State of the Market Report for the ERCOT Wholesale Electricity Markets,” July 2006, available from www.puc.state.tx.us/WMO/documents/annuual_reports/2005annualreport.pdf
– Potomac Economics, Ltd., “Investigation of the wholesale market activities of TXU from June 1 to September 30, 2005,” March 2007, available from www.puc.state.tx.us/about/reports/2005_TXU_Investigation_IMM_Cover.pdf and Revised version, September 2007.

– Public Utility Commission of Texas, Project Number 31972, "Rulemaking on Wholesale Electric Market Power and Resource Adequacy in the ERCOT Power Region, Order Adopting Amendment to section 25.502, New section 25.504 and new section 25.505 as Approved at the August 10, 2006, Open Meeting."
– John Ning Jiang and Ross Baldick, "Distinguishing Design Flaws From Misconduct: A New Approach to Electricity Market Analysis," *IEEE Transactions on Power Systems*, 20(3):1257–1265, August 2005.

Outline

(i) The England and Wales market in 1992–1994,
(ii) The California market in 2000,
(iii) The ERCOT balancing market in 2001–2003,
(iv) The ERCOT balancing market in 2005,
(v) Reconciling the "economic" model to the "commercial" model,
(vi) Distinguishing design flaws from market power.

3.1 The England and Wales market in 1992–1994

(i) England and Wales industry,
(ii) The pool,
(iii) Regulatory oversight,
(iv) Empirical framework,
(v) Marginal cost and capacity data,
(vi) Mark-ups and price-cost margins,
(vii) Why are the prices lower than the theoretical models?
(viii) Effect of price cap,
(ix) Conclusion.

3.1.1 England and Wales industry

- Peak demand in early 1990s of around 50,000 MW.
- Generation and transmission owned and operated by single governmental entity, “Central Electricity Generating Board” (CEGB) until April 1990.
- Subsequently, transmission assets and “pool operation” transferred into the “National Grid Company” (NGC).
- Generation assets divided amongst three companies:
  - Nuclear Energy, about 9,000 MW (remained in public sector for six years),
  - National Power, about 30,000 MW (privatized), and
  - PowerGen, about 20,000 MW (privatized).
- Additional supply from Scotland and France of around three thousand MW.
- Pumped storage hydroelectric facilities owned by NGC.

England and Wales industry, continued

- Is the market concentrated?
- Using "capacity" HHI literally (ignoring pumped storage):
  - Nuclear energy, 0.14 share of capacity,
  - National Power, 0.48 share of capacity,
  - PowerGen, 0.32 share of capacity,
  - Scotland and France combined, 0.05 share of capacity,
- HHI is 0.36 (or 3600 $\%^2$).
- Regulatory authorities in the United States typically use around 0.2 (or 2000 $\%^2$) as the threshold for a concentrated market.
- However, since nuclear and Scotland and France were essentially never marginal, the England and Wales market is often characterized as a duopoly, with National Power and PowerGen as the important protagonists:
  - analogous to situation in homework in on-peak condition if baseload and intermediate units are at capacity so that only peaking units are competing.
- Duopoly-based HHI is even higher!

England and Wales industry, continued

- Twelve "regional" distribution companies remained, renamed "regional electricity companies" (RECs).
- Figure 1 on page 807 of Wolfram shows the change in ownership structure in 1990.
- Focus of paper is the period 1992–1994.
- There was an additional 7,000 MW of new entry between 1990 and 1994, most of it combined-cycle gas turbines.
- During study period of 1992–1994, coal was "marginal" generating source for about 80% of time:
  - very different to California and ERCOT, where gas is usually marginal source,
  - later there was great increase in gas-fired generation in England and Wales, "the dash for gas."
- In 1996 and 1999, there were divestitures (sales) of some of National Power's and PowerGen's capacity to other firms:
  - will consider the effect of this in a later analysis.

3.1.2 The pool

- Day-ahead offer-based economic dispatch based on NGC forecast of demand in each half-hour (day-ahead market abolished in 2000).
- Differences between real-time demand and forecast handled by “adjustment” in real-time, with additional costs “uplifted” to energy price.
- “Vesting contracts” were set up by government between generators and RECs in 1990:
  - these were forward financial contracts with contract periods ranging from one to three years,
  - accounted for about 85% or 90% of National Power and PowerGen capacity.
  - all vesting contracts expired by March 1993.
- Some contracts were subsequently negotiated to replace expiring vesting contracts, but only accounting for about 50% of capacity:
  - contract cover decreased over study period, providing an opportunity to see empirically whether contract cover had an effect on prices.

3.1.3 Regulatory oversight

- Limited direct regulation of the market except through referral to anti-trust enforcement agency.
- In 1994, regulator agreed to *not* refer to anti-trust agency if National Power and PowerGen agreed to:
  - divestiture (sale of capacity to other parties), and
  - price cap.
- A price cap on the *average* price over time was put in place in April 1994:
  - different from typical implementation of price cap in North America where price cap is a limit on highest price in market,
  - somewhat similar to "peaker net margin" provisions in place in ERCOT where average earnings over time are partially limited based on infra-marginal profits for peaker compared to annualized cost for peaker.
- Divestitures occurred in 1996, after study period.

3.1.4 Empirical framework

- Large end-users bought directly from the pool or had contracts with retailers (contracts for differences based on pool prices).
- End-users with on-site generation and flexible operations could respond to prices or projections of prices and therefore are potentially price-responsive demand.
- Recall the results of the Cournot model:
  - Generation-weighted average price cost margin is $L = \sum_j (s_j)^2/e$.
  - If there are n firms each with same costs, a "symmetric Cournot oligopoly," then $s_j = 1/n$ for each firm and:

$$\begin{aligned} L &= \sum_j (s_j)^2/e, \\ &= \sum_j (1/n)^2/e, \\ &= 1/ne. \end{aligned}$$

  - Re-arranging, $\sum_j (s_j)^2 = Le = 1/n$, the generation-weighted average price cost margin times the elasticity.

Empirical framework, continued

- Instead of focusing on the HHI, $\sum_j (s_j)^2$, as a predictor of Le, Wolfram estimates $\theta = Le$ directly based on:
  - observations of prices and data about marginal costs to calculate the price-cost margin L, and
  - estimate of e.
- The value of θ estimates the effective level of competition, given profit maximization by all firms:

  $\theta = 0$ means no price mark-up and therefore perfect competition,
  $\theta = 1$ is consistent with monopoly pricing,
  $\theta = 1/n$ if there are n firms in a "symmetric Cournot oligopoly."
- Still requires estimation of e.

3.1.5 *Marginal cost and capacity data*

- Generator heat rates of much of the capacity is known from historical data published prior to privatization.
- Main exception is that combined-cycle gas turbine plants were built subsequent to privatization and were assumed to have 45% efficiency from thermal energy to electricity (approximately 7,500 Btu/kWh heat rate).
- Coal prices were based on limited information about contract prices.
- Other fossil fuel prices based on published market fuel prices.
- Nuclear assumed to have marginal cost around 11 to 13 £/MWh:
  - apparently above actual operating costs,
  - but even at this cost were always baseloaded.
- Capacity for each generator each month in the model based on the daily "stated" capacity:
  - capacity set equal to the mean plus one-half standard deviation of the daily "stated" capacity during that month.

3.1.6 Mark-ups and price-cost margins

- Figure 2 on page 812 of Wolfram shows observed prices and estimated industry marginal cost curve for January 1993 and July 1993.
- Mark-ups range from around one or two £/MWh to around £30/MWh, with typical values in the £5/MWh to £15/MWh range.
- Table 1 on page 813 of Wolfram shows:

  (i) the empirical generation-weighted average price-cost margin,

  (ii) θ, the empirical generation-weighted average price-cost margin times the elasticity, and

  (iii) theoretical estimate of θ, based on "highest priced SFE" (least-competitive supply function equilibrium, to be discussed in later lectures) and an assumed demand slope of 500 MW/(£/MWh)).

- Demand slope corresponds to elasticity of $e = 0.17$:
  - not based on direct measurement or observation,
  - elasticity below one, so Cournot model predicts incentives to increase prices above observed prices!

Mark-ups and price-cost margins, continued

- Typical price-cost margin of 0.25 in first few entries of first column of Table 1 of Wolfram implies prices about one-third higher than marginal.
  - Roughly consistent with Figure 2 on page 812 of Wolfram.
- Price-cost margin of 0.5 in "March 1994, Above median" entry of first column of Table 1 implies prices about double marginal.
  - Regression of price with respect to marginal cost also suggests prices nearly double marginal.
- Empirical values of $\theta \approx 0.05$ in first few entries of second column of Table 1 of Wolfram correspond to symmetric Cournot oligopoly with $n = 20$, whereas there were only $n = 2$ main generating firms.
- Theoretical values of $\theta \approx 0.3$ from highest priced SFE model in first few entries of third column of Table 1 of Wolfram are higher than empirical values by about an order of magnitude.
- Both theoretical models–symmetric Cournot oligopoly and highest priced SFE–predict much higher prices than actually observed!

