Can Forward Commodity Markets Improve Short-Term Market Performance? Evidence from Wholesale Electricity

Akshaya Jha and Frank A. Wolak*

September 8, 2019

Abstract

Economists traditionally argue that forward commodity markets allow more efficient risk-sharing and information aggregation. However, there is little empirical evidence that commodity markets provide economic benefits to producers and consumers of the commodity. This paper demonstrates that the introduction of financial trading to California's electricity market on February 1st, 2011 improved price discovery and lowered production costs. Specifically, we document that the average, standard deviation and maximum of the differences between day-ahead and real-time electricity prices across California fell after 2/1/2011. Moreover, variable input costs (input energy) per MWh fell by 3% (4%) in high demand hours after 2/1/2011.

JEL Codes: G1, L94, Q48

Keywords: Wholesale Electricity Markets, Financial Trading, Explicit Virtual Bidding, Arbitrage, Information Aggregation, Transaction Costs

---

*Akshaya Jha: H. John Heinz III College, Carnegie Mellon University, 4800 Forbes Avenue, Pittsburgh, PA 15213. Email: akshayaj@andrew.cmu.edu. Frank A. Wolak: Program on Energy and Sustainable Development and Department of Economics, Stanford University, 579 Serra Mall, Stanford, CA 94305-6072, e-mail: wolak@zia.stanford.edu. We would like to thank Severin Borenstein, Jim Bushnell, Lucas Davis, Meredith Fowlie, Bill Hogan, Han Hong, Ali Hortacsu, Paul Joskow, Lynne Kiesling, Erin Mansur, Ignacia Mercadal, Steve Puller, Mar Reguant, Peter Reiss, Andres Santos, and Ali Yurukoglu for their helpful comments on previous drafts.

1 Introduction

Forward commodity markets allow producers and consumers to insure against fluctuations in the spot price of the commodity. Indeed, there is ample empirical evidence across many industries and commodities that hedging is widespread (Basu and Miffre, 2013). However, there is less focus on the information aggregation benefits of forward markets. Specifically, forward commodity markets aggregate information about the future values of spot prices across market participants (Newbery, 2008). Previous work has argued that the information encoded in forward commodity prices can reduce both the costs of producing the commodity and the prices paid by consumers of the commodity (Working (1953); Gray (1964); Cox (1976)). 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 online in two years and produces for 20 years would be profitable.

There is a growing empirical literature demonstrating that forward prices for a commodity provide important information about future spot prices (Cheng and Xiong, 2014). However, forward markets can clear months or even years before the commodity is to be delivered. Further, a wide range of market participants potentially use the information contained in forward prices to facilitate their participation in a variety of different upstream and downstream markets. As a result, it has proven difficult to empirically identify whether producers and consumers of a commodity realize tangible economic benefits from the information provided by a forward market for this commodity.

We argue that electricity markets are an ideal setting to investigate this hypothesis. 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 in both 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 limits the number of market participants and baseline level of liquidity in the forward market. Moreover, 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. This is different from other commodity markets where forward contracts can be traded months or even years before real-time delivery of the commodity. Consequently, relative to other commodities, differences between day-ahead and real-time electricity prices are less likely to be driven by either hedging or inter-temporal discounting.

Finally, purely financial trading is integrated into the physical market clearing process. Specifically, any market participant can take a position in the day-ahead market as a "virtual" electricity supplier or demander by submitting an offer to buy or sell energy at any location on the transmission network. If this offer is accepted in the day-ahead market, then the position must be reversed in the real-time market as a price-taker. The system operator clears day-ahead and real-time markets using the offers submitted by both physical and financial participants.

Clearing these markets requires solving a complex mixed integer problem. This is because market operators are tasked with dispatching hundreds of electricity generation units to meet real-time demand at thousands of locations on the transmission network. In doing so, operators are faced with thousands of different operating constraints tied to factors such as transmission network capacity and the time it takes for different generation units to start up or change their level of output.

Allowing market participants to take financial positions in the day-ahead market could help the system operator find lower cost solutions to this optimization problem and thereby reduce the total cost of serving demand. Specifically, expected profit-maximizing financial traders submit virtual bids to the day-ahead market in an attempt to arbitrage differences between day-ahead and real-time prices. We hypothesize that the mechanism used to clear the day-ahead market aggregates the information encoded in these financial trades, leading to a market equilibrium that satisfies demand at lower cost. Financial trading is likely to be particularly beneficial during high demand hours when system operating constraints are likely to bind.

This paper examines the empirical validity of the assertions made in the previous paragraphs using the introduction of financial trading to California's wholesale electricity market. Specifically, we first provide empirical evidence that the average, variance, and maximum of the difference between day-ahead and real-time electricity prices are lower after financial trading is introduced. This is consistent with purely financial participants increasing liquidity in both day-ahead and real-time markets. These reductions in day-ahead/real-time price spreads also suggest that day-ahead market outcomes better reflect real-time system conditions as a consequence of allowing purely financial participation.

We next show that the average variable cost of fossil-fuel-fired electricity production decreases after the introduction of financial trading in the relatively high demand hours when the operating constraints described above are most likely to bind. This suggests that the day-ahead information on real-time conditions encoded in purely financial trades benefits the producers and consumers of electricity in California.

We first develop a statistical measure of how well forward prices reflect expected spot prices. This measure is based on a hypothetical financial trader with access to 24 assets corresponding to the day-ahead/real-time price spreads for each hour of the day.[1] The financial participant faces per-unit trading cost c associated with buying or selling these assets. She 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 (i.e.: the maximum absolute average price spread). 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 average price spread is greater than c (i.e.: the null hypothesis that profitable arbitrage opportunities exist), and (2) that the maximum absolute average price spread is less than c (i.e.: the null hypothesis that no profitable arbitrage opportunities exist).

We apply this statistical framework 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. Our results indicate that the maximum absolute difference between hourly average day-ahead and real-time prices fell substantially after California implemented financial trading on February 1st, 2011.[2] Moreover, both the average and volatility of day-ahead/real-time price spreads fell after 2/1/2011. Combined, this evidence suggests that day-ahead prices better reflect real-time conditions as a result of allowing purely financial participation.

Of course, one may be concerned that day-ahead/real-time price spreads decreased over time for reasons unrelated to the introduction of financial trading. To assuage this

[1]This formulation is consistent with market rules. Specifically, trading in wholesale electricity markets occurs daily rather than hourly: participants submit virtual bids corresponding to the day-ahead/real-time price spreads at each location on the transmission network for all 24 hours of the following day simultaneously.

[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 California introduced financial trading.

concern, we assess how maximum absolute price spreads change using a difference-in-differences framework. Specifically, before financial trading was introduced, electricity suppliers could arbitrage day-ahead/real-time price spreads by adjusting the day-ahead and real-time offer curves associated with their generation units. This strategy was only feasible at locations with generation units. In contrast, before financial trading was introduced, there was no way for market participants to arbitrage price spreads at locations not associated with generation units (i.e.: "demand withdrawal locations"). Consistent with this intuition, we find that the reduction in maximum absolute price spreads due to the introduction of financial trading is larger for demand withdrawal locations relative to locations with generation units. Moreover, maximum absolute price spreads are not statistically different on average for generation locations versus demand-withdrawal locations after 2/1/2011. This provides further evidence that allowing financial participation resulted in day-ahead electricity prices that better reflect real-time conditions, especially at demand locations where arbitrage was previously unfeasible.

In the second portion of our paper, we quantify the physical efficiency gains from the introduction of financial trading. As noted earlier, the market efficiency gains from financial trading are likely to be largest during high demand hours when transmission and other operating constraints are most likely to bind. We therefore employ a difference-in-differences estimator that compares various market performance measures in high demand hours versus low demand hours before versus after purely financial participation was allowed.

Our results indicate that, in hours with system-wide electricity demand greater than the 90^{th} percentile of the distribution of demand, the introduction of financial trading resulted in a 3.1% (4.0%) decrease in fuel costs (thermal input energy consumed) per MWh of electricity produced from fossil-fuel-fired sources. The annual fuel cost savings and annual reduction in carbon emissions in high demand hours due to financial trading implied by our estimates are roughly 13.2 million dollars and 235,000 tons of CO_2 respectively. Finally, we find that the incidence of generation unit start-ups as well as the ancillary service costs associated with balancing supply and demand decrease in high demand hours after the introduction of financial trading. These findings are inconsistent with the argument that financial traders place undue stress on the physical constraints inherent to producing and distributing electricity.

