
Sequential Markets, Market Power, and Arbitrage

Author(s): Koichiro Ito and Mar Reguant

Source: *The American Economic Review*, JULY 2016, Vol. 106, No. 7 (JULY 2016), pp. 1921-1957

Published by: American Economic Association

Stable URL: <https://www.jstor.org/stable/43861116>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

is collaborating with JSTOR to digitize, preserve and extend access to *The American Economic Review*

Sequential Markets, Market Power, and Arbitrage[†]

By KOICHIRO ITO AND MAR REGUANT*

We develop a framework to characterize strategic behavior in sequential markets under imperfect competition and restricted entry in arbitrage. Our theory predicts that these two elements can generate a systematic price premium. We test the model predictions using microdata from the Iberian electricity market. We show that the observed price differences and firm behavior are consistent with the model. Finally, we quantify the welfare effects of arbitrage using a structural model. In the presence of market power, we show that full arbitrage is not necessarily welfare-enhancing, reducing consumer costs but increasing deadweight loss. (JEL D42, D43, L12, L13, L94, Q41)

A variety of economic goods are traded through sequential markets—a set of forward and spot markets for a good such as treasury bonds, stocks, coal, electricity, natural gas, oil, and agricultural products. The rationale behind establishing sequential markets comes from simple economic theory. For a commodity that has uncertainty in its delivery price or quantity, sequential markets can improve the efficiency of the final allocation. In an oligopolistic setting, sequential markets can also help reduce the extent of market power and improve efficiency (Allaz and Vila 1993).

In a stylized setting, prices in sequential markets should converge in expectation (Weber 1981). Empirical evidence is, however, usually inconsistent with this simple economic theory, with prices in sequential markets often exhibiting systematic price differences.¹ Previous studies provide several potential channels that explain why prices in sequential markets do not converge, such as risk aversion (Ashenfelter 1989; McAfee and Vincent 1993) and asymmetric shocks (Bernhardt and Scoones 1994; Salant 2014). We contribute to the literature studying the role of market power in creating such price differences (Saravia 2003; Borenstein et al. 2008).

* Ito: Harris School of Public Policy, University of Chicago, 1155 East 60th Street, Chicago, IL 60637 (e-mail: ito@uchicago.edu); Reguant: Department of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208 (e-mail: mar.reguant@northwestern.edu). For helpful comments, we thank Simon Board, Severin Borenstein, Thomas Covert, Ali Hortaçsu, Ryan Kellogg, Chris Knittel, Matt Notowidigdo, Nick Ryan, David Salant, Joe Shapiro, and Matt White, as well as seminar participants at several institutions and conferences. Ito thanks the Stanford Institute for Economic Policy Research for financial support. Reguant gratefully acknowledges the support of NSF grant SES-1455084.

[†] Go to <http://dx.doi.org/10.1257/aer.20141529> to visit the article page for additional materials and author disclosure statement(s).

¹ Examples of empirical evidence of systematic price differences include treasury auctions (Coutinho 2013), wine auctions (Ashenfelter 1989; McAfee and Vincent 1993), mutual funds (Zitzewitz 2003), and wholesale electricity markets (Saravia 2003; Longstaff and Wang 2004; Borenstein et al. 2008; Hadsell 2008; Bowden, Hu, and Payne 2009; Jha and Wolak 2014; Birge et al. 2014).

We present a new theoretical framework to examine how the existence of market power can interact with constraints to arbitrage, and prevent full price convergence in sequential markets. More generally, we show that limited arbitrage can arise endogenously in an oligopolistic setting, when firms are asymmetric in size. We then analyze the welfare implications of price arbitrage in the presence of market power. Arbitrage almost always improves welfare in a simple model with perfect competition. However, we show that such implications can change once we take into account the existence of market power.

In the context of sequential electricity markets, we consider a model with two sequential markets: the forward market and the real-time market.² Both markets trade the same commodity, electricity, to be produced at a particular delivery time. A monopolist decides how much to sell in each market. We assume that demand is inelastic and allocated in full in the forward market.³ The monopolist still faces a downward-sloping demand curve in both markets, due to the presence of fringe suppliers, who offer production at their marginal cost.

Figure 1 provides a graphical illustration of the model. The monopolist participates in two markets, facing a downward-sloping residual demand curve. In the first market, the monopolist optimally offers q_1 and receives p_1 . In the second market, the monopolist increases its quantity by q_2 , getting p_2 . The key insight is that, because q_1 has already been committed in the first market, the monopolist's strategic position changes in the second market, creating an incentive to produce more, as can be seen by the shift in the marginal revenue curve. The monopolist anticipates these effects and splits its committed quantity between the two markets, equalizing the marginal revenue in the first market to the market price in the second market, which becomes the relevant opportunity cost.⁴ The monopolist behavior generates a price premium in the day-ahead market, with $p_1 > p_2$.

We show that a price premium in the first market can arise under more general conditions, even when several firms are competing in the market. In the presence of asymmetric firms, we show that some firms endogenously withhold quantity, as the monopolist in the example, whereas others arbitrage the price differences away. In particular, large firms have an incentive to withhold quantity in the first market to increase marginal prices, whereas small firms have an incentive to oversell.

We test our theoretical predictions by analyzing firm behavior in the Iberian electricity market. The Iberian market provides several key advantages for testing the implications from our model. First, the Iberian market allocates hourly electricity production from producers to consumers using a day-ahead auction and seven subsequent intraday auctions. This market structure allows firms to update their sales and purchase positions multiple times during a day. Second, there are ample publicly available data for the Iberian electricity market, which allow us to analyze

²In practice, there can be more than two sequential markets. For example, the Iberian electricity market in our empirical analysis has up to eight sequential markets to allocate hourly electricity production.

³This assumption comes from the fact that market operators in electricity markets typically schedule most or all expected demand in the day-ahead market, and use subsequent markets to allow for modifications in production commitments between suppliers.

⁴In the context of the broader economics literature, the regulation in this market creates something similar to a dynamic price-discriminating monopolist facing naïve consumers, leading to a declining price path (Coase 1972). Indeed, in the presence of infinitely many sequential markets, the monopolist creates a price schedule that mimics first-degree price discrimination.

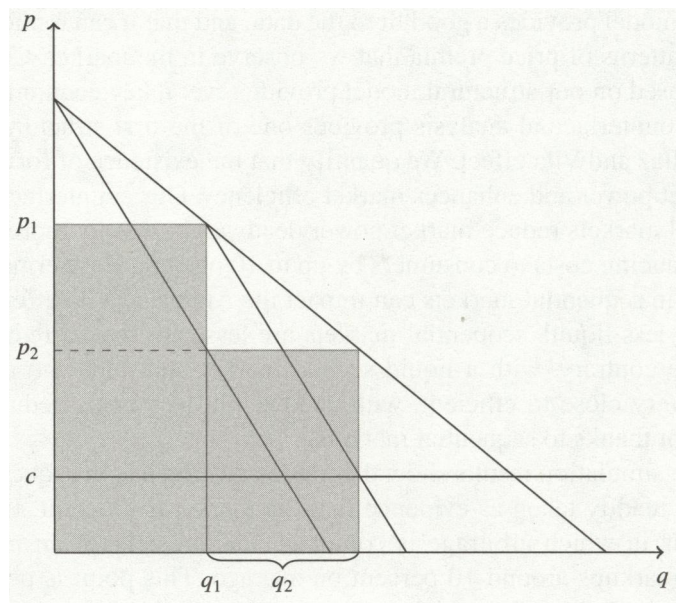


FIGURE 1. A PRICE-DISCRIMINATING RESIDUAL MONOPOLIST

Notes: This figure shows the intuition behind the declining price result (Result 1). A residual monopolist with marginal cost c has an interest in more expensive power plants setting a high price in the first market (p_1). In the second market, the monopolist can regain some of the withheld quantity by lowering the price (p_2).

firms' strategic behavior in sequential markets at high resolution. Third, the Iberian market consists of a few dominant firms that have 70 percent of the market share, as well as many competitive fringe firms, making the exercise of market power relevant. Because our data reveal firms' identities, we can investigate how dominant and fringe firms differently respond to the incentives created by the sequential markets.

Consistent with the predictions from our model, we find a systematic day-ahead price premium in the Iberian electricity market. At the aggregate level, we show that the day-ahead premium correlates with variables that explain the ability to exercise market power, such as total forecasted demand and the elasticity of residual demand. At the micro level, we find that dominant firms undersell or withhold their total electricity production in the forward markets, consistent with our theory. Conversely, fringe producers systematically oversell electricity in the day-ahead market, and update their positions in later markets by purchasing electricity at a lower price. Finally, we also document substantial heterogeneity in overselling and withholding among different generation technologies (e.g., wind, thermal, hydro, and demand), regulatory regimes, and between sophisticated and nonsophisticated firms, providing further support to our theory.

To quantify the welfare effects of sequential markets, market power, and arbitrage, we build a structural Cournot model with a forward and real-time market.⁵ We

⁵The model is a dynamic extension to the static Cournot game in Bushnell, Mansur, and Saravia (2008).

show that the model provides a good fit to the data, and that it can replicate the magnitudes and patterns of price premia that we observe in the market. Counterfactual simulations based on our structural model provide several key economic insights.

First, our counterfactual analysis provides one of the first structural quantifications of the Allaz and Vila effect. We quantify that the existence of forward markets reduces market power and enhances market efficiency. Our counterfactuals suggest that sequential markets reduce market power deadweight loss by more than 60 percent, while reducing costs to consumers by up to 10 percent. Furthermore, we show how liquidity in sequential markets can impact the Allaz and Vila effect. If the second market is less liquid, sequential markets are less effective in reducing market power. On the contrary, with a liquid second market and under no arbitrage, the allocation is very close to efficient, with deadweight loss being reduced by more than 90 percent thanks to sequential markets.

Second, the simulation results show that price convergence in sequential markets should not be readily taken as evidence that the market is efficient. Even in those counterfactuals in which arbitrage is complete, inefficiencies from market power remain, with markups around 10 percent on average. This point is policy-relevant and timely in the context of electricity markets in the United States and other countries. Regulators are considering introducing financial speculators (i.e., virtual bidders) to “improve efficiency.” Whereas arbitrage might improve market outcomes, price convergence in itself is not a sufficient statistic for market efficiency. We believe that such a cautionary tale can be relevant in other policy settings.

We further show that arbitrage in itself is not necessarily welfare improving in the presence of market power.⁶ We find that full arbitrage successfully lowers consumer costs between 2 to 9 percent on average. However, arbitrage reduces production from dominant firms, which in turn can increase production costs, highlighting a key trade-off for policymakers between distributional concerns and efficiency. During our sample period, we estimate increased production costs due to full arbitrage in the order of 11 million euros per year, and the reduction in consumer cost is in the order of 138 million euros per year.⁷ These results suggest that, while arbitrage might harm efficiency, the gains in consumers savings from reduced prices are an order of magnitude larger.

At a more general level, our counterfactual simulations present a quantitative assessment of the role of price discrimination on welfare, a topic that is relevant beyond sequential markets. It is well known that price discrimination may hurt consumers (as consumer rents are extracted more effectively), but that it might increase overall welfare. Using highly detailed data and a structural model, we quantify this trade-off in the context of electricity markets. Our counterfactuals allowing arbitrage (i.e., inducing price convergence between the forward and real-time market) are analogous to experiments assessing the welfare effects of banning price discrimination in other settings.

⁶To emphasize the relationship between the degree of market power and the magnitude of inefficiencies due to arbitrage, we show that arbitrage leads to larger increases of deadweight loss during hours of high market power.

⁷Whereas these magnitudes are not large, average market power levels during our sample period are low due to the economic crisis and the large presence of renewable power, which is cheap at the margin. However, in relative terms, our results suggest that moving from no arbitrage to full arbitrage can increase deadweight loss by more than 20 percent.

Related Literature.—Our paper relates to several literatures. First, it relates to the literature examining sequential markets, market power, and arbitrage, starting with the seminal work by Allaz and Vila (1993). Different than the canonical model, in our setting arbitrage is not competitive, giving rise to endogenous systematic price differences. Also, the forward market arises from the demand side, which is concentrated in the day-ahead market.⁸ Our theoretical framework relates to the literature on durable good monopolies when consumers are not sophisticated or impatient (Coase 1972), and clearance sales (Lazear 1986; Nocke and Peitz 2007). The paper is also related to Coutinho (2013), who shows that similar strategic effects can arise in markets for resale of Treasury bills using a theoretical model of supply function. Our work complements the theoretical literature by quantifying the interaction between market power and arbitrage in sequential markets, testing (and confirming) the predictions from the model empirically.

Second, our theoretical and empirical findings provide key insights into current policy discussions in energy policy. The lack of arbitrage has been documented as a central policy question in many electricity markets in the United States and other countries (Saravia 2003; Longstaff and Wang 2004; Borenstein et al. 2008; Hadsell 2008; Bowden, Hu, and Payne 2009; Jha and Wolak 2014; Birge et al. 2014; Parsons et al. 2015). Our model predicts that a positive day-ahead premium can arise as a result of market power, which we test using data from the Iberian electricity market. The model's prediction is also consistent with price premia found in the literature for other electricity markets including the California, Midwest, New England, New York, and PJM markets.⁹

Our paper is most closely related to Saravia (2003) and Borenstein et al. (2008). Borenstein et al. (2008) show that the price differences in the forward and spot markets in the California electricity market cannot be fully explained by risk aversion, estimation risks, or transaction costs, finding that market power can be a channel driving price premia in the California electricity market. Saravia (2003) shows that regulatory constraints in the New York electricity market generate price premia in the presence of congestion. We show that limited arbitrage can arise endogenously in the presence of asymmetric firms, even in the absence of explicit regulatory constraints or congestion. We also develop a structural model to test and quantify the interaction between market power, arbitrage and welfare using highly disaggregate data.