3.1.7 Why are the prices lower than the theoretical models?

- Neither the symmetric Cournot oligopoly nor highest priced SFE model matches the empirical data:
  - empirically, prices are well above marginal cost, but
  - symmetric Cournot oligopoly and highest priced SFE would predict even higher prices (infinite for Cournot, given elasticity less than one).
- Possible explanations:

  (i) Cournot oligopoly model does not represent the market well:
    - As discussed previously, Cournot ignores issues such as forward contracts.

  (ii) Highest price SFE model does not represent the market well:
    - Will discuss other SFE models of England and Wales in subsequent lectures that may better match the data.

Why are the prices lower than the theoretical models? continued

(iii) Forward contracting not represented:

- However, significant changes in level of contracting over study period corresponded to only small changes in prices,
- Undercuts significance of forward contracting on prices, at least in the context of this market.

(iv) Threat of regulatory action:

- Consistent with changes in prices before and after publication of “regulator’s statements,” but difficult to model.

(v) Threat of entry:

- Average prices are just below capital and operating costs of new combined-cycle gas turbines.

(vi) Errors in base data:

- Results mostly driven by assumptions about coal cost and elasticity of demand.

3.1.8 *Effect of price cap*

- In return for not being referred to the anti-trust authority, National Power and PowerGen agreed to caps on both:

  (i) the simple average of prices, and
  (ii) demand weighted average of prices.

- Caps on average prices in force from March 1994.
- The second cap was not apparently binding.
- To maximize profits, given a binding limit on the simple average of prices, maximize the demand-weighted average of prices:
  - decrease prices at low demands, but
  - increase prices at high demands.
- Figure 3 on page 819 of Wolfram shows schematically this change.
- Consistent with observed changes in prices pre- to post-March 1994:
  - Prior to March 1994, price-cost margin slightly higher at high demand that at low demand.
  - After March 1994, price-cost margin much higher at high demand and very low for low demand.
- Suggests that firms had considerable ability to control prices!

3.1.9 Conclusion

- Empirical evidence suggests prices above marginal costs,
- Theoretical models presented in paper do not match empirical results,
- Ability to adjust prices to maximize profits given price cap suggests that firms kept offers low to avoid regulatory scrutiny.
- Problematic for predictive modelling of firm behavior.

3.2 The California market in 2000

(i) Introduction,
(ii) California market design,
(iii) California supply,
(iv) Divestiture,
(v) Summer supply–demand,
(vi) Retail restructuring,
(vii) Independent generators and demands,
(viii) Focus of study,
(ix) Estimation of competitive prices,
(x) Simulation features and assumptions,
(xi) Results,
(xii) Opposing view,
(xiii) Empirical analysis of withholding,
(xiv) Conclusion.

3.2.1 Introduction

- Review Joskow and Kahn, one of several studies of the California electricity market “crisis” of 2000.
- Focus is on Summer of 2000.
- Wholesale prices in California were persistently many times higher in June 2000 through June 2001 than they had been in 1998 and 1999.
- Changes in supply and demand contributed to the price increase.
- Basic goal of Joskow and Kahn is to see if *simulated* competitive offers and *simulation* of market do or do not explain actual prices.
- Chose to rely on publicly available data for simulation.
- Quantification of some uncertainties in simulation.
- Additionally, characterize withholding given actual prices.
- An opposing view was presented by Harvey and Hogan.

3.2.2 California market design

- The restructured California market opened in April 1998.
- The California ISO (CAISO) operated:
  - ancillary services markets, and
  - "imbalance" energy (that is, "balancing" or "real-time" market), including zonal "congestion management;" that is, adjustment of generation to avoid over-loading inter-zonal transmission constraints.
- The California power exchange (PX) operated day-ahead auction markets for each hour of the next day:
  - auction did not consider start-up issues,
  - generators offers had to be increased above marginal costs to assure recovery of start-up and min-load costs.

California market design, continued

- The bifurcation of the operator of the day-ahead auction from the ISO differs from other markets, including the ERCOT nodal market, where the ISO will operate both the day-ahead and real-time markets:
  - bifurcation makes efficient transmission management extremely difficult since PX auction did not consider transmission,
  - market participants were not "incented" to help with congestion management,
  - incentives, in fact, are to exacerbate congestion in order to be paid to relieve it!

3.2.3 California supply

- In-state:
  - nuclear,
  - hydro,
  - gas steam,
  - gas turbine,
  - cogeneration, wind, and other “qualifying facilities” (QFs).
- Out-of-state:
  - nuclear,
  - hydro
  - coal.

3.2.4 Divestiture

- Three incumbent “investor-owned utilities” (IOUs) owned the in-state nuclear, hydro, and gas generation.
- Gas (about 18,000 MW) represents about half of the in-state capacity and was sold by incumbents to five unaffiliated power companies in 1998 and 1999:
  - Duke,
  - Mirant,
  - AES/Williams,
  - Dynegy,
  - Reliant.
- Incumbents kept the nuclear and hydro and their contracts with QFs:
  - about half of the in-state capacity,
  - but hydro has limited energy capability because of reservoir and inflow characteristics.
- Neither incumbents nor independents completed any new generation from time of divestiture until after 2000.
- Most of the gas generation is from the 1960s and 1970s.

Divestiture, continued

- Borenstein, Bushnell, and Knittel report "capacity" HHIs of around 0.1 (1000 $\%^2$) for the California market:
  - lower than HHIs for ERCOT!
- Using the data from Table 7 on page 23 of Joskow and Kahn for the five independents yields a "capacity" HHI of around 0.2 (2000 $\%^2$).
- Market concentration evidently lower than in England and Wales market based on "capacity" HHI.

3.2.5 Summer supply–demand

- Peak demand in 2000 was about 53,000 MW.
- Imports are necessary from other states during peak to meet demand.
- During Summer, marginal supply is typically gas.
- Figure 1 on page 7 of Joskow and Kahn shows marginal costs for in-state gas resources for two representative costs of gas.
- Marginal cost increases rapidly at "top" of supply.
- Marginal cost increases with cost of NO_x emissions credits.
- Marginal costs (and hence prices) will be high whenever "top" of supply is dispatched:
  - key question is whether market power increases prices above what would be high competitive prices,
  - difficulty of distinguishing high competitive prices from even higher prices if there is exercise of market power.

3.2.6 Retail restructuring

- Retail consumers could either:
  - stay as customer of incumbent IOU with "default service" regulated energy price of around $60/MWh, or
  - become customer of unregulated "electricity service provider" (ESP).
- About 85% of retail demand remained with IOUs:
  - more remained with IOUs than had initially been anticipated,
  - retail restructuring just not that exciting to customers!
- Restructuring rules required IOUs to serve default service customers by purchases from PX and ISO markets.
- IOUs were required to offer their generation into the PX and ISO markets.
- Most wholesale trade took place through PX and ISO.

Retail restructuring, continued

- Until Fall 2000, forward contracting was essentially not permitted under restructuring rules:
  - This left IOUs with large, unhedged retail obligations, as in Stoft five generator example with retailers that were unaffiliated and did not have contracts with generators.
  - Stoft example is stylized version of California crisis!
  - Extremely different to England and Wales where vesting contracts were explicitly set up as part of privatization.
- Short run elasticity of demand nearly zero due to lack of metering and lack of exposure to real-time prices.
- In hindsight, the lack of forward contracts and the lack of short run elasticity was a prescription for disaster!
  - forward contracts are very important to functioning of markets.

3.2.7 Independent generators and demands

- Could enter into forward contracts with entities other than IOUs.
- Could sell energy to and buy energy from PX and ISO.
- Could sell ancillary services to ISO.
- Could self-supply ancillary service obligations.

3.2.8 *Focus of study*

- PX trade during Summer of 2000:
  - prices prior to Summer 2000 were much lower, and
  - prices after Summer 2000 were affected by changes in market rules, increases in gas prices, gas shortages, IOU credit problems, generation outages that complicate analysis.
- Ignore forward markets and real-time markets:
  - prices roughly similar in all markets,
  - essentially no forward contracting by IOUs.
- Ignore transmission congestion:
  - not very significant during this period.
- Ignore ancillary services:
  - in later "withholding" analysis, will consider possibility that generator is not operating at capacity because it is "supplying" ancillary services.

