Can Financial Participants Improve Price Discovery and Efficiency in Multi-Settlement Markets with Trading Costs?

Akshaya Jha and Frank A. Wolak*

May 10, 2019

Abstract

The introduction of purely financial participants into commodity markets is thought to yield forward prices that better reflect future spot prices, and ultimately, more efficient future production and consumption decisions. However, there are sizable transaction costs associated with trading in most commodity markets. This paper develops a statistical test of the null hypothesis that expected forward/spot price spreads cannot be arbitraged even after accounting for these transactions costs. We apply this test to hourly, location-specific day-ahead and real-time prices from California's wholesale electricity market. The implied trading cost required to reject the null hypothesis of no profitable arbitrage opportunities falls significantly after California allowed purely financial participation. Moreover, variable input costs per MWh of electricity produced fell by 3.6% in high demand hours after the introduction of purely financial participants. Combined, our evidence supports the hypothesis that the introduction of purely financial participants into the California wholesale electricity market decreased the average difference and the volatility of the difference between day-ahead and real-time prices, which ultimately lowered the total variable cost of serving demand.

JEL Codes: G1, L94, Q48

Keywords: Wholesale Electricity Markets, Financial Trading, Explicit Virtual Bidding, Arbitrage, Transaction Costs, Market Efficiency

1   Introduction

In multi-settlement markets, participants buy or sell delivery of a commodity at a pre-specified future date in a forward market with the opportunity to close out that position in a subsequent forward market or in the spot market. Each time the market for this commodity and delivery date combination clears, settlement occurs, meaning that purchases and sales of this product are made at a price that reflects the information available to market participants at that time. Many commodities have multiple forward market settlements before delivery ultimately takes place. For example, an airline might purchase jet fuel for delivery on a pre-specified date in the future. It could close out this position by selling it in any subsequent forward market before the delivery date or by taking delivery of the jet fuel and selling it in the spot market.

Forward markets provide both producers and consumers of a product with future price certainty. This increased future spot price certainty can improve the efficiency of the spot market for a commodity. Specifically, forward prices that are accurate estimates of future spot prices provide valuable information to investors in productive capacity to supply this commodity or investors in productive capacity that consumes this commodity. For example, a potential investor in an oil refinery can use forward prices for gasoline and oil to obtain a more accurate estimate of whether an investment in a facility that comes on line in two years and produces for 20 years would be profitable.

Liquid forward markets for a commodity thus have the potential to improve both future spot price transparency and overall market efficiency. However, many commodity markets have a small number of physical participants. Consequently, purely financial market participants that neither produce nor consume the commodity, but simply take financial positions in forward markets, can increase market liquidity. This logic implies that allowing purely financial participants to trade in multi-settlement commodity markets has the potential to improve both future spot price transparency and real-time market performance.

This paper uses the introduction of purely financial participants into California's wholesale electricity market to examine the validity of this logic. Electricity markets are an ideal setting for this investigation. First, wholesale electricity markets are multi-settlement: exactly the same product—electrical energy delivered to a specific location at

a specific hour of the day—is bought and sold using the same market-clearing mechanism in both the day-ahead and real-time markets. Second, participants are required to post significant collateral with the market operator and pay non-trivial transaction costs in order to trade in these markets. This can limit the number of market participants and baseline level of liquidity in the forward market. Moreover, it is extremely costly to store electrical energy and the time lag between buying/selling electricity in the day-ahead market and subsequent sale/purchase in the real-time market is less than one day. Relative to other commodities, this greatly limits the extent to which day-ahead/real-time electricity price differences are due to either a hedging motive or inter-temporal discounting.

Finally, the market efficiency benefits from a more liquid forward market for grain, gold, oil, jet fuel or other commodities are likely to be realized in a multi-year time frame. In contrast, introducing purely financial participants in multi-settlement (day-ahead/real-time) electricity markets has the potential to quickly reduce the cost of serving real-time demand. This is because market/system operators are required to take the offers submitted by generation unit owners and dispatch of hundreds of electricity generation units to meet real-time demand at thousands of locations in a transmission network. In doing this, system operators are faced with a multitude of different operating constraints such as transmission network capacity constraints as well as generation unit start-up, ramping, and minimum runtime constraints. Computing the market-clearing configuration of generation unit output levels given the offers submitted by market participants requires solving an extremely complex mixed-integer programming problem. Allowing purely financial participants to supply or demand virtual energy in the day-ahead market could help system operators find better solutions to this optimization problem and thereby reduce the total cost of serving demand, particularly during high demand hours when transmission, generation unit-level, and other system operating constraints are likely to be most relevant.

This paper has two parts. First, we develop a statistical test of arbitrage between forward and spot markets and apply this test to data from California's wholesale electricity market. The canonical principle of no arbitrage implies that the forward price of a commodity at time t for delivery at time $t + k$ ($P^f_{t,t+k}$) is equal to the expected value of the spot price for the same commodity at time $t + k$ (P^s_{t+k}) conditional on the infor-

mation available to market participants at time t (i.e.: $P^f_{t,t+k} = E_t[P^s_{t+k}]$). Otherwise, risk-neutral arbitrageurs can earn expected profits. Deviations from the no arbitrage condition that $P^f_{t,t+k} = E_t[P^s_{t+k}]$ are typically attributed either to risk preferences or market inefficiencies (Pindyck (2001); Kilian and Lee (2014); Kilian and Murphy (2014); Knittel and Pindyck (2016)). Less explored are the substantial costs associated with trading in many commodity markets (Williams (1987); Hirshleifer (1988); Frazzini, Israel and Moskowitz (2018); Dávila and Parlatore (2019)). Our statistical framework accounts for the non-trivial costs of trading in wholesale electricity markets and a number of other commodity markets.

Our statistical framework is based on a financial trader with access to 24 assets corresponding to the day-ahead/real-time price spreads for each hour of the day.[1] This financial participant maximizes expected profits by buying or selling the asset corresponding to the hour of the day with the largest average price spread in absolute value. However, the maximum operator is not differentiable, so standard asymptotic methods such as the Delta Method do not apply. We instead use the directional derivative method formulated by Fang and Santos (2018) in order to test two separate null hypotheses for a given value of per-unit trading costs c: (1) that the maximum absolute price spread is greater than c (i.e.: the null hypothesis that profitable arbitrage opportunities exist), and (2) that the maximum absolute price spread is less than c (i.e.: the null hypothesis that no profitable arbitrage opportunities exist).

These statistical tests are applied to hourly, location-specific data on day-ahead and real-time prices from California's wholesale electricity market from 4/1/2009-12/31/2012. We find that the trading costs required to reject the null hypothesis that arbitrage opportunities exist fell substantially after California introduced purely financial participation in its market on February 1, 2011.[2] Moreover, prior to the introduction of purely financial participation, electricity suppliers could arbitrage day-ahead/real-time price spreads only at locations where they owned generation units by adjusting the day-ahead and real-time offer curves of their generation units. There was no way for market participants to

---

[1]This formulation is consistent with market rules; trading in wholesale electricity markets occurs daily rather than hourly, with participants simultaneously having the opportunity to take financial positions corresponding to the day-ahead/real-time price spreads at each location in the transmission network in each of the 24 hours of the following day.

[2]Consistent with our empirical results, a simulation study by Li, Svoboda and Oren (2015) finds that the revenues generated by implementing their optimal trading strategy decrease significantly after the introduction of purely financial participants in California.

arbitrage day-ahead/real-time price spreads at individual demand withdrawal locations in the transmission network. Consistent with this intuition, we find that the reduction in implied trading costs due to the introduction of purely financial participants is larger for demand withdrawal locations relative to locations with generation units. Moreover, the average values of the implied trading costs are not statistically different for generation and load-withdrawal nodes after the introduction of purely financial participation.

The second component of our paper is to quantify the physical efficiency gains from the introduction of purely financial participants. As noted earlier, the market efficiency gains from introducing purely financial participants are likely to be largest during high demand hours when transmission and other operating constraints are increasingly relevant. We therefore employ a difference-in-differences estimator that compares market performance measures such as the hourly variable cost of thermal generation and the hourly input fossil-fuel burned by thermal generation units in high demand versus low demand hours before versus after the introduction purely financial participants.

Our results indicate that, in hours with a system demand greater than the 90th percentile of the distribution of demand, the introduction of financial trading resulted in a 3.6% (4.2%) decrease in fuel costs (thermal input energy consumed) per MWh of electricity produced from fossil-fuel fired sources. The implied annual fuel cost savings and environmental benefits from financial trading in high demand hours are roughly 4.2 million dollars and 145,000 tons of CO_2, respectively. Finally, our results indicate that there is no discernible increase after the introduction of purely financial participants in either the number of high demand hours that at least one fossil-fuel unit was required to start up or the ancillary services costs associated with maintaining real-time supply and demand balance. These results are inconsistent with the argument that allowing purely financial participants places undue stress on the physical constraints inherent to electricity production and distribution.

Combined, our results suggest that both the physical efficiency gains from allowing purely financial trading as well as the costs of financial trading vary across both time and space. Market operators can thus potentially increase the market efficiency benefits from allowing purely financial participation by setting different transaction fees for different locations and hours. In particular, our estimates suggest that a 5 cents per MWh reduction in implied trading costs in high demand hours could result in fuel cost savings

4

ranging from 1-2 cents per MWh.

Our results contribute to the ongoing policy debate surrounding the controversial role played by financial traders in commodity markets. Specifically, many argue that financial traders earn revenues primarily at the expense of suppliers and consumers of the commodity.[3] Particularly for wholesale electricity, some have argued that financial participants submit bids intended to profit from the physical constraints inherent to electricity provision, either by taking advantage of rules pertaining to starting up or ramping power plants (Parsons et al., 2015) or by inducing transmission congestion (Birge et al., 2013). On the other hand, previous work documents a substantial day-ahead/real-time price premium in electricity markets without financial participation, due either to market power exercised by suppliers (Ito and Reguant, 2016) or risk preferences (Routledge, Spatt and Seppi (2001); Bessembinder and Lemmon (2002); Longstaff and Wang (2004); Bessembinder and Lemmon (2006)). The introduction of financial participants in wholesale electricity markets has also been linked to decreases in the exercise of unilateral market power (Saravia (2003); Mercadal (2018)), decreases in the volatility of electricity prices (Hadsell (2007)) and grid reliability benefits (Isemonger (2006)). Our results provides empirical evidence from a large wholesale electricity market that financial participants can improve price transparency and market efficiency without putting unduly stress on the physical constraints inherent to electricity production and transmission.

The remainder of the paper proceeds as follows. The next section describes how California and other U.S. wholesale electricity markets operate both before and after the introduction of purely financial participants. Section 3 presents several examples of how purely financial participants can reduce the cost of serving demand. We derive our statistical tests of arbitrage and no arbitrage opportunities that accounts for transaction costs in Section 4. The empirical results from applying this framework to California's wholesale electricity market are presented in Section 5. Section 6 presents our empirical framework and findings pertaining to quantifying the market efficiency benefits of introducing purely financial participants. Finally, we conclude in Section 7 by discussing how our findings can inform the difficult problem of setting transaction fees in order to maximize market efficiency.

---

[3]See "Traders Profit as Power Grid Is Overworked," New York Times, August 14, 2014 for the case of wholesale electricity markets. See "U.S. Suit Sees Manipulation of Oil Trades," New York Times, May 24, 2011 for the case of oil. See "Did Goldman Sachs Rig Commodities Markets?" CNN Business, November 20, 2014 for the case of aluminum.

2 Wholesale Electricity Market Operations With Versus Without Financial Trading

In this section, we first describe how day-ahead and real-time markets operated in California and other U.S. wholesale electricity markets prior to the introduction of purely financial participants through explicit virtual bidding. The next subsection describes explicit virtual bidding (EVB), the mechanism by which purely financial participants and other market participants can arbitrage day-ahead/real-time price differences. This subsection also discusses how virtual bidding impacts day-ahead and real-time system operation and the different transaction costs faced by purely financial participants. The final subsection describes implicit virtual bidding (IVB), the only mechanism available for market participants to arbitrage day-ahead/real-time price differences before the introduction of EVB. This subsection compares IVB to EVB in order to clarify the mechanisms by which purely financial market participants can reduce the mean and variance of day-ahead/real-time price spreads and the variance of real-time prices relative to IVB as well as why the impact of introducing purely financial participants should be different for generation unit locations versus demand withdrawal locations.

2.1 Locational Marginal Pricing in Multi-Settlement Markets

In most markets, products are shipped directly from seller to buyer. Wholesale electricity markets work differently. In these markets, generators inject electricity into the transmission network and this electricity flows according to Kirchhoff's laws.[4] Thus, commitments between buyers and sellers of electricity constitute a financial rather than physical arrangement: the buyer does not withdraw the actual energy injected into the transmission grid by the seller. The only things that can be measured are the amount of energy injected by a generation unit and the amount of energy withdrawn by a load-serving entity.[5]

---

[4]Schweppe et al. (2013) provides an accessible discussion of Kirchhoff's laws governing current flows through an electrical circuit.

[5]As discussed in Schweppe et al. (2013), Where the energy injected by a generation unit goes in the transmission network and where the electricity withdrawn comes from depends on the level of demand at all locations in the transmission network, the output levels of all of the generation units, the configuration of the transmission network, as well as a number of other technical operating conditions.

All electricity supply industries have transmission networks with finite transfer capacity between locations in the grid. This can often prevent the system operator from accepting energy from low-cost generation units located distant from a load withdrawal point and instead accepting energy from higher cost generation units that can meet this demand. Specifically, transmission congestion limits the amount of low-cost energy that can injected at a location on the transmission network to be transported and withdrawn elsewhere on the grid. It has proven extremely difficult for system operators in the United States (US) to predict which specific transmission links will be congested. As a result, all wholesale markets in the US have adopted dispatch and pricing mechanisms that set potentially different prices at all points of injection and withdrawal on the transmission network.

A dispatch and pricing mechanism called *nodal pricing* or *locational marginal pricing* (LMP) sets spatial prices that reflect all relevant transmission network constraints, transmission losses, generation unit start-up and ramping constraints, and other relevant operating constraints on the transmission network. All formal wholesale markets operating in the United States use variants of the LMP algorithm described by Bohn, Caramanis and Schweppe (1984). To compute locational marginal prices in the day-ahead market, suppliers submit 24 hourly generation unit-level offer curves by 10AM of day t and electricity retailers and other load-serving entities submit locational demand curves, typically at a significantly higher level of spatial aggregation than the nodal level.

These offer curves have two parts: a start-up cost offer and an energy supply curve. The start-up cost offer in the day-ahead market is a fixed dollar payment that must be paid to the generation unit owner if the unit is not generating electricity at the start of day $t + 1$ but is accepted to produce a positive output at some point during that day. The energy offer curve for hour h indicates how much electricity the supplier is willing to provide from the unit in hour h of day $t+1$ as a function of the market-clearing day-ahead price at the unit's location. This energy offer curve is a non-decreasing step function, where each offer price-quantity step determines the minimum price the generation unit owner must be paid in order to produce the quantity associated with that step.[6] The sum of the quantity increments for each energy offer curve is restricted to be less than the capacity of the generation unit.

---

[6]In California, generation unit-level offer curves are permitted to have up to ten price-quantity pairs.

Load-serving entities (i.e.: demanders of electricity) similarly submit location-specific willingness-to-purchase functions in the day-ahead market that are non-increasing in the price at that location. This willingness-to-purchase function is composed of price-quantity pairs ordered from highest to lowest price. A load-serving entity (LSE) is willing to increase the amount of electricity it purchases by a given offer quantity increment provided the market-clearing price is at or below the corresponding offer price increment. For the case of California, these willingness-to-purchase functions are submitted by retailers at the utility service territory level.[7] The California Independent System Operator (ISO), the market and system operator for California's electricity supply industry, allocates shares of these price and quantity offers to the hundreds to thousands of load-withdrawal nodes in the utility's service territory using the ISO's estimates of the fraction of the utility's total demand at each node in its service territory.[8]

For the day-ahead market, the California ISO minimizes the total as-offered cost, the offer price times the offer quantity summed over all units accepted, of serving the demand for energy and ancillary services at all locations in the transmission network during all 24 hours of the following day subject to the ISO's best estimate of the configuration of the transmission network the following day, generation unit runtime and ramping constraints, as well as a number of other system operating constraints.[9] The locational marginal price (LMP) at each node in the transmission network is equal to the increase in the minimized value of the objective function from the ISO's as-offered cost minimization problem as a result of increasing the amount of energy withdrawn at that location by 1 MWh. The LMPs and resulting supply and demand schedules at each location in the transmission network for all 24 hours of the following day are computed before the afternoon of the day before the energy is scheduled to be delivered.

All market participants are notified of these LMPs as well as their day-ahead supply and demand obligations at 1PM on the day before the delivery date. These supply and demand obligations are firm financial commitments to sell or buy the quantities of energy that emerge from the day-ahead locational marginal pricing process. For example,

---

[7]Appendix Figure A.1 presents a map of the territories served by each of California's three investor-owned utilities.

[8]Pacific Gas and Electric has more than 1,500 load withdrawals point in its service territory, Southern California Edison approximately 200, and San Diego Gas and Electric approximately 300.

[9]Ancillary services are the collection of operating reserves required by the ISO to maintain a reliable supply of electricity in real-time. Wolak (2019) describes the economic efficiency properties of co-optimizing the procurement of energy and ancillary services.

suppose a supplier sold 50 MWh to be delivered to a given location in the 6PM to 7PM hour of the following day at a price of 40 dollars per MWh. This supplier is guaranteed to be paid \$2,000 (= 50 MWh × \$40/MWh) regardless of the actual production of energy from its generation unit during that hour of the following day. This is the sense in which settlement of the day-ahead market has occurred. Commitments to supply energy and ancillary services are bought and sold at the relevant locational day-ahead price.

Between the close of the day-ahead market and the start of real-time system operation, actual demand at each location on the transmission network is realized. Some generation units must produce more or less than their day-ahead energy schedules in order to meet this actual demand at all of the more than 4,000 load-withdrawal points in California. At least 75 minutes in advance of each hour of real-time system operation, generation unit owners submit offer curves specifying their willingness to increase or decrease their output relative to their day-ahead schedules. Starting with midnight on the delivery date, these offer curves are used to clear the real-time market during each 5-minute interval within the hour to meet actual demand at each location in the transmission network given the real-time configuration of the transmission network and real-time output levels of all generation units.

It is important to emphasize that the configuration of the transmission network in the real-time market and the set of available generation units in real-time can differ significantly from the ISO operator's best estimate of the configuration of the transmission network and the set of available generation units used to determine day-ahead market outcomes. The real-time LMP dispatch and pricing process yields real-time production levels for all generation units that are accepted to supply additional energy or less energy relative to their day-ahead schedules in the real-time market, as well as prices at all nodes in the transmission network.

The 5-minute real-time LMP at each location on the transmission network is equal to the increase in the optimized value of the as-offered cost function as a result of increasing the amount of energy withdrawn at that location by 1 MWh. The hourly real-time price is the average of the twelve 5-minute real-time prices within that hour. Any electricity demander that consumes more than its day-ahead schedule in a given hour pays for this additional consumption at the hourly real-time price. Any electricity demander that consumes less than its day-ahead schedule in a given hour receives the real-time price for

any scheduled energy it does not consume. Similar logic applies to generation units that do not produce the quantity they sold in the day-ahead market. Those that produce more energy receive the real-time price at their location and those that produce less pay the real-time price at their location.

The combination of a day-ahead forward market and a real-time spot market is called a "multi-settlement market" because only real-time deviations from participants' day-ahead schedules are settled at the hourly real-time price. Recall our previous example in which a generation unit sold 50 MWhs of energy in the day-ahead market at a specific location for the 6PM to 7PM hour on day $t+1$ at a price of 40 dollars per MWh. Suppose that this unit actually only produced 30 MWhs of electricity between 6PM and 7PM on day $t+1$. In this case, the owner would have to purchase the remaining 20 MWhs at the hourly time-weighted average real-time price corresponding to the same location between 6PM and 7PM in order to meet its forward market commitment. If the unit instead produced 55 MWhs, then the additional 5 MWhs beyond its 50 MWhs day-ahead schedule is sold at the hourly real-time price for that location/hour-of-the-day.