Finally, our empirical analysis provides a new way to exploit sequential markets to measure market power. Comparing bidding strategies across firms in sequential markets can provide an additional test for competitive behavior, complementing

⁸In Allaz and Vila (1993), the forward motive arises for strategic reasons, as suppliers compete to obtain market share by lowering their price in the earlier markets. In our setting, the forward market arises from the demand side. In this sense, our baseline setup closely resembles a procurement auction in which demand is allocated among the participants, followed by resale.

⁹Similar to our empirical findings in the Iberian electricity market, the literature finds *positive* day-ahead price premia in many electricity markets, including the New York market (Saravia 2003), the PJM market (Longstaff and Wang 2004), the New England market (Hadsell 2008), and the Midwest market (Bowden, Hu, and Payne 2009) for most hours. For the California market, Borenstein et al. (2008) find statistically significant *negative* day-ahead price premia for part of their sample periods. However, they also document evidence of monopsony power during this period that was exercised by buyers in the market.

existing methods that are based on markup estimation.¹⁰ We also show that price differences between sequential markets, which are readily available for many markets, can in themselves be interpreted as a lower bound to markups in the day-ahead market.

I. Model

In this section, we develop a model of sequential markets, inspired by electricity markets. We consider a simple model in which a residual monopolist is deciding production in two stages. For simplicity, and given our main focus on the role of market power, we consider a setting in which there is no uncertainty and no risk aversion.¹¹

The problem of the monopolist is to decide how much commitment to take at the first market (forward or day-ahead market) at a price p_1 , and how much to adjust its commitment in the second market (real-time market) at a price p_2 . Final production is determined by the sum of commitments in each market.

Residual Demand.—Residual demand in the first market (day-ahead) is given by

$$(1) \quad D_1(p_1) = A - b_1 p_1.$$

A represents the total forecasted demand, which is planned for and cleared in the day-ahead market. Whereas demand is inelastically planned for, the monopolist faces a residual demand with slope b_1 . One micro-foundation is that residual demand is the inelastic demand A minus the willingness to produce by fringe suppliers, who are willing to produce with their power plants as long as p_1 is above their marginal cost, $C^{\text{fringe}}(q) = q/b_1$.

In the second market (real-time), commitments to produce can be updated. Therefore, it is a secondary market for reshuffling the agreed production, while the total production remains A . We assume that the residual demand in the second market is given by

$$(2) \quad D_2(p_1, p_2) = (p_1 - p_2)b_2.$$

This residual demand implies that fringe suppliers produce more than their day-ahead commitment if $p_2 > p_1$, and produce less than their day-ahead commitment if $p_2 < p_1$. Consequently, the monopolist increases its quantity in the second market as long as $p_1 > p_2$. For the special case of $b_1 = b_2$, this residual demand implies that fringe suppliers are willing to move along their original supply curve. In the context

¹⁰There is an extensive literature that estimates market power in electricity markets. For example, see Wolfram (1998, 1999); Borenstein, Bushnell, and Wolak (2002); Kim and Knittel (2006); Puller (2007); Bushnell, Mansur, and Saravia (2008); McRae and Wolak (2014); and Reguant (2014). To estimate market power, most papers estimate power plants' marginal costs using engineering estimates or first-order conditions. Instead, we show the presence of market power by showing asymmetric behavior in sequential markets by fringe and dominant firms.

¹¹We extend the model to allow for common uncertainty in the counterfactual experiments in Section IV.

of electricity markets, we assume that $b_2 \leq b_1$, as production tends to be less flexible in the real-time market.¹²

Monopolist Problem.—The monopolist maximizes profits by backward induction. At the second market, p_1 and q_1 have already been realized. The monopolist's problem is

$$(3) \quad \begin{aligned} \max_{p_2} \quad & p_2 q_2 - C(q_1 + q_2), \\ \text{s.t.} \quad & q_2 = D_2(p_1, p_2), \\ & q_1 = D_1(p_1). \end{aligned}$$

The solution to the last stage gives an implicit solution to p_2 and q_2 . In the first stage, the monopolist takes into account both periods. By backward induction, q_2 and p_2 become now a function of p_1 ,

$$(4) \quad \begin{aligned} \max_{p_1} \quad & p_1 q_1 + p_2(p_1) q_2(p_1) - C(q_1 + q_2(p_1)), \\ \text{s.t.} \quad & q_1 = D_1(p_1). \end{aligned}$$

A. Market Equilibrium and the Role of Arbitrage

To gain intuition, we consider equilibrium outcomes for a simplified example with linear residual demand and constant marginal costs, $C(q) = cq$.¹³ Result 1 summarizes some useful comparative statics.

RESULT 1: *Assume that the monopolist is a net seller in this market (i.e., $q_2 > 0$) and there is no arbitrage. In equilibrium,*

- $p_1 > p_2 > c$;
- $p_1 - p_2$ is increasing in A , decreasing in b_1 , and increasing in b_2 ;
- if $b_2 = b_1$, $q_1 = q_2$;
- if $b_2 < b_1$, $q_1 > q_2$.

In equilibrium, the monopolist exercises market power in both markets. However, in the second market, its position from the first market is already sunk. Therefore, it has an incentive to produce some more quantity, whereby lowering the price and its markup. Intuitively, price differences are correlated with market features that give rise to market power (large demand, inelastic fringe supply). One can also see that price differences, $p_1 - p_2$, provide a lower bound to markups in the day-ahead market, $p_1 - c$.

¹²Empirically, we find that b_1 tends to be larger than b_2 by a factor of 5 to 10. Hortaçsu and Puller (2008) find evidence that the supply curve of fringe suppliers is relatively inelastic at the real-time market, which could be explained by a lack of sophistication or adjustment and participation costs.

¹³We provide a full derivation of equilibrium prices and quantities, as well as proofs of the results, in online Appendix A.

It is important to note that this simplified example has been presented under the assumption that the monopolist is a net seller. Under the alternative assumption that the monopolist is a net buyer (i.e., a monopsonist), the results are reversed: in the absence of arbitrageurs, or in the presence of limits on arbitrage, there would be a real-time premium, i.e., $p_2 > p_1$.¹⁴

Arbitrage.—In our example, we have assumed that demand is not elastic and, by construction, is all planned for already in the first market (A). This is motivated by the fact that the electricity day-ahead market is intended to plan for all (or most) forecasted demand. The downward-sloping residual demand comes from the presence of fringe suppliers, which are bidding at their marginal cost.

In equilibrium, fringe suppliers sell more in the first market at a better price, and then reduce their commitments in the real-time market at lower prices, making some profit. However, the equilibrium leaves room for further arbitrage. Given that $p_1 > p_2$, competitive fringe suppliers could oversell even more at the first market, e.g., by offering their production below marginal cost, and trade back those commitments in the second market. The residual demand would no longer be given by total demand minus the marginal cost curve of fringe producers.

We consider the case in which fringe suppliers compete for these arbitrage opportunities until $p_1 = p_2$.¹⁵ Abstracting from changes in the slope of the residual demands (b_1, b_2), consider an arbitrageur that can shift the residual demand at the forward market, by financially taking a position to sell, and buy back the same quantity at the real-time market.¹⁶ An arbitrageur can sell a quantity s in the first market, and buy it back in the second market.¹⁷ The modified residual demands become

$$(5) \quad D_1(p_1, p_2, s) = A - b_1 p_1 - s,$$

$$(6) \quad D_2(p_1, p_2, s) = (p_2 - p_1)b_2 + s.$$

Note that this formulation is analogous to demand not clearing in full in the first market. The effective demand in the first market is $A - s$, while s is added to the residual demand in the second market.

If the costs of arbitrage are relatively small and the arbitrageurs market is competitive, s increases until p_1 converges to p_2 . In the equilibrium, therefore, the presence of such arbitrage erases the forward market price premium.

¹⁴In the context of the California electricity market, Borenstein et al. (2008) find evidence in support of monopsony power.

¹⁵An alternative interpretation is that the demand side could arbitrage by waiting until the second period. We emphasize supplier incentives in the theoretical framework because, empirically, it is where we see most of the arbitrage. Demand-side arbitrage is not as salient for two main reasons. First, the regulator tends to plan for all forecasted demand in the first market, so total cleared demand tends to be quite constant across markets. This implies that demand, in net, has limited contributions to price arbitrage. Second, final electricity consumers are very insensitive to prices and retail choice opportunities (Hortaçsu, Madanizadeh, and Puller 2015). At the same time, such consumers are served by vertically integrated generating companies. These companies profit more from increasing generation prices than by lowering rates to consumers, which has very limited impacts on profit. Therefore, even if such arbitrage opportunities exist, the largest companies do not have incentives to exploit them.

¹⁶Virtual bidders in markets such as Midcontinent Independent System Operator (MISO) and California engage in these type of commitments, which, contrary to our empirical application, are allowed in those markets.

¹⁷Because we do not restrict s to be positive, the arbitrageur could be effectively selling at the first market and buying back at the second market. In equilibrium, however, $s > 0$.

RESULT 2: Assume that the monopolist is a net seller and arbitrageurs are competitive so that, in equilibrium, s is such that $p_1 = p_2$. Then,

- $p_1 = p_2 > c$;
- $p_1(s)$ is decreasing in s and $p_2(s)$ is increasing in s ;
- $q_1(s)$ decreases with s and $q_2(s)$ increases with s ;
- s reduces total output by the monopolist, $q_1 + q_2$.

One important insight that arises from this result is that competitive arbitrage in this market does not lead to competitive prices. The rationale for this result is that the monopolist is still required to produce the output after all sequential markets close. The arbitrageurs are only engaging in financial arbitrage, but do not produce s . One can also see that arbitrage reduces *both* the price and the monopoly quantity in the first period, as the monopolist responds to the arbitrage by withholding some more output.

Explicit Limits to Arbitrage.—In practice, s might not be chosen to equalize prices, e.g., due to some institutional constraints or transaction costs. In electricity markets, it is common to limit participation to agents that have production assets. An arbitrageur cannot take a purely financial position in the market unless it is backed up by an actual power plant.

We introduce arbitrage constraints on fringe suppliers, by introducing an exogenous limit K on s .

RESULT 3: Assume that arbitrageurs are limited in their amount of arbitrage, i.e., $s \leq K$. Then,

- whenever the arbitrage capacity is binding, i.e., $s = K$, then $p_1 > p_2$;
- price differences are more likely to arise when K is lower, all else equal;
- price differences are more likely to arise when A and b_2 are large, and when b_1 is small, all else equal;
- $p_1 - p_2$ is increasing in A , decreasing in b_1 , and increasing in b_2 .

The result shows that a price premia arises whenever arbitrageurs are capacity constrained. In such case, the comparative statics are similar to the case without arbitrage, and the price premia is larger when market conditions are more extreme, e.g., when demand is large or the fringe supply is less elastic.

Strategic Arbitrage.—Even if there are no explicit limits to arbitrage, limited arbitrage can arise endogenously due to limited competition on the arbitrageur side. For example, in electricity markets, participation by purely financial companies is often not allowed. In other markets, entry costs or participation barriers might limit the number of participants. As a polar example, we consider the presence of a single arbitrageur, who has an incentive to arbitrage price differences, but not to close the gap completely. In our setting, and given the limitations to arbitrage in the market, we interpret this special case as representing a scenario in which a limited set of players can engage in price arbitrage. The profits of the arbitrageur are given by $(p_1 - p_2)s$.

We calculate the sequential Cournot equilibrium between the monopolist producer (q_1, q_2) and the strategic arbitrageur (s). In the first stage, they choose q_1 and s simultaneously, in the second stage, the monopolist can adjust to q_2 .

RESULT 4: *Assume that there is a single firm which has the ability to arbitrage, and maximizes profits. Then,*

- $p_1 > p_2 > c$;
- $p_1 - p_2$ is increasing in A , decreasing in b_1 , and increasing in b_2 ;
- $p_1 - p_2$ is smaller than in the absence of strategic arbitrage, i.e., $s > 0$.

In the presence of a strategic arbitrageur, the main predictions of the model hold. The price premium is larger when demand is large, fringe suppliers submit inelastic supplies, and the real-time market is more elastic.

Endogenous Withholding and Arbitrage with Asymmetric Firms.—In our empirical setting, arbitrage is performed by firms that also produce in the market. Suppose there exists a limited set of firms competing in sequential markets. Could this behavior be explained as an equilibrium? As a first reaction, one might think that a firm producing in the market should have no incentive in reducing the day-ahead price. However, if firms are sufficiently different in size, we show that some will have an incentive to withhold, and some will have an incentive to arbitrage.

Consider a situation in which there are two firms, one of which is large and sells substantial amounts of output, and one of which is small. Both of them have the ability to engage in strategic withholding or arbitrage. For the purposes of the empirical exercise, it is useful to think about the large firm as one with several types of production (e.g., coal, gas, nuclear, wind), and the small producer as one with wind farms. As before, there is also a set of nonstrategic fringe firms.

We model the large firm as the monopolist in the model above, with constant marginal cost c . We assume that the marginal cost of the monopolist is low enough that it becomes a large player in equilibrium. For the small firm, we assume that it behaves as the strategic arbitrageur in the previous example, with the main difference that, on top of getting profits from arbitrage, it also gets profits from wind output. The profit function becomes $p_1 q^w + (p_1 - p_2)s$, where q^w is the farm's wind output, which is exogenously given by weather patterns, e.g., wind speed and direction.¹⁸

The presence of wind output attenuates the incentives of the arbitrageur to bring p_1 down. If the wind farm is small enough, it still has a net incentive to arbitrage and increase its profits by overselling in the first markets, i.e., setting $s > 0$. However, if the quantity produced by the wind farm is large enough, the arbitrageur does no longer have an incentive to arbitrage, and behaves in line with the monopolist, driving a price premium. Result 5 summarizes the comparative statics.

¹⁸ Empirically, these results also hold for other technologies, but wind makes the comparison simpler as final production is mostly exogenous. Similar insights arise with the second firm is also choosing quantity endogenously, as long as it is smaller in size due to predetermined factors, e.g., due to capacity constraints or higher marginal costs.