Focus of study, continued

- Table 1 on page 9 of Joskow and Kahn shows:
  - actual demand-weighted average DA PX prices each month in 1998, 1999, and 2000, and
  - forecast of prices by California Energy Commission (CEC) for 2000.
- Actual prices significantly different from CEC forecast for months from Summer 2000 onwards.

3.2.9 Estimation of competitive prices

- Simulate “competitive” energy prices by clearing demand against marginal cost to examine whether actual prices can be explained by increases in demand and gas prices and other “market fundamentals.”
- Ignore revenues from AS markets.
- Ignore distinction between day-ahead and real-time.
- Ignore start-up and shut-down issues.
- Process ignores a number of issues in electricity market:
  - susceptible to criticism that ignored issues may have significantly increased competitive prices compared to estimates.

3.2.10 Simulation features and assumptions

- Demand:
  - modelled as 100 “slices” in each month, ignoring “chronological” issues so that no start-up and shut-down represented,
  - increased by 3% compared to actual to approximately represent demand for AS capacity,
  - higher in Summer 2000 than in previous Summers.
- No planned maintenance of in-state generation during Summer.
- Forced outage rate for gas based on historical data then used to adjust industry marginal cost curve.
- Actual monthly hydro production shared amongst demand slices on basis of:
  - minimum and maximum production levels, and
  - assign hydro to high demand levels “peak shaving” consistent with minimum and maximum levels.
- Less hydro available than in previous years.
- Nuclear modelled as full capacity.

Simulation features and assumptions, continued

- Gas-fired generation:
  - gas prices based on recorded prices (\$3 to \$7 per million Btu).
  - heat rates based on commercial database (up to approximately 17,000 Btu per kWh, with some having higher heat rates).
  - NO_x emissions prices based on estimate of monthly average (from \$1 to \$10 per pound).
  - emissions rates based on public data, regulatory filings, and commercial database (up to approximately 1 pound of NO_x per MWh, with some having higher emissions rates), and
  - forced outage rates based on historical data (from 6% to 13%)
- Nominal capacity derated to approximately account for forced outages.
- Emissions add as much as tens of dollars per MWh or more to marginal generation cost as shown in Figure 1 on page 7 of Joskow and Kahn.

Simulation features and assumptions, continued

- Imports modelled as "elastic supply" referenced to actual realized imports Q^{import} and actual realized import price P^{import}:

$$q^{\text{import}}(P) = \left(P/P^{\text{import}}\right)^{\eta} Q^{\text{import}}.$$

- Note that:

$$\begin{aligned}
\frac{\partial q^{\text{import}}}{\partial P}(P) &= \eta \frac{P^{\eta-1}}{(P^{\text{import}})^{\eta}} Q^{\text{import}}, \\
&= \eta \left(P/P^{\text{import}}\right)^{\eta} Q^{\text{import}}/P, \\
&= \eta q^{\text{import}}(P)/P.
\end{aligned}$$

- That is,

$$\eta = \frac{\partial q^{\text{import}}}{\partial P}(P) \frac{P}{q^{\text{import}}(P)},$$

  is the sensitivity of import quantity to price, normalized by price and quantity.

Simulation features and assumptions, continued

- The parameter η is called the "elasticity of supply" and is analogous to the elasticity of demand, but refers to the supply side.
- Assumed value of $\eta = 0.33$ is very rough estimate!
- Assumption becomes critical when in-state supply insufficient to meet demand since:
  - demand is modelled as not being price responsive, so
  - values of η, Q^{import}, and P^{import} then determine estimate of competitive price.
- Imports were limited in 2000 compared to previous years because of supply-demand balance in other states.

3.2.11 Results

- Table 2 on page 15 of Joskow and Kahn shows estimates of average competitive prices for various possible emissions prices versus average actual prices.
- Data excerpted from Table 2 of Joskow and Kahn, using estimated actual emissions prices:

Month	Average actual price ($/MWh)	Average competitive price estimate ($/MWh)
May	47.23	55.11
June	120.20	67.23
July	105.72	63.25
August	166.24	105.15
September	114.87	88.96

Results, continued

- Actual prices up to double the estimated competitive price:
  - similar to reported highest mark-up of price above marginal cost in England and Wales study.
- Recall “capacity” HHIs:
  - England and Wales around 0.36 or 3600 $\%^2$ (or higher based on duopoly),
  - California around 0.1 to 0.2 or 1000 $\%^2$ to 2000 $\%^2$.
- Despite lower “capacity” HHI in California than England and Wales market, exercise of market power similar to England and Wales.

Results, continued

- Competitive estimates are higher in Summer 2000 than actual prices had been in 1998 and 1999, reflecting effects of gas prices, demand levels, hydro, import levels, and emissions prices:
  - Some of “crisis” was due to “market fundamentals” that would have affected price of electricity whether or not restructuring had occurred in California.
- In addition, significant difference between competitive estimate and actual price:
  - Actual prices in June through September 2000 considerably higher than competitive estimates.
- Joskow and Kahn conclusion is that actual prices were higher than justified by “market fundamentals” and competitive offers:
  - that is, there was withholding that increased prices.

3.2.12 *Opposing view*

- An opposing view was taken by Harvey and Hogan.
- They argue that:
  - estimates of competitive prices require detailed data that is not available from public domain data sources or is only approximate, and
  - errors in simulation model (as opposed to data for model) further contribute to errors.
- Sensitivity of results to these various uncertain issues make the competitive price estimate unreliable:
  - estimates need to include sensitivity information in order to estimate reliability of results,
  - results misleading without explicit sensitivity analysis.
- For example, Joskow and Kahn's gas marginal cost curve is extremely sensitive to both:
  - outages, and
  - withholding.
- Consequently, distinguishing high prices due to withholding from high prices due to the effects of forced outages is difficult.

Opposing view, continued

- As another example, because chronology of demand is not represented, Joskow and Kahn's model does not consider inter-temporal issues such as:
  - start-up and shut-down costs,
  - ramp-rate limits.
- Harvey and Hogan also observe that detailed "production cost modelling" tools have historically been used primarily to calculate sensitivities:
  - calculate the difference in outcomes (such as prices),
  - between two related cases.
- Such sensitivity analysis may provide a valid estimate of a *difference* in outcomes even if the estimates of the outcomes themselves are not accurate.
- The simulation in Joskow and Kahn relies on a simulation that is simplified compared to production cost modelling software:
  - using the simplified simulation to calculate an "absolute" value is problematic.

Opposing view, continued

- Considerable data is required for the model.
- The model simplifies many of the operational decisions compared to reality.
- Consequently, there is some, possibly significant, uncertainty in the competitive benchmark prices.
- The level of uncertainty is difficult to evaluate without considerably more data and simulation effort.
- The competitive benchmark is suggestive but not definitive.
- Joskow and Kahn then performed a withholding analysis to complement the competitive benchmark analysis.

3.2.13 *Empirical analysis of withholding*

- A somewhat different approach takes the actual market prices, focusing on high priced hours, and asks whether gas generators withheld supply from the market when prices were high.
- Alternative hypotheses:

withholding there were gas generators for which price was higher than left-hand marginal cost, yet the generators were not operating at capacity,

scarcity prices high prices were due to "scarcity," meaning that prices had to be above the left-hand marginal cost of all in-state generators in order to clear the market.

- To be conservative, a high estimate of the highest left-hand marginal cost is used, based on:
  - high heat rate of 17,000 Btu/kWh for gas turbine, and
  - high emissions of 1 pound of NO_x per MWh.
- Calculate left-hand marginal cost implied by these characteristics together with the actual gas price and the actual emissions price.

Empirical analysis of withholding, continued

- Table 5 on page 19 of Joskow and Kahn shows this left-hand marginal cost for various months and the number of hours when actual price was above this left-hand marginal cost, both in Southern California zone (SP15) and Northern California zone (NP15).
- Define “output gap” to be the difference between the total gas capacity in a zone minus the actual gas generation in that zone.
- Output gap could be due to:
  - capacity being used for ancillary services,
  - forced outages,
  - (in South) South to North transmission constraints,
  - other technical constraints, or
  - withholding.
- To simplify analysis, hours when transmission constraints were limiting were omitted from subsequent analysis.

Empirical analysis of withholding, continued

- Table 7 on page 23 of Joskow and Kahn shows output gap for June 2000.
- For example, in SP15, there is a 3,913 MW output gap, of which a maximum of 1,326 MW might be due to capacity reserved for AS:
  - non-gas resources such as hydro could also account for some of these AS,
  - so assuming that all AS are supplied by gas under-estimates the true output gap.
- The remainder, an under-estimate of output gap totalling around 2,600 MW, is about 23% of capacity.
- Historical forced outage rates are in range 6% to 13%, so a 23% remainder is much larger than can be accounted for by forced outages.
- To the extent that other technical constraints were not limiting, this circumstantial evidence suggests withholding.