2.2   Explicit Virtual Bidding

All US wholesale electricity markets currently allow for purely financial participation through explicit virtual bidding. With explicit virtual bidding (EVB), every market participant has access to the following purely financial instrument: buy (sell) one MWh of electricity at a given location and hour-of-the day in the day-ahead market if the day-ahead price is below (above) the offer price, with the *obligation* to sell (buy) back one MWh at the same location and hour-of-the-day in the real-time market as a price-taker (i.e.: accepting the prevailing real-time price for closing out this purely financial position in the day-ahead market). These financial offer curves are termed "virtual bids" or "convergence bids" because an expected profit-maximizing purely financial trader will typically take positions at a location in the day-ahead market that reduce the magnitude of the difference between day-ahead and real-time prices at that location.

Although virtual bids and physical bids are separately identified to the system operator, the California ISO treats physical and virtual bids the same when running the day-ahead LMP pricing and dispatch process. However, the California ISO knows that

any day-ahead "virtual energy" sale or "virtual energy" purchase must be reversed in the real-time market. Recall that in the real-time market only physical generation units that submit offers into the real-time market are dispatched to meet the *actual* demand at each location on the transmission network.[10] If a purely financial player sells 10 MWh of virtual energy at a given location in the day-ahead market, she must purchase this 10 MWh back at the real-time price for that location because she cannot actually supply any energy in real-time. Similarly, if a purely financial player buys 10 MWh of virtual energy in the day-ahead market, she must sell 10 MWh at the real-time price at that location because she cannot consume any energy in real-time.

This logic implies that the actions of virtual bidders directly influence day-ahead and real-time market outcomes, typically by closing the gap between day-ahead and real-time prices. For example, submitting a virtual bid to sell (buy) one MWh in the day-ahead (real-time) market earns positive revenues if and only if the day-ahead price is higher than the real-time price. However, submitting this virtual bid increases supply (demand) in the day-ahead market (real-time market), making it less likely that day-ahead prices will be higher than real-time prices.

There are three broad types of transaction costs associated with financial trading in California's wholesale electricity market: collateral, trading fees and uplift. Purely financial participants must post collateral greater than the total value of virtual bids they submit each day. The total value of virtual bids submitted each day is equal to the sum of the product of the absolute value of megawatt-hours offered times the applicable reference price for a virtual bid at that location.[11] This collateral does not earn any rate of return while it is held by the California ISO. Moreover, there can be a lag of more than two weeks between when a market participant requests that some or all of its collateral be returned and when this money is actually returned. Consequently, a purely financial participant is foregoing non-trivial financial returns on any collateral posted with the California ISO in order to engage in virtual bidding.[12]

Purely financial participants have to pay roughly 0.5 cents for each price and quantity pair associated with each incremental or decremental virtual bid they submit. They must

---

[10]Generation units with energy schedules sold in the day-ahead market can either produce this energy from their own unit or purchase the energy from the real-time market.

[11]See the California ISO document, "Convergence bidding, participating in markets, credit policy implications," for a description of the process used to compute nodal reference prices.

[12]See the California ISO document, "California ISO Credit Management," for more background.

also pay 9 cents per MWh *cleared* in fees associated with "market services". For example, consider a virtual bidder that submits a demand curve with 10 price/quantity steps to the day-ahead market. If 50 MWh of her demand bid clears, she must pay trading fees of $0.09 \times 50 + \$0.005 \times 10 = 4.55$ dollars. Finally, all financial participants are required to pay a monthly transaction fee of 1,000 dollars regardless of the volume of virtual bids they submit or clear.[13]

Purely financial participants are also required to pay "uplift" charges when system-wide virtual demand is larger than system-wide virtual supply (i.e.: net virtual demand is greater than zero). Uplift charges compensate suppliers that are forced to quickly start up or ramp up their units in order to meet this excess virtual demand. System-wide uplift charges are allocated to different market participants based on their individual levels of net virtual demand cleared. Figure 1 shows the annual average uplift charge per MWh of electricity demand for the five major Independent System Operators (ISOs) in the United States for 2009-2013. This figure indicates that average uplift charges range from roughly 40 to 60 cents per MWh. However, these annual average values conceal significant volatility in the daily value of these uplift charges throughout the year that impact the profitability of the daily virtual bidding actions of purely financial players.[14] As noted in Federal Energy Regulatory Commission (2014), uplift payments are typically correlated with demand levels, input fuel costs and the magnitude of the divergences between day-ahead and real-time prices. Moreover, Federal Energy Regulatory Commission (2014) notes that the volatility of daily uplift costs have been increasing over time in California and two other US markets. This logic implies that days with sizeable day-ahead/real-time price spreads are also likely to have large uplift charges, and a significant share of these charges are recovered from purely financial participants. This implies that sizeable day-ahead/real-time price differences can be rendered unprofitable by these uplift cost allocations.

This discussion demonstrates that, although the marginal cost of submitting a virtual bid and having it clear is less than one dollar, the annual average cost per MWh of clearing

---

[13]These transaction fees are listed in Session 7 of the Convergence Bidding tutorial published by California's independent system operator (CAISO (2015b)).

[14]Uplift charges reflect payments made to resources whose commitment and dispatch result in a shortfall between the resources offers and the revenue earned from the market. There are three main reasons for market revenue shortfalls that require uplift payments: (1) some of the operating costs and limits of a resource are not reflected in market prices; (2) unmodeled system constraints; and (3) the dispatch and commitment of inflexible resources or the commitment of resources ineligible to set price.

Figure 1: Annual Uplift Charges for the 5 Major ISOs: 2009-2013
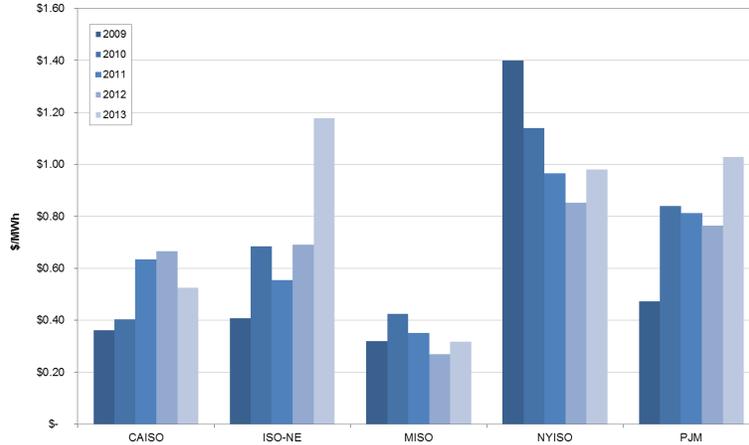
Notes: This figure is taken Federal Energy Regulatory Commission (2014). Annual average uplift charges (in dollars per MWh) are calculated for each independent system operator (ISO) by dividing total annual uplift credits (in dollars) by total annual electricity demand (in MWhs). Total uplift charges and total electricity demand for CAISO for 2009 are based on the nine months of data after 4/1/2009. FERC estimated the total charges and electricity demand for ISO-NE for 2012. Uplift charges for PJM for the years 2012 and 2013 exclude the credits associated with reactive services (these credits amount to approximately 45 million dollars per year).

virtual bids in the California ISO is likely be considerably higher because of the monthly fixed charge to participate in the California ISO markets, the foregone revenues on the collateral posted, and the realized uplift charges during the days that virtual bid MWhs cleared during the year for this market participant. The hourly actual transaction cost can be significantly higher than the annual average because of the volatility in uplift costs.

2.3 Profiting from Price Spreads With Versus Without EVB

Prior to the introduction of financial trading, only suppliers could profit from expected day-ahead/real-time price spreads by selling more or less energy in the day-ahead market than they expected to produce in real-time at the locations that they owned generation units. Using physical bids rather than virtual bids to exploit expected day-ahead/real-time price spreads is termed "implicit virtual bidding". Prior to the introduction of EVB, generation unit owners were only allowed to submit physical day-ahead offers at locations where they inject electricity, their minimum offered quantity must be positive, and their maximum offered quantity must be less than the unit's capacity. Thus, prior to EVB, a supplier that expects day-ahead prices to be far lower than real-time prices could

at best submit zero quantity into the day-ahead market and offer their entire capacity into the real-time market. This supplier also faced no competition for exploiting day-ahead/real-time price spreads at locations where it alone owned generation units. This is different from the case of EVB, where any market participant can submit virtual bids at any allowed node in the transmission network.[15]

Implicit virtual bidding is only feasible at locations with generation units, because load-serving entities are only allowed to submit demand bids at the level of their service territory.[16] For example, California's Independent System Operator (ISO) requires its three large distribution utilities—Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E)—to bid their electricity demand into the day-ahead market at the service-area or load aggregation point (LAP) level. The California ISO then allocates this demand to all nodes in the utility's service territory using load-distribution factors (LDFs).[17] Thus, it is very costly for a utility to implicitly virtual bid one MWh at a single location, because this would effectively require the utility to implicitly virtual bid one MWh at all nodes within its service territory.

3   How Purely Financial Participants Can Improve Market Efficiency

Without EVB, only suppliers were able to participate in day-ahead and real-time electricity markets at the nodal level, and only at the nodes where they inject electricity. Consequently, they face no direct competition from other suppliers that don't inject electricity at these locations. As noted earlier, load-serving entities must submit their demand bids at the LAP level, which typically involves the system operator allocating their service-territory level demand bid to up to thousands of nodes for the three large dis-

---

[15]The California ISO and other US wholesale markets typically restrict the set of nodes at which virtual bidding is allowed. The California ISO eliminates nodes it deems are electrically equivalent to other nodes where virtual bidding is allowed.

[16]The California ISO assigns collections of nodes to load-serving entities and each load-serving entity is charged an hourly price equal to the weighted average of MWhs withdrawn at each of these nodes for their wholesale electricity. This collection of nodes is a called a Load Aggregation Point (LAP) and this price-setting mechanism is called LAP pricing.

[17]For example, if a utility is required to serve 100 MWhs of electricity demand and the ISO computes equal LDFs for the ten nodes in its service area, then the LDFs are equal to 1/10 for each node. The load aggregation point (LAP) level price received by the utility is equal to the LDF-weighted average of the 10 nodal-level prices in its service area.

tribution utilities in California. Thus, the introduction of explicit virtual bidding allows far more opportunities for purely financial participants to contribute their information to the day-ahead and real-time markets at each individual node in the form of virtual bids.[18]

EVB allows both physical market participants and purely financial players to submit virtual bids in the day-ahead market at any location for any hour of the day. The profit-maximizing actions of market participants seeking to exploit day-ahead/real-time price spreads at any location on the grid creates incentives for the day-ahead schedules of generation units to be as close as possible to their real-time output. This outcome means that the California ISO has to accept fewer supply and demand offers from generation units in the real-time market. These actions also create incentives for day-ahead generation schedules to be as close as possible to the real-time output levels that minimize the cost of operating the transmission network in real-time. This section provides several examples of how explicit virtual bidding can lead to market efficiency gains.

As noted earlier, solving for LMPs and generation unit dispatch levels in the day-ahead market is an extremely challenging mixed-integer programming problem, particularly during high demand hours when transmission constraints and generation unit run-time and ramping constraints are likely to be binding. The LMP optimization problem determines which generation units to start up and how much each of these generation units are scheduled to produce. This optimization problem has thousands of choice variables and thousands of constraints and thus has many potential local optima. If virtual bidders are able to figure out lower-cost solutions to meeting demand across the transmission network, they will likely also earn profits from the resulting difference between day-ahead and real-time prices for the reasons described below.

Consider the example of a transmission link between nodes A and B where the virtual bidder believes that 3 MWs more transmission capacity will be made available in real-time than is available in the day-ahead market, so that more energy can flow from A to B in real-time. A virtual bidder can submit a virtual demand bid for 3 MWhs at node A. This demand bid increases the amount of generation scheduled at node A by 3 MWhs. Similarly, she can submit a virtual supply bid at node B for 3 MWhs that reduces the amount of energy supplied at node B by this amount. These two virtual bids create

---

[18]See Appendix Tables A.1 and A.2 for a full list of both the physical and financial participants that are licensed to place virtual bids in California's wholesale electricity market.

3 MWs of virtual transmission capacity between nodes A and B. When the real-time market clears and it turns out that 3 MWs more transmission capacity between A and B becomes available, the additional 3 MWhs of virtual supply provided in real-time at node A will flow to node B, replacing the 3 MWhs of virtual demand provided in real-time at node B. In this case, the virtual bidder is likely to earn a profit on both transactions because she correctly anticipated the existence of the additional transmission capacity in real-time.

It is important to emphasize that the purely financial participant does not have to know that there is a difference between the amount of available transmission capacity between node A and node B in the day-ahead versus real-time markets to find this profitable strategy. She would only need to notice that day-ahead prices at node A are lower than real-time prices and day-ahead prices at node B are higher than real-time prices. This alone would cause her to submit a demand bid in the day-ahead market at node A and submit a supply bid in the day-ahead market at node B.

Another market-efficiency enhancing action by purely financial traders is based on the fact that load-serving entities submit demand bids at the LAP level in the day-ahead market and the California ISO uses load-distribution factors to allocate this LAP-level demand to individual nodes. If a purely financial market participant finds that too much demand is allocated to a specific node in the day-ahead market relative to the quantity of demand that will ultimately show up in real-time, he can submit a supply bid at that node to sell energy in the day-ahead market. Then, when a smaller actual demand shows up in real-time, he gets to buy his virtual supply back at the resulting lower real-time price.

Just as in the previous case, it is not necessary that the purely financial player understand the divergence between the quantity of demand that the retailer serves at this node in real-time and the quantity of demand allocated to the node in the day-ahead market through the LAP-level bidding process. Locations where more demand is allocated in the day-ahead market than is demanded in real-time will have day-ahead prices that are higher than real-time prices. The virtual bidder would thus profit from selling in the day-ahead market and buying back in real-time at these locations.

Our final example concerns the case of whether to start a fossil fuel-fired generation unit in anticipation of producing the next day. Because of the runtime constraints, if

this unit is not committed in the day-ahead market it will not be available to operate in real-time. This energy must instead be produced by units that are more responsive but have higher operating costs. Suppose a purely financial player believes that committing this unit in the day-ahead market provides a lower cost solution to meeting demand in the peak hours of the day than turning on fast-responding but more expensive units. This financial player can submit virtual demand bids at that unit's location in the day-ahead market for the peak hours of the day in order to commit the unit in the day-ahead market, which in turn enables lower cost real-time dispatch. In real-time, the financial player would then sell the virtual energy back at a higher average price for these hours.

Just as in the other two cases, the purely financial player would not need to understand why day-ahead prices are lower than real-time prices at the location of the runtime-constrained fossil fuel generation unit during these hours. Purely financial players would only need to attempt to exploit this profitable arbitrage opportunity by submitting virtual demand bids during peak hours of the day at this location. The implied sale in the real-time market improves the efficiency of the real-time market outcome by dispatching the run-time constrained unit in the day-ahead market.

With the introduction of explicit virtual bidding (EVB), all physical market participants–generation unit owners and load-serving entities–as well as all purely financial participants can submit virtual supply and demand bids at any of the thousands of nodes where virtual bidding is allowed by the California ISO. As the above examples illustrate, the process of exploiting profitable day-ahead/real-time arbitrage opportunities at nodes throughout the transmission network can yield market efficiency gains.

The introduction of purely financial participation through virtual bidding moves a multi-settlement wholesale electricity market using the locational marginal pricing algorithm increases the likelihood that the "separable" condition described by Ostrovsky (2012). This condition is necessary for prices in a multi-settlement market with a finite number of partially informed traders to aggregate all available information. The existence of non-zero trading costs implies that the California ISO wholesale market does not satisfy all of the conditions for the information aggregation result in Ostrovsky (2012). That being said, the basic insight of the "separable" condition is that, for every non-degenerate prior belief about states of the world, there exists a trader who, with positive probability, receives an informative signal. Expanding the number of participants able

to arbitrage day-ahead/real-time price differences at each node in the transmission network increases the likelihood that some market participant receives an informative signal about real-time conditions in the transmission network at that location.

Before concluding this section, we should emphasize that all of the examples of virtual bidding improving market efficiency are significantly more likely to occur during high demand hours when both transmission network constraints and generation runtime constraints are more likely to bind. Even the issue of non-representative load distribution factors (LDFs) is more relevant during stressed system conditions when the differences in day-ahead and real-time prices across nodes within a service territory are likely to be larger. For this reason, our empirical framework for measuring the efficiency benefits from introducing purely financial participants to California's wholesale electricity market relies on differences between high versus low demand periods before versus after the introduction of EVB.

4   Statistical Test of Arbitrage With Trading Costs

This section develops a hypothesis testing framework to determine whether or not a profitable trading strategy exists when accounting for the presence of transactions costs. In this framework, a trader can buy and sell positions in hourly day-ahead/real-time price spreads corresponding to each of the 24 hours-of-the-day. Trading strategies based on the first lag of the vector of day-ahead/real-time price differences are not feasible because market participants submit their offers into the day-ahead market for day t *before* the vector of day-ahead/real-time price differences for date $t-1$ is made public. Market participants can thus condition their trading strategies on price differences from two or more days prior.

In Appendix Section C.1, we formulate a statistical test of the null hypothesis that the elements of the autocorrelation matrices between the current vector of day-ahead/real-time price spreads and the second through tenth lags of this vector are jointly zero. We find little empirical evidence against this null hypothesis, suggesting that traders cannot earn significantly more profits by conditioning on day-ahead/real-time price differences from two or more days prior to the current day. See Appendix Tables C.1 and C.2 for the results when applying this statistical test to price spreads calculated at the service-

area level for the three large investor-owned utilities in California and at the nodal level respectively.

4.1 The Trader's Problem

Our statistical test is based on a market participant who buys or sells hourly positions, a_h, associated with X_{hd}, where $X_{hd} \equiv (P_{hd}^{DA} - P_{hd}^{RT})$ is the difference between the day-ahead and real-time electricity prices in hour-of-the-day h in day-of-sample d. The trader can condition her positions on the hour of the day, so that a_h can take on different values for for $h \in \{1, 2, ..., 24\}$. Moreover, a_h can be positive or negative.[19] Let $a = (a_1, a_2, ..., a_{24})'$ denote the 24×1 vector of hourly positions. Consistent with market rules, she chooses positions for all hours of the day simultaneously. Holding a positive (negative) position earns revenues if and only if the day-ahead price for hour h of day d is higher (lower) than the real-time price in hour h of day d. In other words, a trader earns positive revenue if and only if the position, a_h, and realized price spread, $X_{d,h}$, have the same sign.

Let $\mu_h \equiv E(X_{h,d}) = E(P_{h,d}^{DA}) - E(P_{h,d}^{RT})$ be the unconditional expectation of the day-ahead/real-time price spread for hour h; define μ to be the 24×1 vector composed of $(\mu_1, \mu_2, ..., \mu_{24})'$. The trader faces trading cost c associated with buying or selling one MWh of any combination of these 24 assets. The trader's expected profit-maximization problem is:

$$\underset{a \in R^{24}}{max} \quad \underbrace{a'\mu - c\sum_{i=1}^{24}|a_i|}_{\text{Expected Profits}} \text{ subject to } \sum_{i=1}^{24}|a_i| = 1 \tag{1}$$

The vector of positions $a^*(\mu) \in R^{24}$ denotes the solution to the constrained optimization problem described in Equation 1. We consider both the null hypothesis that profitable trading strategies exist (i.e.: $a^*(\mu)'\mu - c > 0$) and the null hypothesis that no profitable trading strategies exist (i.e.: $a^*(\mu)'\mu - c \leq 0$).

The trader pays the same per-unit trading cost c regardless of whether they buy or sell the asset; this is why overall trading costs are calculated based on the sum of the absolute values of the portfolio weights (i.e.: $c\sum_{i=1}^{24}|a_i|$). Canonical portfolio choice models (Markowitz (1952); Sharpe (1994)) instead impose a normalization based on "net

---

[19]A positive (negative) value of a_h implies selling (buying) energy in the day-ahead market and buying (selling) it back in the real-time market.

position" ($\sum_{i=1}^{24} a_i = 1$). This normalization assumes that a positive position in one asset can be offset by a negative position in another asset, which only makes sense in models with zero transaction costs.