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 producers and consumers of the commodity.[3] Particularly for wholesale electricity, some have argued that financial participants submit bids intended to profit from the physical realities underpinning electricity production and transmission, 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., 2018). 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)). Moreover, increases in financial trading volumes have 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 provide empirical evidence from a large wholesale electricity market that financial traders can improve price transparency and market efficiency without unduly stressing 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 financial trading. Section 3 discusses several examples of how purely financial participants can reduce the cost of serving demand. We present descriptive trends in day-ahead/real-time price spreads in Section 4. We derive a statistical measure of how well forward prices reflect expected spot prices in Section 5; Section 6 presents the results from applying this statistical framework to California's wholesale electricity market. Section 7 describes our empirical framework and findings pertaining to quantifying the market efficiency benefits of introducing financial trading. Finally, we conclude in Section 8 by discussing the policy implications of our findings.

[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 provide details on how day-ahead and real-time markets operated in California and other U.S. wholesale electricity markets prior to the introduction of financial trading (i.e.: 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. We emphasize in this subsection that purely financial bids are directly incorporated into the day-ahead and real-time market clearing processes. The final subsection compares how market participants could arbitrage day-ahead/real-time price spreads before versus after EVB was introduced. This discussion clarifies why we expect financial trading to result in day-ahead prices that better reflect real-time conditions as well as why we expect the impact of financial trading to be most pronounced in: (1) high demand hours and (2) at locations without generation units.

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 electricity markets, generation units inject electricity into the transmission network and this electricity flows according to Kirchhoff's laws (Schweppe et al., 2013). Thus, commitments between buyers and sellers of electricity constitute financial rather than physical arrangements: 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 (i.e.: demander of electricity).[4]

All electricity supply industries have transmission networks with finite transfer capacity between locations in the grid. Due to this, the system operator must sometimes satisfy demand at a given location using higher-cost generation units closer to this demand location rather than lower-cost units located farther away. Put another way, transmission

[4] As discussed in Schweppe et al. (2013), both the quantity of electricity injected by each generation unit as well as where this electricity is withdrawn depends on the level of demand at all locations on 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.

congestion limits the amount of low-cost energy that can be 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 to predict which specific transmission links will be congested. As a result, all U.S. wholesale markets have adopted a dispatch and pricing mechanism that sets potentially different prices at all points of injection and withdrawal on the transmission network.

This dispatch and pricing mechanism is called *nodal pricing* or *locational marginal pricing* (LMP). The LMP algorithm sets a potentially different price at each pricing location (termed a pricing "node") that reflects all relevant transmission network constraints, transmission losses, generation unit start-up and ramping constraints, and other relevant operating constraints on the transmission network (Bohn, Caramanis and Schweppe, 1984). Locational marginal prices in the day-ahead market are determined based on bid curves submitted by suppliers and demanders. Specifically, suppliers submit generation unit-level offer curves and load-serving entities submit locational demand curves for each of the 24 hours of the following day. Market participants must submit all 24 of the hourly bid curves by 10AM on day t for electricity to be delivered on day $t+1$.

Offer curves have three parts: a start-up cost offer, a no-load 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 no-load cost offer is a fixed dollar payment that must be paid to the generation unit owner for each hour that the unit is accepted to produce electricity. Finally, 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 price-quantity step determines the minimum price that the generation unit owner must be paid in order to produce the quantity associated with that step.[5] The sum of the quantity increments for each energy offer curve is restricted to be less than the capacity of the generation unit.

Load-serving entities (i.e.: demanders of electricity) similarly submit location-specific willingness-to-purchase bid curves in the day-ahead market that are non-increasing in the

[5]In California, suppliers are permitted to submit generation unit-level offer curves with up to ten price-quantity pairs.

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 that the market-clearing price is at or below the corresponding offer price increment. In California, LSEs typically submit willingness-to-purchase bid curves at the utility service territory level. The market operator allocates shares of these bid curves to the demand-withdrawal nodes in the utility's service territory.[6] This allocation is based on the market operator's estimate of the fraction of the utility's total demand that is withdrawn from each of the locations in its service territory.

California's Independent System Operator (ISO) clears the day-ahead market based on total as-offered cost, which is the total amount paid to suppliers based on their bid curves to satisfy the demand for energy and ancillary services at all locations on the transmission network during all 24 hours of the following day.[7] Specifically, the ISO minimizes as-offered cost subject to the ISO's best estimate of the configuration of the transmission network the following day, generation unit runtime and ramping constraints, and other system operating constraints. The locational marginal price (LMP) at each node on 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.

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 market clearing process. For example, 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 $\times$ \$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

[6]For example, consider the three major investor-owned electricity distribution utilities in California: Pacific Gas and Electric has more than 1,500 nodes in its service territory, Southern California Edison approximately 200, and San Diego Gas and Electric approximately 300. Appendix Figure A.1 presents a map of the territories served by each of these utilities.

[7]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.

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 electricity 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 real-time demand at all of the more than 4,000 demand-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 system 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 market clearing process results in real-time prices at all nodes on the transmission network as well as deviations between day-ahead scheduled output and real-time production levels for all generation units.

The 5-minute real-time price at each location on the transmission network is equal to the increase in the optimized value of the as-offered cost 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 their day-ahead schedule in a given hour receives the real-time price for any scheduled energy they do not consume.

The combination of a day-ahead forward market and a real-time spot market is called a "multi-settlement market". This is 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 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 day-ahead schedule of 50 MWhs is sold at the hourly real-time price for that location/hour-of-the-day.

2.2 Explicit Virtual Bidding

All U.S. 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.: accept the prevailing real-time price for closing out this purely financial position). 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.

The market operator treats physical and virtual bids the same when running the day-ahead market clearing process. However, the market operator knows that any day-ahead sale or purchase of "virtual energy" must be reversed in the real-time market. Specifically, 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 market earns positive revenues if and only if the day-ahead price is higher (lower) than the real-time price. However, submitting this virtual bid increases supply (demand) in the day-ahead market, making it less likely that day-ahead prices will be higher than real-time prices.

Submitting virtual bids is not costless. Specifically, purely financial participants must post collateral of significant value before being allowed to submit virtual bids. In addition, all U.S. wholesale electricity markets charge transaction fees associated with submitting virtual bids as well as additional fees if these virtual bids are accepted in the day-ahead market. Finally, financial traders must pay "uplift" charges when system-wide virtual demand is larger than system-wide virtual supply (i.e.: net virtual demand is greater than zero). These charges compensate suppliers that are forced to start up or ramp up their units in order to meet net virtual demand. Appendix Section B provides further details on the transaction costs associated with financial trading in California's wholesale electricity market.

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, and they could do so only at locations where they owned generation units. Specifically, a supplier that expected the day-ahead price at one of their locations to be higher (lower) than the corresponding real-time price might sell (buy) more energy in the day-ahead market than they expected to produce in real-time. Using physical bids rather than virtual bids to exploit expected day-ahead/real-time price spreads is termed "implicit virtual bidding".

However, suppliers are only allowed to submit physical day-ahead offers at locations where they inject electricity. Moreover, the physical offer curves submitted by suppliers must have a minimum offered quantity greater than zero and a maximum offered quantity 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 at a location where they inject electricity could at best submit zero quantity into the day-ahead market and offer their entire capacity into the real-time market. In addition, if only one supplier owns generation units at a location, then this supplier alone has the ability to arbitrage expected day-ahead/real-time price spreads at that location. With the introduction of EVB, any market participant can submit virtual bids at any node on the transmission network where financial trading is allowed.[8]

Load-serving entities (LSEs) submit demand bids at the level of their service territory. The California ISO allocates the demand bids submitted by each LSE to the nodes in its service territory using load distribution factors (LDFs).[9] It is thus very costly for a LSE to adjust its physical bid in order to arbitrage expected day-ahead/real-time price spreads at a single location. Specifically, since LSEs must submit their demand bids at the service territory level by rule, any adjustment to these physical bids would have implications for the profits or losses earned at all of the nodes in its service territory.

3 How Financial Trading Can Lower Costs

The previous section established that, prior to the introduction of explicit virtual bidding (EVB), only suppliers could arbitrage expected day-ahead/real-time price spreads, and they could so only at the locations where they inject energy. In contrast, explicit virtual bidding 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.[10] The previous section also discussed that market operators must solve an extremely challenging mixed-integer programming problem in order to determine locational marginal prices and generation unit dispatch levels in the day-ahead market. Specifically, this optimization problem has thousands of choice variables as well as thousands of constraints and thus has many potential local optima. This section provides several examples of ways in which the profit-maximizing trades placed by purely financial participants can help the market operator find lower cost solutions to this optimization problem, ultimately resulting in

[8]U.S. wholesale electricity markets typically restrict the set of nodes at which virtual bidding is allowed. For example, California's ISO does not allow virtual bidding at nodes it deems to be "electrically equivalent" to other nodes where virtual bidding is allowed.