RESULT 5: Assume that there is a single firm which has the ability to arbitrage, who also owns wind farms, and maximizes profits. Then,

- $p_1 > p_2$;
- $p_1 - p_2$ is increasing in A and q^w , decreasing in b_1 , and increasing in b_2 as long as $q^w < \bar{q}^w$;
- $s > 0$ as long as $q^w < \underline{q}^w$, otherwise $s \leq 0$;
- $q_2 > 0$ as long as $q^w < \bar{q}^w$.¹⁹

Under this scenario, a price premium still arises. If the wind farm is small, it has an incentive to arbitrage price differences, i.e., $s > 0$. However, if the wind farm is large enough, then it has no incentive to arbitrage. In fact, it may have an incentive to undersell in the first market, i.e., $s < 0$. The monopolist, on the contrary, does not arbitrage the price differences away as long as it is large relative to the other player. If the monopolist became small enough relative to the wind producer, the roles could eventually revert. The monopolist would behave as an arbitrageur ($q_2 \leq 0$), while the wind farm would create the price premium. For an intermediate range of wind output $q^w \in [\underline{q}^w, \bar{q}^w]$, both strategic players have aligned incentives to increase the premium, i.e., $s < 0$ and $q_2 > 0$. In all cases, $p_1 > p_2$.

Result 5 is important, as it highlights that asymmetric behavior in arbitrage can arise endogenously in a sequential model with asymmetric firms. The larger firms have an incentive to behave as the “monopolist” in our model, withholding in the day-ahead market to increase the marginal price of electricity. On the contrary, the smaller firms have an incentive to arbitrage some of the price differences away.

B. Summary

Results 1 through 5 describe the behavior of a strategic firm under alternative assumptions regarding the extent and nature of arbitrage. Table 1 shows common predictions across alternative models of arbitrage (no arbitrage, full arbitrage, limited arbitrage, strategic arbitrage, and asymmetric oligopoly). One can see that a price premium arises as long as arbitrage is not full, and in such cases it is increasing with demand, decreasing with the slope of the residual demand in the first period, and increasing with the slope of the residual demand in the second period. Furthermore, in all equilibria the largest firm has an incentive to withhold output in the first market, and increase its production commitments in the second one.

We explore these testable implications in the empirical section. Before we proceed to our empirical analysis, we give some more institutional details on how sequential markets are organized in our particular application, the Iberian electricity market.

¹⁹In particular, $\bar{q}^w \equiv \frac{(5b_1^2 + 2b_1b_2 + b_2^2)(A - b_1c)}{17b_1^2 - 6b_1b_2 + b_2^2}$, $\underline{q}^w \equiv b_1 \frac{A - b_1c}{5b_1 - b_2}$, and $\bar{q}^w \equiv \frac{(3b_1 + b_2)(A - b_1c)}{7b_1 - b_2}$, with $\underline{q}^w \leq \bar{q}^w \leq \bar{q}^w$: see online Appendix A for details.

TABLE 1—SUMMARY OF PREDICTIONS BY CASES

	No arbitrage (1)	Full arbitrage (2)	Limited arbitrage (3)	Strategic arbitrage (4)	Asym. firms (5)
Positive premium	✓	—	✓	✓	✓
Increasing with A	✓	—	✓	✓	✓
Decreasing with b_1	✓	—	✓	✓	✓
Increasing with b_2	✓	—	✓	✓	✓*
Undercommitment by dominant	✓	✓	✓	✓	✓
Overcommitment by fringe	—	✓	✓	✓	✓*

Notes: Premium refers to $p_1 > p_2$. Undercommitment by dominant refers to additional quantity in the second period being positive, $q_2 > 0$. Overcommitment by fringe refers to additional quantity supplied by the arbitrageur/smaller firm being positive in the first period, $s > 0$. See Table A.1 in the online Appendix A for numerical expressions under the assumption that $b_1 = b_2$.

* As long as second strategic firm is small enough.

II. Institutions and Data

We leverage the unique market structure of the Iberian electricity market to study strategic behavior in sequential markets.²⁰ We begin by providing institutional details on how the sequential markets are organized in our setting. We then explain what features of a typical deregulated electricity market restrict full arbitrage between the forward market and the spot market. Finally, we describe the data used for our empirical analysis.

A. Sequential Markets in the Iberian Electricity Market

A deregulated electricity market usually consists of a day-ahead forward market and a real-time spot market. Most energy production is first allocated in the day-ahead market. The real-time market is used to ensure the balance between scheduled demand and supply. The Iberian electricity market is organized in a centralized fashion, with a day-ahead market and up to seven intraday (real-time) markets. Figure 2 shows how the sequential markets are structured. In the day-ahead market (day $t - 1$), producers and consumers submit their supply and demand bids for each of the 24 hours of delivery day t , and production for each hour is auctioned simultaneously using a uniform rule, setting a marginal price of electricity for each hour of the day.²¹ The day-ahead market plans for roughly all expected electricity, whereas sequential markets allow for retrading.²² After the clearance of the day-ahead market, the system operator checks congestion in the electricity grid. In the presence of congestion, the system operator may require some changes in the initial commitments, readjusting the position of several units based on their willingness to readjust.

²⁰The Iberian electricity market encompasses both the Spanish and Portuguese electricity markets, and was created in July 2007.

²¹In reality, the auction takes the form of a modified uniform auction, as explained in Reguant (2014).

²²In terms of volume, roughly 80 percent of the electricity that is traded in the centralized market is sold through this day-ahead market. Firms can also have bilateral contracts in addition to their transactions in the centralized market.

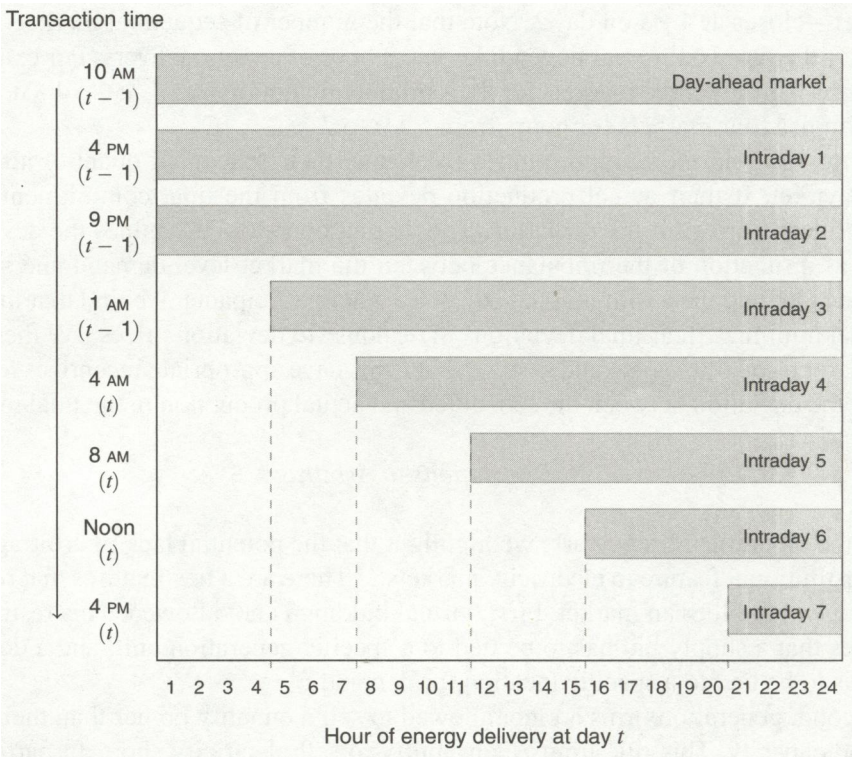


FIGURE 2. SEQUENTIAL MARKETS IN THE IBERIAN ELECTRICITY MARKET

Notes: This figure describes the timeline of sequential markets in the Iberian electricity market during our sample period. For a given hour of their production, firms can bid in the day-ahead market and multiple intraday markets. The position in the last market for a given hour represents their final physical commitment to produce electricity. For example, at noon firms can change their commitments until the fifth intraday market. Their position at the fifth intraday market determines the amount of electricity that they are expected to produce.

After the congestion market, the first intraday market opens, still on day $t - 1$. In the first intraday market, producers and consumers can bid for each of the 24 hours of day t to change their scheduled supply and demand from the day-ahead market. For example, if suppliers want to reduce their commitments to produce, they can purchase electricity in the intraday market. Likewise, if firms want to produce more than the assigned quantity, they can sell more electricity in the intraday market. This means that an electricity supplier can become a net seller or buyer in the intraday market. After the first intraday market, firms have additional opportunities to update their positions through subsequent intraday markets as shown in the figure. In each of the intraday markets, the market clearing price is determined by a set of simultaneous uniform price auctions for each delivery hour.

Sequential markets allow firms to adjust their scheduled production multiple times. For example, consider a firm that wants to deliver electricity for 9 PM on day t . The firm first participates in the day-ahead market at 10 AM on the day before production (day $t - 1$). After realizing the auction outcome of the day-ahead market, the firm can update their position by purchasing or selling electricity in the subsequent seven intraday markets. The final intraday market—the seventh intraday

market—closes at 4 PM on day t . Note that the number of sequential markets available for the firm is different depending on the hour of energy delivery. For example, the firm has only three markets for their production hours from 1 AM to 4 AM, while the firm has four markets for hours from 5 AM to 7 AM.

Firms have no more opportunities to change their scheduled quantity after the final market. If their actual production deviates from the final commitment, they have to pay a price for the deviation. The market operator determines the deviation price as a function of the imbalance between the market-level demand and supply for the hour, and the willingness to adjust by other participants. We find that firms in general minimize their final deviations in response to deviation prices. We therefore do not focus on this aspect and assume that firms have appropriate incentives to minimize the deviation between the scheduled and actual production in the final market.

B. Restrictions to Arbitrage

In the theoretical framework, we highlight that the potential lack of arbitrage is a key institutional feature in electricity markets.²³ There are a few features that restrict arbitrage in the Iberian market. First, virtual bidding is not allowed. This restriction implies that a supply bid has to be tied to a specific generation unit, and a demand bid has to be tied to a specific location for demand.²⁴

Second, generation firms are not allowed to sell a quantity higher than their generation capacity. This rule limits their ability to sell electricity short. In particular, if firms intend to use most of their full generation capacity to produce electricity, this rule implies that such firms have a very limited ability to sell electricity short. Similarly, generation firms cannot purchase electricity in intraday markets if their net production reaches zero. This rule limits their ability to purchase electricity in a market with lower expected price.

Finally, the system operator clears roughly all forecasted demand in the day-ahead market, to plan for how the electricity will flow through the grid and prepare for potential contingencies. This rule limits the arbitrage ability for the demand side. Whereas some demand agents arbitrage, large distribution companies do not appear to engage in such behavior. Indeed, we find that distribution companies often compensate for the arbitrage by other agents, to better match forecasted demand.²⁵

C. Ability and Incentives to Arbitrage

What about the ability and incentive to arbitrage? Given explicit and implicit restrictions on arbitrage, some technologies might be better suited than others. For

²³ See Borenstein et al. (2008) for a description of similar issues in the context of California.

²⁴ While some electricity markets recently started to allow virtual bidding, it is still prohibited in many electricity markets, including the Iberian electricity market. For example, the New York electricity market started to allow virtual bidding in November 7, 2001 (Saravia 2003) and the California electricity market recently started to allow virtual bidding (Jha and Wolak 2014). Although economists are usually in favor of introducing virtual bidding, system operators in many electricity markets are often hesitant about its implementation. They are often concerned that virtual bidding may create large uncertainty in the final deliveries of electricity, which affects the system reliability.

²⁵ As explained above, large retail companies also have limited incentives to look for low electricity prices, as they are vertically integrated and final consumers are very price insensitive. High day-ahead prices, on the contrary, benefit their generation profits.

example, it is relatively difficult to arbitrage at the margin with thermal plants, because, whenever they produce, they tend to use their full capacity (e.g., see Cullen 2015). Therefore, it would be difficult to oversell in the day-ahead market, as firms cannot bid more than their plant capacity. They also have some minimum production requirements to operate safely. Therefore, if they want to offer some output to arbitrage when they are not operating, they need to offer a substantial amount of their capacity (typically 40 percent of their capacity). Thermal power plants also tend to be, for the most part, in the hands of large producers.

As we will show in the data, we find that wind is one of the most active technologies arbitraging in the market. Wind farms have some technological advantages that make them quite attractive to arbitrage. First, wind farms almost never use their maximum capacity because wind does not blow all the time. On average, they use about one-third of their installed capacity. This means that they have greater abilities to sell electricity short in a lower-priced market. Second, wind generation faces substantial uncertainty in their expected wind. Therefore, these units have limited control on their final output, requiring them to update their reports. This means that participation costs in arbitrage might be smaller. Finally, a substantial part of wind farms are not owned by dominant firms. Therefore, from a market structure point of view, a share of wind farms has not only the ability, but also the incentive, to arbitrage.²⁶

D. Data

We construct a dataset using publicly available data from the market operator, Operador del Mercado Ibérico de Energía (OMIE), and the system operator, Red Eléctrica de España (REE), of the Iberian electricity market. Our dataset comes from three main sources.

The first dataset is the bidding data from the day-ahead and intraday markets. On a daily basis, electricity producers submit 24 hourly supply functions specifying the minimum price at which they are willing to produce a given amount of electricity at a given hour of the following day. Similarly, retailers and large electricity consumers submit 24 hourly demand functions specifying the price-quantity pairs at which they are willing to purchase electricity. The market operator orders the individual bids to construct the aggregate supply and demand functions for every hour, and the intersection of these two curves determines the market clearing price and quantities allocated to each bidder. Sellers (buyers) receive (pay) the market clearing price times their sales (purchases). Accordingly, for each of the 24 hours of the days in the sample, we observe the price-quantity pairs submitted by each firm for each of their power plants. We also observe all the price-quantity pairs submitted by the buyers. Importantly, we observe each bidding unit's curves both at the day-ahead and the intraday markets. For each of the bidding units, we know whether their identity and

²⁶During our main sample period, production from wind farms receives the market price plus a flat subsidy. The subsidy is based on final production, so overstating production in forward markets does not increase revenues from the subsidy. This structure was changed in 2013, one year after our main sample period, which removed the incentives to arbitrage. We exploit this policy change in Section IIIB.

type (buyers, traditional power producers (thermal, hydro) or “special regime” producers, such as renewable production, biomass, cogeneration).