The revenues earned by solving the optimization problem presented in Equation 1 are:

$$\phi(\mu) \equiv a^*(\mu)'\mu \;=\; \begin{array}{c} max \\ h \in \{1, ..., 24\} \end{array} |\mu_h| \tag{2}$$

In words, the trader simply buys or sells 1 MWh of the asset with the highest expected pay-off in absolute value.

4.2   Implementation of our Statistical Test of Arbitrage

We implement our hypothesis test using data on realized day-ahead/real-time price spreads for each hour h of each day-of-sample d. Let $\mathbf{X_d} = (X_{d,1}, X_{d,2}, ..., X_{d,24})'$ be the 24×1 data vector composed of realized day-ahead/real-time price spreads for day-of-sample d. Then, our estimate of the unconditional expectation of day-ahead/real-time price spreads for each hour-of-the-day is simply the sample average (i.e.: $\hat{\mu} \equiv \overline{X} = \frac{1}{N}\sum_{d=1}^{N} \mathbf{X_d}$). Our estimate of the revenue generated from the trader's optimal strategy, presented in Equation (2), is the element of $\overline{X}$ that is largest in absolute value (i.e.: $\phi(\overline{X}) \equiv \begin{array}{c} max \\ h \in \{1,...,24\} \end{array} |\overline{X}_h|$). Our test statistic is based on the difference between $\phi(\overline{X})$ and per-unit trading cost c.

However, the maximum operator is not differentiable; we thus cannot use the Delta Method to derive the asymptotic distribution of $\phi(\overline{X})$. Instead, we use the method developed by Fang and Santos (2018) for testing hypotheses involving directionally differentiable functions of a regular parameter estimate. This method is applicable because $\phi(\mu)$ is a directionally differentiable function of the parameter vector μ and sample average $\overline{X}$ is a regular estimator of μ_0 (i.e.: $\sqrt{N}(\overline{X} - \mu_0)$ is asymptotically normally distributed). Fang and Santos (2018) propose a modified bootstrap estimator for the asymptotic distribution of $\sqrt{N}(\phi(\overline{X}) - \phi(\mu))$.

To implement this estimator, we simulate the distribution of $\phi(\overline{X})$ using a procedure based on numerical derivatives developed by Hong and Li (2018). For this procedure, we first compute moving blocks bootstrap re-samples of $\overline{X}$ with block size equal to the

largest integer less than or equal to $N^{1/3}$ (Kunsch et al. (1989)).[20] Let the sample average calculated from the bth bootstrap re-sample be denoted $\overline{X}^b$. We next construct:

$$Z^b = \frac{\phi(\overline{X} + \sqrt{N}(\overline{X}^b - \overline{X})\epsilon) - \phi(\overline{X})}{\epsilon} \tag{3}$$

for $b = 1, 2, ..., B$. Hong and Li (2018) demonstrates that the asymptotic distribution of $\sqrt{N}(\phi(\overline{X}) - \phi(\mu))$ can be approximated by the bootstrap distribution of Z^b provided that, as sample size N goes to infinity, ϵ tends to zero but $\sqrt{N}\epsilon$ tends to infinity. To satisfy these conditions, we set $\epsilon = N^{-1/3}$, which is the value recommended by Hong and Li (2018).

This estimate of the asymptotic distribution of $\sqrt{N}(\phi(\overline{X}) - \phi(\mu))$ allows us to test for the existence of profitable arbitrage opportunities for any per-unit trading cost c. However, setting c equal to the explicit transaction fee charged by the market operator ignores both the monthly fixed charge to participate in California's electricity market and the opportunity cost of the money posted as collateral with the system operator. Using the transaction fee for c also assumes that there is no opportunity cost associated with the time and effort of the individual undertaking the trades as well as no operating cost associated with scaling up a trading strategy. For these reasons, we instead use our hypothesis testing procedure to calculate the per unit trading costs implied by just rejecting the null hypothesis of profitable arbitrage as well as just rejecting the null of no profitable arbitrage.

To do this, we estimate the distribution of $\phi(\overline{X})$ using moving blocks bootstrap; in particular, the bth re-sample gives us:

$$\phi(\overline{X})^b = \phi(\overline{X}) + \frac{Z^b}{\sqrt{N}}. \tag{4}$$

This bootstrap distribution is used to compute two values. The first, c_{lower}, is the smallest value of the trading cost, c, that would cause rejection of the null hypothesis that a

---

[20]Given a sample $\{\mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_N}\}$, each moving blocks bootstrap re-sample $b \in \{1, 2, ..., B\}$ is constructed as follows. First, we partition the data into K *non-overlapping* blocks of size M: $\{\mathbf{B_1}, \mathbf{B_2}, ..., \mathbf{B_K}\} \equiv \{\mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_M}\}, \{\mathbf{X_{M+1}}, ..., \mathbf{X_{2M}}\}, ..., \{\mathbf{X_{M(K-1)+1}}, ..., \mathbf{X_{KM}}\}$. Next, let S be a discrete uniform variable over the integers $\{0, 1, ..., K\}$; we construct the bth bootstrap re-sample by drawing K integers from S independently and identically and merging together blocks based on these draws. For example, if we draw $\{2, 5, K, ..., 5\}$, then the bootstrap sample would be $\{\mathbf{B_2}, \mathbf{B_5}, \mathbf{B_K}, ..., \mathbf{B_5}\}$. When implementing this procedure, we set $M = floor(N^{1/3})$ and $K \equiv floor(\frac{N}{M})$, where $floor(X)$ is X rounded down to the nearest integer.

profitable trading strategy exists (i.e.: reject the null hypothesis that $a^*(\mu)'\mu - c > 0$). Therefore, c_{lower} is equal to the 5th percentile of the bootstrapped distribution of $\phi(\overline{X})$. The second magnitude, c_{upper}, is the largest value of the trading cost that results in rejection of the null hypothesis that no profitable trading strategy exists (i.e.: reject the null hypothesis that $a^*(\mu)'\mu - c \leq 0$). As a result, c_{upper} is simply the 95th percentile of the distribution of $\phi(\overline{X})$. Our estimates are thus lower bounds on the true c_{lower} and c_{upper} required to reject the null hypotheses of arbitrage and no arbitrage respectively to the extent that we are under-estimating the trading revenues that a market participant can earn by exploiting expected day-ahead/real-time price spreads.

5   Estimates of Implied Trading Costs

This section describes the results from applying our statistical test of arbitrage with transaction costs to hourly day-ahead and real-time electricity prices from California's wholesale electricity market.

5.1   Data

We use hourly data on day-ahead and real-time wholesale electricity prices at all pricing locations (i.e.: nodes) in the transmission grid run by California's Independent System Operator (CAISO). Our data span from April 1, 2009 to December 31, 2012.[21] There are over 5,000 nodes, all with potentially different prices, depending on the level of demand, configuration of the transmission network, set of available generation units and other system operating constraints. We also consider the load aggregation point (LAP) level prices faced by each of California's three major electricity distribution utilities: Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E). The LAP-level real-time price for a given hour is computed by taking the quantity-weighted average over the real-time prices at all nodes in the utility's service area with a positive amount of energy withdrawn in that hour. LAP-level day-ahead prices are computed the same way using day-ahead quantities.

Figure 2 presents a comparison by hour-of-the-day of the average difference between

---

[21]California introduced nodal pricing on April 1, 2009.

the day-ahead and real-time prices faced by PG&E, SCE, and SDG&E before versus after the introduction of financial trading (i.e.: explicit virtual bidding) on February 1, 2011. This figure provides descriptive evidence that day-ahead/real-time price spreads are larger in absolute value before the introduction of explicit virtual bidding (EVB) for each of the three distribution utilities.[22] For example, day-ahead prices for PG&E are much lower than real-time prices on average for the hours of 8PM to 12AM. We demonstrate in Appendix Section C.2 that the reduction in average day-ahead/real-time price differences after the introduction of purely financial participants is statistically significant.

We next perform a joint test of the null hypothesis that the average differences between day-ahead and real-time prices are zero for all 24 hours of the day. This test is performed separately for each LAP before versus after the introduction of explicit virtual bidding. We compute our chi-squared test statistics accounting for the potential for up to 14 days of autocorrelation using the autocorrelation-consistent covariance matrix estimator developed by Newey and West (1987).[23] The test statistics corresponding to these tests, presented in Table 1, are quite large. We reject the null hypothesis that all 24 of the hourly price difference means are zero for every LAP both before and after EVB.[24]

However, this result does not imply that traders can profitably arbitrage day-ahead/real-time price spreads after the introduction of EVB. This is because the statistical tests presented in Table 1 do not account for the potentially sizable transaction costs faced by purely financial participants in the California ISO. In the next subsection, we present results from a statistical test of arbitrage that accounts for the fact that day-ahead/real-time price differences can differ from zero simply due to transaction costs.

## 5.2 Implied Trading Cost Estimates

We first implement the statistical test of arbitrage described in Section 4 using data on the day-ahead and real-time electricity prices faced by each of the three major electricity

---

[22]Appendix Figure A.2 plots these hourly average day-ahead/real-time spreads along with their pointwise 95% confidence intervals. This figure indicates that we can reject the null hypothesis that expected price spreads are zero for some hours-of-the-day even after the introduction of financial trading.

[23]The Newey-West covariance matrix estimator is: $\hat{\Sigma} = \hat{\Lambda}_0 + \sum_{j=1}^{m} w(j,m)(\hat{\Lambda}_j + \hat{\Lambda}_j{}')$, where $\hat{\Lambda}_j = \frac{1}{T}\sum_{t=j+1}^{T}(X_t - \overline{X})(X_{t-j} - \overline{X})'$, $\overline{X} = \frac{1}{T}\sum_{t=1}^{T} X_t$, and $w(j,m) = 1 - \frac{j}{m+1}$. We account for $m = 14$ days of autocorrelation.

[24]The upper $\alpha = 0.05$ critical value for the χ^2_{24} distribution is 36.415.

Figure 2: Day-Ahead/Real-Time Price Spreads By Hour-of-Day: Before and After EVB
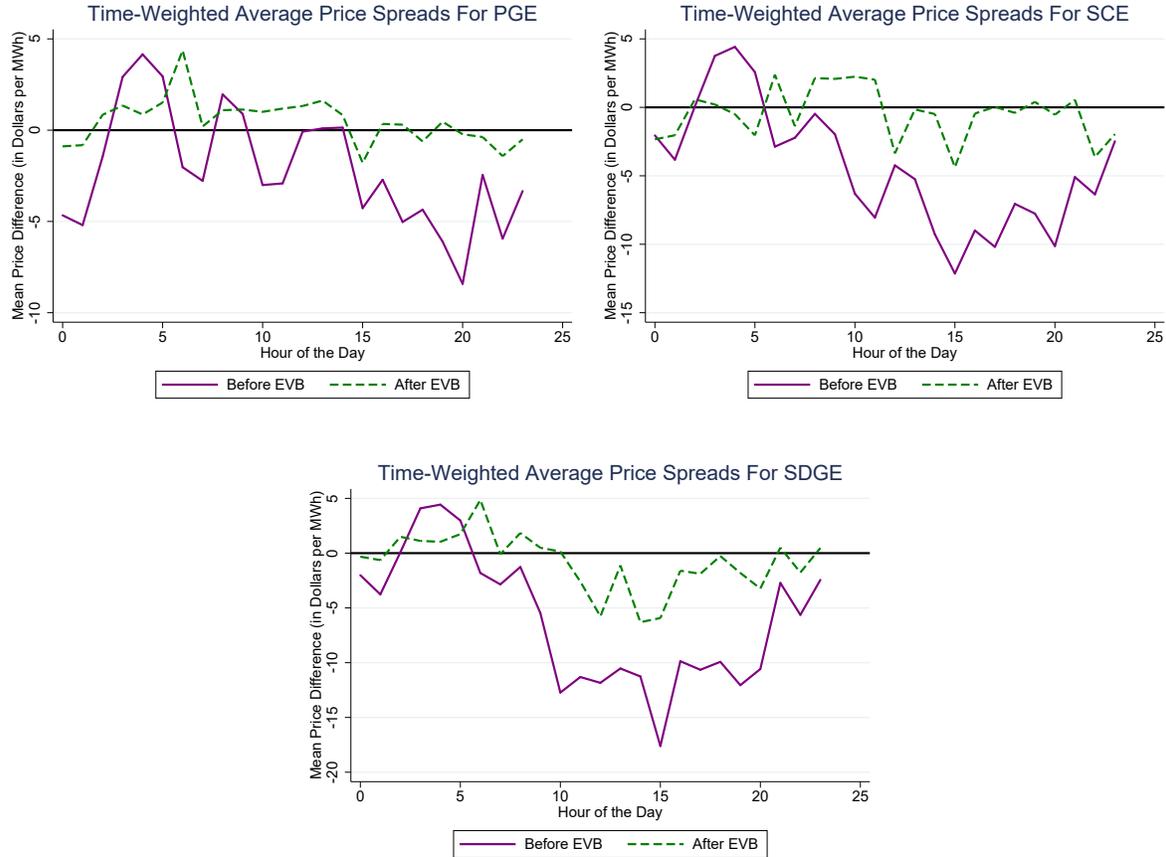
Notes: This figure presents the average for each hour-of-the-day of the difference between day-ahead and real-time electricity prices for each of three major load aggregation points (LAPs) in California. We plot this average price spread separately for the sample periods before versus after the introduction of explicit virtual bidding (EVB). The three LAPs considered in this figure correspond to the territories served by Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E).

Table 1: Test Statistics for the Joint Test of Zero Mean Price Differences

	Before EVB	After EVB
PG&E	140.175	194.317
SCE	137.659	164.065
SDG&E	151.456	260.327

Notes: This table presents the chi-squared test statistic corresponding to the null hypothesis that all 24 hour-of-the-day average day-ahead/real-time price spreads are equal to zero. This statistical test is performed separately for each load aggregation point (LAP) before and after the introduction of explicit virtual bidding (EVB). The three LAPs correspond to the territories served by Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E). We compute these test statistics accounting for the potential for up to 14 days of autocorrelation using the autocorrelation-consistent covariance matrix estimator developed by Newey and West (1987). We reject the null hypothesis if the test statistic is larger than 36.415, which is the upper $\alpha = 0.05$ critical value for the χ^2_{24} distribution.

Table 2: LAP-Level Implied Trading Costs: c_{lower} and c_{upper}

	LAP	Before EVB	After EVB
Lower 5% C.I. (c_{lower})	PG&E	8.290	4.509
	SCE	11.728	4.662
	SDG&E	16.028	6.594
Upper 95% C.I. (c_{upper})	PG&E	15.211	7.387
	SCE	21.307	8.827
	SDG&E	36.281	10.740

Notes: This table presents the implied trading costs from our statistical framework, estimated separately for each load aggregation point (LAP) for sample periods before the introduction of explicit virtual bidding (4/1/2009-2/1/2011) versus after the introduction of EVB (2/1/2011-12/31/2012). The three LAPs considered in this table correspond to the territories served by California's three major electricity distribution companies: Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E). c_{lower} is the smallest value of per-unit trading costs c for which we can reject the null hypothesis that a profitable trading strategy exists while c_{upper} is the largest value of c such that we can reject the null hypothesis that no profitable trading strategy exists.

distribution utilities in California: Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E). Specifically, Table 2 reports our estimated implied trading costs both before and after the introduction of explicit virtual bidding for each load aggregation point (LAP). Recall that c_{lower} is the smallest value of per-unit trading costs, c, for which we can reject the null hypothesis that a profitable strategy exists while c_{upper} is the largest value of c for which we can reject the null hypothesis that no profitable trading strategy exists. Table 2 demonstrates that our estimates of c_{lower} and c_{upper} fell substantially after the introduction of explicit virtual bidding for all three LAPs.[25] This result is consistent with the logic outlined in Section 2.

In order to more formally compare implied trading costs before versus after the introduction of explicit virtual bidding (EVB), Figure 3 plots the bootstrap distribution of the *difference* in implied trading costs (i.e.: $\phi(\overline{X}^{pre}) - \phi(\overline{X}^{post})$) for each LAP before versus after EVB. The left vertical line in this figure is the 5th percentile of the distribution of $c_{pre} - c_{post}$ and the right vertical line is the 95th percentile of this distribution. If the 5th percentile of the distribution of bootstrap estimates of $c_{pre} - c_{post}$ is greater than zero, then we can reject the null hypothesis that $c_{pre} \leq c_{post}$ at a 5% significance level.

---

[25]Appendix Figure A.3 plots the bootstrap distributions of implied trading costs for the pre-EVB and post-EVB sample periods for each of the three LAPs.

If the 95th percentile of the bootstrapped distribution of estimated $c_{pre} - c_{post}$ is less than zero, then we can reject the null hypothesis that $c_{pre} \geq c_{post}$ at a 5% significance level. For both SCE and SDG&E, we reject the null hypothesis that our implied trading costs are higher post-EVB relative to pre-EVB, but fail to reject the null hypothesis that our implied trading costs are higher pre-EVB relative to post-EVB. However, we do not have the statistical power to reject either null hypothesis for PG&E. Summarizing, Figure 3 provides statistical evidence that the implied costs of trading LAP-level price spreads fell after the introduction of explicit virtual bidding for the SCE and SDG&E LAPs.

We also compute c_{lower} and c_{upper} for each pricing node in the California ISO control area. Figure 4 plots the values of c_{lower} and c_{upper} for each node before and after the introduction of EVB. This figure plots the across-node distributions of c_{lower} and c_{upper} separately for nodes associated with generation units ("Gen Nodes") and nodes not associated with generation units ("Non-Gen Nodes"). We see from Figure 4 that the distributions of both c_{lower} and c_{upper} are shifted downward post-EVB relative to pre-EVB. Thus, for any fixed value of trading costs c, we reject the null hypothesis that a profitable trading strategy exists for more nodes in the post-EVB sample period relative to the pre-EVB sample period. Similarly, our estimated across-node distribution of c_{upper} indicates that the null hypothesis that no profitable trading strategies exist for a given trading cost c can be rejected for more nodes before EVB relative to after EVB.