[9]The California ISO assigns collections of nodes to each load-serving entity; this collection of nodes is called a Load Aggregation Point (LAP). Each load-serving entity is charged an hourly price equal to the LDF-weighted average of the nodal prices in their LAP. For example, if the ISO estimates that demand at each of the ten nodes in a utility's service area is the same, then the LDFs are equal to 1/10 for each node. The load aggregation point (LAP) level price faced by the utility is equal to the LDF-weighted average of the 10 nodal-level prices in its service area.

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

reductions in production costs.

Our first example concerns the case of whether to start a fossil-fuel-fired generation unit in anticipation of producing the next day. Because of runtime constraints, if this "long-start" unit is not committed in the day-ahead market, it will not be available to operate in real-time. If this unit does not operate, real-time demand must instead be satisfied by "fast-start" units that are more responsive but have higher operating costs. Suppose a purely financial player believes that committing a "long-start" unit in the day-ahead market provides a lower cost solution to meeting demand in the peak hours of the day than relying on fast-responding but more expensive units. This financial player can submit virtual demand bids at the long-start 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. If the financial player is correct, her actions result in lower cost real-time dispatch: the long-start, lower cost unit produces electricity rather than the fast-start, higher cost unit. Moreover, the financial player is likely to profit from her accurate prediction: day-ahead prices at the long-start unit's location will be lower due to the trader's virtual demand bid and real-time prices will be higher as a consequence of the unit operating.

It is important to emphasize that the purely financial player would not need to understand why day-ahead prices are lower than real-time prices at the long-start unit's location. Market participants would only need to exploit this profitable arbitrage opportunity by submitting virtual demand bids at this location during peak hours of the day. These purely financial trades can directly affect which generation units are dispatched, potentially decreasing the cost of satisfying the same levels of real-time demand across the transmission grid.

Adding a layer of complexity to the first example, consider a transmission link between two locations A and B. Suppose a virtual bidder believes that 3 MWs more transmission capacity will be made available on this transmission link in real-time than is available in the day-ahead market. If this is true, more energy can flow from A to B in real-time than the operator estimates when clearing the day-ahead market. Consequently, a financial trader can submit a virtual demand bid for 3 MWhs at node A and a virtual supply bid for 3 MWhs at node B to the day-ahead market. These purely financial bids can have potentially sizable market efficiency benefits. For example, suppose that there is a long-start, low-cost unit at node A and a fast-start, high-cost unit at node B. In the absence of the trader's virtual bids, the market operator believes in the day-ahead market that transmission will not be available between A and B. He thus calls upon the fast-start unit at node B to satisfy demand at B rather than the long-start unit at node A. In contrast, as a consequence of the virtual bids placed by the trader, the market operator in the day-ahead market instead calls upon the long-start unit to satisfy virtual demand at A. Moreover, he doesn't dispatch the fast-start unit because there is already virtual supply at node B. If the virtual bidder's prediction is correct, this change in day-ahead dispatch lowers costs. Specifically, if there is actually transmission between A and B in real-time, it would be a good decision to ensure that the lower-cost, long-start unit at A is available to satisfy demand at B.

It is important to emphasize that the trader is likely to earn a profit on both of her virtual bids because she correctly anticipated the existence of additional transmission capacity in real-time. That being said, the purely financial participant does not have to know that there is a difference in 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/real-time price spreads at node A are negative and day-ahead/real-time price spreads at node B are positive. 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.

Finally, load-serving entities (i.e.: electricity demanders) are obligated by rule to submit demand bids at the service territory level. The California ISO uses load distribution factors to allocate this LAP-level demand to individual nodes. If a financial trader believes that too much demand is allocated to a specific node in the day-ahead market relative to real-time demand, she can submit a virtual supply bid at that node in the day-ahead market. If real-time demand is actually smaller than day-ahead allocated demand as she predicted, real-time prices are likely to be lower than day-ahead prices. In this case, the trader's virtual bid would result in positive revenues. As with the previous examples, day-ahead prices that better reflect real-time conditions allow the market operator to find better solutions to his price and dispatch optimization problem, which in turn enables suppliers to better plan which units to commit to production in the day-ahead market.

As the above examples illustrate, expected profit-maximizing financial trades intended to arbitrage day-ahead/real-time price spreads can yield market efficiency gains. These examples are contingent on traders bringing information to the market in the form of virtual bids and the market clearing process aggregating this information. We should note that economic theory does not guarantee that this information aggregation will occur. Specifically, the existence of non-zero trading costs implies that California's wholesale electricity market does not satisfy all of the theoretical assumptions required for the information aggregation result in Ostrovsky (2012). That being said, the basic insight of the "separable" condition described in Ostrovsky (2012) is that, for every non-degenerate prior belief about states of the world, there exists a trader who receives an informative signal with positive probability. Expanding the number of participants that are able to arbitrage day-ahead/real-time price spreads at each node in the transmission network increases the likelihood that at least one market participant receives an informative signal about real-time conditions at some node on the transmission network.

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

4 Descriptive Trends in Price Spreads

We use hourly data on the day-ahead and real-time wholesale electricity prices at all pricing locations (i.e.: nodes) on the transmission grid run by California's Independent System Operator (ISO). Our data span from April 1, 2009 to December 31, 2012.[11] There are over 5,000 nodes; nodal-level day-ahead and real-time prices are set by the market operator based on the levels of demand at all nodes, the configuration of the transmission network, the set of available generation units and other system operating constraints.

[11] California introduced nodal pricing on April 1, 2009 (Wolak, 2011).

We also consider the load aggregation point (LAP) level prices faced by each of California's three major investor-owned 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 faced by each utility in each 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 in the same way using day-ahead quantities.

Figure 1 presents a comparison by hour-of-the-day of the average difference between 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 1st, 2011. This figure documents that day-ahead/real-time price spreads are larger in absolute value before the introduction of explicit virtual bidding (EVB) for all three of the utilities.[12] 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.1 that the reduction in average day-ahead/real-time price differences after EVB is introduced is statistically significant. Further, in Appendix Section C.2, we present results indicating that the volatility of both day-ahead/real-time price spreads and real-time prices fell after the introduction of financial trading. The reduction in both the mean and volatility of day-ahead/real-time price spreads after 2/1/2011 suggests that day-ahead prices better reflect expected real-time conditions as a consequence of financial trading.

To more formally assess the information on real-time conditions contained in day-ahead prices, the next two sections consider a hypothetical risk-neutral trader seeking to profit from expected day-ahead/real-time price spreads. This statistical test is based on the maximum over hours-of-the-day of the absolute value of the 24×1 vector of hourly average price spreads. In the extreme case with zero transaction costs, if expected price spreads are not zero for all hours, then not all forecast-able information on real-time prices is reflected in day-ahead prices. In general, the level of transaction costs required to prevent this hypothetical trader from arbitraging expected day-ahead/real-time price spreads is our measure of information content. We emphasize that this is **not** a measure of the profits that actual traders might be able to earn; indeed, before the introduction

[12] Appendix Figure A.2 plots hourly average day-ahead/real-time spreads along with their pointwise 95% confidence intervals.

Figure 1: 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) on 2/1/2011. 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).

of EVB, purely financial traders were not allowed to participate in the market.

Finally, one might be concerned that the descriptive trends discussed in this section are not driven by the introduction of EVB, but rather, by changes over time in economic conditions such as electricity demand or investment in renewables. For this reason, we formulate a difference-in-differences estimator based on comparing our measure of information content across locations before versus after purely financial participation was allowed.

5 Statistical Test of Arbitrage With Trading Costs

In this section, we develop a hypothesis testing framework to determine whether or not a profitable trading strategy exists when accounting for the presence of transaction costs. 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 to 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 only condition their trading strategies on realized price differences from two or more days prior.

In Appendix Section C.3, we formulate a statistical test of the null hypothesis that the elements of the autocovariance 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 fail to reject this null hypothesis, suggesting that traders cannot earn significantly more profits by conditioning on realized price spreads from two or more days prior to the current day.[13] For this reason, our statistical test focuses on trading strategies that condition only on the hour-of-the-day. It is important to emphasize again that our goal is not to test whether traders can actually make profits in California's wholesale electricity market. Instead, our goal is to construct a statistical measure of how well day-ahead prices reflect real-time conditions.

[13]See Appendix Tables C.5 and C.6 for the results when applying this statistical test to price spreads at the service-area level and the nodal level respectively.

5.1 The Trader's Problem

Let $X_{h,d} \equiv P^{DA}_{h,d} - P^{RT}_{h,d}$ be the difference between day-ahead and real-time electricity prices in hour-of-the-day h in day-of-sample d. Our statistical test is based on a market participant who buys or sells hourly positions a_h associated with $X_{h,d}$. Since the trader can condition her positions on the hour of the day, a_h can take on a different value for each $h \in \{1, 2, ..., 24\}$. Moreover, a_h can be positive or negative.[14] 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 her position a_h has the same sign as realized price spread $X_{d,h}$.