The second dataset includes planning and production outcomes from the system operator. These system operator data include market clearing prices, aggregate demand and supply from each type of generation, demand forecast, wind forecast, and weather forecasts. The dataset also includes production commitments at each sequential market at the unit level. One advantage of the system operator data is that we can separate production commitments from wind, solar, and other renewable technologies, whereas in the bidding data these units are often aggregated into a single bidding entity, due to their smaller size. One limitation of the system operator data, however, is that they come from the Spanish system operator, and therefore do not include Portuguese production units. Our results are very similar whether we focus on the Spanish electricity market (using these more detailed operational data), or the Iberian electricity market as a whole (using only bidding data).

The third dataset, which is particularly important for our welfare counterfactual analysis, includes plant characteristics, such as generation capacity, type of fuel, thermal rates, age, and location, for conventional power plants (nuclear, coal, and gas). Combining these data with fuel cost data, we can obtain reasonable estimates of the marginal cost of production at the unit level. We also obtain CO₂ emissions prices and emissions rates at the plant level. As shown in Fabra and Reguant (2014), firms in the Spanish electricity markets fully internalize emissions costs. Therefore, we add them to the unit level marginal costs.

We use data from January 2010 to December 2012. During this period, the four largest generating firms were Iberdrola, Endesa, EDP, and Gas Natural. Their generation market share was on average 68 percent during this period (22, 19, 13, and 11 percent respectively).²⁷ These firms own a variety of power plants from thermal plants to wind farms. In the empirical analysis, we define these four largest firms as *dominant firms*. The market also includes many new entrants that own wind farms or new combined cycle plants. We define them as *fringe firms*.

Table 2 shows the summary statistics of the bidding data and market outcomes, where each variable is associated to its closest analogue in the theoretical model. There are 26,304 hour-day observations in the sample, with an average market price of 44.7/MWh (megawatt hours) in the day-ahead forward market and 43.8/MWh in the spot market. On average, there is a day-ahead market premium of 0.9/MWh. Whereas the premium is not large on average, there is substantial heterogeneity across days and hours, as discussed below. The table also reports the slopes of the residual demand curves that are used in the following sections. The slope is systematically larger for the day-ahead market, as the day-ahead market tends to be have more participants. Finally, the average forecasted wind production is 5.0 GWh (gigawatt hours), being on average approximately 17 percent of total demand.

²⁷Figure D.1 in online Appendix D shows the evolution of their market shares over the sample.

TABLE 2—SUMMARY STATISTICS OF MAIN VARIABLES

	Mean	SD	P25	P50	P75
Price day-ahead (p_1)	44.7	14.1	38.6	48.0	53.5
Price intraday 1 (p_2)	43.8	13.9	38.0	46.2	52.5
Day-ahead premium ($p_1 - p_2$)	0.9	4.0	−0.4	0.5	2.6
Average slope of DA res. demand (b_1)	343.2	102.9	281.9	316.4	369.9
Average slope of I1 res. demand (b_2)	69.9	24.6	54.5	66.2	80.7
Demand forecast (A)	29.3	5.2	24.8	29.4	33.3
Wind forecast (q'')	5.0	2.8	2.8	4.5	6.7

Notes: Prices in euros/MWh. Slopes in MWh/euro. Demand and wind forecasts in GWh. Slope of residual demand computed for the four biggest firms in the market. Number of observations: 26,304.

III. Evidence of Market Power and Arbitrage

In the theory section, we developed a model that characterizes how market power, arbitrage, and constraints for arbitrageurs influence market equilibrium prices in sequential markets. In this section, we provide empirical evidence for the theoretical predictions by analyzing firm behavior in the Iberian wholesale electricity market.

A. Forward Market Price Premium and Market Power

We begin by documenting systematic forward market price premia in the Iberian wholesale electricity market. Our theory predicts that a forward market price premium could emerge if a net-seller firm has market power and market participants have limited arbitrage abilities. This prediction is consistent with the forward price premium observed in the Iberian wholesale electricity market.²⁸ Figure 3 shows average market prices (euros/MWh) for each of the eight sequential markets (the day-ahead market and seven intraday markets), in which the horizontal axis shows hours for electricity delivery. The figure indicates that there is a systematic positive day-ahead price premium—the day-ahead prices are higher than intraday market prices. The prices also appear to be declining in time. This is particularly true for the last intraday market. For example, see prices for hours 12 to 15. The fifth intraday market has a particularly lower price than the prices in the other markets for the same hours.²⁹

Our theory indicates that several key factors can influence the price premium. The results summarized in Table 1 predict that the day-ahead price premium should

²⁸Our theory is also consistent with empirical evidence of price premia in the US electricity markets documented by previous studies. For example, Saravia (2003) documents a forward market price premium in the New York electricity market, which is similar to our finding in the Iberian electricity market. Borenstein et al. (2008) and Jha and Wolak (2014) find a spot-market price premium in the California electricity market, which is still consistent with our theoretical prediction because monopsony power is likely to be an important factor for the price premium in the California market, as documented by Borenstein et al. (2008).

²⁹In addition to the average market prices presented in this figure, we provide the twenty-fifth, fiftieth, and seventy-fifth percentiles of the day-ahead price premium in Table D.1 in the online Appendix. The table suggests that the positive average day-ahead price premium in the figure is not an artifact of some price outliers. The evidence is particularly strong for the afternoon and evening hours, in which the median day-ahead price premium is above 1 euro/MWh across the sequential markets. For hours after midnight, the median day-ahead premium is zero, but the distribution of the price premium is systematically shifted to the right, still giving a positive day-ahead premium on average.

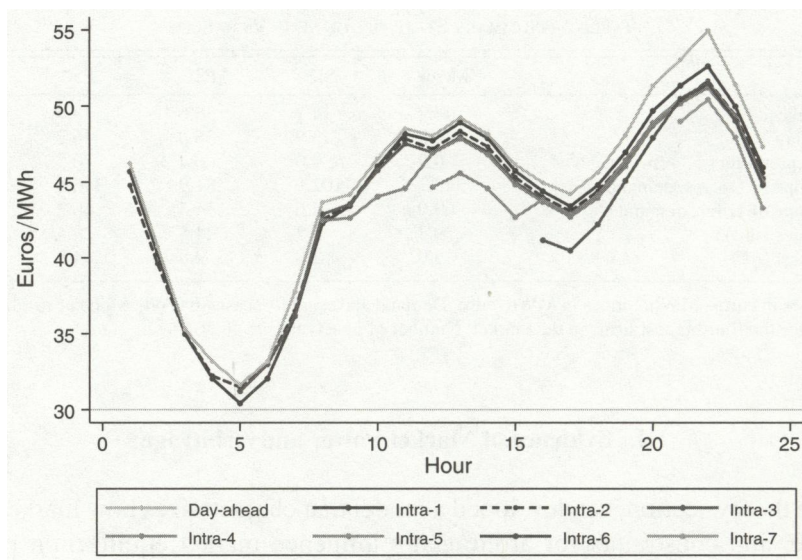


FIGURE 3. MARKET CLEARING PRICE IN THE DAY-AHEAD AND INTRADAY MARKETS

Notes: This figure shows the average market clearing price (euros per MWh) in the day-ahead and intraday markets, in which the horizontal axis shows hours for electricity delivery. Day-ahead market tends to exhibit prices that are on average higher than in the subsequent sequential markets.

be increasing in demand A , decreasing in the slope of the residual demand in the day-ahead market b_1 , and increasing in the slope of the residual demand in the intraday market b_2 . An important advantage of our micro-level bidding data is that we can directly calculate the slopes of residual demand (b_1 and b_2) from the bidding data because we observe plant-level supply and demand bids for every hour in every market. For hour h , day t , and market k , we calculate a residual demand curve for the four dominant firms, Iberdrola (IBEG), Endesa (ENDG), Gas Natural Fenosa (GASN), and EDP/HC (HCENE). We then calculate the slopes of the residual demand at the market clearing price.³⁰ We then test our theoretical predictions by estimating an ordinary least squares (OLS) regression,

$$(7) \quad \Delta p_{ht} = \alpha + \beta A_{ht} + \gamma_1 b_{1ht} + \gamma_2 b_{2ht} + \phi \mathbf{X}_{ht} + u_{ht},$$

where Δp_{ht} is the day-ahead price premium (euros/MWh) for hour h and day t , A_{ht} is the day-ahead demand forecast (GWh), b_{1ht} and b_{2ht} are the slopes of residual demand for the day-ahead market and for the first intraday market. The parameters of interest, β , γ_1 , and γ_2 , describe how the demand forecast and the slopes of the residual demand are associated with the day-ahead price premium. For the control

³⁰We use two methods to calculate the slope at the market clearing price. The first approach is to fit a quadratic function to the residual demand curve and obtain a local slope at the market clearing price. The second approach is to fit linear splines with knots at $\{0, 10, 20, 30, 40, 50, 60, 70, 90\}$ euros/MWh to the residual demand curve. We use the first approach for our main results. Because the two approaches produce similar local slopes, our regression results change very little if we use the second approach.

variables in X_{ht} , we include week of sample fixed effects and hour fixed effects.³¹ We cluster the standard errors at the week of sample.

Table 3 shows our regression results for equation (7). We begin by including only the demand forecast as the independent variable. Column 1 shows that an increase in the demand forecast is associated with an increase in the price premium. In column 2 and 3, we include the slopes of residual demand at the day-ahead market and the first intraday market. Consistent with our theoretical predictions in Result 3, we find that (i) more elastic residual demand in the day-ahead market is associated with a decrease in the price premium, and (ii) more elastic residual demand in the intraday market is associated with an increase in the price premium. In column 4, we include wind forecast. Large wind forecast implies that wind farms operate at closer to their generation capacities. This means that, if wind farms are major arbitrageurs, their arbitrage capacity is lower when there is more wind forecast.

In addition, Result 5 suggests that the presence of wind output may attenuate the incentives to arbitrage, as wind farms become larger. For these two reasons, we expect a positive sign for the effect of the wind forecast on the day-ahead price premium. The result in column 4 indicates that we find empirical evidence consistent with the theoretical prediction. Finally, a potential concern for the OLS regression is that the slopes of residual demand can be endogenous, which is likely to produce attenuation bias for the relationship between the price premium and the slopes.³² To address this concern, we instrument the slopes of residual demand with hourly weather variables (temperature, dew point, and humidity) in column 5. The estimates from the IV regression provide evidence consistent with our theoretical predictions, except for the effect of the demand forecast.³³

B. Arbitrage by Fringe and Dominant Firms

The findings in the previous section imply that (i) there are systematic day-ahead price premia in the Iberian sequential electricity market, and that (ii) market power plays an important role in creating the price premia. With such arbitrage opportunities, our theory predicts that fringe firms should engage in arbitrage, but dominant firms that exercise market power may not arbitrage, and more generally, have an incentive to withhold output in the forward markets. In this section, we examine how fringe firms and dominant firms respond to the price arbitrage opportunities in sequential markets.

The first part of this section investigates graphical evidence of arbitrage. We focus on arbitrage by wind farms and arbitrage by all technologies as a whole. As explained in Section IIC, wind farms have technological advantages that make them quite attractive to arbitrage. Importantly, these advantages are common to wind farms owned by dominant firms and those owned by fringe firms. Therefore, wind provides empirical advantages for us to test if dominant and fringe firms respond

³¹Including alternative dimensions of time fixed effects (e.g., day-of-sample fixed effects or month-of-sample fixed effects) does not significantly change the results.

³²Note that the demand forecast is exogenous and predetermined because it is publicly available to firms before the day-ahead market.

³³Weather explains electricity consumption very well, potentially reducing the remaining variation in the demand forecast.

TABLE 3—DAY-AHEAD PRICE PREMIUM, DEMAND FORECAST, AND SLOPE OF RESIDUAL DEMAND

	(1)	(2)	(3)	(4)	(5)
Demand forecast (GWh)	0.132 (0.025)	0.135 (0.025)	0.103 (0.025)	0.098 (0.024)	−0.002 (0.039)
Slope of residual demand in day-ahead market		−0.019 (0.003)	−0.024 (0.003)	−0.040 (0.003)	−0.090 (0.014)
Slope of residual demand in intraday market			0.050 (0.008)	0.065 (0.009)	0.241 (0.050)
Wind forecast (GWh)				0.365 (0.039)	0.786 (0.121)
Observations	26,145	26,145	26,145	26,145	26,093
IV	No	No	No	No	Yes

Notes: This table shows the estimation results of equation (7). The dependent variable is the day-ahead price premium in euros/MWh. All regressions include hour of the day fixed effects and week fixed effects. The standard errors are clustered at the week of sample. For the IV regression, we use average daily temperature, maximum daily temperature, minimum daily temperature, hourly temperature, dew points, and humidity interacted with the hour of the day dummy variables to instrument the slopes of the residual demand for the day-ahead market and the intraday market.

to arbitrage opportunities differently. The second part of this section leverages our microdata at the plant level to provide statistical evidence on heterogeneity across production technologies. Players in electricity markets are heterogeneous in their technologies such as wind, cogeneration, demand, thermal, hydroelectric, and solar. We examine how firms use these technologies differently to arbitrage in sequential markets.