Empirical analysis of withholding, continued

- Figure 1 on page 7 of Joskow and Kahn suggests that withholding of as little as 1000 MW of capacity (or equivalently, increasing demand by 1000 MW in the absence of withholding) could increase prices by 50%:
  - withholding of 2,600 MW likely to be profitable in absence of forward contracts and therefore constitutes exercise of market power.

Empirical analysis of withholding, continued

- Table 7 on page 23 of Joskow and Kahn shows that amongst five unaffiliated generators, Duke consistently has a much smaller gap than other firms:
  - Duke had forward contracted about 90% of its capacity,
  - so withholding not profitable for Duke.
- Circumstantial evidence is very strong that:
  - the four firms that could profit from withholding did, in fact, withhold, while
  - the firm that could not profit from withholding did not, in fact, withhold.
- Similar results for July through September 2000 as shown in Table 8 on page 25 of Joskow and Kahn:
  - output gap much larger than capacity reserved for AS,
  - difference between output gap and capacity reserved for AS is much larger than historical forced outage rate, and
  - the only forward contracted firm, Duke, has the smallest output gap, consistent with historical forced outage rates.

3.2.14 Conclusion

- Joskow and Kahn performed two basic analyses:
  - (i) comparison of competitive benchmark simulation to actual prices, and
  - (ii) withholding analysis.
- Both analyses are susceptible to criticism that either data or model compromise results.
- Withholding analysis makes a stronger argument since it asks, for each particular generator, was that generator withholding during high price periods:
  - avoids need to have access to considerable data,
  - focuses on characteristics of a particular generator for which it is easier to check assumptions and data.
- Evidence suggests that profitable withholding was undertaken by four of five generating firms in California in 2000.
- Fifth firm was forward contracted so that withholding was not profitable and firm did not withhold.

3.3 The ERCOT balancing market in 2001–2003

(i) Goal of analysis,
(ii) ERCOT,
(iii) Profit maximization by firm,
(iv) Conclusion.

3.3.1 Goal of analysis

- In our models so far and in our homework we have explicitly or implicitly assumed that each firm is a profit maximizer with full access to all information about all firms.
- Hortaçsu and Puller:
  - consider empirically whether or not firms behave as profit maximizers in choosing offers into the ERCOT balancing, and
  - discuss decision-making in the context of information actually available to a firm versus the ideal case of full access to all information about all firms.
- Also consider forward contract position in profit function.

3.3.2 ERCOT

- Market initially opened to wholesale competition in 1996.
- Wholesale and retail market restructured in 2001:
  - will be restructured again for nodal market in December 2010.
- Restructuring in 2001 allowed for both wholesale and (partly) retail competition.
- Retail customers of vertically integrated incumbents were allowed to choose a new retailer.
- Incumbents were required to serve remaining retail customers at a regulated price, the "price-to-beat:"
  - lower than previous regulated rates, but
  - high enough to allow for profitable entry by retailers,
  - price adjusted with changes in gas price index.

ERCOT, continued

- Considerable generation owned by non-incumbents:
  - most new capacity built in late 1990s was combined-cycle gas turbine.
- Table 1 of Hortaçsu and Puller shows generation ownership of most market participants, expressed in terms of total installed capacity.
- "Capacity" HHI for whole of ERCOT is only 1376 $\%^2$, assuming all "Others" owned by a single firm:
  - using typical regulatory threshold of 1800 or 2000, suggests un-concentrated market.
- "Capacity" HHI of particular geographic areas considerably higher.

ERCOT, continued

- Market participants can enter into forward contracts.
- Each day, market participants “schedule” their generation and demand for each 15 minute interval of next day with ERCOT ISO.
- Schedule could deviate from forward contract.
- “Balancing market” run by ERCOT to deal with deviations of actual demand from scheduled, actual generation from scheduled, and to keep flows between “zones” within limits:
  - each firm makes an “up balancing energy service” (UBES) offer for generating more from its portfolio than scheduled in a given zone,
  - makes a “down balancing energy service” (DBES) offer for generating less than schedule in a zone,
  - up and down offers are piecewise linear, with offers interpolated between price-quantity pairs.
- Balancing market cleared every 15 minutes.
- About 2-5% of total demand traded in balancing market.

ERCOT, continued

- Transmission:
  - If “inter-zonal” transmission limits binding then prices differ in each zone:
    - focus of study is on times without inter-zonal congestion.
  - “Intra-zonal” transmission limits dealt with in separate process:
    - interaction of inter-zonal and intra-zonal processes not explicitly modelled.
- Other constraints:
  - A number of other constraints, such as ramp-rate constraints, can limit the amount of capacity that can actually be dispatched in the balancing market:
    - focus of study is on 6:00–6:15 pm interval when ramp-rate constraints not expected to be extremely limiting,
    - ramp-rate constraints may have been much more limiting in other intervals.

3.3.3 Profit maximization by firm

- Basic economic assumption is that firm i maximizes its profit:

$$\pi_i(s_i) = \text{revenue}(s_i) - \text{costs}(s_i),$$

- where s_i is the "strategic variable" that can be adjusted by the firm:
  - for example, $s_i = Q_i$, the quantity generated,
  - more generally, s_i could be a vector with entries specifying the parameters of an offer (such as the intercept a_i and slope b_i of the offer in the homework problem, so that $s_i = \begin{bmatrix} a_i \\ b_i \end{bmatrix}$),
  - even more generally, s_i could specify a function, such as the offer function.
- As before, under suitable conditions, maximum profit is characterized by differentiating profit with respect to s_i and setting the derivative (or vector of partial derivatives) equal to zero.
- Equivalent to setting derivative of revenue equal to derivative of costs.

Profit maximization by firm, continued

- For the moment, we will focus on the quantity generated, Q_i, as the strategic variable:
  - we will see that a *collection* of quantities generated and the corresponding prices can be assembled to specify an offer,
  - will then re-interpret this collection of quantity–prices as specifying the offer function, so that the strategic "variable" will be the offer function itself.

Profit maximization by firm, continued

- To evaluate revenue as a function of quantity, we must evaluate the price as a function of the quantity:
  - we also need to know the forward contract quantity, but we will initially ignore that issue,
  - effects of forward contracts are included in reported numerical results.
- Residual demand (RD) for firm i is the difference between the demand minus the supply of all other firms, as a function of price.
- The inverse of the residual demand, p^{d}_{-i}, specifies the price as a function of the residual demand.
- To clear the market, the generation Q_i by firm i must equal the residual demand:
  - therefore, $p^{\mathrm{d}}_{-i}(Q_i)$, the inverse residual demand function evaluated at quantity Q_i, yields the clearing price versus generation Q_i.

Profit maximization by firm, continued

- Previously, we called the derivative of costs with respect to Q_i the marginal costs.
- Similarly, consider revenue as a function of Q_i:

$$\text{revenue}(Q_i) = Q_i p^{\text{d}}_{-i}(Q_i),$$

where p^{d}_{-i} is the inverse residual demand.

- Recall that the derivative of revenue with respect to Q_i is called the marginal revenue (MR):

$$\text{MR}(Q_i) = Q_i \frac{\partial p^{\text{d}}_{-i}}{\partial Q_i}(Q_i) + p^{\text{d}}_{-i}(Q_i).$$

- The first-order necessary conditions for profit maximization correspond to finding a quantity where marginal costs and marginal revenue are equal.

Profit maximization by firm, continued

- What is an appropriate model of demand?
  - dependence on *price* due to price-responsiveness, possibly very small, and
  - dependence on *weather* and other issues that are uncertain at time the offer is specified (and do not depend on price).
- Ignoring price-responsiveness, we might model demand with a random variable ε that represents the uncertain "demand shock."
- What is an appropriate model of the supply of the other firms?
  - dependence on price, and
  - uncertainties due to operational and other considerations.
- Ignoring uncertainties in supply, we might model the supply of other firms as depending on price:
  - alternatively, to approximately include uncertainty in supply, we might model the net demand minus supply shock by ε.
- Residual demand faced by firm i is the difference between demand minus supply of other firms.

Profit maximization by firm, continued

- In reality, firm i will not be able to know supply of other firms until *after* market clears, but we will initially assume that supply of other firms is known to firm i:
  - profit maximizing we describe could really only be carried out *ex post*, (after the fact), but firms actually need to make offers *ex ante* before market clears!
  - will consider effects of using "old" data from previous days to construct next offer based on some of the *ex ante* information available.