Let $\mu_h \equiv E(X_{h,d}) = E(P^{DA}_{h,d}) - E(P^{RT}_{h,d})$ 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 transaction cost c associated with buying or selling one MWh of any combination of these 24 assets. The trader's expected profit-maximization problem is:

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

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 position" (i.e.: $\sum_{i=1}^{24} a_i = 1$). 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

[14] 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.

models with zero transaction costs.

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

$$\phi(\mu) \equiv a^*(\mu)'\mu \ = \ \max_{h \in \{1, ..., 24\}} |\mu_h| \tag{2}$$

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

5.2 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. Specifically, 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.: $\phi(\overline{X}) \equiv \max_{h\in\{1,...,24\}} |\overline{X}_h|$). 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 population average 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)).[15] Let the sample average

[15] 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:

calculated from the b^{th} 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 trading strategies for any per-unit trading cost c. However, setting c equal to the 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. Plugging in the posted 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 implementing 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 b^{th} re-sample gives us:

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

We use this bootstrap distribution to compute two values. The first, c_{lower}, is the smallest value of the trading cost that would cause rejection of the null hypothesis that a profitable trading strategy exists (i.e.: reject the null hypothesis that $a^*(\mu)'\mu - c > 0$). Therefore, c_{lower} is equal to the 5^{th} percentile of the bootstrapped distribution of $\phi(\overline{X})$. The second

$\{\mathbf{B_1}, \mathbf{B_2}, ..., \mathbf{B_K}\} \equiv \{\mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_M}\}, \{\mathbf{X_{M+1}}, ..., \mathbf{X_{2M}}\}, ..., \{\mathbf{X_{M(K-1)+1}}, ..., \mathbf{X_{KM}}\}$. Next, let S be a discrete uniform variable over the integers $\{1, 2, ..., K\}$; we construct the b^{th} 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.

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 95^{th} percentile of the distribution of $\phi(\overline{X})$. These estimates are 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.

6 Estimates of Implied Trading Costs

This section discusses 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. The first subsection presents descriptive trends in the maximum over hours-of-the-day of the absolute value of the 24×1 vector of hourly average price spreads. We term these maximum absolute average price spreads "implied trading costs" because, as discussed in the previous section, these magnitudes reflect a hypothetical trader's ability to arbitrage price spreads given a fixed per-unit transaction cost.

The second subsection examines how implied trading costs change before versus after the introduction of explicit virtual bidding (EVB) at locations with versus without electricity generation units. Combined, the evidence presented in this section indicates that the day-ahead/real-time price spreads after EVB are more difficult for a hypothetical trader to arbitrage than the price spreads before EVB. This in turn suggests that day-ahead prices better reflect real-time conditions as a consequence of allowing purely financial participation.

6.1 Descriptive Trends in Implied Trading Costs

We first implement the statistical test of arbitrage described in Section 5 using data on the day-ahead and real-time electricity prices faced by Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E). Specifically, Table 1 reports our estimated implied trading costs both before and after the introduction of explicit virtual bidding for each load aggregation point (LAP). Recall

Table 1: LAP-Level Implied Trading Costs: c_{lower} and c_{upper}

	LAP	Before EVB	After EVB
Lower 5% C.I.	PG&E	8.576	5.683
(c_{lower})	SCE	12.125	6.064
	SDG&E	16.423	7.618
Upper 95% C.I.	PG&E	14.462	9.652
(c_{upper})	SCE	20.223	11.797
	SDG&E	32.521	14.976

Notes: This table presents the implied trading costs from the statistical framework discussed in Section 5, estimated separately for each load aggregation point (LAP) for the 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). For each LAP in each sample period, c_{lower} (c_{upper}) is the 5^{th} (95^{th}) percentile of the bootstrapped distribution of the maximum over hours-of-the-day of the absolute value of the $24{\times}1$ vector of hourly average day-ahead/real-time price spreads.

that c_{lower} is the smallest value of per-unit trading costs for which we can reject the null hypothesis that a profitable strategy exists while c_{upper} is the largest value of trading costs for which we can reject the null hypothesis that no profitable trading strategy exists. Table 1 demonstrates that our estimates of c_{lower} and c_{upper} are substantially lower after the introduction of explicit virtual bidding (EVB) for all three LAPs.[16]

In order to more formally compare implied trading costs before versus after EVB was introduced, Figure 2 plots the bootstrap distribution of the *difference* in implied trading costs (i.e.: the difference in maximum absolute average price spreads) for each LAP before versus after EVB. The left vertical line in this figure is the 5^{th} percentile of the distribution of $c_{pre} - c_{post}$ and the right vertical line is the 95^{th} percentile of this distribution. If the 5^{th} percentile of this distribution is greater than zero, then we can reject the null hypothesis that $c_{pre} \leq c_{post}$ at a 5% significance level. Similarly, we can reject the null hypothesis that $c_{pre} \geq c_{post}$ at a 5% significance level if the 95^{th} percentile of the bootstrapped distribution of $c_{pre} - c_{post}$ is less than zero. For all three LAPs, we reject the null hypothesis that implied trading costs are higher post-EVB relative to pre-EVB, but fail to reject the null hypothesis that implied trading costs are higher pre-EVB relative to post-EVB. Put another way, Figure 2 provides statistical evidence that

[16] 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.

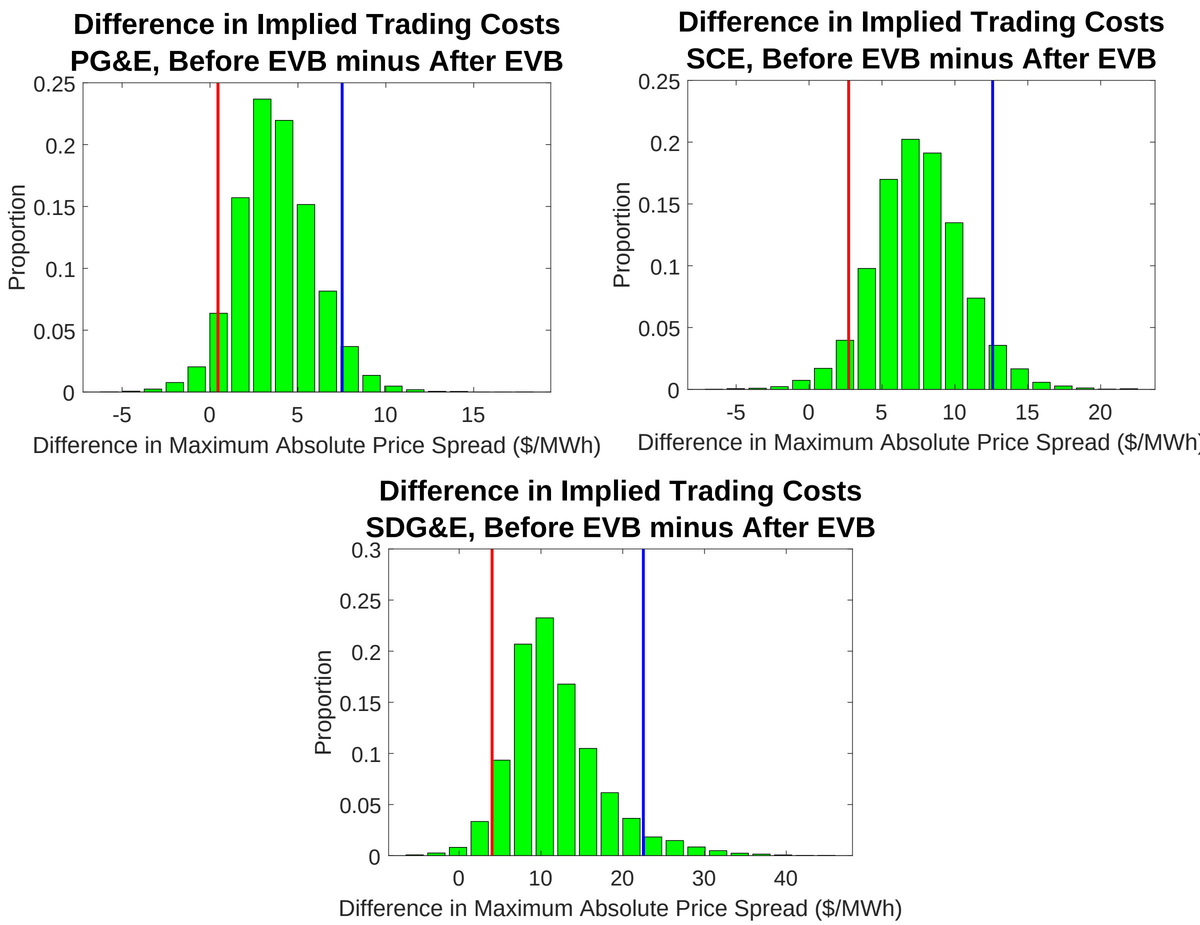
maximum absolute average price spreads fell after the introduction of EVB for all three LAPs.