Aggregate Patterns.—We begin by showing graphical evidence from the raw data. We examine how fringe and dominant firms update their positions (i.e., commitment to produce or purchase a certain amount of electricity for a given hour) through the sequential markets. Consider electricity production from wind farms (q^w). We aggregate plant-level production to the total production for two groups: (i) fringe firms and (ii) dominant firms. The dominant firms include the four largest firms in the market—IBEG, ENDG, GASN, and HCENE.³⁴ Recall that firms have up to seven markets to update their positions before their final position is determined. We use subscript k to denote each market: the day-ahead market ($k = 0$), the first intraday market ($k = 1$), ..., and the seventh intraday market ($k = 7$). For each of the two groups, we define the position at a given market relative to the final position by

(8) $\Delta q_{ghtk}^w = q_{ghtk}^w - q_{ght,final}^w$, with $g = \{\text{fringe, dominant}\}$,

where q_{ghtk}^w is group g 's position at market k for electricity production for hour h on day t , and $q_{ght,final}^w$ is its final position. Therefore, Δq_{ghtk}^w shows the extent to which group g oversells in market k relative to the final position. Similarly, we

³⁴During our sample period, about 30 percent of wind generation came from the wind farms owned by the four dominant firms.

study how firms update their positions for Q , which is their total production from all types of power plants, including wind, thermal, hydroelectric, and others. $\Delta Q_{ghkt} = Q_{ghkt} - Q_{ghk,final}$ shows the extent to which group g oversells its total production in market k relative to its final position. We calculate the means of Δq_{ghk}^w and ΔQ_{ghk} for group g , hour h , and market k , during our sample period, which is from January 2010 to December 2012.

Figure 4 shows the mean of Δq_{ghk}^w in panel A and the mean of ΔQ_{ghk} in panel B. For fringe wind farms, we find substantial overselling in the forward markets. They oversell in forward markets and gradually adjust their positions toward their final positions. This gradual adjustment reflects the option values to adjust positions in the sequential markets. This evidence is not an artifact of their portfolio composition because panel B shows the same evidence for fringe firms' aggregate production, which include production from all technologies. On aggregate across production technologies, fringe firms commit to produce more energy at the forward markets than what they actually deliver.

The evidence is particularly compelling at the discontinuous differences in Δq_{ghk}^w between the sequential markets for hour 5, 8, 12, 16, and 21. These discontinuities are consistent with the market structure. For example, at hour 12, firms have five intraday markets to update their positions. The overselling is largest at the first market and decreases over time. In particular, there is a discontinuous drop between the fourth and fifth intraday markets. This is because firms have no more opportunity to correct their positions after the last market. In the last market, they set their positions nearly equal to their actual final production.³⁵ In contrast, the overselling behavior is very different for hour 11. First, firms do not oversell in the fourth intraday market. This is because the fourth intraday market is the last market for hour 11. Second, they oversell less in the first, second, and third markets for hour 11 relative to the amount of overselling for hour 12. This is because hour 11 has a smaller number of available markets, which creates different option values compared to option values in hour 12.

The data show notably different evidence for dominant firms. Panel A shows that there is almost no significant amount of overselling with wind by these large firms. The difference between their positions in the forward markets and the final production is much smaller than that for fringe wind farms. Furthermore, panel B shows that dominant firms *undersell* in the forward markets with their overall portfolio. They withhold sales in the forward markets and sell more in the later markets, as suggested by our theory. This evidence is consistent with our theoretical prediction (Result 5)—dominant firms that exercise market power have an incentive to withhold output in the forward markets.³⁶

One potential concern is that there is slight overselling by dominant wind farms for the day-ahead market. However, the nature of overbidding appears to be quite

³⁵Note that wind farms in this market have incentives to minimize the deviation between their final commitment quantities and actual production because there are "deviation prices" that penalize errors between final commitments and actual output. Although we do not focus on their responses to the deviation prices in this paper, we find very strong evidence that wind farms generally minimize last-minute deviations.

³⁶When we compare the overselling quantities by fringe firms and those by dominant firms, a potential concern is that the levels of production are different. To examine this point, we construct the figures based on the natural log of production quantities in the online Appendix D, in which we find consistent evidence.

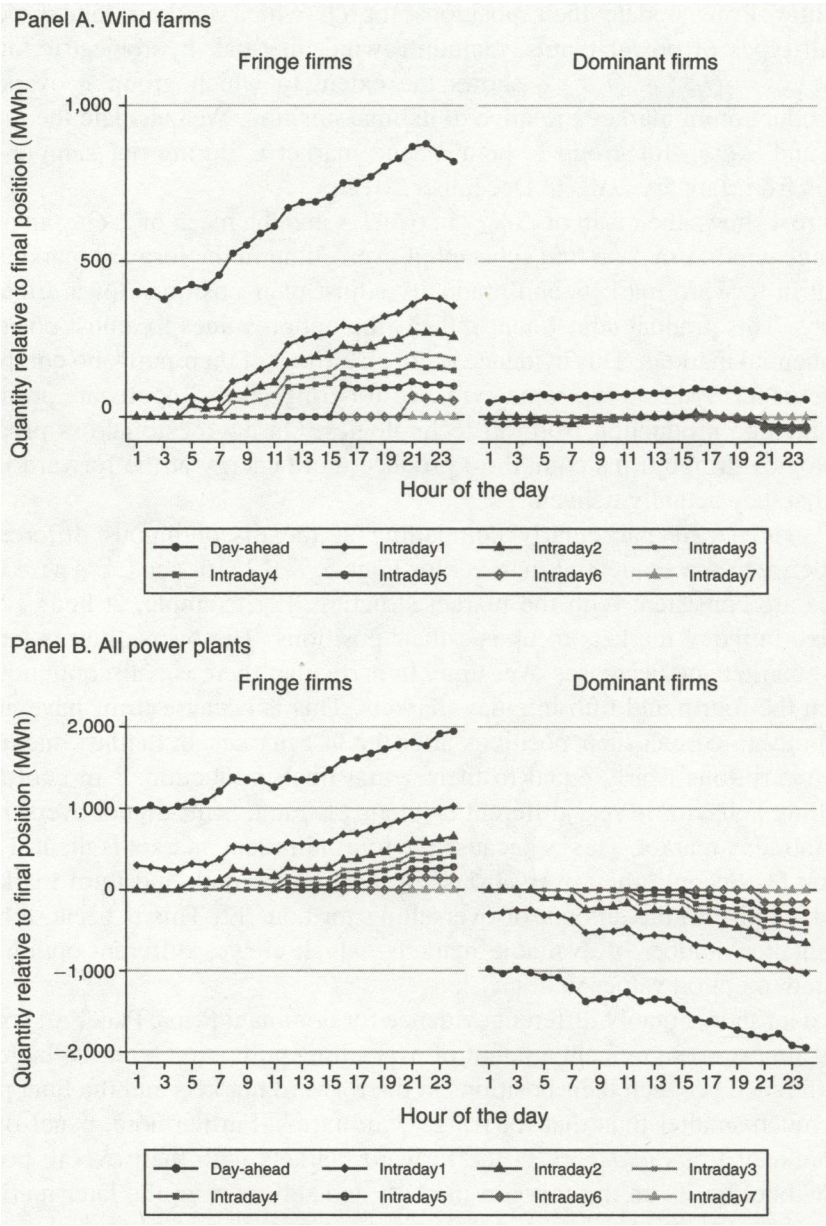


FIGURE 4. SYSTEMATIC OVERSELLING AND UNDERSELLING IN FORWARD MARKETS

Notes: This figure shows average changes in fringe and dominant positions between a given market and their final commitment. Positive values imply that a group is promising more production than it actually delivers after all markets close.

different, as it is flat across hours, while overselling by fringe wind farms appears to correspond to the price arbitrage opportunities. The most likely reason for this behavior is the congestion market, which happens between the day-ahead and the first intraday market. Dominant firms appear overstate wind production in the day-ahead market to reshuffle their production after the congestion market, even

though in net they are withholding output, as shown in panel B.³⁷ In fact, we see no overselling by dominant wind farms in all of the intraday markets, which open after the congestion market. In online Appendix D, we present additional graphs, in which we show the position of each of the four biggest firms, both for wind farms and their overall portfolio. The graphs confirm that congestion induces substantial reshuffling across the dominant firms.³⁸ After congestion is controlled for, behavior in the intraday markets is very consistent across firms, and in line with the predictions of our model.

Further Evidence from a Policy Change in 2013.—Starting from January 2013, the electricity price for wind farms became a rate that was not linked to prices in the sequential markets. This new policy made wind farms have no incentive to arbitrage in the sequential markets. We exploit this quasi-experiment to test if fringe wind farms stopped engaging in arbitrage in 2013, which is the year after our main data period. Figure 5 shows the overselling quantities of fringe wind farms by calendar year. There is systematic forward market overselling in 2010, 2011, and 2012. However, there is no more significant overselling in 2013. Note that the amount of wind production by fringe wind farms is similar between 2012 and 2013. Therefore, this result is not driven by a change in wind production. To explore this point further, we also provide the same figure using the changes in the log of wind production in online Appendix D. We find consistent evidence that the fringe wind farms stopped engaging in arbitrage in response to the policy change in 2013. Furthermore, we find that the price premia in 2013 are slightly larger than those in the previous years, which is consistent with our theory because less arbitrage should produce larger price premia.

Arbitrage by Sophisticated and Nonsophisticated Bidders.—Hortaçsu and Puller (2008) find differences in bidding behavior between “sophisticated” bidders and others in the Texas electricity market. In our context, we find that fringe wind farms engage in arbitrage most, but a potentially interesting possibility is that, even among fringe wind farms, there can be sophisticated and nonsophisticated bidders who exploit arbitrage opportunities differently.³⁹ To examine this question, we exploit our bidding data at the firm level, in which we observe hourly bids by each bidder. We find that a few bidders manage bids for a large capacity of fringe wind farms, where others manage bids for relatively smaller capacity of fringe wind farms.⁴⁰ We test if the large bidders, who are likely to be more sophisticated bidders, arbitrage

³⁷ Importantly, the congestion market does not typically ration wind generation in itself, as it is given priority in the grid. The Spanish wind association reports “In 2012, curtailment on wind power generation reached 0.25 percent of total possible generation, above the 0.18 percent of the previous year” (Spanish Wind Energy Association 2013, p. 44).

³⁸ Congestion is particularly relevant for GASN and ENDG. ENDG appears to be overselling with its portfolio, but this is because some of its power plants are in constrained regions. GASN, on the other hand, appears to massively undersell in the day-ahead market, which is again driven by congestion in the opposite direction. Unfortunately, these congestion patterns are very persistent, and therefore it is difficult to find a period with no congestion during our sample. Most of the flows in the congestion market are traded among these two firms, although IBEG and HCENE also experience some congestion events during the sample, which involve smaller amount of energy.

³⁹ We thank a referee who suggested this analysis.

⁴⁰ Some bidders act as an aggregator for multiple wind farms. Therefore, the owner of a wind farm is not necessarily the one who manages bids in the market. Our bidding data allow us to identify the bidders, which enables us to do the analysis in this section. These bidders offer fixed price contracts to the farm owners, and therefore have

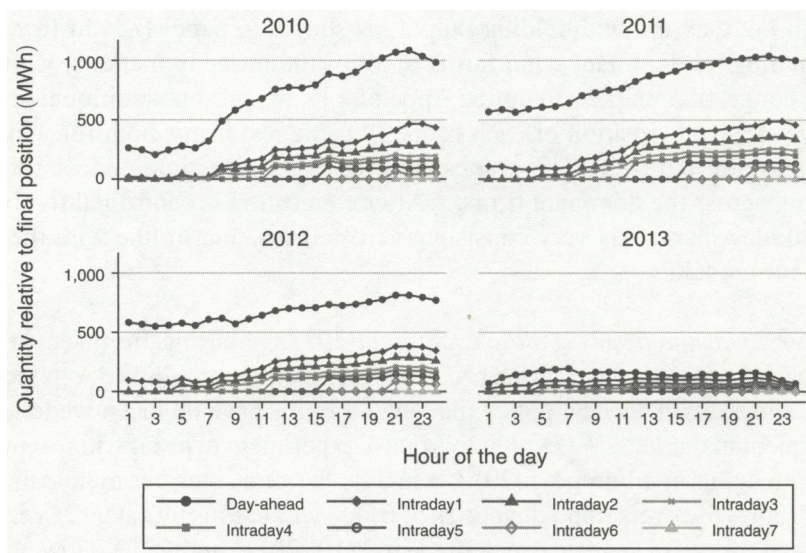


FIGURE 5. EFFECT OF POLICY CHANGE IN 2013 ON FORWARD MARKET OVERSELLING BY FRINGE WIND FARMS

Notes: This figure shows the forward market overselling quantities for fringe wind farms by calendar year. Also see notes in Figure 4.

differently than the small bidders. In Table D.3 in the online Appendix, we show three findings. First, both small and large bidders show systematic overselling in 2010, 2011, and 2012. Second, both types of bidders do not show systematic overselling in 2013, which is because of the policy change discussed in the previous section. Third, large bidders oversell more strongly than small bidders. These findings suggest that sophistication in bidding is a key factor to explain heterogeneity in arbitrage among fringe wind farms.

Heterogeneity in Arbitrage by Production Technologies.—The aggregate patterns provide strong evidence that fringe and dominant firms respond to the incentives in the sequential markets in a way that is consistent with our theoretical predictions. Our theory also suggests that the amount of arbitrage should be positively associated with price premia that are forecastable from market fundamentals such as demand forecasts. In addition, arbitrage can differ by production technologies because the ability to arbitrage depends on power plant types.

To test these predictions, we leverage our microdata at the firm level by production technologies—wind, cogeneration, demand, thermal, hydroelectric, solar, and all technology as a whole. Because the level of production is very different by firms and production technologies, we examine log deviations. For firm j , production technology s , hour h , and day t , we define the change in the firm's position from the day-ahead market to the final commitment by $\Delta \ln q_{jht, DA}^s = \ln q_{jht, DA}^s - \ln q_{jht, FI}^s$ and the day-ahead price premium by $\Delta p_{ht, DA} = p_{ht, DA} - p_{ht, FI}$.

an incentive to maximize market wind farm profits. As shown in Results 5, if they are small enough, they will still have an incentive to arbitrage.