Next, we compute the bootstrap distribution of estimates of $c_{pre} - c_{post}$ for each of the more than 4,000 nodes in the California ISO control area; "pre" refers to the sample period before the introduction of explicit virtual bidding (4/1/2009-2/1/2011) and "post" refers to the sample period after the introduction of explicit virtual bidding (2/1/2011-12/31/2012). The first row of Table 3 reports the proportion of nodes for which we reject the null hypothesis that implicit trading costs increased after the introduction of EVB (i.e.: $c_{pre} \leq c_{post}$), separately for generation nodes ("Gen Nodes") versus non-generation nodes ("Non-Gen Nodes"). The second row of Table 3 reports the proportion of nodes for which we reject the null hypothesis that implicit trading costs decreased after the introduction of EVB (i.e.: $c_{pre} \geq c_{post}$), once again separately for generation nodes versus non-generation nodes. We reject the null hypothesis that implicit trading costs increased after the introduction of EVB for more than 70 percent of the nodes; the percentage of rejections at non-generation nodes is slightly higher than at generation nodes. In

Figure 3: Bootstrap Distribution of the Difference in Trading Costs
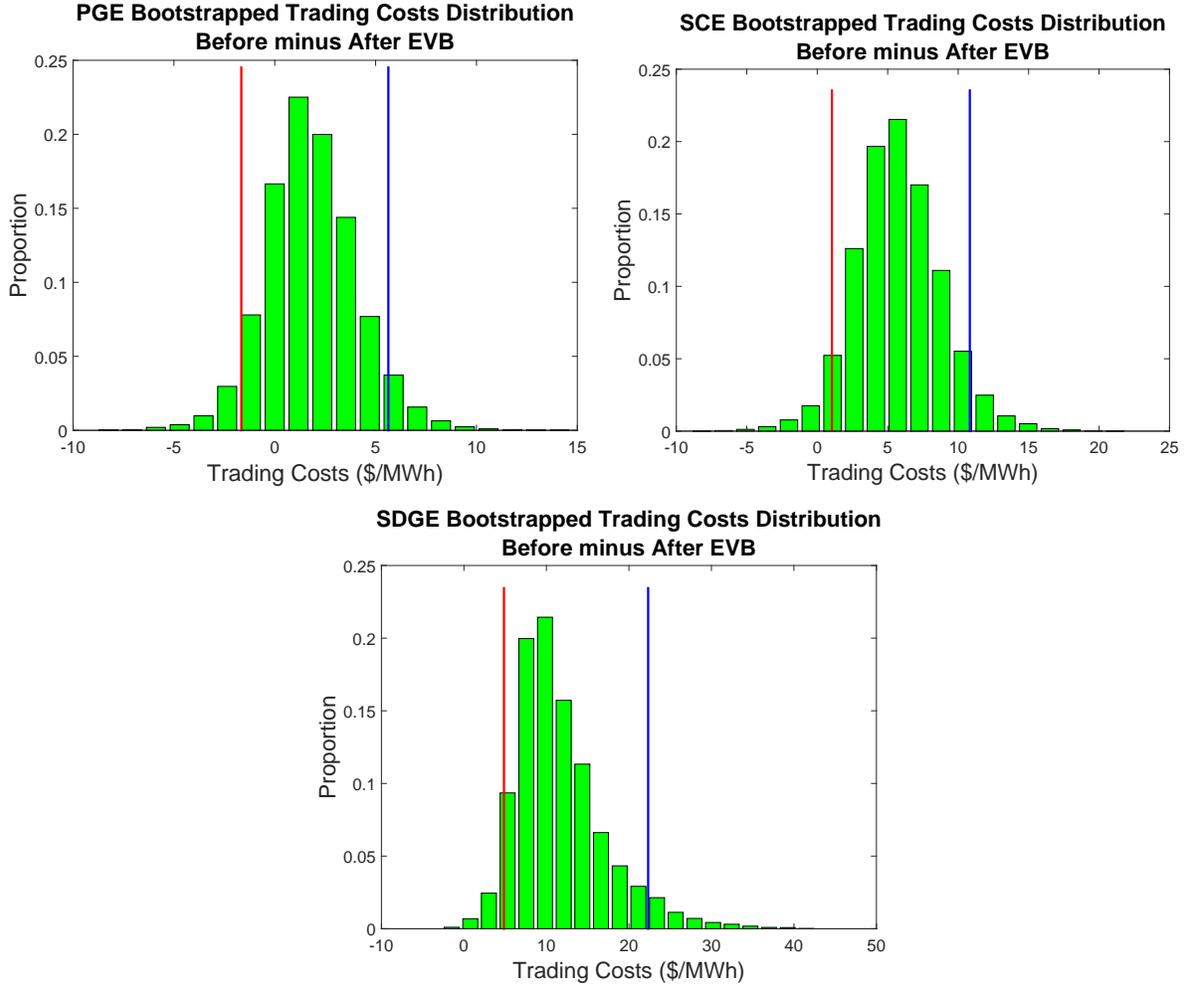
Notes: This figure plots the bootstrap distribution of the difference in trading costs (i.e.: $(\phi(\overline{X}^{pre}) - \phi(\overline{X}^{post}))$, where "pre" indicates the sample period before the introduction of explicit virtual bidding (4/1/2009-2/1/2011) and "post" indicates the sample period after the introduction of explicit virtual bidding (2/1/2011-12/31/2012). We plot this distribution separately for each of three LAPs; these LAPs correspond to the territories served by California's three major electricity distribution utilities: Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E). The left vertical line on the graph is the 5th percentile of the distribution of $c_{pre} - c_{post}$ and the right vertical line is the 95th percentile of the distribution of $c_{pre} - c_{post}$.

Figure 4: Nodal-Level Distribution of Implied Trading Costs: Before and After EVB
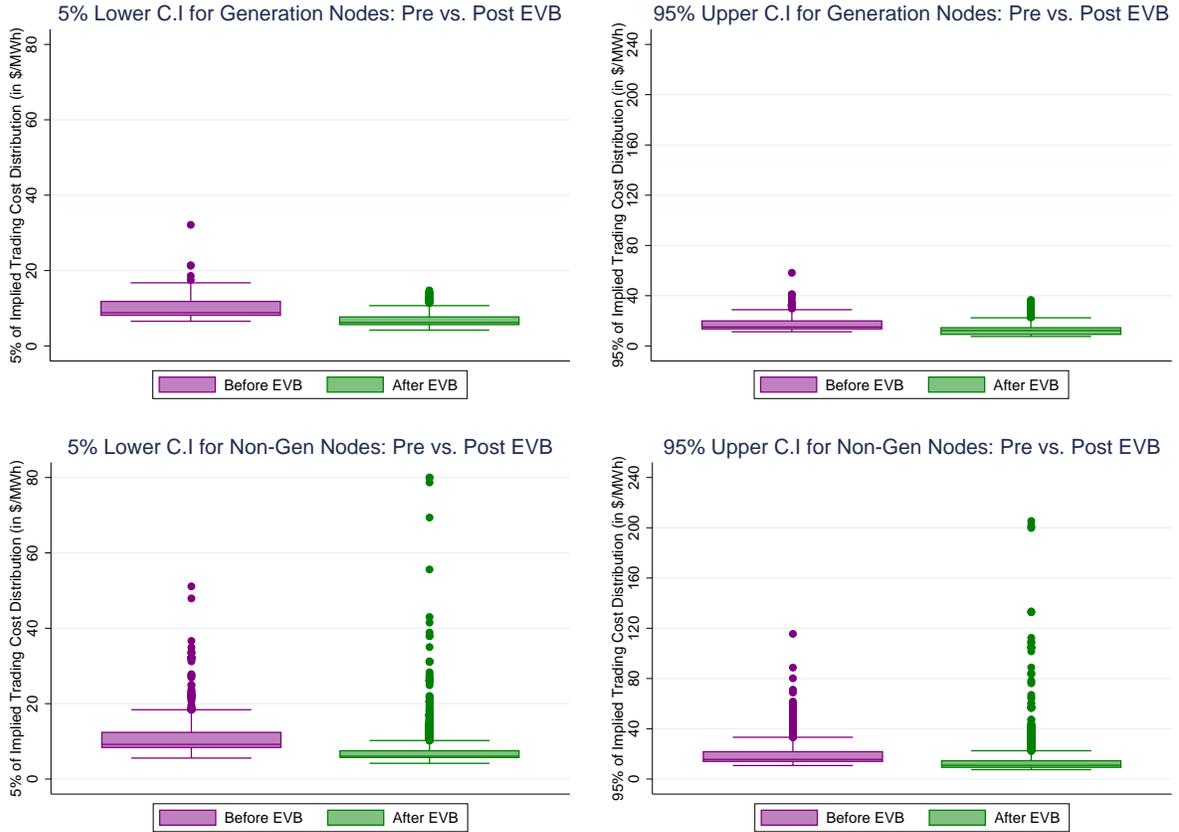
Notes: This figure plots the values of c_{lower} and c_{upper} for each pricing location (i.e.: node) estimated separately for the sample period before the introduction of explicit virtual bidding (4/1/2009-2/1/2011) versus the sample period after the introduction of explicit virtual bidding (2/1/2011-12/31/2012). c_{lower} is the smallest value of per-unit trading costs c for which we can reject the null hypothesis that a profitable trading strategy exists while c_{upper} is the largest value of c such that we can reject the null hypothesis that no profitable trading strategy exists. We plot the across-node distributions of c_{lower} and c_{upper} separately for nodes associated with generation units and nodes not associated with generation units. The box portion of this box and whiskers plot contains all nodes within the 25% through 75% quantiles of the distribution of our implied trading costs. The whiskers include nodes in the $1.5 \times IQR$ of the distribution of implied trading costs, where IQR (inter-quartile range) is the distance between the 25% and 75% quartiles of a distribution. Finally, the remaining nodes are outliers outside of the $1.5 \times IQR$ of the distribution of implied trading costs.

Table 3: Proportion of Nodes that Reject $c_{pre} \leq c_{post}$ or $c_{pre} \geq c_{post}$

	Total	1(Gen Node)	1(Non-Gen Node)
$H_0 : c_{pre} \leq c_{post}$	0.707	0.659	0.711
$H_0 : c_{pre} \geq c_{post}$	0.042	0.076	0.039
Number of Nodes	4,316	355	3,961

Notes: This first row of this table reports the proportion of nodes for which we can reject the null hypothesis that implied trading costs are larger after the introduction of explicit virtual bidding (EVB) relative to before the introduction of EVB (i.e.: $c_{pre} \leq c_{post}$), separately for generation nodes ("Gen Nodes") versus non-generation nodes ("Non-Gen Nodes"). The second row of this table reports the proportion of nodes for which we can reject the null hypothesis that implied trading costs are larger pre-EVB relative to post-EVB (i.e.: $c_{pre} \geq c_{post}$), once again separately for generation nodes and versus non-generation nodes.

contrast, we reject the null hypothesis that trading costs fell after the introduction of EVB for less than 5 percent of the nodes. A rejection frequency of 5% is consistent with the null hypothesis being true for all nodes because the size of each hypothesis test is $\alpha = 0.05$.

As discussed in Section 2, before the introduction of explicit virtual bidding (EVB) suppliers could only exploit expected differences between day-ahead and real-time prices using their physical bids at the locations where their generation units inject electricity ("generation nodes"). Load-serving entities could only place physical bids at the service territory or load aggregation point (LAP) level, making it extremely costly for load-serving entities to exploit expected day-ahead/real-time price spreads using their physical bids (termed "implicit virtual bidding"). Based on this, we expect implied trading costs to be higher at non-generation nodes relative to generation nodes before EVB because neither suppliers nor demanders could implicitly virtual bid at non-generation nodes. The introduction of explicit virtual bidding allowed any market participant to place virtual bids at any node; thus, we expect the mean reduction in implied trading costs after EVB to be larger for non-generation nodes relative to generation nodes.

To test these two hypotheses, we regress the value of c_{lower} at each node both before and after the introduction of EVB on a constant, an indicator variable for whether the node is associated with a generation unit (i.e.: 1(Gen Node), an indicator variable for whether the implied trading cost is from the post-EVB period (i.e.: 1(Post EVB)), and an indicator variable for whether the observation is from a generation node during the post-EVB period. The unit of observation for this regression is thus a node in the pre-

EVB or post-EVB sample period. Heteroskedasticity-consistent standard errors are in parentheses. Finally, we run the same regression with c_{upper} as the dependent variable.

Table 4 presents the results of estimating this difference-in-differences specification for both the 5% and 95% percentiles of the bootstrapped distribution of implied trading costs for each node before versus after EVB. We see that the coefficient on Post EVB is negative for both c_{lower} and c_{upper}, indicating that the average level of implied trading costs across locations fell after the implementation of EVB. Moreover, the results indicate that both c_{lower} and c_{upper} are significantly lower for generation nodes relative to non-generation nodes prior to EVB. However, this difference across generation versus non-generation nodes is essentially eliminated after the introduction of explicit virtual bidding. Specifically, we fail to reject the null hypothesis that the sum of the coefficient estimates corresponding to the variables 1(Gen Node) and 1(Post EVB) × 1(Gen Node) is zero for both c_{lower} and c_{upper}. As expected, the difference in implied trading cost before versus after explicit virtual bidding (i.e: $c^{pre} - c^{post}$) fell more for non-generation nodes than for generation nodes.

Summarizing, Table 4 provides statistical evidence consistent with all three of our hypotheses: (1) implied trading costs are lower for generation nodes versus non-generation nodes prior to EVB, (2) implied trading costs fell after EVB, and (3) implied trading costs fell more for non-generation nodes relative to generation nodes after EVB.

Table 4: Implied Trading Costs Before vs. After EVB For Gen. versus Non-Gen Nodes

Dependent Variable	(1) c_{lower}	(2) c_{upper}
1(Post EVB) × 1(Gen Node)	0.532	1.421
	(0.174)	(0.431)
1(Post EVB)	-3.527	-5.404
	(0.075)	(0.193)
1(Gen Node)	-0.654	-1.765
	(0.119)	(0.250)
Constant	10.72	19.16
	(0.054)	(0.118)
Mean of Dep. Var.	8.840	16.221
Std. Dev. of Dep. Var.	3.847	9.306
Number of Obs.	9,791	9,791
R^2	0.202	0.080

Notes: This table reports the results from our difference-in-differences specification comparing implied trading costs before versus after the introduction of explicit virtual bidding (EVB) for pricing locations (i.e.: nodes) associated with generation units (i.e.: "Gen Nodes") versus not associated with generation units (i.e.: "Non-Gen Nodes"). The unit of observation for these regressions is a node before versus after the introduction of EVB. We report heteroskedasticity-consistent standard errors in parentheses. This specification includes an indicator 1(Post EVB) for implied trading costs estimated for the sample period after the introduction of explicit virtual bidding (2/1/2011-12/31/2012); the pre-EVB sample period is 4/1/2009-2/1/2011. We also include an indicator 1(Gen Node) that's one if and only if the node is associated with a generator. Finally, we include the interaction between these two variables (1(Post EVB) × 1(Gen Node)). We consider two dependent variables: c_{lower} (Column 1) and c_{upper} (Column 2). c_{lower} is the smallest value of per-unit trading costs c for which we can reject the null hypothesis that a profitable trading strategy exists and c_{upper} is the largest value of c for which we can reject the null hypothesis that no profitable trading strategy exists.

6 Quantifying Market Efficiency Benefits in High versus Low Demand Hours

This section describes the data, methodology, and results of our empirical analysis demonstrating that the introduction of purely financial participants resulted in market efficiency benefits in high demand hours relative to low demand hours. As discussed in Section 3, the market efficiency benefits from introducing financial trading are likely to be largest in the highest demand hours.

6.1 Data

We utilize hourly data on generator-level output from the California ISO, daily natural gas prices for each of Northern and Southern California from SNL Financial[26], and daily fuel oil prices from the Energy Information Administration. These data span the sample period 4/1/2009-3/31/2012. We also have information for each electricity generation unit in the California ISO control area on its capacity, fuel type (i.e: natural gas, oil, wind, etc.), total amount of thermal energy required to start up the unit, and its heat rate *curve*. This heat rate curve tells us: "if a generation unit is currently utilizing X% of their generation capacity, we need Y Million BTU (MMBTU) of thermal input in order to produce one more MWh of electrical energy". Combined, these data allow us to compute the total quantity of thermal energy consumed by each generation unit during each hour of our sample period.

We define generation unit i as starting in hour t if its output in hour t-1 is zero ($Q_{i,t-1} = 0$) and its output in hour t is greater than zero Define the thermal energy required to start up unit i to be E_i^S. Each unit also submits the heat rate HR_i^k relevant for each of $k = 1, 2, ..., 10$ quantity steps. In particular, the thermal energy required to produce $Q_{i,t}$ is:

$$E_{i,t}^P = \sum_{k=1}^{10} HR_i^k (Q_{i,t} - \sum_{j=1}^{k-1} \overline{Q}_i^j)^+$$

where the notation $y^+ \equiv y \times 1(y > 0)$, $\overline{Q}_i^j$ corresponding to step j is the maximum level of output for which heat rate HR_i^j is applicable. Then, the total quantity of thermal

---

[26]We use the city-gate natural gas price corresponding to the Pacific Gas and Electric delivery point and the Southern California Gas Citygate delivery point respectively.

energy consumed by unit i in hour t is:

$$E_{i,t} = E_{i,t}^P + E_i^S 1(Q_{i,t} > 0)1(Q_{i,t-1} = 0)$$

The total fuel costs associated with this quantity of energy consumed is simply:

$$TC_{i,t} = E_{i,t} \times P_{i,t}^F$$

where $P_{i,t}^F$ is the daily fuel price (either natural gas or oil) applicable to that unit. For natural gas fired units, we use the PG&E City-gate natural gas price if the unit is located in Pacific Gas and Electric's service territory and we use the SCE City-gate price if the unit is located in either Southern California Edison's service territory or San Diego Gas and Electric's service territory. We calculate the costs of the relatively small amount of distillate fuel oil burned by generation units in California using the diesel fuel price paid in Los Angeles.[27] We sum the total fuel costs, total heat energy used, and total number of starts over fossil-fired units for each hour-of-sample in order to construct our first three market-level outcome variables. In particular, we assess how the introduction of financial trading impacted the hourly aggregate fuel costs per MWh of electricity produced by fossil-fuel fired units, the hourly aggregate thermal energy used per MWh, and whether any fossil fuel fired unit was forced to start up in each hour.

The California ISO incurs ancillary service costs in order to procure sufficient "regulation reserves," "spinning reserves" and "non-spinning reserves" to make certain that electricity supplied equals electricity demanded at every instant in time even in the face of generation and tranmission deratings and outages. Policymakers and regulators have expressed concern that financial traders take advantage of physical constraints such as transmission congestion in order to make profits, driving up these ancillary service costs. Thus, we consider the ancillary service costs paid by CAISO per MWh of fossil-fuel-fired electricity produced in each hour-of-sample as one of our market outcome measures.

Finally, we use hourly data on day-ahead and real-time electricity prices paid by each of California's three large, investor-owned utilities: Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E). In particular, our last three outcome measures are the absolute differences between hourly

---

[27]In particular, we use data from the Energy Information Administration (EIA) on the "ultra-low sulfur CARB diesel spot price" relevant to Los Angeles, California.

day-ahead and real-time prices paid by the three utilities. This allows us to explore the extent to which day-ahead/real-time price spreads converged after the introduction of financial trading in relatively high demand hours.

6.2 Difference-in-Differences Methodology

For each market outcome Y_t in hour-of-sample t, we employ the following difference-in-differences framework:

$$Y_t = \alpha_m + \gamma_h + \theta_w + X_t\phi\beta_0\text{HIGH}_t^p + \delta_{DD}\text{HIGH}_t^p \times \text{POSTEVB}_t + u_t \tag{5}$$

where we include month-of-sample fixed effects (α_m), hour-of-the-day fixed effects (γ_h), and an indicator for whether the day-of-sample is a weekday versus weekend (θ_w). This specification also controls for a host of factors X_t: the log of total electricity demand, two separate controls for the logs of the citygate natural gas prices paid in PG&E and SCE, as well as separate controls for the log of total hourly production from (1) wind and solar sources, (2) nuclear sources, and (3) hydro sources.

The indicator variable POSTEVB_t is equal to one if the day-of-sample is after the introduction of financial trading in 2/1/2011 and is equal to zero otherwise; this variable is not included separately in the regression specification because it is absorbed by the month-of-sample fixed effects. The indicator variable HIGH_t^p is equal to one if and only if aggregate electricity demand in hour-of-sample t is larger than the pth percentile of the distribution of aggregate hourly demand across our 4/1/2009 to 3/31/2012 sample period. Although our primary specifications are based on the $p = 90^{th}$ percentile of demand, we show that our results are qualitatively similar for the 50, 75, 95, and 99 percentiles in a sensitivity analyses. The independent variable of interest, $\text{HIGH}_t^p \times \text{POSTEVB}_t$, captures how the introduction of financial trading impacts each hourly market-level outcome Y_t in high demand hours relative to low demand hours. Finally, we cluster standard errors by week-of-sample.

6.3 Primary Findings

Table 5 presents the results from estimating the difference-in-differences specification described in the previous subsection. We discuss the robustness checks, such as the pre-trend analysis, pertaining to these results in the next subsection. Column (1) of Table 5 indicates that the introduction of financial trading to California's wholesale electricity market resulted in a 3.6% reduction in average fuel costs per MWh in high demand hours relative to low demand hours. Similarly, we see a 4.2% reduction in average input heat energy per MWh due to financial trading in relatively high demand hours. Translating these effects to dollars, a 3.6% reduction in fuel costs per MWh across all of the power plants in our sample results in a 4.2 million dollar reduction in the annual fuel costs paid in high demand hours. Similarly, a 4.2% reduction in thermal energy per MWh, implies a reduction in CO_2 emissions of 145,000 tons, focusing only on high demand hours.