Profit maximization by firm, continued

- Figure 1 of Hortaçsu and Puller shows residual demand (RD) for two different values of the random variable ε describing demand shock, say ε_1 and ε_2:
- Since supply of others is assumed to be known and fixed:
  - varying values of random demand shock ε result in the residual demand shifting to right and left,
  - the two residual demand curves RD_1 and RD_2 differ "horizontally" by $\varepsilon_1 - \varepsilon_2$.
- If we think of quantity as independent variable, then the inverses of RD_1 and RD_2 specify the inverse residual demand for the demand shocks ε_1 and ε_2, respectively.
- Actual offer into market by firm *i* is S_i^o:
  - if the value of the random demand shock was ε_1 so that the residual demand is RD_1,
  - then the market would clear at point D.

Profit maximization by firm, continued

- Given the residual demand, it is possible to calculate the marginal revenue:
  - the curves labelled MR_1 and MR_2, thought of as having independent variable Q_i, are the marginal revenues corresponding to demand shocks ε_1 and ε_2, respectively,
  - evaluation of marginal revenue requires *slope* of inverse residual demand curve.
- Marginal cost of firm i is denoted MC_i:
  - we have previously used c_i' for marginal cost.

Profit maximization by firm, continued

- Now we find the quantity where marginal cost and marginal revenue are equal.
- Re-arranging condition, we obtain a similar relationship for mark-up of price above marginal cost as in the Cournot derivation:

$$p^{\mathrm{d}}_{-i}(Q_i) - \mathrm{MC}_i(Q_i) = -Q_i \frac{\partial p^{\mathrm{d}}_{-i}}{\partial Q_i}(Q_i).$$

- For demand shock ε_1, we find the intersection of MR_1 with MC_i:
  – the corresponding quantity, Q_{i1} maximizes the profit.
- How do we construct an offer that achieves this quantity Q_{i1}?
  – the resulting price is the price P_{i1} on the curve RD_1 at quantity Q_{i1},
  – make an offer such that, at Q_{i1}, the corresponding offer price P_{i1} causes the residual demand to be Q_{i1}.
  – that is, the offer must have price P_{i1} at quantity Q_{i1}, shown as point B.
- Similarly, for demand shock ε_2, the offer must have price P_{i2} at quantity Q_{i2}, shown as point C.

Profit maximization by firm, continued

- If we trace out various values of ε_k, we can find the corresponding prices P_{ik} and quantities Q_{ik} that maximize profits, given the demand shock ε_k.
- The collection of price-quantity pairs forms an offer function $\mathrm{S}_i^{\mathrm{xpo}}$ that is the *ex post* profit-maximizing response to residual demand due to demand shock and fixed (slope) offers of other firms.
- Key observation:
  - given the offers of the other firms and *any* actual value of ε,
  - the offer $\mathrm{S}_i^{\mathrm{xpo}}$ will arrange for profit maximization by firm i.
- Market rules require that the offer price be non-decreasing versus quantity:
  - that is, the inverse of $\mathrm{S}_i^{\mathrm{xpo}}$ must be non-decreasing in order to correspond to a valid offer,
  - problematic when there are capacity constraints, but not for the cases analyzed here.

Profit maximization by firm, continued

- How do we incorporate forward contracts?
- One possibility is to assume that the schedule matches the forward contract:
  - however, there is no obligation for this to be case,
  - so, must estimate forward contract position from actual data.
- Suppose that the firm has forward contract with quantity QC_i and price PC_i:
  - forward contract quantity QC_i is not necessarily equal to scheduled generation.
- Revenue and marginal revenue become:

$$\begin{aligned}
\mathrm{revenue}(Q_i) &= (Q_i - \mathrm{QC}_i)\, p^{\mathrm{d}}_{-i}(Q_i) + \mathrm{QC}_i \times \mathrm{PC}_i, \\
\mathrm{MR}(Q_i) &= (Q_i - \mathrm{QC}_i)\frac{\partial p^{\mathrm{d}}_{-i}}{\partial Q_i}(Q_i) + p^{\mathrm{d}}_{-i}(Q_i).
\end{aligned}$$

- For evaluation of marginal revenue including forward contracts:
  - must evaluate forward contract quantity QC_i, but
  - forward contract price PC_i is irrelevant.

Profit maximization by firm, continued

- How to estimate QC_i?
- If the firm generates, in total, $Q_i > \mathrm{QC}_i$ then:
  - it is selling quantity QC_i at the forward price,
  - it is selling $Q_i - \mathrm{QC}_i$ at the market price.
- For the sale of quantity $Q_i - \mathrm{QC}_i$ at the market price to be more profitable to the firm than just generating QC_i, market price must be at or above marginal cost:
  - to ensure profitability, offer price must be no lower than the marginal cost evaluated at Q_i for each $Q_i > \mathrm{QC}_i$.
- If the firm generates, in total, $Q_i < \mathrm{QC}_i$ then:
  - it is selling quantity QC_i at the forward price,
  - it is buying $\mathrm{QC}_i - Q_i$ at the market price.
- For the purchase of quantity $Q_i - \mathrm{QC}_i$ at the market price to be more profitable to the firm than just generating QC_i, market price must at or below marginal cost:
  - to ensure profitability, offer price must be no higher than marginal cost evaluated at Q_i for each $Q_i < \mathrm{QC}_i$.

Profit maximization by firm, continued

- To summarize, assuming rational behavior, offer will change from being at or below marginal cost to being at or above marginal cost as quantity increases from below to above the forward quantity QC_i:
  - this observation should be true whether or not offer achieves *ex post* optimal profits,
  - simply a manifestation of not wanting to lose money compared to generating QC_i!
- If the actual offer $\mathrm{S}_i^{\mathrm{o}}$ and the marginal costs are observed, the forward quantity can be estimated as the quantity where they intersect:
  - point A in Figure 1 of Hortaçsu and Puller,
  - if firm had offered competitively then $\mathrm{S}_i^{\mathrm{o}}$ and the marginal costs would be coincident!
- Forward quantity should be included in profit maximization calculation:
  - complication in case of retail obligations is that forward commitment at fixed price may be for whatever demand occurs!
  - firm may have good estimate of retail obligation at time of making offer.

3.3.4 Analysis of observed offers

- Focus is on 6:00–7:00 pm hours in 2001–2003 without inter-zonal transmission congestion.
- Consider the 6:00–6:15 pm interval each such day:
  - only use data from that interval if there was no congestion throughout the hour.
- In ERCOT balancing market, offers are piecewise linear.
- Up-balancing and down-balancing offers represent offers to deviate from day-ahead scheduled quantities:
  - "zero" offer quantity corresponds to day-ahead scheduled quantity,
  - scheduled quantity for firm i can be different from forward contract position QC_i.

Analysis of observed offers, continued

- Marginal costs are utilized for units that are dispatchable:
  - include gas and coal,
  - do not include nuclear and wind.
- Operational constraints such as ramp-rate constraints that might further limit dispatch are ignored:
  - the 6:00–6:15 pm interval was chosen because Hortaçsu and Puller anticipated that ramp-rate constraints would not be binding during this interval.
- Hortaçsu and Puller model assumes piecewise constant offers:
  - residual demand would be piecewise constant,
  - smooth out actual offers to estimate average slope of residual demand for use in marginal revenue calculation.
- Hortaçsu and Puller found that their estimates of prices were within about 5% of actual prices, despite:
  - erroneous model of offers, and
  - ignoring ramp-rate constraints.

Analysis of observed offers, continued

- Figure 2 of Hortaçsu and Puller shows an example marginal cost function, MC, for Reliant, together with actual offer, residual demand, and the *ex post* optimal offer:
  - zero offered quantity corresponds to day-ahead scheduled quantity,
  - marginal cost and actual offer intercept at the forward contract position QC_i (about 500 MW), assuming rational offer behavior.
- Reliant's actual offer tracks *ex post* optimal offer, except for quantities:
  - above its forward contract position, and
  - more than 1000 MW below day-ahead scheduled quantity.
- Figure 7 of Hortaçsu and Puller shows similar graphs for:
  - TXU: actual offer tracks *ex post* optimal offer for quantities between day-ahead scheduled quantity and forward contract position of around 600 MW,
  - Calpine: does not track *ex post* optimal offer, and
  - Guadalupe Power Partners (a relatively small firm): does not track *ex post* optimal offer, and evidently does not want to change its production from its day-ahead schedule.

Analysis of observed offers, continued

- As a quantitative measure of performance in market, compare actual profits to *ex post* optimal profits.
- Measure is:

$$\text{PercentAchieved} = \frac{\pi^{\text{Actual}} - \pi^{\text{Avoid}}}{\pi^{\text{XPO}} - \pi^{\text{Avoid}}},$$

- where:

π^{Actual} are the profits with the actual offer,
π^{XPO} are the *ex post* optimal profits, and
π^{Avoid} are the profits with offers chosen to avoid selling any generation or backing off any generation compared to schedule.