Similarly, we define the same variables for the change in the firms' positions and the price premium between the first intraday market and the final market: $\Delta \ln q_{jht, I1}^s = \ln q_{jht, I1}^s - \ln q_{jht, FI}^s$ and $\Delta p_{ht, I1} = p_{ht, I1} - p_{ht, FI}$. Given that we want to test how firms change their positions in response to price premia that are forecastable at the time of bidding, we need to construct forecastable price premia $\Delta \hat{p}_{htk}$.

To obtain a purely predetermined forecastable relationship between the price premia and the demand forecast, we use hourly data in 2009 (the year before our main data period) to regress Δp_{htk} on the demand forecast, which is a key explanatory variable for the variation of price premia and publicly available at the time of bidding. We then use the regression coefficients to obtain forecastable price premia $\Delta \hat{p}_{htk}$ for the 2010–2012 period. The idea behind this approach is that firms can forecast the price premia by knowing the relationship between the price premia and the demand forecast from the past.⁴¹ We estimate the following equation by OLS, separately for each technology s and each market k ,

$$(9) \quad \Delta \ln q_{jhtk}^s = \alpha + \beta \Delta \hat{p}_{htk} + u_{htk}, \quad \text{with } k = \{DA, I1\},$$

where β shows the percentage change in the arbitrage with respect to a change in the forecastable price premium by 1 euro/MWh. We include firm fixed effects and month of sample fixed effects to regression (9).⁴² We calculate bootstrapped standard errors to account for sampling variation of \hat{p}_{htk} .⁴³

Table 4 presents the regression results for the day-ahead market in panel A and the first intraday market in panel B. In each panel, we present our main results based on data for 2010–2012 as well as the results for 2013, which was after the policy change. Each cell in the table shows the point estimate and standard error of β from a separate regression of (9) for particular technology type and firm type (fringe or dominant). For example, the estimate in the top-left cell implies that 1 euro/MWh increase in the forecastable forward market price premium is associated with an increase in the log arbitrage by a 0.067 percentage point for fringe wind farms.⁴⁴ The results across technology types indicate that fringe firms engage in arbitrage by using a variety of technologies, including wind, cogeneration, demand, and hydroelectric. Among them, wind shows the largest response to the price arbitrage opportunities. This is consistent with our theoretical prediction about the ability to arbitrage—wind has the most flexibility to arbitrage because their capacity constraints tend to be less binding.

⁴¹ If firms can forecast the price premia better than our approach, our approach attenuates β . That is, the estimates from our approach provide a lower bound (in absolute values) for the responses to the forecastable price premia. We explore alternative forecasting models (e.g., including nonlinear terms and interaction terms with hour dummy variables or week dummy variables), but we find that these alternative methods produce very similar estimates to our main results.

⁴² Including these fixed effects has almost no effects on the point estimate of β . We also run the regression with different dimensions of fixed effects such as hour fixed effects and find that our results are robust to different dimensions of time fixed effects.

⁴³ Our bootstrapping is based on 1,000 bootstrapped samples clustered at the day of sample level. Using more than 1,000 bootstrapped samples barely changes the standard errors from the ones obtained from 1,000 bootstrapped samples.

⁴⁴ To interpret the magnitude of our estimates, it is useful to report the distribution of the right-hand side variables. For $\Delta p_{ht, DA}$, we have -2.7 (p10), -0.42 (p25), 0.05 (p50), 2.56 (p75), and 5 (p90). For $\Delta p_{ht, I1}$, we have -3.3 (p10), -1.14 (p25), 0 (p50), 1.55 (p75), and 3.95 (p90), all in euros/MWh. Therefore, considering "1 euro/MWh increase" is reasonable given the variation in the price premium from the data.

TABLE 4—HETEROGENEITY IN ARBITRAGE BY FRINGE AND DOMINANT FIRMS ACROSS TECHNOLOGIES

	Wind	Cogen	Demand	Thermal	Hydro	Solar	All tech
<i>Panel A. Day-ahead market</i>							
Fringe	0.067 (0.010)	0.029 (0.004)	0.007 (0.002)	−0.002 (0.003)	0.025 (0.004)	0.005 (0.004)	0.040 (0.006)
Dominant	0.014 (0.004)	0.002 (0.001)	−0.004 (0.001)	−0.017 (0.006)	−0.005 (0.002)	0.006 (0.002)	−0.063 (0.014)
Fringe (2013)	0.004 (0.010)	0.006 (0.004)	0.014 (0.002)	0.008 (0.003)	−0.001 (0.004)	0.011 (0.004)	0.006 (0.006)
Dominant (2013)	0.001 (0.004)	0.001 (0.001)	−0.009 (0.001)	−0.021 (0.006)	−0.002 (0.002)	−0.005 (0.002)	−0.046 (0.014)
<i>Panel B. First intraday market</i>							
Fringe	0.098 (0.025)	0.027 (0.009)	0.026 (0.009)	−0.006 (0.004)	0.034 (0.008)	0.007 (0.010)	0.057 (0.016)
Dominant	0.006 (0.005)	−0.000 (0.003)	0.000 (0.001)	−0.024 (0.013)	−0.003 (0.003)	0.006 (0.005)	−0.131 (0.052)
Fringe (2013)	0.025 (0.025)	0.019 (0.009)	0.029 (0.009)	−0.009 (0.004)	−0.002 (0.008)	0.031 (0.010)	0.023 (0.016)
Dominant (2013)	0.000 (0.005)	0.001 (0.003)	−0.001 (0.001)	−0.039 (0.013)	0.002 (0.003)	−0.011 (0.005)	−0.147 (0.052)

Notes: This table shows the estimation results of equation (9). The dependent variable is the log deviation between the forward market position and the final position (panel A) and the log deviation between the first intraday market position and the final position (panel B). The independent variable is the forecastable price premium defined in the text. All regressions include firm fixed effects and month of sample fixed effects. The main regressions use data in 2010–2012. The regressions labeled as (2013) use data in 2013 to test the effect of the 2013 policy change that is described in the main text. We calculate bootstrapped standard errors to account for sampling variation of \hat{p}_{htk} .

A useful comparison is solar. Similar to wind, electricity production from solar is volatile and often does not reach its capacity, which provides an advantage for arbitrage. An important difference to wind is that most solar plants do not participate in the intraday markets and receive a flat tariff. Therefore, solar plants do not have an incentive to arbitrage in the sequential markets. This is consistent with our finding that fringe solar plants show statistically insignificant responses to the arbitrage opportunities. Another useful comparison is wind in 2013. As explained in the previous section, wind plants lost their incentive to arbitrage after the policy change in 2013. We find statistically insignificant responses by fringe wind farms in 2013, which is consistent with the fact that the new policy made them have no incentive to arbitrage.

In contrast to the results for fringe firms, we find little evidence of arbitrage for dominant firms. For most technology types, we find insignificant estimates, which indicate that dominant firms do not change their positions in the sequential markets in response to the forecastable arbitrage opportunities. Importantly, we find a statistically significant *negative* coefficient for demand, hydroelectric, and thermal plants, and all technology as a whole. It implies that dominant firms respond to the price premia in the opposite direction, as compared to the responses by fringe firms that engage in price arbitrage. This finding is consistent with our theory, which suggests that dominant firms that exercise market power withhold sales in the forward markets and sell more in the later markets.⁴⁵ Finally, the results for 2013 indicate that

⁴⁵ Note that the negative relationship between the forecastable price premia and the deviation does not necessarily imply that dominant firms reduce the deviation *because* they expect lower price premia. The findings from the

the withholding by dominant firms persisted after the policy change. This evidence is consistent with the fact that the 2013 policy change did not have any direct effects on dominant firms' incentives to withhold production in the forward markets.

Summary.—To summarize, we find that fringe firms engage in profitable arbitrage, whereas dominant firms do not. It is important to note, however, that we do not show that the amount of arbitrage is optimal, and certainly not enough to fully close the price differences. This fact could be explained by several reasons: transaction costs, institutional constraints on the amount of arbitrage (most likely of regulatory nature, as capacity constraints are usually not binding for wind, even with such levels of arbitrage), and strategic arbitrage. We explore these hypotheses in the counterfactual experiments in the next section, by comparing the observed amount of arbitrage to the equilibrium levels under full arbitrage and strategic arbitrage.

IV. Counterfactual Experiments

We find evidence that there is a systematic day-ahead premium in the Iberian electricity market, and that fringe wind farmers appear to arbitrage some of these differences away. How much does this behavior contribute to closing the price gap? What are the welfare implications?

Consider the simple example in Section I under two polar cases, one with no arbitrage (Result 1) and one with full arbitrage (Results 2). From the equilibrium analysis, it follows that the total quantity produced by the monopolist is lower when there is full arbitrage, as p_1 decreases, but p_2 increases. Given that p_2 determines the final allocation, the quantity produced by the monopolist is further away from the first best under full arbitrage. The intuition is that full arbitrage removes the ability of the monopolist to dynamically exercise market power across sequential markets. Under full arbitrage, the monopolist exercises relatively less market power in the first market, but more market power in the last market. These results suggest that introducing full arbitrage in this market is not necessarily welfare-enhancing, as it reduces consumer costs at the expense of lower productive efficiency.

In this section, we quantify this trade-off between consumer surplus and productive efficiency. To do so, we construct a counterfactual model that allows us to empirically assess the interaction between market power and arbitrage. To make the counterfactual experiments empirically relevant, we extend the theoretical model to accommodate for several strategic firms, a flexible marginal cost function, and demand uncertainty.

A. Model for Counterfactual Simulations

We construct an empirical model to simulate the effects of alternative arbitrage policies in this market. The model extends the simple framework in several ways.

previous sections indicate that it is likely to be the opposite. When the demand forecast is high, dominant firms withhold their sales in the forward markets, which increases the price premia. Our regression (9) examines the relationship between the forecastable price premia and the deviation, which does not necessarily imply the causal relationship.

We consider a model with two markets and N strategic firms that are playing a Cournot Nash equilibrium. Firms have capacity constraints. Each firm has a marginal cost curve that is piece-wise linear and continuous. The residual demand that the strategic firms face is also piece-wise linear. Demand in the second period is uncertain, with a commonly known distribution.

We solve the model by backward induction. In the real-time market, firms choose their optimal output levels given their previous commitments, which are the state variable of the game. We solve the last stage as a complementary problem, as in Bushnell, Mansur, and Saravia (2008), for a given quantity sold in the day-ahead market. For the cases in which there is arbitrage, firms take the amount of arbitrage as given. See online Appendix B.1 for the equation details.

In the first stage, firms decide how much energy to sell in the day-ahead market, taking into account the strategic impacts to second-stage payoffs. We solve the optimal quantity in the first market with an iterated best-response algorithm in which firms are maximizing their joint profits between the first and the second market. See online Appendix B.2 for the pseudo-code of the iteration.

We consider four different regimes for our simulations:⁴⁶

- **No Arbitrage:** We consider the case in which the oligopolists participate in sequential markets, and arbitrage is not allowed. Fringe firms passively offer their production at marginal cost.
- **Wind Arbitrage:** We consider the case in which wind farms are arbitraging price differences by, on aggregate, overbidding 20 percent of their actual expected production. We do not take a stand on whether such 20 percent is optimal.
- **Strategic Arbitrage:** We consider the case in which there is an arbitrageur who is strategic. It maximizes its profit by extracting rents from arbitrage without fully closing the price gap. Limited arbitrage arises as an equilibrium outcome.
- **Full Arbitrage:** We consider the case in which there is full arbitrage. The arbitrageurs engage in arbitrage so that the price in the first market equals the expected price in the second market.

Additionally, we consider a first-best counterfactual, in which costs are minimized, as well as a counterfactual in which there is a single market.

We use data from the Iberian electricity market to build the baseline model and assess the welfare implications of these various counterfactuals. To set the different parameters, we exploit the richness of our empirical setting, and use direct empirical analogues in the data. Given that our model is admittedly stylized, this approach allows us to avoid using the model to fit the parameters, providing an extra check on the validity of the framework.⁴⁷

Dominant Firms.—We can obtain a reasonable estimate of the marginal cost of production at the generation unit level. We collect unit-level technology parameters, such as heat rates, from the regulatory report by the market operator. We also obtain

⁴⁶The online Appendix B details how these counterfactuals mathematically affect the simulation procedure.

⁴⁷Here we provide an overview of the steps that we follow to estimate the different parameters. The interested reader can find more details in the online Appendix C.

daily fuel cost data for gas-, coal-, and oil-fired plants and nuclear power plants from the Bloomberg database. Using engineering cost functions for each type of units, we calculate the constant marginal cost for each unit for each day. Based on this procedure, we can construct an increasing step function of the marginal cost curve for each firm that includes their thermal and nuclear power plants.

There are a few important factors to be considered when constructing the marginal cost curve. First, we focus on the marginal cost curve of thermal and nuclear plants owned by dominant firms. The units included in our cost curve produce on average around 40 percent of all electricity generation in the market. Second, not all power plants are available for a given day. For example, a plant is unavailable when it has a scheduled maintenance. We exclude these units to create the marginal cost curves based on available units for a given day. Third, power plants often have bilateral contracts in addition to their production through the centralized markets. Our data include bilateral contracts, which we take into account when building firms' net position.⁴⁸

Finally, we make a simplifying assumption on congestion. As explained above, between the day-ahead and the intraday markets, the system operator adjusts for the congestion by asking firms to change their production, which can give rise to local market power. Modeling the strategic incentives that arise from congestion, by endogenizing network flows in this market, is beyond the scope of this paper.