One source of the efficiency gains from purely financial participants described in Section 3 is the incentive for these participants create for a lower cost dispatch of available of generation units. Indeed, we see from Column (3) that at least one unit has to start up in 4.2% less high demand hours after financial trading on average. Both the sign and lack of statistical significance of this effect assuages concerns that physical and financial participants are placing trades in order to profit from the physical constraints such as start-up or ramping. To more directly address this concern, we consider total "ancillary services" payments made to electricity suppliers in order to ensure that supply meets demand at every instant in time. If the introduction of purely financial participants resulted in an exacerbation of transmission, start-up, ramping, or other system operating constraints, we would expect a marked increase in the ancillary services costs. This does not turn out to be the case; Column (4) of Table 5 demonstrates that there is no statistical difference in the ancillary costs per MWh paid out in relatively high demand hours before versus after financial trading. In fact, the point estimate is negative, suggesting that financial participation results in *lower* ancillary services costs per MWh in relatively high demand hours.

Finally, the last three columns of Table 5 consider the effect of financial trading on the absolute difference between the day-ahead and real-time electricity prices faced by each of California's three major distribution utilities: PG&E, SCE, and SDG&E. All three columns indicate that the average absolute day-ahead/real-time price spread fell

Table 5: Diff-in-Diff: High Demand vs. Low Demand Hours Before vs. After EVB

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1(Demand > 90%)	-0.036	-0.042	-0.042	-0.071	-5.715	-13.902	-21.568
× 1(Post EVB)	(0.012)	(0.012)	(0.028)	(0.139)	(3.003)	(5.665)	(9.864)
1(Demand > 90%)	0.032	0.036	0.008	0.123	2.496	6.231	7.882
	(0.005)	(0.005)	(0.022)	(0.055)	(2.145)	(4.395)	(5.891)
R^2	0.903	0.362	0.125	0.590	0.041	0.039	0.049
Avg. DV in Levels	32.63	8.048	0.686	1.704	13.199	15.055	16.426
Number of Obs.	26,277	26,277	26,277	26,276	26,277	26,277	26,277

Notes: The unit of observation for these regressions is hour-of-sample. Standard errors are clustered by week-of-sample and are reported in parentheses. The dependent variables considered in this table are: (1) the log of fuel cost per MWh, (2) the log of input energy per MWh, (3) an indicator of whether any fossil fuel fired units started up, (4) the log of ancillary services per MWh, (5) the absolute value of the day-ahead/real-time prices in PG&E, (6) the absolute value of the day-ahead/real-time prices in SCE, and (7) the absolute value of the day-ahead/real-time prices in SDG&E. The row titled "Avg. DV in Levels" reports the mean of the dependent variable in levels even if the dependent variable is logged when running the regression. For all of the specifications, we control for month-of-sample fixed effects, hour-of-the-day fixed effects, an indicator for whether the day-of-sample is a weekday versus weekend, the log of total electricity demand, two separate controls for the logs of the citygate natural gas prices paid in PG&E and SCE as well as separate controls for the log of total hourly production from (1) wind and solar sources, (2) nuclear sources, and (3) hydro sources.

Table Description: This table presents the difference-in-differences results pertaining to the effect of explicit virtual bidding (i.e.: financial trading) on market outcomes in high demand hours relative to low demand hours. The data used for this table span the sample period 4/1/2009-3/31/2012. California introduced explicit virtual bidding (EVB) on 2/1/2011; the "Post EVB" indicator is thus one if and only if the day-of-sample is on or after 2/1/2011. Hours-of-sample with system demand greater than the 90% of the distribution of hourly demands across our 4/1/2009-3/31/2012 sample period are deemed to be "high demand" hours.

significantly after 2/1/2011 in relatively high demand hours. Moreover, we demonstrate in Appendix Section C.3 that the volatility in both day-ahead/real-time price spreads and real-time prices fell after the introduction of purely financial participants. Combined, our findings suggests that day-ahead electricity prices better predict expected real-time prices, providing both electricity suppliers and demanders with the information to make their day-ahead market outcomes more closely reflect real-time system conditions.

6.4 Robustness Checks

This subsection discusses the various robustness checks pertaining to our difference-in-differences analysis. First, Appendix Table B.1 indicates that there is no statistical difference in the *trends over time* in average outcomes in high demand hours versus low demand hours prior to the introduction of financial trading on 2/1/2011. This provides evidence that our results are not driven by pre-existing trends in how outcomes vary across high versus low demand hours.

Next, the empirical results remain similar if we remove the 28 days before and after 2/1/2011 (see Appendix Table B.2); our findings do not seem to be due to any short-run adjustments to the policy change. In addition, Appendix Table B.3 presents the results using only data from the 6 months before and after the introduction of financial trading on 2/1/2011. These empirical results are quantitatively quite similar to our primary results from Table 5.

Appendix Table B.5 presents our specifications estimated on data aggregated to the daily level. The results at the daily-level are broadly consistent with those presented in Table 5 with the exception of the absolute day-ahead/real-time price spreads. This in unsurprising given that day-ahead prices are far more likely to deviate substantially from real-time prices during certain hours of the day; financial trading is especially beneficial in reducing price spreads in those hours.

Finally, Appendix Table B.6 shows that our estimated effect of financial trading on average fuel costs per MWh in relatively high demand hours remains quantitatively similar if "high demand" is defined to hours-of-sample above the 50, 75, 95, and 99 percentiles of hourly demand rather than the $p = 90^{th}$ percentile as considered in our primary specifications. We use this effect in order to calculate the efficiency benefits

37

from slightly lowering the per-unit transaction costs in relatively high demand hours in the next section. Thus, it is comforting that our physical efficiency benefit per MWh of fossil-fuel-fired electricity production doesn't change significantly based on our definition of "relatively high demand".

7   Implications for the Design of Electricity Markets

Allowing purely financial participants in multi-settlement LMP wholesale electricity markets is not without controversy. Many have argued that financial traders take advantage of physical constraints such as transmission or the start-up/ramping of units in order to profit at the expense of electricity consumers. We assess the empirical importance of these concerns by comparing market outcomes in high versus low demand hours before versus after California introduced financial trading to its electricity market on February 1st 2011. Using this difference-in-differences framework, we find that, in relatively high demand hours, the introduction of financial trading resulted in a 3.6% (4.2%) reduction in fuel costs (thermal energy) per MWh. This reduction in fuel costs per MWh implies an annual fuel cost savings in relatively high demand hours of roughly 4.2 million dollars. Moreover, the annual environmental benefit from the reduction in fuel used during high demand hours due to financial trading is 145,000 tons of CO_2. These market efficiency gains do not appear to come with a deterioration in physical market conditions. Specifically, we find no statistical difference before versus after February 1st 2011 in either the number of high demand hours in which at least one thermal unit was required to start up or the per MWh ancillary services costs associated with ensuring that supply meets demand in high demand hours.

Of course, the market efficiency gains from financial trading are directly tied to the financial volumes traded in the market. This trading volume in turn is a function of the transaction costs associated with financial trading. We develop a statistical framework to estimate the implied per-unit trading costs faced by traders in California's wholesale electricity market. These implied trading costs are defined to be the trading cost that just causes a failure to reject the null hypothesis of no profitable trading strategy. Our empirical findings indicate that nodal-level implied trading costs decreased significantly after California introduced purely financial participation in 2/1/2011. We find substantial

heterogeneity across nodes in the distribution of implied trading costs; the 5% (95%) quantile of the distribution of implied trading costs across locations is \$5.29 (\$12.84) per MWh.

Our results thus indicate that both implied trading costs and the physical efficiency benefits from financial trading vary over both time and space. This suggests the scope for extracting additional efficiency benefits from financial trading by setting trading costs that vary by location and hour-of-the-year. To underscore this point, we consider the following back-of-the-envelope calculation in which we slightly increase (decrease) per-unit transaction costs in low (high) demand hours. First, the 25% quantile (75% quantile) of our across node distribution of implied trading costs is 2.47 (4.11) dollars per MWh lower after the introduction of financial trading for generation nodes. The corresponding 25% and 75% quantiles of the reductions in implied trading costs across non-generation nodes are 2.65 and 4.86 dollars per MWh respectively. Our market efficiency estimates imply that the reduction in implied trading costs after 2/1/2011 results in a roughly 1.12 dollar per MWh reduction in fuel costs in relatively high demand hours. Dividing these two estimates and multiplying by 0.05 gives us that a 5 cents per MWh reduction in implied trading costs in high demand hours results in fuel cost savings ranging from 1-2 cents per Mwh in these hours depending on which location is considered.[28]

However, the operator's ability to influence market outcomes using transaction fees is limited by the extent to which the cost of trading is based on implicit factors such as the opportunity cost of traders' time and effort. Indeed, even after 2/1/2011, our average estimated implied trading cost (\$7.18 per MWh) is far higher than the roughly \$0.50-\$0.70 per MWh in transaction fees charged by the market operator. That being said, our back-of-the-envelope calculation suggests that even a 5 cents per MWh reduction in transaction fees in high demand hours can result in economically sizable efficiency benefits. There is thus substantial scope for improvements to electricity market design even without solving the difficult problem of setting transaction fees subject to equilibrium responses by financial participants.

---

[28]As an example, financial trading reduced implied trading costs by 4.11 dollars per MWh at the 75% quantile in the distribution of non-generation nodes. Assume that this 4.11 dollars per MWh reduction in implied trading costs yielded a 1.12 dollar per MWh reduction in fuel costs savings in relatively high demand hours. This back-of-the-envelope calculation implies that a 5 cent reduction in implied trading costs would result in $1.4 = 0.05 \times \frac{1.12}{4.11}$ cents per MWh in fuel cost savings.

References

Bessembinder, Hendrik, and Michael L Lemmon. 2002. "Equilibrium pricing and optimal hedging in electricity forward markets." *the Journal of Finance*, 57(3): 1347–1382.

Bessembinder, Hendrik, and Michael L Lemmon. 2006. "Gains from trade under uncertainty: The case of electric power markets." *The Journal of Business*, 79(4): 1755–1782.

Birge, John, Ali Hortaçsu, Ignacia Mercadal, and Michael Pavlin. 2013. "The role of financial players in electricity markets: An empirical analysis of miso."

Bohn, Roger E, Michael C Caramanis, and Fred C Schweppe. 1984. "Optimal pricing in electrical networks over space and time." *The Rand Journal of Economics*, 360–376.

California Independent System Operator. 2015. "Convergence Bidding: Settlement Implications and Convergence Bidding Fees." *https://www.caiso.com/Documents/ConvergenceBiddingSession7-ConvergenceBiddingFees.pdf* .

California Independent System Operator. 2019. "List of SCs, CRR Holders, Convergence Bidding Entities." *https://www.caiso.com/Documents/ListofSchedulingCoordinatorsCRRHoldersandConvergenceBiddingEntities.pdf* .

Dávila, Eduardo, and Cecilia Parlatore. 2019. "Trading costs and informational efficiency." *National Bureau of Economic Research.*

Fang, Zheng, and Andres Santos. 2018. "Inference on directionally differentiable functions." *The Review of Economic Studies*, 86(1): 377–412.

Federal Energy Regulatory Commission. 2014. "Staff Analysis of Uplift in RTO and ISO Markets." *https://www.ferc.gov/legal/staff-reports/2014/08-13-14-uplift.pdf* .

Frazzini, Andrea, Ronen Israel, and Tobias J Moskowitz. 2018. "Trading costs."

Hadsell, Lester. 2007. "The impact of virtual bidding on price volatility in New York's wholesale electricity market." *Economics Letters*, 95(1): 66–72.

Hirshleifer, David. 1988. "Residual risk, trading costs, and commodity futures risk premia." *The Review of Financial Studies*, 1(2): 173–193.

Hong, Han, and Jessie Li. 2018. "The numerical delta method." *Journal of Econometrics*, 206(2): 379–394.

Isemonger, Alan G. 2006. "The benefits and risks of virtual bidding in multi-settlement markets." *The Electricity Journal*, 19(9): 26–36.

Ito, Koichiro, and Mar Reguant. 2016. "Sequential markets, market power, and arbitrage." *American Economic Review*, 106(7): 1921–57.

Kilian, Lutz, and Daniel P Murphy. 2014. "The Role of Inventories and Speculative Trading in the Global Market for Crude Oil." *Journal of Applied Econometrics*, 29(3): 454–478.

Kilian, Lutz, and Thomas K Lee. 2014. "Quantifying the Speculative Component in the Real Price of Oil: The Role of Global Oil Inventories." *Journal of International Money and Finance*, 42: 71–87.

Knittel, Christopher R, and Robert S Pindyck. 2016. "The Simple Economics of Commodity Price Speculation." *American Economic Journal: Macroeconomics*, 8(2): 85–110.

Kunsch, Hans R, et al. 1989. "The jackknife and the bootstrap for general stationary observations." *The Annals of Statistics*, 17(3): 1217–1241.

Li, Ruoyang, Alva J Svoboda, and Shmuel S Oren. 2015. "Efficiency impact of convergence bidding in the california electricity market." *Journal of Regulatory Economics*, 48(3): 245–284.

Longstaff, Francis A, and Ashley W Wang. 2004. "Electricity forward prices: a high-frequency empirical analysis." *The journal of finance*, 59(4): 1877–1900.

Markowitz, Harry. 1952. "Portfolio selection*." *The journal of finance*, 7(1): 77–91.

Mercadal, Ignacia. 2018. "Dynamic competition and arbitrage in electricity markets: The role of financial players."

Newey, Whitney K, and Kenneth D West. 1987. "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix." *Econometrica: Journal of the Econometric Society*, 703–708.

Ostrovsky, Michael. 2012. "Information aggregation in dynamic markets with strategic traders." *Econometrica*, 80(6): 2595–2647.

Parsons, John E, Cathleen Colbert, Jeremy Larrieu, Taylor Martin, and Erin Mastrangelo. 2015. "Financial arbitrage and efficient dispatch in wholesale electricity markets."

Pindyck, Robert S. 2001. "The Dynamics of Commodity Spot and Futures Markets: A Primer." *The Energy Journal*, 1–29.

Routledge, Bryan R, Chester S Spatt, and Duane J Seppi. 2001. "The" spark spread:" An equilibrium model of cross-commodity price relationships in electricity."

Saravia, Celeste. 2003. "Speculative trading and market performance: the effect of arbitrageurs on efficiency and market power in the New York electricity market." *Center for the Study of Energy Markets*.

Schweppe, Fred C, Michael C Caramanis, Richard D Tabors, and Roger E Bohn. 2013. *Spot pricing of electricity.* Springer Science & Business Media.

Sharpe, William F. 1994. "The sharpe ratio." *Journal of portfolio management,* 21(1): 49–58.

Williams, Jeffrey. 1987. "Futures markets: A consequence of risk aversion or transactions costs?" *Journal of Political Economy,* 95(5): 1000–1023.

Wolak, Frank A. 1989. "Local and Global Testing of Linear and Nonlinear Inequality Constraints in Nonlinear Econometric Models." *Econometric Theory,* 51(1): 1–35.

Wolak, Frank A. 2019. "The Role of Efficient Pricing in Enabling A Low-Carbon Electricity Sector." *http://web.stanford.edu/group/fwolak/cgi-bin/sites/default/files/ efficient_pricing_wolak_march_31_2019.pdf* .

Appendices

List of Appendix Figures

List of Appendix Tables

A   Additional Tables and Figures

Appendix Figure A.1 presents a map of the territories served by each of California's investor-owned utilities.

Appendix Tables A.1 and A.2 list the market participants that are registered to place financial trades in California's wholesale electricity market. Suppliers and demanders that inject or withdraw electricity ("participants that schedule electricity") are listed separately from purely financial players ("participants that don't schedule electricity"). This list suggests that a sizable number of both physical and financial participants place financial bids in this market.

Appendix Figure A.2 shows hourly average day-ahead/real-time spreads, along with their pointwise 95% confidence intervals, for the PG&E, SCE, and SDG&E LAPs before versus after the introduction of explicit virtual bidding (EVB). We can reject that the average day-ahead/real-time price spread is zero for some hours of the day for all three load-serving entities. Importantly, this by itself does not imply that financial traders can profitably arbitrage LAP-level day-ahead/real-time price differences. One must also take into account the potentially sizable transaction costs associated with these price spreads.

Appendix Figure A.3 plots the bootstrap distributions of implied trading costs for the pre-EVB and post-EVB sample periods for each of the three LAPs. The solid vertical lines correspond to the values of c_{lower} and c_{upper} for the pre-EVB sample period and the dotted vertical lines on each graph correspond to the values of c_{lower} and c_{upper} for the post-EVB sample. This figure indicates that both c_{lower} and c_{upper} fell substantially after the introduction of EVB. We consider a formal test of the null hypothesis that c_{lower} and c_{upper} remained the same before versus after EVB in Figure 3.

Figure A.1: California's Investor-Owned Utility Service Areas

Notes: This figure is a map of the territories served by each of California's investor-owned electric utilities. This map is provided by the California Energy Commission; see https://www.ferc.gov/market-oversight/mkt-electric/california.asp.

Figure A.2: Hourly Graphs of Price Differences with 95% C.I: Before and After EVB
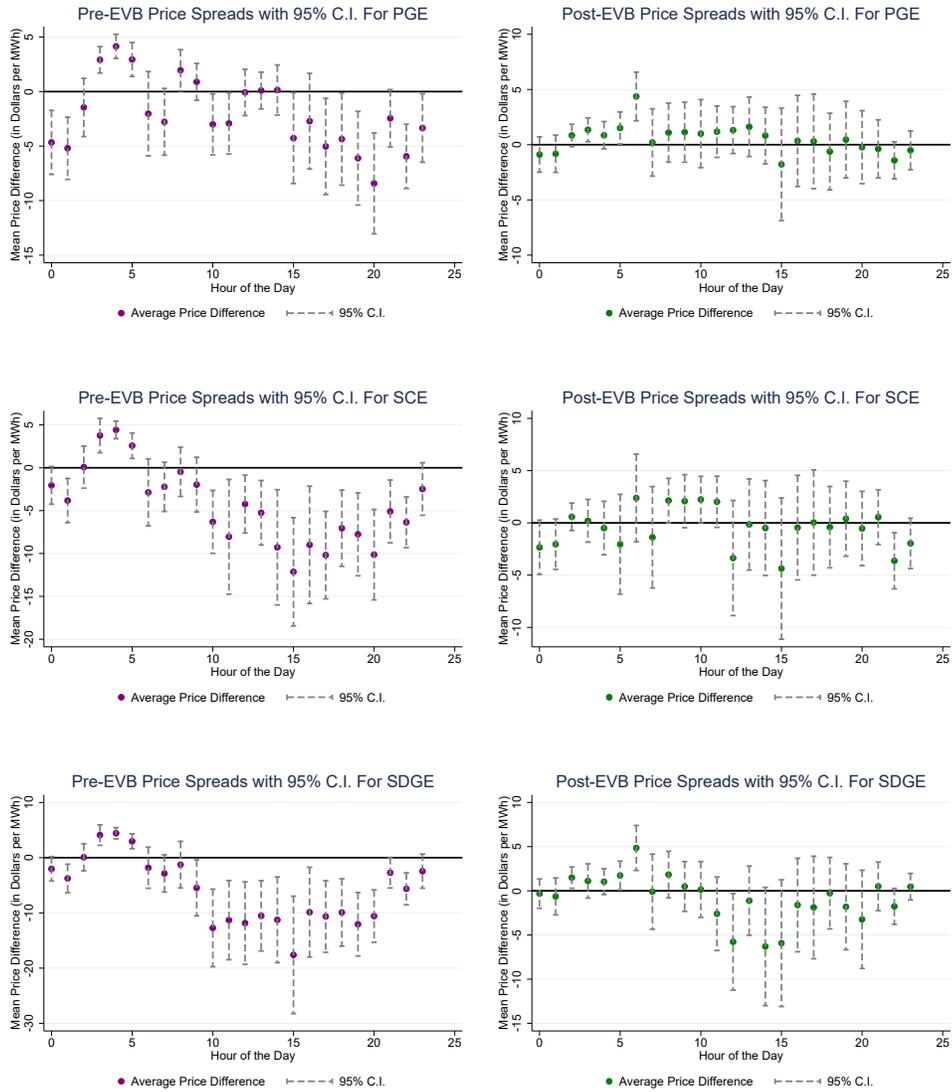
Notes: This figure presents average day-ahead/real-time price spreads for each hour-of-the-day and each load aggregation point (LAP), separately for the sample periods before versus after the introduction of explicit virtual bidding (EVB). The three LAPs correspond to Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E). This figure also includes pointwise 95% confidence intervals around the day-ahead/real-time price spread averages for each hour-of-the-day.