Analysis of observed offers, continued

- Column (1) of Table 2 of Hortaçsu and Puller shows PercentAchieved.
- The differences $\pi^{\text{Actual}} - \pi^{\text{Avoid}}$ and $\pi^{\text{XPO}} - \pi^{\text{Avoid}}$ are shown in columns (3) and (5), respectively.
- Many firms appear to be foregoing profits from balancing market:
  - actual profits are less than *ex post* optimal profits, particularly for smaller firms,
  - larger firms achieve closer to *ex post* optimal profits.
- Note that this analysis ignores profits from sales outside of the balancing market:
  - profits from balancing market, with 2–5% of total market sales, may be dwarfed by profits from other sales.

Analysis of observed offers, continued

- What are possible explanations for actual profits being lower than *ex post optimal*?

  (i) Market participants are offering competitively.

  (ii) Profit from *ex post* optimal offer is optimistic compared to what could actually be achieved, given knowledge at time offer was made and detailed offer rules.

  (iii) Market participants do not want to adjust their generation compared to schedule or there are costs to adjustment that are not represented in model.

  (iv) Anticipation of possibility of transmission congestion affected offer strategy (even though study period considered only intervals when transmission congestion did not occur).

  (v) Collusion.

  (vi) Bounded rationality or transaction costs of formulating offers is too high relative to profit.

Analysis of observed offers, continued

(vii) Balancing market is not important enough to market participants in terms of overall profitability.

(viii) Market participants needed to learn about making profit-maximizing offers.

Analysis of observed offers, continued

(i) Competitive offers?

- Competitive profits (offers matching marginal costs) are lower than *ex post* optimal.
- However, actual offers are far from competitive!
- Prices for "up balancing energy service" (UBES) are far above marginal and, particularly for smaller market participants, also above the *ex post* optimal offer price.
- Prices for "down balancing energy service" (DBES) are far below marginal and, particularly for smaller market participants, also below the *ex post* optimal offer price.
- Price versus quantity relationship of offers is far "steeper" than competitive and, particularly for smaller market participants, steeper than *ex post* optimal:
  - Figure 7 of Hortaçsu and Puller for Guadalupe Power Partners.

Analysis of observed offers, continued

- Participants are withholding from the market:
  - Offers well above marginal cost indicate *economic withholding*.
  - Moreover, Figure 41 of the “2005 State of the Market Report for the ERCOT Wholesale Electricity Markets” indicates that over half of the available capacity is not being offered into the balancing market, so there is also considerable *physical withholding* in addition to the economic withholding!
  - Physical withholding may be response to imperfections in ERCOT representation of portfolio ramp-rate constraints and other issues.
- Ironically, economic withholding is *not* profitable for many of the small firms.
- If small market participants had offered more of their capacity at prices somewhat closer to competitive they would have both:
  - improved profits, and
  - reduced overall actual production costs!
- Production costs are approximately $6000/h higher than optimal based on economically dispatching the offered capacity using marginal costs.

Analysis of observed offers, continued

(ii) Are *ex post* optimal profits optimistic?

- By definition, offers of everyone else are not available until after offer of firm must be submitted.
- Information about aggregate offers is available to market participants with an approximately two day lag.
- In practice, instead of using actual offer of everyone else to construct best response, use the most recent available offer information from everyone else.
- Would be best response if everyone else maintained the same offer:
  - “naive best response” under naive assumption that other offers do not change and that demand conditions have not changed,
  - uses information actually available *ex ante* at time offer must be submitted,
  - ignores additional information about weather etc that might also be available at this time.

Analysis of observed offers, continued

- Column (2) of Table 2 of Hortaçsu and Puller shows PercentAchieved using the naive best response based on most recent offers of everyone else.
- Naive best response profit is very close to *ex post* best response.
- Market participants could have achieved close to *ex post* best response by simply finding the best response to the most recent available offers of everyone else:
  - similar to updating offers in homework based on what everyone else did in the most recent week.
- Wolak identifies *theoretical* deficiencies in the Hortaçsu and Puller model that may undercut the position that the actual offers are not profit maximizing.

Analysis of observed offers, continued

(iii) Are there adjustment costs that invalidate the calculation of profit?

- Most firms have automatic controls on generation and presumably can adjust gas-fired production fairly easily:
  - gas-fired generation is marginal in most intervals in study,
  - ramp-rate constraints not expected to have been tightly binding during the 6:00–6:15 pm interval, but may have affected the amount of "dispatchable" supply during other intervals.
- "Bid-ask spreads" (the difference between the offer price for UBES minus the offer price for DBES):
  - range from \$2/MWh to \$30/MWh, despite similar technology amongst most participants,
  - suggest that some firms just do not want to participate in balancing market!
  - bid-ask spreads decrease over time, suggesting increased willingness to participate in balancing market.

Analysis of observed offers, continued

(iv) Did anticipated transmission constraints affect behavior?

- Statistical analysis suggests little effect of anticipation of transmission.

(v) Could the steep offers of some smaller participants be evidence that they are colluding to make offers that result in monopoly level of pricing?

- But overall payoffs of these firms is low since they sell and buy only a small amount from the balancing market.
- Not consistent with smaller firms colluding to maximize profits!

Analysis of observed offers, continued

(vi) Could it be too hard to figure out the best offers or not worthwhile to do so?

- Calculation of naive best response requires putting together a spreadsheet or a Matlab program.
- Some participants formulate offers based on heuristics that may be inconsistent with profit maximization:
  - including "sunk costs" such as debt payments into offer prices.
- Recall that operating profit maximization depends on marginal costs only, not on "sunk costs" that do not change with production.

(vii) Balancing market not important enough?

- It may be that the profits from the balancing market are too small to justify effort, particularly for smaller market participants.
- Risks due to uncertainties may exceed profits from more active participation in balancing market.

(viii) Learning?

- It may be that the firms need to learn about the market in order to improve their profits.

3.3.5 Conclusion

- Some larger firms achieve a significant fraction of *ex post* optimal profits:
  - consistent with basic economic hypothesis of maximizing profits.
- Other firms, particularly smaller firms, offer to avoid the market:
  - reduced profits, and
  - less efficient dispatch overall.
- Some learning over study period:
  - anecdotal evidence is that market participants have learned more in time since.

3.4 The ERCOT balancing market in 2005

(i) Independent Market Monitor report,
(ii) ERCOT balancing energy market,
(iii) Price spike intervals,
(iv) Exercising versus abusing market power,
(v) Assessment of TXU offers,
(vi) Conclusion.

3.4.1 Independent Market Monitor report

- Review the results of 2007 “Independent Market Monitor” (IMM) report into TXU actions in Summer 2005.
- IMM report was requested by Public Utility Commission of Texas (PUCT):
  - assess whether TXU “abused” (that is, (egregiously) exercised) market power,
  - released March 2007,
  - PUCT staff recommended large fine on TXU in response to report findings.
- The IMM report was revised in September 2007:
  - have incorporated corrections.
- As with California analysis, considerable data and analysis involved in study:
  - some data for this ERCOT analysis is not publicly available.

Independent Market Monitor report, continued

- IMM report follows up on IMM 2005 “2005 State of the Market Report for the ERCOT Wholesale Electricity Markets” (2005 SOM).
- The 2005 SOM identified that the two largest generator asset owners in ERCOT were “pivotal” in the balancing market for many pricing intervals in 2005:
  - one of these owners is TXU,
  - if their offers had been withdrawn from the balancing market then there would have been insufficient remaining offers to meet the balancing market demand,
  - ignoring any price responsiveness of demand.
- In homework, if groups 1 or 2 withdrew from the market and did not offer then there would be insufficient offers from other groups to meet the demand in interval 3 (ignoring price responsiveness).
- Balancing market represented only approximately 5% of total ERCOT demand in 2005.

3.4.2 ERCOT balancing energy market

- As discussed in part previously in relation to Hortaçsu and Puller, features of the ERCOT balancing market include:
  - offer-based economic dispatch considering inter-zonal transmission constraints, cleared each 15 minutes based on ERCOT forecast of net deviation from day-ahead "schedules,"
  - uniform clearing price in each zone, "Market Clearing Price for Energy" (MCPE),
  - served about 5% of energy in ERCOT on average,
  - sometimes met over 10% of demand, particularly during peak.

3.4.3 Price spike intervals

- IMM report focuses on “price spike intervals,” defined as intervals where:
  - MCPE is more than
  - an estimate of high marginal cost gas units specified by:
    - the “Houston Ship Channel” natural gas price index, multiplied by
    - 20,000 Btu per kWh (20 million Btu per MWh).
- As in California study, this threshold of MCPE is chosen to exceed the marginal cost of most generators:
  - NO_x emissions not internalized by price.
- “Study Period” is hours 10 through 23 of each day of June 1 through September 30, 2005:
  - 657 price spike intervals in Study Period,
  - two-thirds of price spike intervals occurred in July and August,
  - TXU was pivotal in 554 of the 657 price spike intervals,
  - over 6,000 intervals in total in Study Period.
- Summary results in IMM report are for all of Study Period.