Residual Demand.—We approximate the residual demand curve faced by dominant firms from the data. Using the bidding data from fringe firms and the approach used in Section IV, we obtain the residual demand curve for the largest four dominant firms. We then calculate b_{1t} and b_{2t} , the slopes of the residual demand curves at the market clearing prices for the day-ahead market and the first intraday market, respectively. It is important to note that our residual demand slopes also take into account any elasticity coming from demand bids. They are also specific to each day and hour, so the slopes capture varying conditions in the competitive environment.

We take the residual demand slopes around the observed equilibrium price, and approximate residual demand linearly. In order to estimate the demand intercept, A , we use day-ahead clearing prices and quantities. For a given day-ahead price p_{1t} , dominant production q_{1t} , and residual demand slope b_{1t} , we calculate $A_t = q_{1t} + b_{1t}p_{1t}$.⁴⁹ The resulting estimates of the term A cannot be directly interpreted. Rather, the term A is an auxiliary construct that allows us to fit the residual demand around equilibrium prices in a parsimonious way. In our approach, this approximation approach is valid as long as our counterfactuals are of local nature, so that the slope estimate for the residual demand is still meaningful.⁵⁰

The empirical evidence from the previous sections shows that wind farms oversell in the day-ahead market. For our counterfactual analysis, we consider arbitrage as a shift in the residual demand curve. We assume that in the data, firms are overstating

⁴⁸ We treat bilateral contracts as a financial position, but we obtain similar results if we impose that firms always produce at least enough electricity to cover their bilateral contracts, as the constraint is almost never binding in equilibrium.

⁴⁹ A similar approach is used in Bushnell, Mansur, and Saravia (2008).

⁵⁰ We have decided to keep the residual demand as locally linear so that our computational model closely matches the theoretical framework. A linear demand also makes the computational simulations less demanding.

their wind output by 20 percent in the day-ahead market. Therefore, we set arbitrage to be 20 percent of the public wind forecast. In other counterfactuals, we investigate alternative market outcomes that endogenize the amount and nature of the arbitrage.

Finally, to model changes in forecasted net demand between the day-ahead market and the real-time market, we use the distribution of changes in scheduled demand minus scheduled wind production. We compute the standard deviation of these changes, on top of other forecastable differences, using a prediction model. We find that the residual uncertainty is approximated very well by a normal distribution, and thus use a normal distribution with mean zero and standard deviation equal to 350.⁵¹ In the simulations, and in order to reduce computation time, we approximate such a distribution with 15 representative draws, which are weighted according to their densities.

B. Baseline Results

We present results from the counterfactual model for the period between January 2010 and December 2011. For our baseline results, we simulate the Cournot equilibrium for the case in which wind farms are overbidding, which is the closest to the actual observed behavior.

Figure 6 presents the day-ahead price distribution and day-ahead premium against the actual data over different hours. One can see that the model does a fairly good job at capturing the main patterns in the data. The price distribution is comparable to that observed in our data, in spite of missing some price spikes.⁵² The model also predicts a distribution of price premium that resembles the one in the data. The distribution shows that, in the presence of market power and limited arbitrage, a positive price premium can arise in equilibrium. Table 5 below also shows that the baseline case (wind arbitrage with $b_2 < b_1$) fits well the observed market quantities in the data.

We can also use our simulations to compare predicted arbitrage under alternative assumptions. Our model computes arbitrage outcomes under several alternative hypothesis: 20 percent of wind production, strategic arbitrage, and full arbitrage. Because a price premium is present in the market, we know that wind arbitrage is not full. Yet, how far is it from full arbitrage? How does it compare to the amount of arbitrage that a strategic arbitrageur would do?

Figure 7 presents the distribution of arbitrage amounts under the different counterfactuals considered for the more realistic case in which $b_2 < b_1$. The figure shows that the actual amount of arbitrage in this market is larger than what a single strategic arbitrageur would do, consistent with firms competing, to some extent, for these arbitrage opportunities. However, the arbitrage amount is less than what would be needed for prices to converge. There are several potential explanations. First, there could be some costs to arbitrage, such as having a person in charge of preparing optimal strategies. Second, such large amounts of arbitrage may be discouraged by

⁵¹ The estimated distribution has little skewness and a kurtosis near 3. See online Appendix C for details.

⁵² Our simplified model does not have startup costs. It is well known that abstracting from startup costs will limit the ability to generate price spikes (Bushnell, Mansur, and Saravia 2008; Reguant 2014). Consistent with the literature, our model overpredicts production at hours of high demand, and thus underpredicts prices.

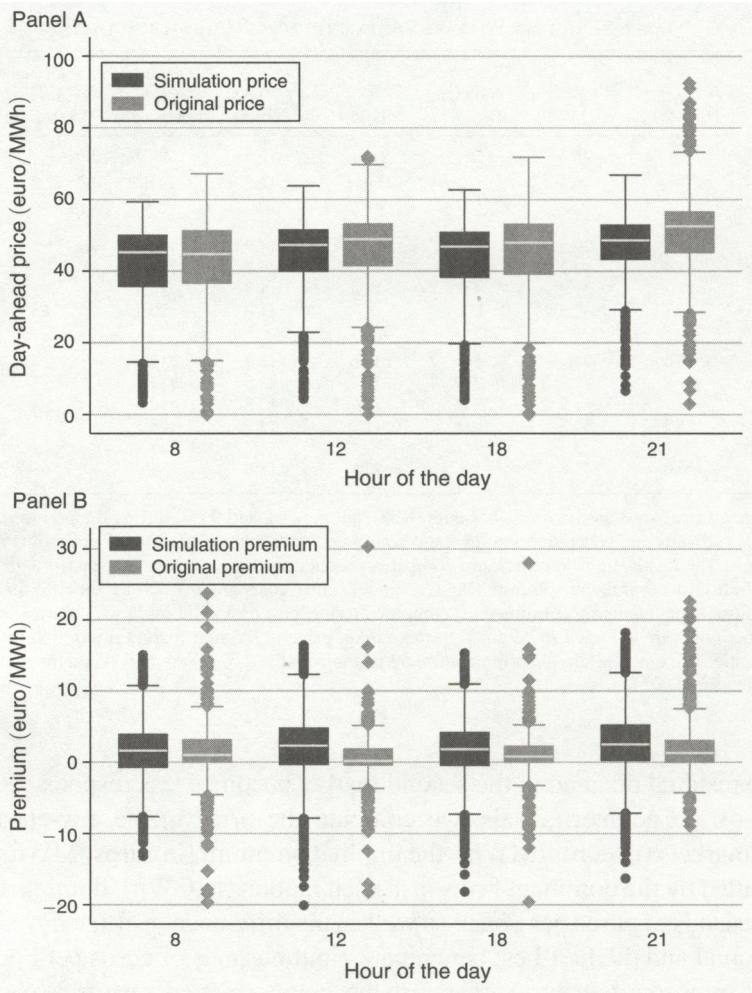


FIGURE 6. BASELINE SIMULATION RESULTS FOR $b_2 < b_1$

Notes: This figure shows the distribution of prices and premia in the data (light gray) as compared to the predictions of our structural model (dark gray). The distribution of market outcomes is captured relatively well by our stylized model across different hours of the day.

the regulator. Finally, whereas fringe firms engage in arbitrage, only a few sophisticated ones exploit the most profitable arbitrage strategies. As shown in the theoretical model, if only few firms participate in the market, they may have little incentives to fully close the price gap. In this sense, whereas arbitrage is not monopolistic, it becomes is limited for strategic reasons.

C. The Role of Sequential Markets and Arbitrage

We compare the performance in terms of welfare under four different arbitrage regimes (baseline, full arbitrage, no arbitrage, and strategic arbitrage), together with a spot-only counterfactual (no sequential markets) and a first-best counterfactual (least cost). For each of the counterfactuals, we consider two cases: one in which the residual demand has the same slope in the second market ($b_2 = b_1$), and one in

TABLE 5—HOURLY WELFARE COMPARISON ACROSS COUNTERFACTUALS

	p_1 (E/MWh)	p_2 (E/MWh)	Premium (E/MWh)	Q_1 (GWh)	$Q_1 + Q_2$ (GWh)	Dominant profit (000s E/h)	Deadweight loss (000s E/h)	Δ Cons. cost from FB (000s E/h)
First-best (b_1)	—	38.2	—	—	15.3	60.5	—	—
Spot only (b_1)	—	46.5	—	—	12.8	123.2	17.2	265.5
Case $b_2 = b_1$								
No arbitrage	45.1	39.5	5.6	13.2	14.9	122.0	1.3	221.8
Str. arbitrage	44.6	40.2	4.4	12.0	14.7	119.0	1.7	204.3
Wind 20%	44.7	39.9	4.9	12.4	14.8	116.4	1.5	210.3
Full arbitrage	42.5	42.5	0.0	7.7	14.0	100.7	4.8	138.0
Case $b_2 < b_1$								
No arbitrage	44.0	38.7	5.3	13.6	13.9	112.3	5.8	186.1
Str. arbitrage	43.8	40.3	3.5	13.1	13.8	111.4	6.2	180.8
Wind 20%	43.7	41.5	2.2	12.7	13.8	110.0	6.4	178.4
Full arbitrage	43.5	43.5	0.0	12.2	13.7	108.3	7.1	170.3
Original data	46.0	44.8	1.3	12.1	13.8	—	—	—

Notes: Welfare comparisons use a sample of hours (8 AM, noon, 6 PM, and 9 PM) during the period January 2010 to December 2011. Profits and costs represent average hourly costs. Profits are the sum of net profits across the four dominant firms. The deadweight loss measure compares production costs between a scenario with market power and the first best cost-minimizing solution. Changes in consumer costs are also with respect to the first best, and represent additional payments by consumers as compared to the least cost solution in which price equals marginal cost. Consumer costs are assumed to be equal to the market price in the first market multiplied by total demand, which is inelastic. The outcomes in the original data are taken from the day-ahead market and the first intraday market in the same sample of hours.

which the residual demand in the second market becomes less responsive ($b_2 < b_1$). For each of the counterfactuals, we compute the price in the day-ahead and the real-time markets (in euros/MWh), the implied premium (in euros/MWh), the quantity scheduled by the dominant firms in at each market (in GWh), dominant firm profits (in thousands of euros per hour), as well as the difference in efficiency between the counterfactual and the first-best benchmark (in thousands of euros per hour).

Table 5 presents hourly averages of the counterfactual results.⁵³ Several findings come out from the counterfactual simulations. First, it is important to note that sequential markets, independent from the form of arbitrage, perform better than a single market. Prices tend to be lower in the presence of the two markets, and the deadweight loss is also lower. This difference in performance is isolating the role of sequential markets in reducing firms’ market power, as pointed out by Allaz and Vila (1993).⁵⁴ One can see that market power and its associated deadweight loss go down very substantially, by more than 60 percent. This is due to the fact that firms have an incentive to compete more aggressively in the second market, leading to real-time prices that are closer to the first best. In the spot only case, deadweight loss from market power is about 17,200 euros per hour, but it goes down to 7,100 or less in the presence of sequential markets. Whereas these hourly welfare measures might appear to be small, such inefficiencies can add up to annual amounts of roughly 140 million and 60 million euros, respectively, which represents a deadweight loss

⁵³ We also show the results for each hour separately in online Appendix tables D.3 through D.7.

⁵⁴ In our model, we allow the single market to clear under best conditions, after uncertainty has been revealed and with slope equal to b_1 . Therefore, the differences in efficiency that we document are only driven by attenuation of market power.

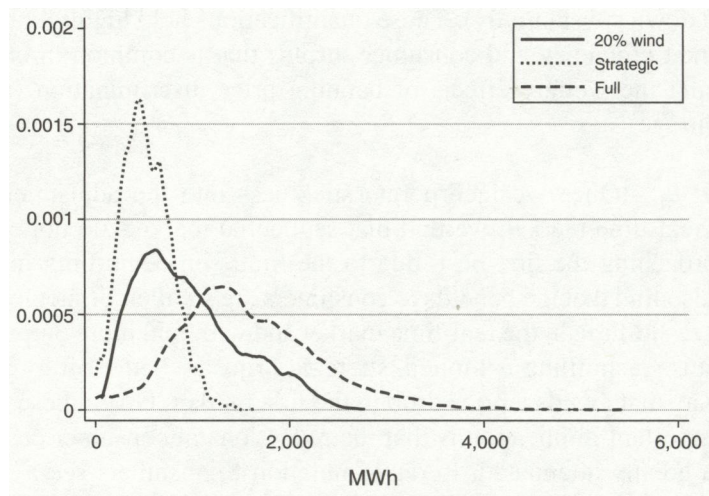


FIGURE 7. STRATEGIC VERSUS FULL ARBITRAGE FOR $b_2 > b_1$

Notes: This figure shows the distribution of arbitrage under alternative assumptions regarding arbitrage behavior. One can see that the amount of arbitrage set as 20 percent of wind production appears to be between the amount of arbitrage a single arbitrageur would do, and that under perfect competition.

of 1 to 2 percent of the economic volume of the market. The last column of the table shows that the reduction in market power also translates into significant reductions of consumer payments.

Whereas sequential markets improve the final allocation, firms are still able to substantially extract rents, as the discriminating monopolist in the model. Strategic firms are better off in the case of no arbitrage, as compared to the one with full arbitrage, as seen in Table 5. Interestingly, in the absence of arbitrage, firms make almost as much profit as in the case with a single market, highlighting their ability to extract rents. Conditional on having sequential markets, what is the role for arbitrage?

Case $b_2 = b_1$.—Full arbitrage closes the gap between p_1 and p_2 , and results in the lowest p_1 among the four regimes. Even though the hourly price reduction might seem small in levels, it represents savings of 5 to 8 percent for consumers. As seen in Table 5, such price differences represent substantial hourly consumer savings, which can add up to thousands of euros per hour. At the same time, however, arbitrage increases total production costs because the quantity produced by the strategic firms inefficiently goes down, due to increased withholding. We find that full arbitrage more than doubles market deadweight loss when compared to no arbitrage. In a simple model with perfect competition, arbitrage *improves* market efficiency. However, our result shows that such implication can change once we take into account the existence of market power in sequential markets.