Figure A.3: Bootstrap Distribution of $\phi(\overline{X})$ with 95% C.I: Before and After EVB
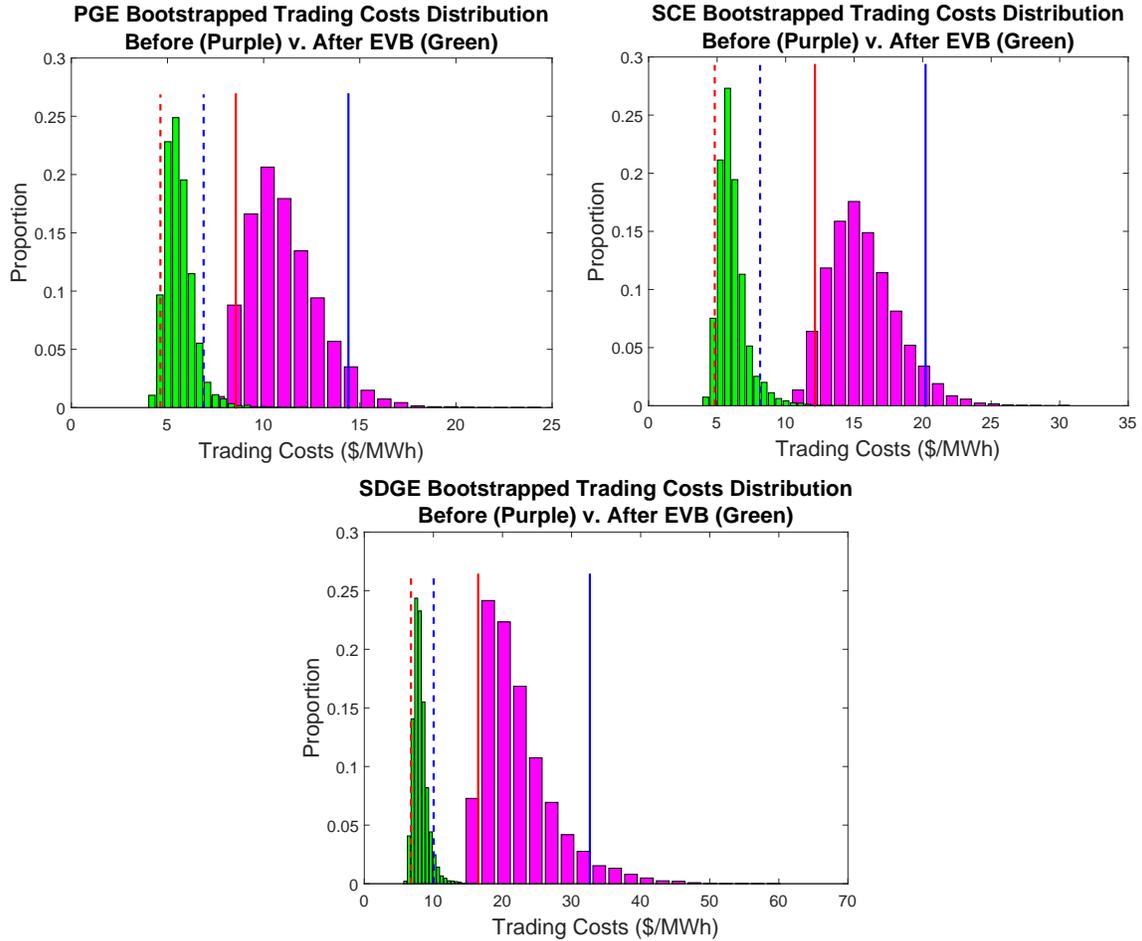
Notes: This figure plots the bootstrap distributions of implied trading costs separately for sample periods before the introduction of explicit virtual bidding (4/1/2009-2/1/2011) versus after the introduction of EVB (2/1/2011-12/31/2012) for each of three load aggregation points (LAPs). These LAPs correspond to territories served by each of California's three major electricity distribution utilities: Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E). The solid vertical lines correspond to the values of c_{lower} and c_{upper} for the pre-EVB sample period while the dotted vertical lines correspond to the values of c_{lower} and c_{upper} for the post-EVB sample. c_{lower} is the smallest value of per-unit trading costs c for which we can reject the null hypothesis that a profitable trading strategy exists while c_{upper} is the largest value of c for which we can reject the null hypothesis that no profitable trading strategy exists.

Table A.1: List of Market Participants Registered to Explicit Virtual Bid: Part (1)

	Participants that Schedule Electricity	Participants that Don't Schedule Electricity
(1)	J. Aron and Company, LLC	Amber Power
(2)	Brookfield Energy Marketing LP	Appian Way Energy Partners West, LLC
(3)	BP Energy Company	ATNV Energy
(4)	Engelhart CTP (US), LLC	Automated Algorithms LLC
(5)	Calpine Energy Services, LP	Bilton Wong Power, Inc.
(6)	California Department of Water Resources	Blackout Power Trading, Inc.
(7)	Just Energy	Calicot Energy LLC
(8)	Citigroup Energy, Inc.	Clear Power LLC
(9)	ConocoPhillips Company	Cumulus Master Fund
(10)	Shell Energy North America (US), L.P.	Darby Energy, LLLP
(11)	CWP Energy, Inc.	Dynasty Energy California Inc.
(12)	DC Energy California, LLC	Dynamis Capital, LLC
(13)	DTE Energy Trading Inc.	Eagle's View Partners, Ltd
(14)	EDF Trading North America, LLC	EDP Renewables North America LLC
(15)	Dynegy Marketing and Trade, LLC	ETC Endure Energy LLC
(16)	CP Energy Marketing (US), Inc.	ETRACOM, LLC
(17)	Boston Energy Trading and Marketing, LLC (formerly Edison Mission Marketing and Trading)	FANTODS, LLC
(18)	Exelon Generation Company, LLC	Freepoint Commodities, LLC
(19)	NextEra Energy Marketing, LLC	Golden Dome LLC (previously Ecesis LLC)
(20)	Guzman Energy, LLC	Gridmatic Inc.
(21)	Rubicon NYP Corp	Heartland Power Inc.
(22)	Castleton Commodities Merchant Trading L.P.	Hemsworth Capital Midwest LP
(23)	Macquarie Energy LLC	High Resolution Energy LLC
(24)	MAG Energy Solutions, Inc.	Hopewell Capital Partners, LP
(25)	Modesto Irrigation District	Inertia Power VII, LLC
(26)	GenOn Energy Management, LLC	LTSTE Investments, LLC

Notes: This table presents the first part of the list of market participants that are registered to place financial bids in California's wholesale electricity market (CAISO (2019a)). This table is split into two columns: the first column lists physical participants who actually inject of withdraw electricity from the grid while the second column lists purely financial participants that don't actually inject or withdraw electricity.

Table A.2: List of Market Participants Registered to Explicit Virtual Bid: Part (2)

	Participants that Schedule Electricity	Participants that Don't Schedule Electricity
(27)	Morgan Stanley Capital Group Inc.	Mercuria Energy America, Inc.
(28)	Constellation NewEnergy, Inc.	MET West Trading, LLC
(29)	NRG California South, LP	Monterey CA, LLC
(30)	NRG Power Marketing LLC	NDC Partners LLC
(31)	Pacific Gas and Electric Company (PG&E - Trading)	NorthStar SW Ltd.
(32)	Portland General Electric Company	Precept Power LLC
(33)	Avangrid Renewables, LLC	Red Wolf CT, LLC
(34)	Public Service Company of Colorado (Xcel Energy)	Saracen Energy West, LP
(35)	Powerex Corp.	Sesco Caliso, LLC
(36)	Royal Bank of Canada	Sirius Power Trading LLC
(37)	Rainbow Energy Marketing Corporation	Solios Power, LLC
(38)	City of Roseville (Roseville Electric)	Tios Capital, LLC
(39)	Southern California Edison Company	Tommy Energy Solutions Corp
(40)	San Diego Gas and Electric Company	Triolith Energy Fund, LP
(41)	Calpine Energy Solutions	TrueLight Energy Fund, LP
(42)	Direct Energy Business, LLC	Trumpet Trading, LLC
(43)	Sempra Gas and Power Marketing, LLC	Tungsten Power LP
(44)	Sacramento Municipal Utility District	Tyne Hill Investments LP
(45)	TransCanada Energy Sales Ltd.	Uncia Energy LP - Series C
(46)	City of Tacoma Department of Public Utilities, Light Division	Velocity American Energy Master I, LP
(47)	The Energy Authority, Inc	XO Energy CAL, LP
(48)	TEC Energy Inc.	Yuma Electric, LLC
(49)	TransAlta Energy Marketing (U.S.) Inc.	
(50)	Tenaska Power Services Co.	
(51)	Valley Electric Association, Inc.	
(52)	Vitol, Inc.	
(53)	Western Area Power Administration Sierra Nevada Region (WAPA)	
(54)	ZGlobal Inc.	

Notes: This table presents the second part of the list of market participants that are registered to place financial bids in California's wholesale electricity market (CAISO (2019a)). This table is split into two columns: the first column lists physical participants who actually inject of withdraw electricity from the grid while the second column lists purely financial participants that don't actually inject or withdraw electricity.

B Difference-in-Differences Robustness Checks

This Appendix section describes the different robustness checks pertaining to the difference-in-differences results presented in Table 5 in Section 6.

B.1 Pre-Trends Analysis

This subsection argues that the results presented in Table 5 are not driven by pre-existing trends in outcomes between high versus low demand hours. Namely, in order to interpret the findings from our difference-in-differences analysis as causal, trends over time in outcomes should not be different across low and high demand outcomes prior to the introduction of explicit virtual bidding (i.e.: financial trading). We test this assumption in two ways. First, we show that the first difference in hourly outcomes prior to explicit virtual bidding (EVB) is the same across high versus low demand hours. The second approach is to estimate an event study framework, demonstrating that: (1) effects estimated separately for each quarter-of-sample are not significantly different from zero prior to EVB, and (2) there is no pre-existing trend in these effects prior to EVB.

B.1.1 Statistical Test on First-Differences

The definition of "common pre-existing trends" is that the slope over time in outcomes is the same across high versus low demand hours. The "slope over time" is simply the first difference in outcomes: $\frac{Y_{t+1}-Y_t}{t+1-t} = Y_{t+1} - Y_t$. Thus, to formally test the "common pre-existing trends" assumption, we estimate the following regression model using only data from hours-of-sample before the introduction of financial trading on February 1st, 2011:

$$Y_t - Y_{t-1} = \alpha_m + \gamma_h + \theta_w + X_t\phi\beta_0\text{HIGH}_t^p + u_t \tag{6}$$

where we include month-of-sample fixed effects (α_m), hour-of-the-day fixed effects (γ_h), and an indicator for whether the day-of-sample is a weekday versus weekend (θ_w). This specification also controls for a host of factors X_t: the log of total electricity demand, two separate controls for the logs of the citygate natural gas prices paid in PG&E and SCE, as well as separate controls for the log of total hourly production from (1) wind and solar sources, (2) nuclear sources, and (3) hydro sources. Finally, we cluster standard errors

Table B.1: Diff-in-Diff Robustness Check: Pre-Trends Analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1(Demand > 90%)	0.001	0.001	-0.007	0.002	0.848	-0.113	-1.173
	(0.001)	(0.001)	(0.009)	(0.009)	(0.855)	(1.823)	(1.480)
R^2	0.031	0.018	0.017	0.111	0.013	0.011	0.009
Average DV in Levels	33.396	8.002	0.677	1.229	13.402	15.66	17.275
Number of Obs.	16,078	16,078	16,078	16,078	16,078	16,078	16,078

Notes: The unit of observation for these regressions is hour-of-sample. Standard errors are clustered by week-of-sample and are reported in parentheses. The dependent variables considered in this table are the first differences of: (1) the log of fuel cost per MWh, (2) the log of input energy per MWh, (3) an indicator of whether any fossil fuel fired units started up, (4) the log of ancillary services per MWh, (5) the absolute value of the day-ahead/real-time prices in PG&E, (6) the absolute value of the day-ahead/real-time prices in SCE, and (7) the absolute value of the day-ahead/real-time prices in SDG&E. The row titled "Avg. DV in Levels" reports the mean of the dependent variable in levels even if the dependent variable is logged when running the regression. For all of the specifications, we control for month-of-sample fixed effects, hour-of-the-day fixed effects, an indicator for whether the day-of-sample is a weekday versus weekend, the log of total electricity demand, two separate controls for the logs of the citygate natural gas prices paid in PG&E and SCE as well as separate controls for the log of total hourly production from (1) wind and solar sources, (2) nuclear sources, and (3) hydro sources.

Table Description: This table presents the pre-trends analysis corresponding to the difference-in-differences results presented in Table 5; in particular, we estimate the specification documented in Equation (6) in Appendix Section B.1. The data used for this table span the sample period 4/1/2009-1/31/2011, noting that California introduced explicit virtual bidding (i.e.: financial trading) on 2/1/2011. Hours-of-sample with system demand greater than the 90% of the distribution of hourly demands across our 4/1/2009-3/31/2012 sample period are deemed to be "high demand" hours.

at the week-of-sample.

B.1.2 Event Study Framework

We also assess the common pre-existing trends assumption using an event study framework. In particular, we estimate the following regression specification:

$$
Y_t = \alpha_m + \gamma_h + \theta_w + X_t \phi \beta_0 \text{HIGH}_t^p
$$
$$
+ \sum_{\tau=-6}^{4} \beta_\tau D_t^\tau + \beta_{\tau<-6} D_t^{\tau<-6} + \beta_{\tau>6} D_t^{>6} + u_t
\tag{7}
$$

where, as before, we include month-of-sample fixed effects (α_m), hour-of-the-day fixed effects (γ_h), an indicator for whether the day-of-sample is a weekday versus weekend (θ_w) as well as a host of other variables X_t: the log of total electricity demand, two separate controls for the logs of the citygate natural gas prices paid in PG&E and SCE,

as well as separate controls for the log of total hourly production from (1) wind and solar sources, (2) nuclear sources, and (3) hydro sources. Standard errors are clustered at the week-of-sample level.

D_t^τ is an indicator function that is equal to one if and only if: (1) hour-of-sample t is a "high demand hour" and (2) the difference between the quarter-of-sample q and the first quarter of 2011 (denoted $f(q)$) is equal to τ (i.e.: $\tau = q - f(q)$). For example, the quarter in which financial trading was introduced (i.e.: the first quarter of 2011) would correspond to $\tau = 0$. $D_t^{\tau < -6}$ ($D_t^{\tau > 6}$) is equal to one if and only if: (1) hour-of-sample t is a "high demand hour", and (2) the difference between quarter of sample q and the first quarter of 2011 is less than -6 (greater than 6). Finally, we normalize to zero the coefficient corresponding to the quarter prior to the introduction of financial trading (i.e.: $\beta_{\tau = -1} = 0$).

The results of this event study analysis are plotted up in two separate figures. Appendix Figure B.1 focuses on the effects for the log of fuel costs per MWh, the log of thermal energy per MWh, whether thermal units started up in the hour, and the log of ancillary services costs per MWh. Appendix Figure B.2 plots the event study coefficients corresponding to the absolute day-ahead/real-time price spreads for each of California's three major electricity distribution companies: Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E). In both cases, the coefficient estimates corresponding to the first quarter of 2011 as well as four quarters prior (i.e.: $\tau = 0$ and $\tau = -4$) are missing simply because there are no high demand hours in those two quarters. Moreover, the coefficient estimates corresponding to the last quarter of 2009 are generally substantially noiser than the other coefficient estimates, reflecting the relatively small number of high demand hours in this quarter. That being said, across all seven panels, we do not see a substantial trend in effect in the quarters prior to the introduction of financial trading. This provides comforting evidence in favor of the common trends assumption required to interpret the results of Table 5 as causal.

B.2   Additional Robustness Checks

This subsection considers our remaining sensitivity analyses pertaining to the difference-in-differences results presented in Table 5 in Section 6. First, Appendix Table B.2 shows

Figure B.1: Event Study Estimates: Market Outcomes
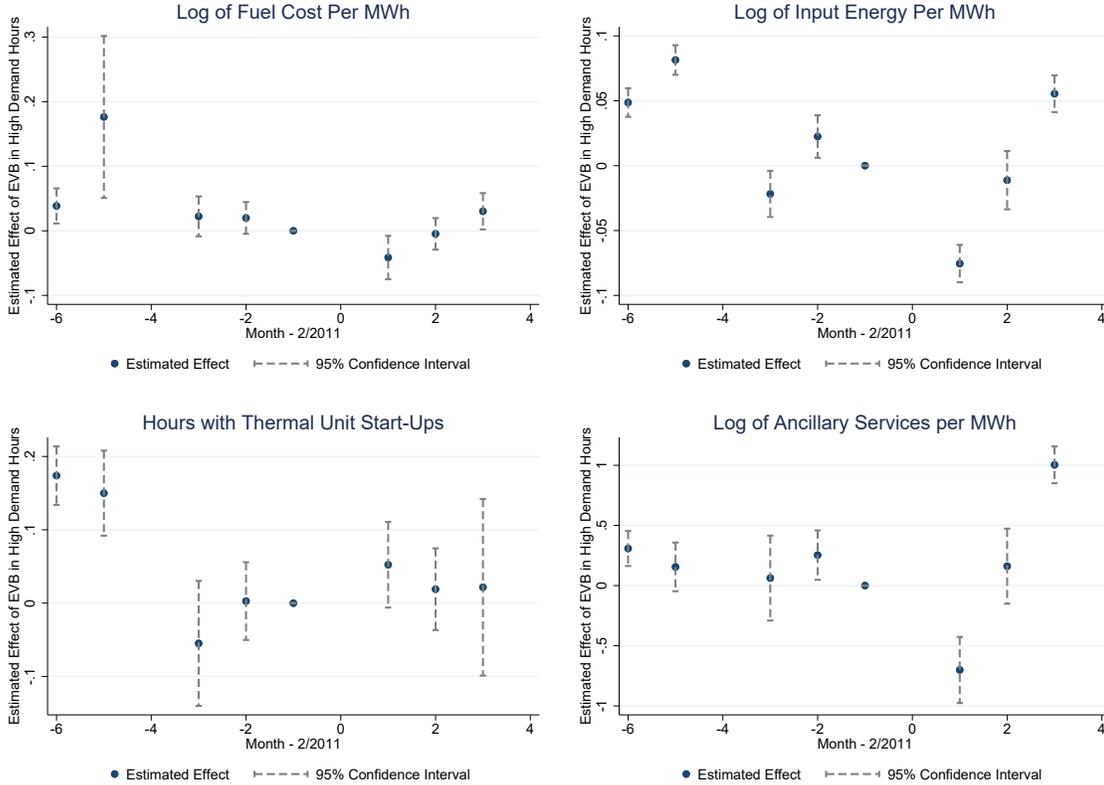
Notes: This figure presents the coefficient estimates and 95% confidence intervals corresponding to the event study framework described in Equation (7). The dependent variables considered in this figure are: (top left panel) the log of fuel cost per MWh, (top right panel) the log of input energy per MWh, (bottom left panel) an indicator of whether any fossil fuel fired units started up, and (bottom right panel) the log of ancillary services per MWh. The coefficient estimates reflect the average outcomes in high relative to low demand hours τ quarters away from the introduction of financial trading on February 1st, 2011; hours-of-sample with system demand greater than the 90% of the distribution of hourly demands across our 4/1/2009-3/31/2012 sample period are deemed to be "high demand" hours. The x-axis simply plots the number of quarters away from the first quarter of 2011 (i.e.: τ). All coefficient estimates are relative to the quarter before the first quarter of 2011 (i.e.: $\beta_{\tau=-1}$ is normalized to be zero). For all of the event study specifications, we control for month-of-sample fixed effects, hour-of-the-day fixed effects, an indicator for whether the day-of-sample is a weekday versus weekend, the log of total electricity demand, two separate controls for the logs of the citygate natural gas prices paid in PG&E and SCE as well as separate controls for the log of total hourly production from (1) wind and solar sources, (2) nuclear sources, and (3) hydro sources.

that our empirical results remain similar if we estimate Equation (5) removing the 28 days before and after 2/1/2011. Thus, our findings do not seem to be due to any short-run adjustments to the policy change.