Price spike intervals, continued

- Figure 1 of IMM report shows the distribution of price spike intervals:
  - mostly occur in hours ending 10 through 23,
  - intervals where total ERCOT demand is highest.
- In many intervals, TXU offered “up balancing energy service” (UBES) at prices above “generic” estimates of its marginal costs.
- TXU offered on average about 2000 MW of UBES during price spike intervals.
- TXU was a net seller in many of these intervals.
- Figure 2 of IMM report shows that the offer prices exceeded “generic” estimates of marginal costs by varying amounts during price spike intervals:
  - 58.9% of TXU offers were less than \$50/MWh above estimated marginal costs, but
  - some TXU offers were more than \$200/MWh above estimated marginal costs.

Price spike intervals, continued

- Figure 3 of IMM report shows an example of balancing market operation during a price spike interval:
  - interval ending 1200 on July 11, 2005,
  - balancing market demand is 3,152 MW,
  - offered "dispatchable" supply is 3,772 MW,
  - TXU made all of the offers having price above $150/MWh,
  - TXU made none of the 2923 MW of offers having price below $150/MWh,
  - offers from other companies are insufficient to meet demand,
  - so TXU was pivotal in this interval,
  - clearing price is $252.48/MWh.

Price spike intervals, continued

- "Dispatchable" supply means only those offers associated with generation:
  - that was available and on-line or was quick-start,
  - that was not prevented from increasing due to transmission constraints, and
  - that was within ramp-rate limits.
- Availability of all offer data to IMM and the use of ERCOT's "Scheduling, Pricing, and Dispatch" (SPD) model avoids uncertainties in the other empirical studies about the amount of dispatchable supply available in the balancing market:
  - for example, effect of ramp-rate constraints in Hortaçsu and Puller study could only be considered indirectly by choosing an interval where it was assumed that the ramp-rate constraints would not be severely limiting.

Price spike intervals, continued

- Relaxing ramp-rate and other constraints would allow for more supply and lower clearing prices:
  - Figure 4 of IMM report compares the dispatchable supply (the "Constrained Offer Curve") to the total supply that was available and on-line (the "Constraints Relaxed Offer Curve.")
  - If transmission constraints and ramp-rates could be ignored then prices would have been much lower!
- As Harvey and Hogan argued in the context of simulation of California, representation of these details is essential to determining absolute price levels.
- Engineering constraints have significant influence on market prices.

3.4.4 Exercising versus abusing market power

- At the time of the study period there was no PUCT definition of market power.
- IMM report defines "market power as the ability for a market participant to profitably raise prices significantly above competitive levels:"
  - IMM report not very explicit about "significant" but quantitatively assesses several statistics such as excess profits above competitive, which could be compared to a standard of "significant."
  - Analyzes whether market power was exercised.
- Long discussion in report about definition of market power underscores lack of a clear definition being provided by PUCT for this time period.
- Definition of market power includes discussion of:

  (i) Relevant product market,
  (ii) Relevant geographic market,
  (iii) "Pivotal analysis."

Exercising versus abusing market power, continued

- PUCT has subsequently made definitions of "market power" and "market power abuse" explicit (PUCT rulemaking in project Number 31972, amendment to Public Utility Commission Rule section 25.502):
  - market power is the "ability to control prices or exclude competition in a relevant market," and
  - "market power abuse" includes practices that "unreasonably...reduce the level of competition" and include "withholding of production, precluding entry, and collusion."
- PUCT definition:
  - omits fundamental "economic" issue of profitability,
  - distinguishes between having and exercising (that is, "abusing"),
  - market power abuse includes qualification "unreasonably" that is evidently synonymous with "significant," but fails to define "unreasonably" in a quantitative manner.
- We will follow IMM definition, recognizing that it also omits explicit definition of "significant."
- Quantitative definition of "significant" is a policy decision.

3.4.4.1 Relevant product market

- What generation is available to offer into the balancing market?
- All in-service capacity in ERCOT?
  - But most off-line capacity cannot be brought on-line in the 15 minute time frame of the balancing market.
- All on-line generators and quick-start generators?
  - But even quick-start can only be partially deployed in a 15 minute interval, so only consider the energy that could be deployed in the 15 minute interval.
- Only the generation actually offered into the balancing market?
  - As shown in Figure 41 of the "2005 State of the Market Report for the ERCOT Wholesale Electricity Markets," over half of the available capacity is not being offered into the balancing market.
  - Reasons include the ERCOT portfolio dispatch process.
  - As discussed in Hortaçsu and Puller, for the capacity that *is* offered, many UBES offer prices are far above marginal costs.
- IMM report uses actual offered capacity to analyze withholding by TXU:
  - similar to Joskow and Kahn withholding analysis.

3.4.4.2 Relevant geographic market

- When transmission constraints are binding, competition is more limited:
  - unlike previous studies, the IMM report includes intervals when transmission constraints are binding.
- During study period, most commonly binding inter-zonal constraints were:

  (i) South to Houston,
  (ii) South to North, and
  (iii) North to Houston.

- Most TXU capacity is in North zone.
- The intra-zonal constraint management process can affect the capacity available for inter-zonal constraint management:
  - inefficiency due to two-step inter-zonal (that is, "zonal") and intra-zonal (that is, "local") congestion management process in ERCOT.

3.4.4.3 Pivotal analysis

- Consider those price spike intervals when TXU offers were needed to meed inelastic demand, so that TXU was "pivotal."
- Price spike intervals are typically associated with net "up" demand.
- Figures 5 through 8 show quantities of:
  - TXU "up balancing energy service" (UBES) offers,
  - other UBES offers, and
  - "up" demand that had to be met by the UBES offers.
- Figures 5 through 8 show that TXU was pivotal during most price spike intervals.
- During intervals when TXU is pivotal, it has the ability to increase prices above competitive levels:
  - TXU presumably also has this ability in many intervals when it is not pivotal, as in the case of groups 1 & 2 in intervals 1 & 2 in homework.
- To determine if TXU exercised market power:
  - must assess whether TXU offers were actually at prices that are above competitive, and, if so,
  - must assess whether offers at above competitive prices were profitable.

3.4.5 Assessment of TXU offers

(i) Were TXU offers above marginal costs?
(ii) Did TXU withhold?
(iii) What is impact on prices?
(iv) Did TXU profit from withholding?

3.4.5.1 TXU balancing energy offers

- In 2004, TXU implemented an offer strategy called “Rational Bidding Strategy” (RBS).
- In 2005, TXU used RBS:
  - for more of its generation offers,
  - over longer periods of time, and
  - particularly during periods of high demand.
- According to TXU, RBS involves setting offers equal to the lesser of:
  - “a self-imposed regulatory offer cap,” and
  - capital and operating costs,
- apparently whenever such offers would be accepted.
- When the resulting offer price would not be accepted, TXU offers closer to marginal cost.

TXU balancing energy offers, continued

- RBS evidently means:
  - offering above marginal cost when TXU has market power, and
  - offering close to marginal cost when TXU does not have market power.
- That is, RBS involves exercising market power:
  - may forbear from *maximizing* operating profits in anticipation that a level of operating profits that roughly covers capital and other costs would be tolerated by PUCT.
- As discussed previously, such forbearance is very difficult to analyze!
  - PUCT response to forbearance was also apparently unexpected by TXU!

TXU balancing energy offers, continued

- Generic marginal costs were estimated for TXU on-line generators in each interval.
- Simulated competitive offers were then based on the estimated marginal costs.
- Figure 9 shows a simulated competitive offer for TXU North zone generation for 5:30pm on July 19, 2005:
  – The simulated competitive offer closely matches the actual TXU offer for this interval.
  – Presumably TXU did not have “significant” market power during this interval and offered close to competitively to maximize operating profits.

TXU balancing energy offers, continued

- Figure 10 shows a simulated competitive offer for TXU North zone generation for 5:30pm on July 20, 2005:
  - The simulated offer does not match the actual TXU offer for this interval.
  - Similar units were online on July 19 and July 20.
  - TXU offer prices are substantially higher than on July 19 and set price well above the level consistent with a competitive offer from TXU.
- Figure 11 summarizes distribution of above marginal cost offers from TXU.