We also compute c_{lower} and c_{upper} for each pricing node in California's wholesale electricity market. Figure 3 plots the values of c_{lower} and c_{upper} for each node before versus 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"). Figure 3 shows that the across-node distribution of c_{lower} is shifted downward post-EVB relative to pre-EVB. This implies that, for any fixed value of trading costs, we reject the null hypothesis that a profitable trading strategy exists for more nodes in the post-EVB sample period than the pre-EVB sample period. Similarly, the distribution of c_{upper} also shifts downward after EVB is introduced. This indicates that, fixing a transaction cost, the null hypothesis that no profitable trading strategies exist can be rejected for more nodes before EVB relative to after EVB.

Next, we compute the bootstrap distribution of estimated $c_{pre} - c_{post}$ for each of the more than 4,000 nodes that exist both before and after EVB. The first row of Table 2 reports the proportion of nodes for which we reject the null hypothesis that implied 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 2 reports the proportion of nodes for which we reject the null hypothesis that implied 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.

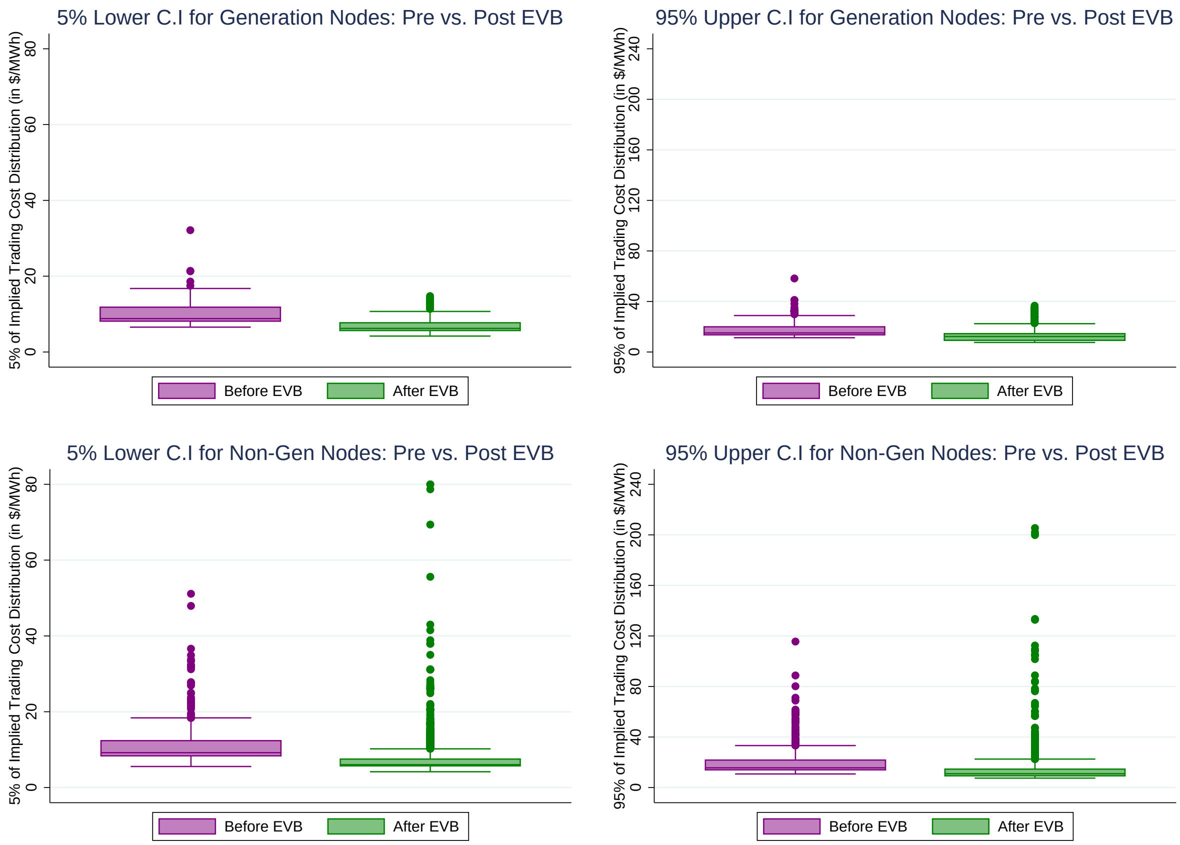
We reject the null hypothesis that implied trading costs increased after the introduction of EVB for more than 70 percent of nodes. In contrast, we reject the null hypothesis that trading costs fell after EVB was introduced for less than 5 percent of nodes. A rejection frequency of 5% is consistent with the null hypothesis being true for all nodes because the size of each hypothesis test is $\alpha = 0.05$. Summarizing, for the majority of nodes, we reject the null hypothesis that maximum absolute price spreads increased after EVB but fail to reject the null hypothesis that maximum absolute price spreads decreased after EVB. This suggests that day-ahead prices better reflect real-time conditions after purely financial participation was allowed.

Figure 2: Bootstrap Distribution of the Difference in Trading Costs



Notes: This figure plots the bootstrap distribution of the difference in "implied 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). Note that $\phi(\overline{X}) \equiv \underset{h\in\{1,\ldots,24\}}{max} |\overline{X}_h|$ is the maximum over hours-of-the-day of the absolute value of the 24×1 vector of hourly average day-ahead/real-time price spreads. We plot this bootstrap distribution separately for the day-ahead/real-time price spreads faced by each of California's three major investor-owned 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 in red is the 5^{th} percentile of the distribution of $\phi(\overline{X}^{pre}) - \phi(\overline{X}^{post})$ and the right vertical line in blue is the 95^{th} percentile of the distribution of $\phi(\overline{X}^{pre}) - \phi(\overline{X}^{post})$.

Figure 3: Nodal-Level Distribution of Implied Trading Costs: Before and After EVB



Notes: This figure plots the nodal-level distribution of estimates of c_{lower} and c_{upper} for each pricing location (i.e.: node). These "implied trading costs" are estimated separately for each node 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). For each node in each sample period, c_{lower} (c_{upper}) is the 5^{th} (95^{th}) percentile of the bootstrapped distribution of the maximum over hours-of-the-day of the absolute value of the 24×1 vector of hourly average day-ahead/real-time price spreads. We plot the across-node distributions of c_{lower} and c_{upper} separately for nodes associated with generation units versus nodes not associated with generation units. The box portion of this box and whiskers plot contains all nodes within the 25^{th} through 75^{th} percentiles of the nodal-level distribution of implied trading costs. The bottom (top) whisker is defined by the smallest (largest) value that is within $1.5 \times IQR$ of the 25^{th} (75^{th}) percentile of the distribution of implied trading costs, where IQR (inter-quartile range) is the distance between the 25^{th} and 75^{th} percentiles of a distribution. The remaining points are outliers.

Table 2: 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 increase after explicit virtual bidding (EVB) is introduced (i.e.: $c_{pre} \leq c_{post}$), separately for nodes associated with generation units ("Gen Nodes") versus nodes not associated with generation units ("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 decrease after the introduction of EVB (i.e.: $c_{pre} \geq c_{post}$), once again separately for generation nodes versus non-generation nodes. This statistical test is based on the bootstrap distribution of the difference between the maximum absolute average day-ahead/real-time price spread before versus after the introduction of EVB at each node (i.e.: $\phi(\overline{X}^{pre}) - \phi(\overline{X}^{post})$). Specifically, $\phi(\overline{X}) \equiv \underset{h \in \{1,\ldots,24\}}{max} |\overline{X}_h|$ is the maximum over hours-of-the-day of the absolute value of the 24×1 vector of hourly average price spreads.

6.2 Difference-in-Differences Framework

Before the introduction of explicit virtual bidding (EVB), only suppliers could exploit expected differences between day-ahead and real-time prices. They could do so only by adjusting their physical bids at the locations where their generation units inject electricity ("generation nodes"). In contrast, load-serving entities (i.e.: electricity demanders) can only submit physical bids at the load aggregation point (LAP) level. This makes 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, in the absence of EVB, we expect implied trading costs to be higher at non-generation nodes relative to generation nodes because no market participant can implicitly virtual bid at non-generation nodes. EVB allows any market participant to place virtual bids at any node; thus, we expect the reduction in implied trading costs after the introduction of EVB to be larger for non-generation nodes relative to generation nodes.