Next, Appendix Table B.3 presents the results using only data from the 6 months before and after the introduction of financial trading on 2/1/2011. The empirical results are quantitatively quite similar to our primary results from Table 5. Further, Appendix Table B.4 presents our findings with the standard errors calculated using the Newey-West

Table B.2: Diff-in-Diff Robustness Check: Excluding the 28 Days Before and After 2/1/2011

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1(Demand > 90%)	-0.036	-0.042	-0.042	-0.062	-5.833	-14.090	-21.664
× 1(Post EVB)	(0.012)	(0.012)	(0.027)	(0.138)	(3.009)	(5.668)	(9.878)
1(Demand > 90%)	0.030	0.035	0.003	0.142	3.205	6.831	8.395
	(0.005)	(0.005)	(0.022)	(0.054)	(2.089)	(4.345)	(5.765)
R^2	0.906	0.366	0.127	0.610	0.041	0.039	0.050
Avg. DV in Levels	32.485	8.043	0.684	1.731	12.987	14.888	16.264
Number of Obs.	24,909	24,909	24,909	24,908	24,909	24,909	24,909

Notes: The unit of observation for these regressions is hour-of-sample. Standard errors are clustered by week-of-sample and are reported in parentheses. The dependent variables considered in this table are: (1) the log of fuel cost per MWh, (2) the log of input energy per MWh, (3) an indicator of whether any fossil fuel fired units started up, (4) the log of ancillary services per MWh, (5) the absolute value of the day-ahead/real-time prices in PG&E, (6) the absolute value of the day-ahead/real-time prices in SCE, and (7) the absolute value of the day-ahead/real-time prices in SDG&E. The row titled "Avg. DV in Levels" reports the mean of the dependent variable in levels even if the dependent variable is logged when running the regression. For all of the specifications, we control for month-of-sample fixed effects, hour-of-the-day fixed effects, an indicator for whether the day-of-sample is a weekday versus weekend, the log of total electricity demand, two separate controls for the logs of the citygate natural gas prices paid in PG&E and SCE as well as separate controls for the log of total hourly production from (1) wind and solar sources, (2) nuclear sources, and (3) hydro sources.

Table Description: This table presents the difference-in-differences results pertaining to the effect of explicit virtual bidding (i.e.: financial trading) on market outcomes in high demand hours relative to low demand hours. The data used for this table span the sample period 4/1/2009-3/31/2012; in contrast with Table 5, we estimate Equation (5) removing the 28 days before and after 2/1/2011. California introduced explicit virtual bidding (EVB) on 2/1/2011; the "Post EVB" indicator is thus one if and only if the day-of-sample is on or after 2/1/2011. Hours-of-sample with system demand greater than the 90% of the distribution of hourly demands across our 4/1/2009-3/31/2012 sample period are deemed to be "high demand" hours.

Figure B.2: Event Study Estimates: Absolute Price Spreads
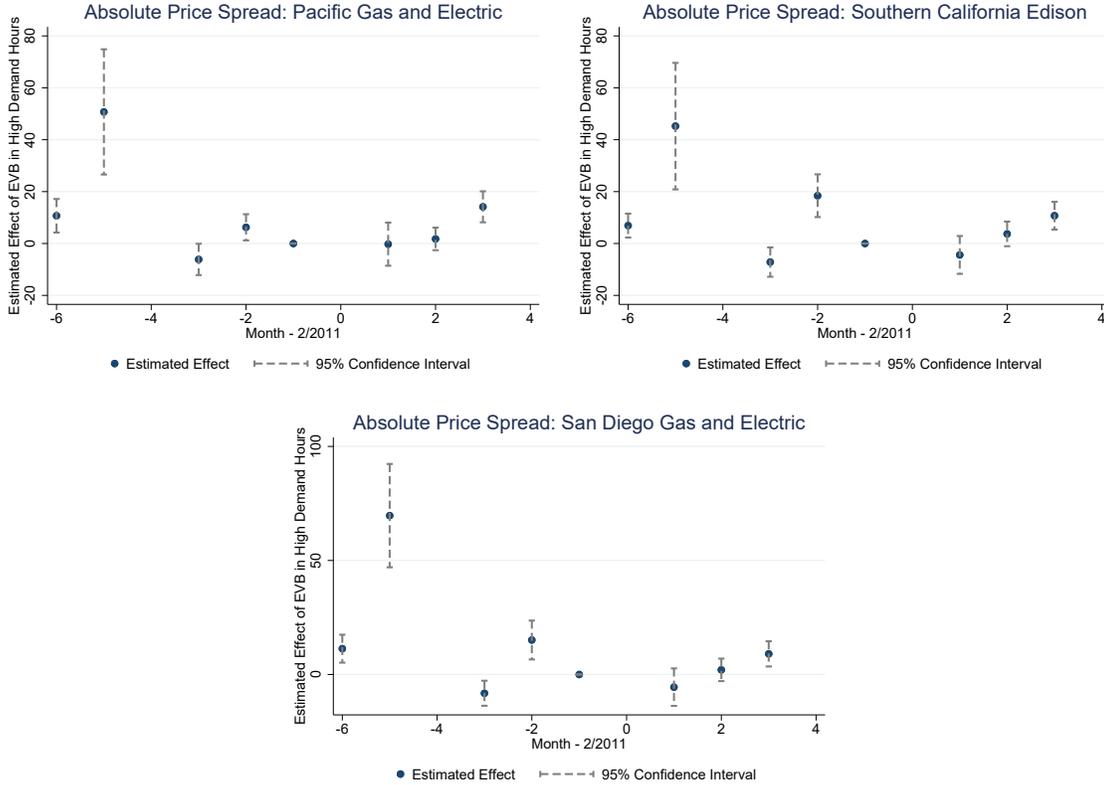
Notes: This figure presents the coefficient estimates and 95% confidence intervals corresponding to the event study framework described in Equation (7). The dependent variables considered in this figure are the absolute differences between day-ahead and real-time prices corresponding to each of California's three major investor-owned utilities: Pacific Gas and Electric (top left panel), Southern California Edison (top right panel), and San Diego Gas and Electric (bottom middle panel). The coefficient estimates reflect the average outcomes in high relative to low demand hours τ quarters away from the introduction of financial trading on February 1st, 2011; hours-of-sample with system demand greater than the 90% of the distribution of hourly demands across our 4/1/2009-3/31/2012 sample period are deemed to be "high demand" hours. The x-axis simply plots the number of quarters away from the first quarter of 2011 (i.e.: τ). All coefficient estimates are relative to the quarter before the first quarter of 2011 (i.e.: $\beta_{\tau=-1}$ is normalized to be zero). For all of the event study specifications, we control for month-of-sample fixed effects, hour-of-the-day fixed effects, an indicator for whether the day-of-sample is a weekday versus weekend, the log of total electricity demand, two separate controls for the logs of the citygate natural gas prices paid in PG&E and SCE as well as separate controls for the log of total hourly production from (1) wind and solar sources, (2) nuclear sources, and (3) hydro sources.

formula (Newey and West, 1987) accounting for 168 hours (i.e.: one week) of autocorrelation in the outcome variables. Our conclusions remain the same when analyzing Appendix Table B.4: financial trading improves market outcomes such as fuel costs per MWh and absolute price spreads without exacerbating market constraints such as unit start ups or ancillary services.

Appendix Table B.5 presents our specifications estimated on data aggregated to the

Table B.3: Diff-in-Diff Robustness Check: 6 Months Before and After 2/1/2011

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1(Demand > 90%)	-0.056	-0.062	-0.006	-0.463	-9.740	-14.367	-12.471
× 1(Post EVB)	(0.016)	(0.016)	(0.030)	(0.167)	(4.879)	(6.300)	(6.105)
1(Demand > 90%)	0.062	0.064	-0.032	0.303	0.092	2.245	-1.367
	(0.011)	(0.011)	(0.032)	(0.099)	(4.347)	(5.681)	(5.968)
R^2	0.675	0.375	0.120	0.665	0.038	0.040	0.038
Avg. DV in Levels	34.635	8.226	0.706	2.588	15.298	17.023	18.27
Number of Obs.	9,503	9,503	9,503	9,503	9,503	9,503	9,503

Notes: The unit of observation for these regressions is hour-of-sample. Standard errors are clustered by week-of-sample and are reported in parentheses. The dependent variables considered in this table are: (1) the log of fuel cost per MWh, (2) the log of input energy per MWh, (3) an indicator of whether any fossil fuel fired units started up, (4) the log of ancillary services per MWh, (5) the absolute value of the day-ahead/real-time prices in PG&E, (6) the absolute value of the day-ahead/real-time prices in SCE, and (7) the absolute value of the day-ahead/real-time prices in SDG&E. The row titled "Avg. DV in Levels" reports the mean of the dependent variable in levels even if the dependent variable is logged when running the regression. For all of the specifications, we control for month-of-sample fixed effects, hour-of-the-day fixed effects, an indicator for whether the day-of-sample is a weekday versus weekend, the log of total electricity demand, two separate controls for the logs of the citygate natural gas prices paid in PG&E and SCE as well as separate controls for the log of total hourly production from (1) wind and solar sources, (2) nuclear sources, and (3) hydro sources.

Table Description: This table presents the difference-in-differences results pertaining to the effect of explicit virtual bidding (i.e.: financial trading) on market outcomes in high demand hours relative to low demand hours. In contrast with Table 5, the data used for this table span the sample period 2/1/2010-8/31/2011 (i.e.: the 12 months before and after 2/2011). California introduced explicit virtual bidding (EVB) on 2/1/2011; the "Post EVB" indicator is thus one if and only if the day-of-sample is on or after 2/1/2011. Hours-of-sample with system demand greater than the 90% of the distribution of hourly demands across our 4/1/2009-3/31/2012 sample period are deemed to be "high demand" hours.

Table B.4: Diff-in-Diff Robustness Check: Accounting for 168 hours of Autocorrelation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1(Demand > 90%)	-0.036	-0.042	-0.042	-0.071	-5.715	-13.902	-21.568
× 1(Post EVB)	(0.012)	(0.012)	(0.028)	(0.138)	(3.029)	(5.631)	(9.505)
1(Demand > 90%)	0.032	0.036	0.008	0.123	2.496	6.231	7.882
	(0.005)	(0.005)	(0.023)	(0.054)	(2.126)	(4.418)	(5.756)
R^2	0.903	0.362	0.125	0.590	0.041	0.039	0.049
Avg. DV in Levels	32.63	8.048	0.686	1.704	13.199	15.055	16.426
Number of Obs.	26,277	26,277	26,277	26,276	26,277	26,277	26,277

Notes: The unit of observation for these regressions is hour-of-sample. Standard errors are calculated using the Newey-West estimator (Newey and West, 1987) accounting for 168 hours (i.e.: one week) of autocorrelation and are reported in parentheses. The dependent variables considered in this table are: (1) the log of fuel cost per MWh, (2) the log of input energy per MWh, (3) an indicator of whether any fossil fuel fired units started up, (4) the log of ancillary services per MWh, (5) the absolute value of the day-ahead/real-time prices in PG&E, (6) the absolute value of the day-ahead/real-time prices in SCE, and (7) the absolute value of the day-ahead/real-time prices in SDG&E. The row titled "Avg. DV in Levels" reports the mean of the dependent variable in levels even if the dependent variable is logged when running the regression. For all of the specifications, we control for month-of-sample fixed effects, hour-of-the-day fixed effects, an indicator for whether the day-of-sample is a weekday versus weekend, the log of total electricity demand, two separate controls for the logs of the citygate natural gas prices paid in PG&E and SCE as well as separate controls for the log of total hourly production from (1) wind and solar sources, (2) nuclear sources, and (3) hydro sources.

Table Description: This table presents the difference-in-differences results pertaining to the effect of explicit virtual bidding (i.e.: financial trading) on market outcomes in high demand hours relative to low demand hours. The data used for this table span the sample period 4/1/2009-3/31/2012. California introduced explicit virtual bidding (EVB) on 2/1/2011; the "Post EVB" indicator is thus one if and only if the day-of-sample is on or after 2/1/2011. Hours-of-sample with system demand greater than the 90% of the distribution of hourly demands across our 4/1/2009-3/31/2012 sample period are deemed to be "high demand" hours.

daily level. In particular, we take the daily sum over hours-of-the-day of electricity demand, fuel costs, thermal input energy, electricity production by type (thermal, nuclear, hydro, and renewables), number of thermal units started up and ancillary services costs; daily total fuel costs, thermal input energy and ancillary services costs are the divided by daily total electricity production from fossil-fuel-fired units in order to construct the outcome variables considered in Appendix Table B.5. In contrast with Table 5, we consider the log of the total number of thermal units that started up in the day rather than an indicator of whether any thermal units started up in the hour. This is because at least one fossil fuel fired unit started up in every day-of-sample. Finally, the hourly day-ahead and real-time electricity prices for each service area are averaged to the daily-level; we take the absolute difference between these day-ahead and real-time prices for our last three outcome measures.

Our difference-in-differences specification thus considers days-of-sample above versus below the 90% of the distribution of daily total electricity demand before versus after the introduction of financial trading. The results of this analysis, presented in Appendix Table B.5, are broadly consistent with those presented in Table 5 with the exception of the absolute day-ahead/real-time price spreads. This in unsurprising given that day-ahead prices are far more likely to deviate substantially from real-time prices during certain hours of the day; financial trading is especially beneficial in reducing price spreads in those hours.

Finally, Appendix Table B.6 shows that our estimated effect of financial trading on average fuel costs per MWh in relatively high demand hours remains quantitatively similar if "high demand" is defined to hours-of-sample above the 50, 75, 95, and 99 percentiles of hourly demand rather than the $p = 90^{th}$ percentile as considered in our primary specifications. We use this effect in order to calculate the efficiency benefits from slightly lowering the per-unit transaction costs in relatively high demand hours in the next section. Thus, it is comforting that our physical efficiency benefit per MWh of fossil-fuel-fired electricity production doesn't change significantly based on our definition of "relatively high demand".

Table B.5: Diff-in-Diff Robustness Check: Daily-Level

	Log of Average Fuel Cost Per MWh						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1(Demand > 90%)	-0.022	-0.030	-0.176	-0.276	-2.810	-2.451	-2.293
× 1(Post EVB)	(0.011)	(0.012)	(0.055)	(0.119)	(2.543)	(2.743)	(2.997)
1(Demand > 90%)	0.037	0.043	0.131	0.250	3.239	-1.022	-4.640
	(0.006)	(0.005)	(0.044)	(0.100)	(2.243)	(2.771)	(5.444)
R^2	0.983	0.750	0.544	0.841	0.124	0.103	0.173
Avg. DV in Levels	32.444	8.003	45.275	1.685	8.353	9.725	10.815
Number of Obs.	1,095	1,095	1,095	1,095	1,095	1,095	1,095

Notes: The unit of observation for these regressions is day-of-sample. Standard errors are clustered by week-of-sample and are reported in parentheses. The dependent variables considered in this table are: (1) the log of fuel cost per MWh, (2) the log of input energy per MWh, (3) an indicator of whether any fossil fuel fired units started up, (4) the log of ancillary services per MWh, (5) the absolute value of the day-ahead/real-time prices in PG&E, (6) the absolute value of the day-ahead/real-time prices in SCE, and (7) the absolute value of the day-ahead/real-time prices in SDG&E. The row titled "Avg. DV in Levels" reports the mean of the dependent variable in levels even if the dependent variable is logged when running the regression. For all of the specifications, we control for month-of-sample fixed effects, an indicator for whether the day-of-sample is a weekday versus weekend, the log of total electricity demand, two separate controls for the logs of the citygate natural gas prices paid in PG&E and SCE as well as separate controls for the log of total hourly production from (1) wind and solar sources, (2) nuclear sources, and (3) hydro sources.

Table Description: This table presents the difference-in-differences results pertaining to the effect of explicit virtual bidding (i.e.: financial trading) on market outcomes in high demand days relative to low demand days. The data used for this table span the sample period 4/1/2009-3/31/2012. California introduced explicit virtual bidding (EVB) on 2/1/2011; the "Post EVB" indicator is thus one if and only if the day-of-sample is on or after 2/1/2011. Days-of-sample with system demand greater than the 90% of the distribution of daily demands across our 4/1/2009-3/31/2012 sample period are deemed to be "high demand" days.

Table B.6: Diff-in-Diff Robustness Check: By Percentage of Demand

| | Log of Average Fuel Cost Per MWh | | | | |
	(1)	(2)	(3)	(4)	(5)
1(Demand > Cut-off)	-0.011	-0.025	-0.036	-0.044	-0.022
× 1(Post EVB)	(0.006)	(0.010)	(0.012)	(0.013)	(0.018)
1(Demand > Cut-off)	-0.005	0.001	0.032	0.062	0.077
	(0.003)	(0.004)	(0.005)	(0.006)	(0.009)
Demand Cut-Off	50%	75%	90%	95%	99%
R^2	0.903	0.903	0.903	0.904	0.904
Avg. DV in Levels	32.63	32.63	32.63	32.63	32.63
Number of Obs.	26,277	26,277	26,277	26,277	26,277

Notes: The unit of observation for these regressions is hour-of-sample. Standard errors are clustered by week-of-sample and are reported in parentheses. The row titled "Avg. DV in Levels" reports the mean of fuel costs per MWh in levels even though fuel costs per MWh are logged when running the regression. For all of the specifications, we control for month-of-sample fixed effects, hour-of-the-day fixed effects, an indicator for whether the day-of-sample is a weekday versus weekend, the log of total electricity demand, two separate controls for the logs of the citygate natural gas prices paid in PG&E and SCE as well as separate controls for the log of total hourly production from (1) wind and solar sources, (2) nuclear sources, and (3) hydro sources.

Table Description: This table presents the difference-in-differences results pertaining to the effect of explicit virtual bidding (i.e.: financial trading) on fuel costs per MWh in high demand hours relative to low demand hours. The data used for this table span the sample period 4/1/2009-3/31/2012. California introduced explicit virtual bidding (EVB) on 2/1/2011; the "Post EVB" indicator is thus one if and only if the day-of-sample is on or after 2/1/2011. Hours-of-sample with system demand greater than the 50%, 75%, 90%, 95%, and 99% of the distribution of hourly demands across our 4/1/2009-3/31/2012 sample period are deemed to be "high demand" hours for the specification estimated in Columns 1, 2, 3, 4, and 5 of this table respectively.

C   Additional Empirical Results

C.1   Is there autocorrelation in daily price differences beyond the first lag?

Our statistical test of arbitrage opportunities accounting for trading costs only considers trading strategies that vary by hour-of-the-day. We do not allow for trading strategies to be updated based on information from past days. However, trading strategies in practice cannot be a function of information from the prior day because the values of the 24×1 vector of real-time prices for day $d-1$ are not known before offers are submitted to the day-ahead market for day d. Thus, any trading strategies involving portfolios of the 24×1 vector of day-ahead/real-time price differences can only condition on realized day-ahead/real-time price differences from $k \geq 2$ days ago ($\mathbf{X_{d-k}}$). Our analysis is restricted to trading strategies that do not condition on any past values of $\mathbf{X_{d-k}}$; this restriction is valid if the autocorrelation matrices for the $\mathbf{X_d}$ process are zero beyond the first lag, noting again that traders cannot utilize the realization $\mathbf{X_{d-1}}$ in setting their trading strategy for day d.