3.4.5.2 Impact of TXU on balancing energy market

- Simulate the effect of changing TXU's offer to being at generic marginal costs; that is, "competitive."
- Simulation used ERCOT's "Scheduling, Pricing, and Dispatch" (SPD) model, which is used in the balancing market:
  - TXU offers replaced by generic marginal costs, but
  - all other simulation data used actual offer data.
- Estimate excess profits above competitive offers:
  - by using actual offer data for everyone besides TXU and using SPD, avoids drawbacks of other empirical studies.
  - As in Joskow and Kahn study, focus on characteristics of a particular portfolio of generation makes it is easier to check assumptions and data.

Impact of TXU on balancing energy market, continued

- Assessment of excess profits above competitive *excludes* 55 intervals when balancing market demand exceeded dispatchable generation in either the actual or simulated cases:
  - assuming no physical withholding, this condition corresponds to scarcity,
  - competitive prices *should* be high during scarcity to reflect “value of lost load.”
- During any intervals of scarcity, prices were probably well below competitive levels:
  - no attempt in IMM report to estimate overall excess profits above competitive, including any scarcity intervals,
  - for example, if scarcity occurred in all 55 intervals, if prices were $1,500/MWh below competitive in these intervals, and if TXU was net selling 1000 MW in these intervals, then total *depressed* profits for TXU would be over $20 million!

Impact of TXU on balancing energy market, continued

- Figures 12 and 13 show, for North Zone and Houston Zone, respectively:
  - the daily average actual MCPE, minus
  - the daily average simulated MCPE.
- Figure 14 shows the monthly average actual MCPEs and the monthly average simulated MCPEs.
- Monthly average actual MCPEs were 3.3% to 19.2% higher than simulated MCPEs:
  - Figure 15 shows the MCPE increases multiplied by the total quantity sold in the balancing market to estimate the increased energy purchase costs.
  - Total increased balancing energy purchase costs over Study Period are about $57,000,000.
- Average actual MCPE was about 11% higher than simulated MCPEs over entire Study Period:
  - Given a, for example, 5% threshold for "significant," this would be viewed as a significant increase.
- All analysis excludes the 55 intervals of possible scarcity.

3.4.5.3 Assessment of TXU's net position

- Figure 16 shows TXU net sales in the balancing energy market.
- Calculating the corresponding profits and subtracting the profits obtained in the simulation of competitive TXU offers yields the excess profits above competitive offers ("Daily Net Profit Increase") shown in Figure 17:
  - includes effect of "transmission congestion rights" (TCRs), retail obligations, and other known contractual obligations, and
  - sales and purchases that are priced at MCPE,
  - but again ignores below competitive prices in 55 excluded intervals of possible scarcity.
- Total TXU excess profits above competitive levels over Study Period are about $18,800,000.
  - Note that this is roughly the same as the estimate of the *depressed* profits during 55 scarcity intervals.
- Table 1 indicates that, as a result of RBS offers, TXU generated 252,000 MWh less than it would have if it were offering competitively.

3.4.6 Conclusion

- TXU had the ability to increase prices above competitive levels.
- RBS involved offering at higher than competitive prices: "economic withholding."
- TXU was a net seller in the balancing market.
- TXU increased its profits compared to competitive offer prices.
- That is, TXU exercised market power.
- Much of data for study is not available publicly.
- Analysis omits intervals of apparent scarcity, so calculations of excess profits above competitive offers are uncertain by a potentially significant amount.

3.5 Reconciling the "economic" model to the "commercial" model

- The empirical studies rely, in part, on modelling the economic incentives facing market participants.
- Incentives to market participants depend on many details in the market rules:
  - the "commercial" model.
- With the exception of the IMM report, which used the actual "Scheduling, Pricing, and Dispatch" (SPD) algorithm and the actual offer data for companies besides TXU, the empirical models we have investigated all depend on "economic" models that abstract from many of the details of the "commercial" model.
- In addition to uncertainties in data, deviations of the economic model from the commercial model contribute to lack of certainty in conclusions of the various empirical studies:
  - uncertain effect of ramp-rate and other constraints.

Reconciling the “economic” model to the “commercial” model, continued

- The IMM report avoids the data problem for the non-TXU data, but that data is not available publicly.
- The IMM report avoids the economic modelling problem by using the SPD model, but not clear that other market participants could reproduce the model conveniently even if data were available:
  - SPD model is not apparently available commercially to market participants.
- We will discuss the issue of the “economic” and “commercial” models further in the context of theoretical modelling, where we will also consider the underlying “physical” model.

3.6 Distinguishing design flaws from market power

- A related issue is that the “commercial” model may itself fail to represent all the engineering constraints:
  - ERCOT zonal balancing market includes representation of inter-zonal constraints, but omits intra-zonal (local) constraints,
  - portfolio offers in balancing market include representation of portfolio-wide ramping constraints that may not match actual ramp-rate constraint of units deployed in balancing market.
- Consequently, even “competitive” offers may result in deviations from efficient operation of system:
  - intra-zonal constraints are enforced in a separate process to balancing market,
  - actual ramp-rate constraints may prevent generator from deploying to level requested by ERCOT.

Distinguishing design flaws from market power, continued

- If the deviation of the commercial model from efficient dispatch entails large cost increases then the deviation is a market design flaw:
  - such design flaws may *exacerbate* market power issues,
  - this complicates assessment of market power and market design flaws.
- Market design flaws introduce a further bias into the assessment of the "absolute" price level under competitive offers.
- As Harvey and Hogan argue, if a competitive benchmark does not represent the effect of these market design flaws then the calculated competitive benchmark prices may be different to what would actually occur with "competitive" offers:
  - for example, ignoring ramp-rate constraints will result in lower competitive benchmark price estimates if lower priced offers are associated with slower ramping capacity,
  - as in comparison in IMM report of using only "dispatchable" offer versus "Constraints Relaxed Offer Curve."

Distinguishing design flaws from market power, continued

- One way to assess this is to estimate correlations between variables and compare these to "benchmarks" based on engineering and economic models.
- For example, a basic model of scheduling in the ERCOT commercial model (prior to introduction of "relaxed balanced scheduling" in November 2002) might be that the scheduled demand is an unbiased forecast of the actual demand:
  - Under ERCOT scheduling rules, scheduled generation equals scheduled demand,
  - Balancing market demand equals difference between actual demand minus scheduled demand (together with any deviations of generation from schedule),
  - So, if actual generation follows scheduled generation then balancing market demand should be uncorrelated with scheduled generation.

Distinguishing design flaws from market power, continued

- Testing this over the period January 2002 to May 2002 yielded:
  - negative, but close to zero, correlation when all intervals considered,
  - negative correlation when off-peak intervals were considered.
- More “down” balancing was required off-peak, reflecting effect of generation at minimum production levels:
  - reflects discrepancy between commercial model and engineering minimum capacity constraints,
  - not (directly) a market power issue.

3.7 Summary

(i) The England and Wales market in 1992–1994,
(ii) The California market in 2000,
(iii) The ERCOT balancing market in 2001–2003,
(iv) The ERCOT balancing market in 2005,
(v) Reconciling the “economic” model to the “commercial” model,
(vi) Distinguishing design flaws from market power.

Homework exercise: Due Tuesday, April 6, by 10pm

- For next week, we will go back to allowing offers to vary for three peak pricing periods with demand:
  - 4150 MW,
  - 4200 MW, and
  - 4250 MW.
- That is, a different offer will be used for each of three pricing periods.
- Suppose that the cost functions for the last homework exercise stayed exactly the same.
- Again assume that the "top" 400 MW of demand in each period will be price responsive, with willingness-to-pay varying linearly from $500/MWh down to $100/MWh.
- Update your offers for the peak demand period to try to improve your profits compared to your previous offers:
  - submit offers for all periods, all three offers will be considered.

Homework exercise: Due Tuesday, April 13, by 10pm

- For next week, we will again allow offers to vary for three peak pricing periods with demand:
  - 4150 MW,
  - 4200 MW, and
  - 4250 MW.
- That is, a different offer will be used for each of three pricing periods.
- Suppose that the cost functions for the last homework exercise stayed exactly the same.
- Again assume that the “top” 400 MW of demand in each period will be price responsive, with willingness-to-pay varying linearly from $500/MWh down to $100/MWh.
- Update your offers for the peak demand period to try to improve your profits compared to your previous offers:
  - submit offers for all periods, all three offers will be considered.

Homework exercise: Due Tuesday, April 20, by 10pm

- For next week, we will again allow offers to vary for three peak pricing periods with demand:
  - 4150 MW,
  - 4200 MW, and
  - 4250 MW.
- That is, a different offer will be used for each of three pricing periods.
- Suppose that the cost functions for the last homework exercise stayed exactly the same.
- Again assume that the "top" 400 MW of demand in each period will be price responsive, with willingness-to-pay varying linearly from \$500/MWh down to \$100/MWh.
- Update your offers for the peak demand period to try to improve your profits compared to your previous offers:
  - submit offers for all periods, all three offers will be considered.