To test these two hypotheses, we regress our estimate of the implied trading cost c_{lower} at each node before and after the introduction of EVB on a constant, an indicator variable that's equal to one if the node is associated with a generation unit, an indicator variable that's equal to one if the implied trading cost is from the post-EVB period, and an indicator variable that's equal to one if the observation is from a generation node during the post-EVB sample period.[17] The unit of observation for this regression is thus a node in the pre-EVB versus post-EVB sample period. Heteroskedasticity-consistent standard errors are in parentheses. Finally, we run the same regression with our estimate of c_{upper} for each node before and after EVB as the dependent variable.

Table 3 presents the results of estimating this difference-in-differences specification. We see that the coefficient estimate on 1(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 coefficient estimates corresponding to 1(Gen Node) indicate that both c_{lower} and c_{upper} are significantly lower for generation nodes relative to non-generation nodes prior to EVB. This difference across generation versus non-generation nodes is essentially eliminated after the introduction of EVB. Specifically, we fail to reject the null hypothesis that the sum of the coefficients corresponding to the variables 1(Gen Node) and 1(Post EVB) $\times$ 1(Gen Node) is zero for both c_{lower} and c_{upper}. Put another way, the reduction in implied trading costs after EVB is introduced is larger for non-generation nodes relative to generation nodes.

Table 3 thus provides statistical evidence consistent with all three of our hypotheses: (1) implied trading costs are lower for generation nodes relative to non-generation nodes prior to EVB, (2) implied trading costs fell after EVB was introduced, and (3) the reduction in implied trading costs after the introduction of EVB was smaller for generation nodes relative to non-generation nodes. This indicates that maximum absolute average day-ahead/real-time price spreads (i.e.: implied trading costs) fell as a consequence of allowing purely financial participation rather than secular trends over time in factors such as demand or investment in renewables. This in turn suggests that day-ahead prices better reflect real-time conditions due to the introduction of financial trading. In the next section, we quantify the physical market efficiency benefits from financial trading; we focus on high demand hours because day-ahead information on real-time conditions is likely to be most valuable during these hours.

[17] Recall that c_{lower} (c_{upper}) is the 5^{th} (95^{th}) percentile of the bootstrapped distribution of maximum absolute average price spreads.

Table 3: Implied Trading Costs Before vs. After EVB For Gen versus Non-Gen Nodes

Dependent Variable	c_{lower}	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 variable 1(Post EVB) that's equal to one if the observation corresponds to the sample period after EVB is introduced (2/1/2011-12/31/2012); the pre-EVB sample period is 4/1/2009-2/1/2011. We also include an indicator variable 1(Gen Node) that's equal to one if and only if the node is associated with a generation unit. Finally, we include the interaction between these two variables (i.e.: 1(Post EVB) × 1(Gen Node)). We consider two dependent variables: c_{lower} in Column 1 and c_{upper} in Column 2. For each node in each sample period, c_{lower} (c_{upper}) is the 5^{th} (95^{th}) percentile of the bootstrapped distribution of the maximum over hours-of-the-day of the absolute value of the 24×1 vector of hourly average day-ahead/real-time price spreads.

7 Benefits from EVB in High Demand Hours

This section describes the data, methodology, and results of our empirical analysis demonstrating that the introduction of explicit virtual bidding (EVB) 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 EVB are likely to be largest in the highest demand hours.

7.1 Data

We utilize hourly data on generation-unit-level output from California's Independent System Operator (ISO), daily natural gas prices for Northern and Southern California from SNL Financial, 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 on the characteristics of each electricity generation unit in the California ISO control area. These characteristics include the unit's capacity, fuel type (i.e.: natural gas, oil, wind, etc.), the total amount of thermal energy required to start up the unit, as well as the unit's 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 BTUs of thermal input in order to produce one more MWh of electrical energy". Combined, our data allow us to compute the total quantity of thermal energy consumed by each generation unit during each hour of our sample period.

To do this, we first define generation unit i as starting in hour t if its output in hour t-1 is zero and its output in hour t is greater than zero (i.e.: $Q_{i,t-1} = 0$ and $Q_{i,t} > 0$). Let E_i^S be the thermal energy required to start up unit i. Next, firms are required to report a heat rate curve for each of their thermal units to the market operator. This heat rate curve consists of up to 10 heat-rate/output steps, allowing us to compute the quantity of thermal energy required to produce 1 MWh of electricity (i.e.: the heat rate) for any level of output. Formally, for each unit i, we know the heat rate HR_i^k relevant for each of $k = 1, 2, ..., 10$ intervals of output $[\overline{Q}_i^{k-1}, \overline{Q}_i^k]$.[18]

[18] By convention, $\overline{Q}_i^0$ is equal to zero and $\overline{Q}_i^{10}$ is equal to the unit's capacity.

Combined, the thermal energy required to produce output $Q_{i,t}$ is:

$$E_{i,t} = \underbrace{\sum_{k=1}^{10} HR_i^k (Q_{i,t} - \sum_{j=0}^{k-1} \overline{Q}_i^j)^+}_{\text{Fuel Used for Production}} + \underbrace{E_i^S 1(Q_{i,t} > 0) 1(Q_{i,t-1} = 0)}_{\text{Start-Up Fuel}}$$

where the notation $y^+ \equiv y \times 1(y > 0)$. 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 the unit. For natural gas fired units, we use the Pacific Gas and Electric (PG&E) citygate natural gas price if the unit is located in PG&E's service territory. We use the Southern California Gas (SCG) citygate price if the unit is located in the territories served by either Southern California Edison or San Diego Gas and Electric. We calculate the costs associated with the relatively small amount of distillate fuel oil burned by generation units in California using the diesel fuel price paid in Los Angeles.[19] We sum the total fuel costs, total heat energy used, and total number of starts over fossil-fuel-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 impacts hourly aggregate fuel costs per MWh of electricity produced by fossil-fuel-fired units, hourly aggregate thermal energy used per MWh, and an indicator variable that's equal to one if at least one fossil-fuel-fired unit started up in the hour.

Our fourth market outcome is the hourly ancillary service costs paid by the California ISO (CAISO) per MWh of fossil-fuel-fired electricity production. CAISO incurs ancillary service costs in order to make certain that electricity supply equals electricity demand at every instant even in the face of generation unit and tranmission deratings and outages. It is especially important to test how ancillary service costs respond to the introduction of EVB because policymakers and regulators have expressed concerns that financial traders take advantage of physical constraints such as transmission congestion in order to make profits.

Finally, we assess how EVB impacts the hourly day-ahead and real-time electricity

[19]Specifically, we collect data from the Energy Information Administration on the ultra-low sulfur CARB diesel spot price relevant to Los Angeles.

prices paid by each of California's three major investor-owned utilities: Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E). Specifically, our last three outcomes are the absolute differences between the hourly 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.

7.2 Difference-in-Differences Methodology

We estimate the following difference-in-differences specification for each market outcome Y_t:

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

where t indexes hour-of-sample. 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, the log of net electricity imports, two separate controls for the logs of the natural gas prices paid at the PG&E and SCG citygates, 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 on 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 is equal to one if and only if aggregate electricity demand in hour-of-sample t is larger than the 90^{th} percentile of the distribution of aggregate hourly demand across our 4/1/2009-3/31/2012 sample period. That being said, we show in a sensitivity analysis that our results are similar for the 50^{th}, 75^{th}, 95^{th}, and 99^{th} percentiles of demand. The independent variable of interest, $\text{HIGH}_t \times \text{POSTEVB}_t$, captures how the introduction of financial trading impacts each each outcome in high demand hours relative to low demand hours. Finally, we cluster standard errors by day-of-sample.

7.3 Primary Findings

Table 4 presents the results from estimating the difference-in-differences specification described in the previous subsection. We discuss robustness checks pertaining to these results, such as the statistical test for common pre-existing trends, in the next subsection. Column 1 of Table 4 indicates that the introduction of explicit virtual bidding (i.e.: financial trading) to California's wholesale electricity market resulted in a 3.1% reduction in average fuel costs per MWh in high demand hours relative to low demand hours. Aggregating across all of the power plants in our sample, this 3.1% decrease in fuel costs per MWh corresponds to a 13 million dollar reduction in the annual fuel costs paid in high demand hours. Similarly, we find a 4% reduction in average input heat energy per MWh in relatively high demand hours due to financial trading (see Column 2 of Table 4). A 4% reduction in thermal energy per MWh translates into an annual reduction in CO_2 emissions of 235,000 tons in high demand hours.