We formulate a statistical test of the null hypothesis that the first R autocorrelation matrices associated with $\mathbf{X_d}$ beyond the first lag are equal to zero. In particular, consider the τth auto-correlation matrix (which is 24×24): $\Gamma(\tau) = E[(X_t - \mu)(X_{t-\tau} - \mu)']$. Consistent with our above discussion, we expect $\Gamma(1)$ to be non-zero but restrict $\Gamma(\tau) = 0$ for all $\tau > 1$. Thus, we consider the Null hypothesis:

$$H : \Gamma(2) = 0, \Gamma(3) = 0, ..., \Gamma(R) = 0$$

for a fixed value of R. Empirically, we test using $R = 10$. This hypothesis test is implemented by first defining $\xi \equiv [vec(\Gamma(2))', vec(\Gamma(3))', ..., vec(\Gamma(L))']'$, where the $vec(.)$ operator takes each 24×24 auto-correlation matrix and stacks it column-wise to create a 576×1 vector. Therefore, ξ has $5760 = 576 \times 10$ elements, all of which must equal zero under the null hypothesis. We use the moving block bootstrap described more fully in the previous subsection in order to estimate the 5760×5760 covariance matrix associated with $\hat{\xi}$. Our Wald statistic $TS = \hat{\xi}' \hat{\Sigma}_{\xi,boot}^{-1} \hat{\xi}$ is asymptotically chi-squared distributed with $24^2 \times (R-1)$ degrees of freedom under the null hypothesis. We perform

Table C.1: Test Statistics for Autocorrelation ($1 < L \leq 10$) in Daily Price Differences

	Before EVB	After EVB
PG&E	2862.2	2767.0
SCE	2789.2	2842.6
SDG&E	3082.1	2700.7

Notes: This table presents the chi-squared test statistics for each load aggregation point (LAP) before and after the introduction of explicit virtual bidding (EVB) corresponding to the null hypothesis that the second through tenth autocorrelation matrices of the 24×1 vector of daily day-ahead/real-time price differences are zero; in math, we are testing the null hypothesis that $\Gamma(2) = \Gamma(3) = ... = \Gamma(10) = 0$. We consider the three LAPs corresponding to California's three major load-serving entities (read: electricity demanders): Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E). To conduct this statistical test, we first estimate $(\Gamma(2), \Gamma(3), ..., \Gamma(10))$, which are each 24×24, pre-EVB and post-EVB for each LAP and stack the elements column-wise; this results in a $5760 = 24^2 \times 9$ element vector. We use a moving block bootstrap procedure in order to estimate the covariance matrix associated with this 5760×1 vector. The upper $\alpha = 0.05$ critical value for these test statistics is $\chi^2_{5184} = 5352.6$.

this test separately for the day-ahead/real-time price differences associated with each location for sample periods before versus after the implementation of explicit virtual bidding.

Our statistical test of arbitrage is based on calculating the most profitable strategy to exploit expected differences between day-ahead and real-time prices that conditions only on hour-of-day. In order to provide statistical evidence that more complex trading strategies do not generate significantly higher revenues, we test the null hypothesis that the second through tenth autocorrelation matrices associated with the 24×1 vector of daily day-ahead/real-time price spreads are zero: $\Gamma(2) = \Gamma(3) = ... = \Gamma(10) = 0$. We do not include the first autocorrelation matrix $\Gamma(1)$ in this test because market participants are required to submit offers into the day-ahead market cleared for day d (for energy to be delivered in day $d + 1$) *prior* to the clearing of the real-time market for day d; thus, traders cannot utilize information on realized day-ahead/real-time price spreads for day d when forming trading strategies for day $d + 1$. Our statistical test for autocorrelation is conducted separately for each of three load aggregation points (LAPs) both before and after the introduction of explicit virtual bidding. We consider the three LAPs corresponding to California's three major load-serving entities (i.e.: electricity demanders): Pacific Gas and Electric (PG&E), Southern California Edison (SCE) and San Diego Gas and Electric (SDG&E). Test statistics are reported in Table C.1; the upper $\alpha = 0.05$ critical value for these test statistics is $\chi^2(5184) = 5352.6$.

63

Table C.2: Proportion of Autocorrelation Tests that Fail to Reject ($\alpha = 0.05$): Number of Locations in Brackets

	Before EVB	After EVB
Non-Generation Node	0.299	0.912
	[4,031]	[4,386]
Generation Node	0.265	0.932
	[669]	[673]

Notes: This table presents the proportion of nodes (read: pricing locations) for which we fail to reject a size $\alpha = 0.05$ test of the null hypothesis that the second through tenth autocorrelation matrices of the 24×1 vector of daily day-ahead/real-time price spreads are zero; in math, we are testing the null hypothesis that $\Gamma(2) = \Gamma(3) = ... = \Gamma(10) = 0$. We perform this hypothesis test separately for each node before and after the introduction of explicit virtual bidding (EVB). We estimate covariance matrices $(\Gamma(2), \Gamma(3), ..., \Gamma(10))$, which are each 24×24, pre-EVB and post-EVB for each node and stack the elements column-wise; this results in a $5760 = 24^2 \times 9$ element vector. We use a moving block bootstrap procedure in order to estimate the covariance matrix associated with this 5760×1 vector. The upper $\alpha = 0.05$ critical value for these test statistics is $\chi^2(5184) = 5352.6$. Generators inject electricity at some nodes; these nodes are called "generation nodes". The remaining nodes are categorized as "non-generation" nodes. The number of nodes in each cell is reported in brackets.

We fail to reject the null hypothesis that the second through tenth autocorrelation matrices are zero for any LAP at the 5% level, either before or after the introduction of explicit virtual bidding. This lends strong evidence in favor of our assertion that daily price differences have zero autocorrelation matrices past the first lag. We repeat these same autocorrelation tests at the nodal level in Table C.2, finding that the null hypothesis that $\Gamma(2) = \Gamma(3) = ... = \Gamma(10) = 0$ can be rejected at a 5% level at approximately 70 percent of nodes before the introduction of explicit virtual bidding (EVB); these rejections occur less frequently at locations where electricity generation is injected (i.e: "generation nodes"), which is consistent with the fact that suppliers could implicitly virtual bid only at the nodes where their generators were located. However, we reject the same null hypothesis at only approximately 7-9% percent of both generation and non-generation nodes after the implementation of explicit virtual bidding; this is consistent with the intuition that financial traders would take advantage of any systematic auto-correlation between day-ahead/real-time price spreads post-EVB. Summarizing, as traders cannot condition on the previous day's real-time price realizations when submitting into the day-ahead market, the results from this subsection help to justify our focus on trading strategies that do not condition on past lags of the daily price difference vector.

C.2 Absolute Average Price Spreads Before Versus After Financial Trading

This subsection describes our statistical test of whether expected day-ahead/real-time price spreads decrease in absolute value after the introduction of explicit virtual bidding (EVB). In particular, we formulate the statistical test of the null hypothesis that $|\mu_{pre}| > |\mu_{post}|$, where $|\mu_J|$ for $J \in \{pre,\ post\}$ is a 24×1 vector composed of the absolute values of the average day-ahead/real-time price differences for each hour-of-the-day $h \in \{1, 2, ..., 24\}$ for a given location computed using the sample period before (after) explicit virtual bidding for $J = $ pre ($J = $ post). We implement this statistical test separately for each pricing location. This statistical framework can also be used to test whether average day-ahead/real-time price spreads *increase* after the introduction of explicit virtual bidding. Failing to reject the null hypothesis that $|\mu_{pre}| > |\mu_{post}|$ but rejecting the null hypothesis that $|\mu_{post}| > |\mu_{pre}|$ gives us statistical evidence that arbitrage opportunities were present in the post-EVB sample period only if they were also present in the pre-EVB sample period *for a given transaction cost c.*

These two multivariate nonlinear inequality constraints tests are implemented using the methodology derived in Wolak (1989). In particular, we compute the following test statistic in order to test the null hypothesis that $|\mu_{pre}| > |\mu_{post}|$:

$$TS = \frac{min}{\theta \geq 0}(|\overline{X}^{pre}| - |\overline{X}^{post}| - \theta)'\hat{V}^{-1}(|\overline{X}^{pre}| - |\overline{X}^{post}| - \theta)$$

where all objects with a "pre" ("post") superscript are based on the 4/1/2009-2/1/2011 (2/1/2011-12/31/2012) period before (after) the introduction of explicit virtual bidding (EVB). $\overline{X}^{pre}$ ($\overline{X}^{pre}$) is a 24×1 vector of the average day-ahead/real-time price differences for each hour-of-the-day for the pre-EVB (post-EVB) sample. We calculate the covariance matrix $\hat{V}$ as follows:

$$\hat{V} = \frac{1}{N^{pre}}diag[SIGN(\overline{X}^{pre})]'\hat{\Sigma}^{pre}diag[SIGN(\overline{X}^{pre})]$$

$$+\frac{1}{N^{post}}diag[SIGN(\overline{X}^{post})]'\hat{\Sigma}^{post}diag[SIGN(\overline{X}^{post})]$$

where the $diag[Z]$ operator takes a vector Z and returns a diagonal matrix with the elements of Z on the diagonal. N^{pre} (N^{post}) is the number of days in the sample period

Table C.3: LAP-level P-values for the Absolute Difference Tests

	$|\mu_{pre}| > |\mu_{post}|$	$|\mu_{post}| > |\mu_{pre}|$
PG&E	0.704	0.141
SCE	0.907	0.006
SDG&E	0.689	0.040

Notes: This table reports the p-value for each load aggregation point (LAP) associated with the statistical test of the null hypothesis that $|\mu_{pre}| > |\mu_{post}|$ (Column 1) and the null hypothesis that $|\mu_{post}| > |\mu_{pre}|$ (Column 2). μ_{pre} (μ_{post}) is a 24×1 vector composed of the average day-ahead/real-time price spreads for a given LAP for each hour-of-the-day for the sample period before (after) the introduction of explicit virtual bidding (EVB); the pre-EVB sample period is 4/1/2009-2/1/2011 while the post-EVB sample is 2/1/2011-12/31/2012. We consider the three LAPs corresponding to California's three major load-serving entities (read: electricity demanders): Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E).

before (after) the introduction of explicit virtual bidding. $\hat{\Sigma}^{pre}$ ($\hat{\Sigma}^{post}$) is a 24×24 estimate of the asymptotic covariance matrix corresponding to $\overline{X}^{pre}$ ($\overline{X}^{post}$); we compute these covariance matrices using the autocorrelation consistent estimator proposed by Newey and West (1987) with $m = 14$ days of lagged data. We reject the null hypothesis that $|\mu_{pre}| > |\mu_{post}|$ if and only if

$$\sum_{h=1}^{24} w(24, 24 - h, \hat{V}) Pr[\chi^2_{(h)} > TS] < \alpha$$

where $\chi^2_{(h)}$ is a chi-squared random variable with h degrees of freedom, $w(24, 24 - h, \hat{V})$ are the weights defined in Wolak (1989), and α is the size of the hypothesis test.

The p-values corresponding to these statistical tests for each load aggregation point (LAP) are presented below in Table C.3. From this table, we see that we cannot reject the null hypothesis that $|\mu_{pre}| > |\mu_{post}|$ for any of the three LAPs, while we can reject the null hypothesis that $|\mu_{post}| > |\mu_{pre}|$ at the 5% level for SCE and SDG&E.

We also run the statistical test of the null hypothesis that $|\mu_{pre}| > |\mu_{post}|$ separately for each node; Column 1 of Table C.4 provides the proportion of nodes for which we fail to reject this null hypothesis, separately for generation versus non-generation nodes. We see from this table that we fail to reject the Null hypothesis that average hourly absolute differences are higher pre-EVB relative to post-EVB for roughly 98% of nodes; there is not a sizable difference in the proportion of nodes for which we fail to reject the null hypothesis for generation versus non-generation nodes. Column 2 of Table C.4 presents the proportion of nodes for which we fail to reject the null hypothesis that $|\mu_{post}| > |\mu_{pre}|$.

Table C.4: Proportion of Nodes for which we fail to reject the $\alpha = 0.05$ sized Absolute Difference Test

	$\lvert\mu_{pre}\rvert > \lvert\mu_{post}\rvert$	$\lvert\mu_{post}\rvert > \lvert\mu_{pre}\rvert$
Gen Node	0.983	0.015
Non-Gen Node	0.988	0.013

Notes: This table reports the proportion of nodes (read: pricing locations) for which we fail to reject the null hypothesis that $\lvert\mu_{pre}\rvert > \lvert\mu_{post}\rvert$ (Column 1) and the null hypothesis that $\lvert\mu_{post}\rvert > \lvert\mu_{pre}\rvert$ (Column 2). Note that μ_{pre} (μ_{post}) is a 24×1 vector composed of the average day-ahead/real-time price differences for a given node for each hour-of-the-day for the sample period before (after) the introduction of explicit virtual bidding. 653 generation nodes and 3,961 non-generation nodes are in both the pre-EVB sample (4/1/2009-2/1/2011) and post-EVB sample (2/1/2011-12/31/2012).

We fail to reject this null hypothesis for only roughly 1.5% of nodes; as before, our results are similar across generation versus non-generation nodes. Combining these two findings, we have strong evidence that nodal-level absolute average day-ahead/real-time price spreads fell after the introduction of financial trading. Interpreting this result in another way, if market participants could trade the financial instrument at a fixed value of trading costs c both before and after EVB, then traders would make profits from exploiting expected day-ahead/real-time price spreads post-EVB only if they would also make profits from exploiting expected day-ahead/real-time price spreads pre-EVB.

C.3 Volatility in Prices Before versus After Financial Trading

As described in Section 2, we also expect financial participation in wholesale electricity markets to reduce day-ahead uncertainty regarding real time prices. We therefore expect both the variance of day-ahead/real-time price differences and the variance of real-time prices to fall after the introduction of explicit virtual bidding (EVB). To test this hypothesis, let Λ^{pre} (Λ^{post}) be the 24×24 contemporaneous covariance matrix corresponding to the 24×1 random vector of hourly realized day-ahead/real-time price spreads for the time period before (after) explicit virtual bidding. The variance of day-ahead/real-time price differences is larger before versus after the introduction of EVB if and only if $\Lambda^{pre} - \Lambda^{post}$ is a positive semi-definite matrix. We construct a statistical test of this null hypothesis by finding the eigenvalues $\hat{\omega}_j$ ($j = 1, 2, ..., 24$) of $\hat{\Lambda}^{diff} \equiv \hat{\Lambda}^{pre} - \hat{\Lambda}^{post}$; we test the joint null hypothesis that all of these eigenvalues are greater than or equal to zero using the same multivariate inequality constraints test employed in the previous subsection.

Our test statistic is:

$$TS \;=\; \min_{z \,\geq\, 0} (\hat{\Lambda}_{diff} \;-\; z)'[Var(\hat{\Lambda}_{diff})]^{-1}(\hat{\Lambda}_{diff} \;-\; z)$$

where the covariance matrix $Var(\hat{\Lambda}_{diff})$ is estimated using a moving-block bootstrap procedure.[29] As described in the previous subsection, the test statistic TS is asymptotically distributed as the weighted sum of chi-squared random variables under the null hypothesis. We also perform this test for the null hypothesis H_2 that $\Lambda^{post} - \Lambda^{pre}$ is a positive semi-definite matrix, which is equivalent to testing the null hypothesis that day-ahead/real-time price spreads are more volatile after EVB relative to before EVB. Failing to reject H_1 but rejecting H_2 would provide us with statistical evidence that the introduction of explicit virtual bidding reduced the variance of day-ahead/ real-time price spreads. Finally, we use the same testing procedure considering real-time prices instead of day-ahead/real-time price spreads in order to provide evidence against the oft-cited concern that EVB increases the volatility of real-time prices.

The p-values from this statistical test, conducted separately for each load aggregation point (LAP), are documented in Table C.5.

From this table, we see that we fail to reject the null hypothesis that daily day-ahead/real-time price differences are more volatile before versus after the introduction of explicit virtual bidding for all three LAPs; similarly, the null hypothesis that daily real-time prices are more volatile pre-EVB versus post-EVB can be rejected for all three LAPs. In addition, we reject the opposite null hypothesis that the volatility of day-ahead/real-time price differences is lower pre-EVB versus post-EVB for all three LAPs; the null hypothesis that the volatility of real-time prices is lower pre-EVB versus post-EVB can be rejected for PG&E and SCE, but not SDG&E.

We also conduct these statistical tests at the nodal level; Table C.6 presents the proportion of nodes for which we *fail to* reject each of our aforementioned null hypotheses, separately for nodes corresponding to locations with generators ("Gen Nodes") versus nodes that don't correspond to locations with generators ("Non-Gen Nodes"). First,

---

[29]In particular, we construct L moving-block re-samples separately for sample periods before versus after EVB. For each re-sample $b \in \{1, 2, ..., L\}$, we estimate $\hat{\Lambda}_b^{pre}$ and $\hat{\Lambda}_b^{post}$. This allows to compute $\hat{\Lambda}_b^{diff} \equiv \hat{\Lambda}_b^{pre} - \hat{\Lambda}_b^{post}$ as well as the eigenvalues associated with $\hat{\Lambda}_b^{diff}$ (which we denote $\hat{\omega}_b$). Finally, we find the empirical covariance of $\hat{\omega}_b$ across the L re-samples in order to obtain our estimate of the covariance matrix for $\hat{\omega}$ (which we denote $Var(\hat{\Lambda}_{diff})$).

Table C.5: LAP-level P-values for Volatility Tests

	LAP	Price Difference	Real-Time Price
	PGE	0.284	0.516
Pre - Post	SCE	0.509	0.697
	SDGE	0.476	0.647
	PGE	0.001	0.016
Post - Pre	SCE	0.001	0.034
	SDGE	0.028	0.165

Notes: This table reports the p-values associated with our statistical test regarding the volatility of day-ahead/real-time price spreads as well as the volatility of real-time prices before versus after the introduction of explicit virtual bidding (EVB). We implement these tests separately for each load aggregation point (LAP); the three LAPs we consider correspond the California's three major load-serving entities: Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E). To implement our test, we estimate the covariance matrix of the 24×1 vector of day-ahead/real-time electricity price spreads as well as the covariance matrix of the 24×1 vector of real-time prices separately for each LAP for the sample period before EVB (4/1/2009-2/1/2011) and after EVB (2/1/2011-12/31/2012). For both day-ahead/real-time price spreads and real-time prices, we test the null hypothesis that the pre-EVB minus post-EVB difference in covariance matrices is positive semi-definite (top panel) as well as the null hypothesis that the post-EVB minus pre-EVB difference in covariance matrices is positive semi-definite (bottom panel).

note that there aren't marked differences across the proportion of nodes for which we reject any null hypothesis for generation versus non-generation nodes. Thus, focusing on the rows corresponding to overall proportions (labeled "Total"), we see from Table C.6 that the proportion of nodes for which we reject the null hypothesis that the volatility of day-ahead/real-time price spreads is higher (lower) pre-EVB relative to post-EVB is 0.737 (0.166). Similarly, the proportion of nodes for which we reject the null hypothesis that the volatility of real-time prices is higher (lower) pre-EVB relative to post-EVB is 0.752 (0.128).

Combined, our statistical evidence indicates the volatility of day-ahead/real-time price spreads fell after the introduction of explicit virtual bidding. This is consistent with the intuition that the introduction of purely financial participants resulted in day-ahead market outcomes that more closely resembled real-time market conditions. In addition, we provided statistical evidence that the volatility of real-time prices is smaller post-EVB relative to pre-EVB; this result runs counter to the claim that financial trading increases the volatility of real-time market outcomes, causing undue stress on the electricity transmission network.

Table C.6: Proportion of Nodes for which we fail to reject the $\alpha = 0.05$ sized Volatility Test

	Node Type	Price Difference	Real-Time Price
	1(Gen Node)	0.736	0.751
Pre-Post	1(Non-Gen Node)	0.744	0.758
	Total	0.737	0.752
	1(Gen Node)	0.168	0.132
Post-Pre	1(Non-Gen Node)	0.156	0.104
	Total	0.166	0.128

Notes: This table reports the proportion of nodes (read: pricing locations) for which we *fail to* reject different null hypotheses regarding the volatility of day-ahead/real-time price spreads as well as the volatility of real-time prices before versus after the introduction of explicit virtual bidding (EVB). We report these proportions separately for nodes where generators are located ("Gen Nodes") versus nodes without generators ("Non-Gen Nodes"); we also report the overall proportions aggregated across all pricing locations ("Total"). To implement our tests, we estimate the covariance matrix of the 24×1 vector of day-ahead/real-time electricity price spreads as well as the covariance matrix of the 24×1 vector of real-time prices separately for each node for the sample period before EVB (4/1/2009-2/1/2011) and after EVB (2/1/2011-12/31/2012). For both day-ahead/real-time price spreads and real-time prices, we test the null hypothesis that the pre-EVB minus post-EVB difference in covariance matrices is positive semi-definite (top panel) as well as the null hypothesis that the post-EVB minus pre-EVB difference in covariance matrices is positive semi-definite (bottom panel).