These findings relate to the role of price discrimination on welfare, a topic that is relevant beyond sequential markets. It is a well-known result that price discrimination may hurt consumers (as consumer rents are extracted more effectively), but that it might increase overall welfare. In our setting, we find that full arbitrage (no discrimination) reduces efficiency, but to the benefit of consumers (consumer

payments go down substantially). These quantifications help highlight the trade-off between market efficiency and consumer surplus that is common in other settings when assessing the welfare effects of banning price discrimination (in our case, allowing arbitrage).

Case $b_2 < b_1$.—Once we incorporate stickiness into the adjustments that can occur in the real-time market, we find that sequential markets do not contribute as much at approaching the first best, due to the limits on reshuffling in the second market. We also find that the benefits to consumers are smaller. Strategic firms anticipate limited reshuffling in the real-time market and withhold more output in the first market. Because reshuffling is limited, strategic firms are better off by withholding output in a way that avoids a price drop in the first market. Given these anticipation effects, an important implication is that the reductions in consumer costs from full arbitrage are greatly attenuated. In our simulations, consumers see a reduction in prices of less than 50c/MWh even under full arbitrage, which translates into more modest consumer savings.

When assessing the role of arbitrage on efficiency, we find similar results than in the previous case. The case with full arbitrage is the least efficient, due to its interaction with market power. We find that deadweight loss is on average 7,100 euros per hour under full arbitrage, while it is 5,800 euros under no arbitrage. These average results mask important heterogeneity across hours. In electricity markets, the degree of market power varies by hours and days, as firms can exercise more market power in hours of high demand. To examine the relationship between the degree of market power and the magnitude of the welfare change, we compare the inefficiencies from arbitrage (deadweight loss under full arbitrage versus deadweight loss under no arbitrage) across hours in Figure 8. Increased deadweight loss from arbitrage tends to be largest at 9 PM, which is the hour of peak demand in the Iberian market. In addition, the skewness of the distribution highlights that the losses in efficiency due to arbitrage are particularly large in hours of extremely high demand, in which firms can exercise substantial market power. No arbitrage is particularly beneficial for these hours because it allows firms to extract rents without sacrificing the efficiency of the final allocation. In relative terms, arbitrage increases market deadweight loss by about 20 percent, compared to the case with no arbitrage.

Arbitrage and Market Liquidity.—Our results are also useful to think about the role of having a responsive secondary market (large b_2). Comparing the case in which $b_2 = b_1$ versus the case in which $b_2 < b_1$, one can see that sequential markets are most efficient at reducing market power when both markets are well participated. First, all arbitrage counterfactuals showcase less deadweight loss than in the case in which the second market is less responsive. One can also see that the amount of deadweight loss is extremely low for the case in which the secondary market is responsive and there is no arbitrage, being only 1,300 euros per hour on average. Additionally, the price effects are substantial, especially when full arbitrage is implemented. This suggests that arbitrage, as a measure to reduce consumer costs, will be most effective when fringe firms actively participate in both the day-ahead and the real-time markets. On the contrary, when the second market is not responsive, the scope for reducing consumer costs through arbitrage is much more limited.

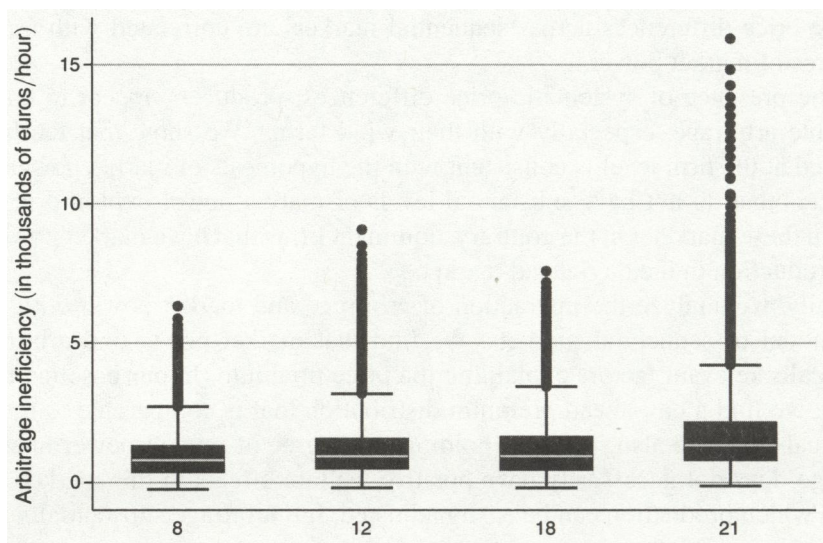


FIGURE 8. INCREASED DEADWEIGHT LOSS FROM ARBITRAGE (*Full versus No Arbitrage*)

Notes: This figure shows the difference in hourly costs between full arbitrage and no arbitrage. Market outcomes are less efficient under full arbitrage, especially during hours of high demand (hour 21).

Discussion.—Finally, it is important to discuss our assumptions on the costs and benefits of arbitrage in itself, which we assume to be zero. On the benefits side, we assume that arbitrage is not productive in itself. One could argue that arbitrage provides additional benefits, e.g., if arbitrageurs have better information, which could improve the counterfactual outcomes under full arbitrage. Arbitrage could also encourage market participation and potentially make secondary markets more liquid. On the cost side, we assume that arbitrage comes at no additional cost, i.e., arbitrage is frictionless and entails no transaction costs. To the extent that arbitrage entails some real costs, the counterfactual welfare outcomes of full arbitrage would be worsened.⁵⁵ Whereas modeling these elements is not the focus of our paper, it is important to keep in mind that they might increase or reduce the attractiveness of arbitrage in practice.

V. Conclusions

We study price differences in sequential markets. In the context of electricity markets, we find that a declining price path can arise in equilibrium under imperfect competition and limited arbitrage, even in the absence of other potential explanations

⁵⁵For example, whether wind farms or financial agents perform the arbitrage can have real consequences, potentially increasing efficiency (Jha and Wolak 2014). In our setting, arbitrage by wind farms could generate dynamic inefficiencies if the system operator is not planning for the right amount of wind production when scheduling reserves. Whereas system operators can circumvent some of these issues by using centralized forecasts, arbitrage can interact with reliability planning more generally. Parsons et al. (2015) discuss other potential ways in which virtual bidding might be counterproductive when interacting with reliability and ramping constraints.

playing a role, such as information updating or risk aversion. Empirically, we show that the price differences across sequential markets are correlated with traditional measures of market power.

In the presence of systematic price differences, producers appear to engage in profitable arbitrage, especially with their wind farms. We show that the behavior observed at the firm level is consistent with the hypothesis of market power. Fringe producers that do not have substantial levels of market power exploit price differences in these market. On the contrary, dominant firms that have market power withhold production in the day-ahead market.

Finally, we analyze the interaction of arbitrage and market power with a structural model of sequential markets. We find that market power and arbitrage are empirically relevant factors explaining the price premium. In our baseline counterfactual, we find a day-ahead premium distribution that is comparable to the one in the actual data. We also find that, holding the degree of market power unchanged, arbitrage does not necessarily have positive welfare effects in this market. For the case in which production can be easily adjusted, full arbitrage substantially reduces day-ahead prices, but at the expense of reduced productive efficiency.

REFERENCES

- Allaz, Blaise, and Jean-Luc Vila. 1993. "Cournot Competition, Forward Markets and Efficiency." *Journal of Economic Theory* 59 (1): 1–16.
- Ashenfelter, Orley. 1989. "How Auctions Work for Wine and Art." *Journal of Economic Perspectives* 3 (3): 23–36.
- Bernhardt, Dan, and David Scoones. 1994. "A Note on Sequential Auctions." *American Economic Review* 84 (3): 653–57.
- Birge, John, Ali Hortaçsu, Ignacia Mercadal, and Michael Pavlin. 2014. "The Role of Financial Players in Electricity Markets: An Empirical Analysis of MISO." <http://home.uchicago.edu/hortacsu/recency.htm> (accessed April 29, 2016).
- Borenstein, Severin, James Bushnell, Christopher R. Knittel, and Catherine Wolfram. 2008. "Inefficiencies and Market Power in Financial Arbitrage: A Study of California's Electricity Markets." *Journal of Industrial Economics* 56 (2): 347–78.
- Borenstein, Severin, James B. Bushnell, and Frank A. Wolak. 2002. "Measuring Market Inefficiencies in California's Restructured Wholesale Electricity Market." *American Economic Review* 92 (5): 1376–1405.
- Bowden, Nicholas, Su Hu, and James Payne. 2009. "Day-Ahead Premiums on the Midwest ISO." *Electricity Journal* 22 (2): 64–73.
- Bushnell, James B., Erin T. Mansur, and Celeste Saravia. 2008. "Vertical Arrangements, Market Structure, and Competition: An Analysis of Restructured US Electricity Markets." *American Economic Review* 98 (1): 237–66.
- Coase, R. H. 1972. "Durability and Monopoly." *Journal of Law & Economics* 15 (1): 143–49.
- Coutinho, Paulo Braulio. 2013. "When-Issued Markets and Treasury Auctions." PhD diss., UCLA.
- Cullen, Joseph. 2015. "Dynamic Response to Environmental Regulation in Electricity Markets." www.josephcullen.com/resources/dynamic-response.pdf (accessed April 29, 2016).
- Fabra, Natalia, and Mar Reguant. 2014. "Pass-Through of Emissions Costs in Electricity Markets." *American Economic Review* 104 (9): 2872–99.
- Hadsell, Lester. 2008. "Day-Ahead Premiums on the New England ISO." *Electricity Journal* 21 (4): 51–57.
- Hortaçsu, Ali, and Steven L. Puller. 2008. "Understanding Strategic Bidding in Multi-Unit Auctions: A Case Study of the Texas Electricity Spot Market." *RAND Journal of Economics* 39 (1): 86–114.
- Hortaçsu, Ali, Seyed Ali Madanizadeh, and Steven L. Puller. 2015. "Power to Choose? An Analysis of Consumer Inertia in the Residential Electricity Market." National Bureau of Economic Research Working Paper 20988.
- Ito, Koichiro, and Mar Reguant. 2016. "Sequential Markets, Market Power, and Arbitrage: Dataset." *American Economic Review*. <http://dx.doi.org/10.1257/aer.20141529>.

- Jha, Akshaya, and Frank A. Wolak.** 2014. "Testing for Market Efficiency with Transactions Costs: An Application to Convergence Bidding in Wholesale Electricity Markets." http://web.stanford.edu/group/fwolak/cgi-bin/sites/default/files/files/CAISO_VB_draft_V8.pdf (accessed April 29, 2016).
- Kim, Dae-Wook, and Christopher R. Knittel.** 2006. "Biases in Static Oligopoly Models? Evidence from the California Electricity Market." *Journal of Industrial Economics* 54 (4): 451–70.
- Lazear, Edward P.** 1986. "Retail Pricing and Clearance Sales." *American Economic Review* 76 (1): 14–32.
- Longstaff, Francis A., and Ashley W. Wang.** 2004. "Electricity Forward Prices: A High-Frequency Empirical Analysis." *Journal of Finance* 59 (4): 1877–1900.
- McAfee, R. Preston, and Daniel Vincent.** 1993. "The Declining Price Anomaly." *Journal of Economic Theory* 60 (1): 191–212.
- McRae, Shaun D., and Frank A. Wolak.** 2014. "How Do Firms Exercise Unilateral Market Power? Evidence from a Bid-Based Wholesale Electricity Market." In *The Manufacturing of Markets: Legal, Political and Economic Dynamics*, edited by Eric Brousseau and Jean-Michel Glachant, 390–420. Cambridge, UK: Cambridge University Press.
- Nocke, Volker, and Martin Peitz.** 2007. "A Theory of Clearance Sales." *Economic Journal* 117 (522): 964–90.
- Parsons, John E., Cathleen Colbert, Jeremy Larrieu, Taylor Martin, and Erin Mastrangelo.** 2015. "Financial Arbitrage and Efficient Dispatch in Wholesale Electricity Markets." MIT Center for Energy and Environmental Policy Research Working Paper 2015-002.
- Puller, Steven L.** 2007. "Pricing and Firm Conduct in California's Deregulated Electricity Market." *Review of Economics and Statistics* 89 (1): 75–87.
- Reguant, Mar.** 2014. "Complementary Bidding Mechanisms and Startup Costs in Electricity Markets." *Review of Economic Studies* 81 (4): 1708–42.
- Salant, David J.** 2014. "Sequential Auctions and Auction Design." <https://sites.google.com/site/dsalant/recent-papers/sequential-auction-aug23.pdf?attredirects=0&d=1> (accessed April 29, 2016).
- Saravia, Celeste.** 2003. "Speculative Trading and Market Performance: The Effect of Arbitrageurs on Efficiency and Market Power in the New York Electricity Market." University of California Berkeley CSEM Working Paper 121.
- Spanish Wind Energy Association.** 2013. *Wind Power '13: Spanish Wind Energy Association*. http://www.aeeolica.org/uploads/Wind_Power_2013_AEE.pdf (accessed April 29, 2016).
- Weber, Robert J.** 1981. "Multiple-Object Auctions." In *Auctions, Bidding, and Contracting: Uses and Theory*, edited by Richard Englebrecht-Wiggans, Martin Shubik, and Robert M. Stark, 165–91. New York: New York University Press.
- Wolfram, Catherine D.** 1998. "Strategic Bidding in a Multiunit Auction: An Empirical Analysis of Bids to Supply Electricity in England and Wales." *RAND Journal of Economics* 29 (4): 703–25.
- Wolfram, Catherine D.** 1999. "Measuring Duopoly Power in the British Electricity Spot Market." *American Economic Review* 89 (4): 805–26.
- Zitzewitz, Eric.** 2003. "Who Cares about Shareholders? Arbitrage-Proofing Mutual Funds." *Journal of Law, Economics, and Organization* 19 (2): 245–80.