We see from Column 3 of Table 4 that, on average, at least one unit has to start up in 3.6% less high demand hours after explicit virtual bidding (EVB) is introduced. This assuages concerns that market participants are submitting trades in order to profit from exacerbating physical constraints such as start-up or ramping. To more directly address this concern, we consider the "ancillary service" payments made to electricity suppliers in order to ensure that supply meets demand at every instant in time. If allowing purely financial participation resulted in the exacerbation of transmission, start-up, ramping, or other system operating constraints, we would expect a marked increase in ancillary service costs. This does not turn out to be the case; Column 4 of Table 4 demonstrates that there is no statistical difference in ancillary service costs per MWh before versus after EVB in relatively high demand hours.

The last three columns of Table 4 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: Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E). All three columns indicate that average absolute day-ahead/real-time price spreads fell significantly after 2/1/2011 in relatively high demand hours. Combined with the evidence presented in Sections 4 and 6, our findings suggest that day-ahead electricity prices better predict expected real-time prices due to the introduction of financial trading. We also find

Table 4: Diff-in-Diff: High versus Low Demand Hours Before versus After EVB

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1(Demand > 90%) × 1(Post EVB)	-0.031***	-0.040***	-0.036*	-0.036	-5.742**	-14.024***	-21.292***
	(0.006)	(0.007)	(0.022)	(0.070)	(2.701)	(4.661)	(7.273)
1(Demand > 90%)	0.033***	0.041***	0.022	0.211***	2.428	5.925*	8.575
	(0.003)	(0.004)	(0.017)	(0.036)	(1.897)	(3.463)	(5.611)
R^2	0.900	0.370	0.127	0.605	0.041	0.039	0.049
Mean of Dep. Var.	3.583	2.082	0.686	0.014	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 day-of-sample and are reported in parentheses. The dependent variables considered in this table are: (1) the log of fuel costs per MWh, (2) the log of input energy per MWh, (3) an indicator variable that's equal to one if and only if at least one fossil-fuel-fired unit started up, (4) the log of ancillary service costs per MWh, (5) the absolute value of the day-ahead/real-time price spread in PG&E, (6) the absolute value of the day-ahead/real-time price spread in SCE, and (7) the absolute value of the day-ahead/real-time price spread in SDG&E. The row titled "Mean of Dep. Var." reports the mean of the dependent variable. All of the regressions listed in this table 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, the log of total electricity imports, two separate controls for the logs of the natural gas prices paid at the PG&E and SCG citygates 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 equal to one if and only if the day-of-sample is on or after 2/1/2011. The indicator variable 1(Demand > 90%) is equal to one if and only if system-wide electricity demand in the hour is greater than the 90^{th} percentile of the distribution of hourly demands across our 4/1/2009-3/31/2012 sample period.

Table 5: Diff-in-Diff Robustness Check: By Percentage of Demand

	Dependent Variable: Log of Average Fuel Cost Per MWh				
	(1)	(2)	(3)	(4)	(5)
1(Demand > Cut-off) × 1(Post EVB)	-0.012***	-0.023***	-0.031***	-0.038***	-0.018**
	(0.003)	(0.004)	(0.006)	(0.007)	(0.009)
1(Demand > Cut-off)	-0.005***	0.003	0.033***	0.061***	0.074***
	(0.002)	(0.002)	(0.003)	(0.004)	(0.006)
Demand Cut-Off	50%	75%	90%	95%	99%
Mean of Dep. Var.	3.583	3.583	3.583	3.583	3.583
R^2	0.900	0.900	0.900	0.901	0.900
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 day-of-sample and are reported in parentheses. The row titled "Mean of Dep. Var." reports the mean of the log of fuel costs per MWh. All of the regressions listed in this table 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, the log of total electricity imports, two separate controls for the logs of the natural gas prices paid at the PG&E and SCG citygates 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 equal to one if and only if the day-of-sample is on or after 2/1/2011. The indicator variable 1(Demand > Cut-off) is equal to one if and only if system-wide electricity demand in the hour is greater than Xth percentile of the distribution of hourly demands across our 4/1/2009-3/31/2012 sample period; X is equal to the 50th, 75th, 90th, 95th, or 99th percentile depending on whether we're considering the specification estimated in Columns 1, 2, 3, 4, or 5 respectively.

empirical evidence supporting the assertion that the information encoded in day-ahead prices as a consequence of financial trading benefits electricity consumers and producers. Specifically, allowing purely financial participation lowers the costs of providing electricity in relatively high demand hours, as evidenced by Columns 1-4 of Table 4.

Finally, Table 5 shows that our estimated effect of financial trading on average fuel costs per MWh in high demand hours remains quantitatively similar if "high demand" is defined based on hours-of-sample above the 50th, 75th, 95th, or 99th percentiles of hourly demand. This provides evidence that our results are not an artifact of using the 90th percentile of demand in our primary specifications. It is especially important that the effect estimated for average fuel costs per MWh is similar across different definitions for "high demand" because we use this effect to calculate the annual aggregate efficiency gains from financial trading in high demand hours.

7.4 Robustness Checks

This subsection discusses robustness checks pertaining to our difference-in-differences analysis. Appendix Section D.1 focuses on pre-trend analysis. First, Appendix Figures D.1 and D.2 present trends in monthly average residualized outcomes separately for high versus low demand hours. These figures suggest that the trends in outcome in high versus low demand hours are similar prior to the introduction of EVB. However, we test this "common trends" assumption more formally by regressing the first difference of each outcome on an indicator for high demand hours. The results of this analysis, presented in Appendix Table D.1, indicate that there is no statistical difference in the time trend of outcomes in high versus low demand hours prior to the introduction of financial trading. This provides evidence that the results provided in Table 4 are not due to pre-existing differences in how outcomes evolve over time in high versus low demand hours.

Next, we show that the empirical results remain similar if we estimate Equation (5) after dropping the 28 days before and after the introduction of financial trading on 2/1/2011 (see Appendix Table D.2). This indicates that our findings are not driven by short-run adjustments to the policy change. In addition, Appendix Table D.3 presents the results when estimating Equation (5) using only data from the 6 months before and after 2/1/2011. These empirical results are quantitatively quite similar to our primary findings listed in Table 4. Our estimated effects also remain statistically significant if we calculate standard errors using the Newey-West formula (Newey and West, 1987) accounting for 168 hours (i.e.: one week) of autocorrelation (see Appendix Table D.4). Finally, our empirical results remain quantitatively similar if we consider day-of-sample fixed effects rather than month-of-sample fixed effects (see Appendix Table D.5).

Appendix Table D.6 presents the results from estimating Equation (5) on data aggregated to the daily level. The results at the daily-level are broadly consistent with those presented in Table 4. The one exception is that the estimated effects for absolute day-ahead/real-time price spreads are no longer statistically significant. This is unsurprising given that the reductions in price spreads due to financial trading are likely to be largest for the hours of the day with the highest average demand.

Finally, we consider an alternative specification based on daily total number of starts. Specifically, we define a day-of-sample as treated if daily total number of starts is above the 90^{th} percentile of the distribution of daily starts. This specification isolates the impact of financial participation on market outcomes due to changes in which units start up in order to meet peak demand (i.e.: the maximum level of electricity demand across the 24 hours of the day). The results from this specification are presented in Appendix Section D.3. These results tell the same story as our primary findings documented in Table 4: the introduction of explicit virtual bidding (EVB) reduces fuel costs per MWh, input heat energy per MWh, ancillary service costs per MWh, and absolute day-ahead/real-time price spreads on days with a large number of starts relative to days with a small number of starts. Moreover, the estimated effect of EVB on fuel costs per MWh on days with a relatively large number of starts is similar regardless of whether "high start" days are defined based on the 50^{th}, 75^{th}, or 95^{th} percentiles of the distribution of daily starts rather than the 90^{th} percentile.

8 Implications for Electricity Market Design

Economists have long speculated that the information provided by forward commodity markets benefits producers and consumers of the commodity (Working (1953); Gray (1964); Cox (1976)). We empirically test this assertion 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. We find that the introduction of financial trading resulted in a 3% (4%) reduction in fuel costs (thermal energy) per MWh in relatively high demand hours. This 3% reduction in fuel costs per MWh implies that the annual total fuel costs incurred during high demand hours are 13.2 million dollars lower as a consequence of allowing purely financial participation. Moreover, our estimated 4% reduction in thermal energy per MWh implies that roughly 235,000 less tons of CO_2 are emitted each year in high demand hours due to the introduction of financial trading.

Financial trading in this context thus comes with substantial economic and environmental benefits. However, allowing purely financial participation in wholesale electricity markets is not without controversy. Specifically, many have expressed concerns that financial traders take advantage of physical constraints such as transmission congestion or the start-up times of generation units in order to profit at the expense of electricity producers and consumers. Our evidence suggests that allowing purely financial participation does not significantly exacerbate the physical constraints inherent to electricity production and distribution. In particular, we find that the incidence of generation unit start-ups *fell* in high demand hours as a consequence of allowing purely financial participation. Moreover, there is no statistical difference before versus after 2/1/2011 in the per-MWh ancillary service costs associated with ensuring that supply meets demand in high demand hours.

Previous work has argued that forward commodity markets provide benefits to consumers and producers in part because they aggregate information about the future values of spot prices across market participants (Newbery, 2008). Consistent with this intuition, we find that both the average and variance of day-ahead/real-time price spreads fell after explicit virtual bidding (EVB) was introduced. Moreover, the maximum over hours-of-the-day of the absolute value of the 24×1 vector of hourly average price spreads is also smaller post-EVB relative to pre-EVB. We call this maximum absolute average price spread an "implied trading cost" because, under some assumptions, the distribution of this magnitude can be used to statistically test whether a hypothetical trader can arbitrage expected price spreads if faced with per-unit trading costs c.

We more formally test whether implied trading costs fell as a consequence of allowing purely financial participation using a difference-in-differences framework. Specifically, prior to the introduction of EVB, only suppliers could arbitrage expected day-ahead/real-time price spreads. They could do so by altering the physical bids associated with their generation units. This "implicit virtual bidding" strategy is thus feasible only at locations with generation units. In contrast, explicit virtual bidding allows any market participant the opportunity to arbitrage expected price spreads at any location on the transmission grid. Consistent with this intuition, we show that the average reduction in implied trading costs due to EVB is larger at locations without generation units relative to locations with generation units. This suggests that day-ahead prices better reflect real-time conditions as a consequence of allowing purely financial participation rather than secular trends over time in factors such as demand or investment in renewables.

Finally, we use our estimates to provide a back-of-the-envelope calculation of the benefits from reducing the per-unit costs of trading in California's wholesale electricity market. Specifically, our results indicate that the average reduction in implied trading costs after EVB was introduced was \$3.45 per MWh. We also estimate that fuel costs per MWh decrease by \$1.09 in high demand hours relative to low demand hours due to EVB. Combined, this implies that a 5 cent reduction in transaction costs corresponds to a roughly 1.6 ($= 5 \times \frac{1.09}{3.45}$) cent reduction in fuel costs per MWh in high demand hours on average. However, there is substantial heterogeneity across locations and hours in our estimated costs and benefits from financial trading. This suggests that the efficiency gains from reducing transaction fees are likely to be even larger if these fees are allowed to vary based on the expected benefits from financial trading at different locations and hours.

References

Basu, Devraj, and Joëlle Miffre. 2013. "Capturing the risk premium of commodity futures: The role of hedging pressure." *Journal of Banking & Finance*, 37(7): 2652–2664.

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 R, Ali Hortaçsu, Ignacia Mercadal, and J Michael Pavlin. 2018. "Limits to Arbitrage in Electricity Markets: A case study of MISO." *Energy Economics*, 75: 518–533.

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*.

Cheng, Ing-Haw, and Wei Xiong. 2014. "Financialization of commodity markets." *Annu. Rev. Financ. Econ.*, 6(1): 419–441.

Cox, Charles C. 1976. "Futures trading and market information." *Journal of Political Economy*, 84(6): 1215–1237.

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*.

Gray, Roger W. 1964. "The attack upon potato futures trading in the United States." *Food Research Institute Studies*, 4(1387-2016-116207): 97–121.

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

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.

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."

Newbery, David M. 2008. "Futures markets, hedging and speculation." *The New Palgrave Dictionary of Economics: Volume 1–8*, 2327–2332.

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."

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.

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. 2011. "Measuring the benefits of greater spatial granularity in short-term pricing in wholesale electricity markets." *The American Economic Review*, 101(3): 247–252.

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* .

Working, Holbrook. 1953. "Hedging reconsidered." *Journal of Farm Economics*, 35(4): 544–561.

List of Figures

List of Tables

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 submit purely financial bids to California's wholesale electricity market. Physical market participants (i.e.: electricity 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 submit financial bids to California's electricity market.

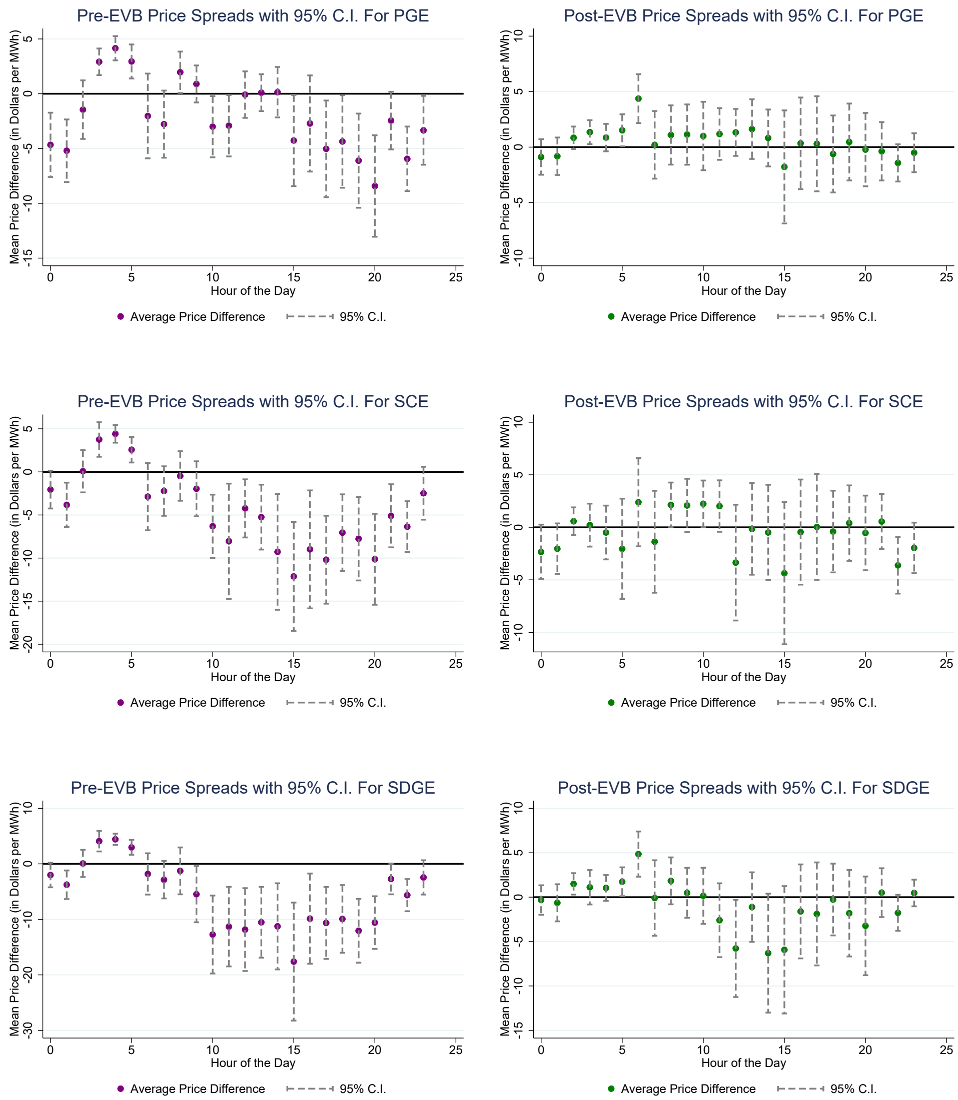
Appendix Figure A.2 presents hourly average day-ahead/real-time electricity price spreads, along with their pointwise 95% confidence intervals, for the load aggregation points (LAPs) corresponding to the territories served by three major investor-owned utilities in California: Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E). We provide separate plots for the sample periods before versus after the introduction of explicit virtual bidding (EVB). Prior to the introduction of EVB, we can reject that the average day-ahead/real-time price spread is zero for many hours of the day for all three utilities. After EVB is introduced, we can reject the null hypothesis of zero average price spread for relatively few hours-of-the-day. Importantly, rejecting this null hypothesis does not imply that financial traders can profitably arbitrage LAP-level price spreads. One must also take into account the sizable transaction costs associated with submitting virtual bids in order to arbitrage these price spreads.

Appendix Figure A.3 plots the bootstrap distribution of the maximum over hours-of-the-day of the absolute value of the 24×1 vector of hourly average day-ahead/real-time price spreads, which we term "implied trading costs". We plot separate distributions for the pre-EVB versus post-EVB sample periods in purple and green respectively. The solid vertical lines on each graph in this figure denote our estimated values for c_{lower} (in red) and c_{upper} (in blue) for the pre-EVB sample period while the dotted vertical lines denote our estimated values for c_{lower} and c_{upper} for the post-EVB sample. Recall that c_{lower} is the smallest value of per-unit transaction costs for which we can reject the null hypothesis

Figure A.1: Territories Served by California's Three Major Investor-Owned Utilities

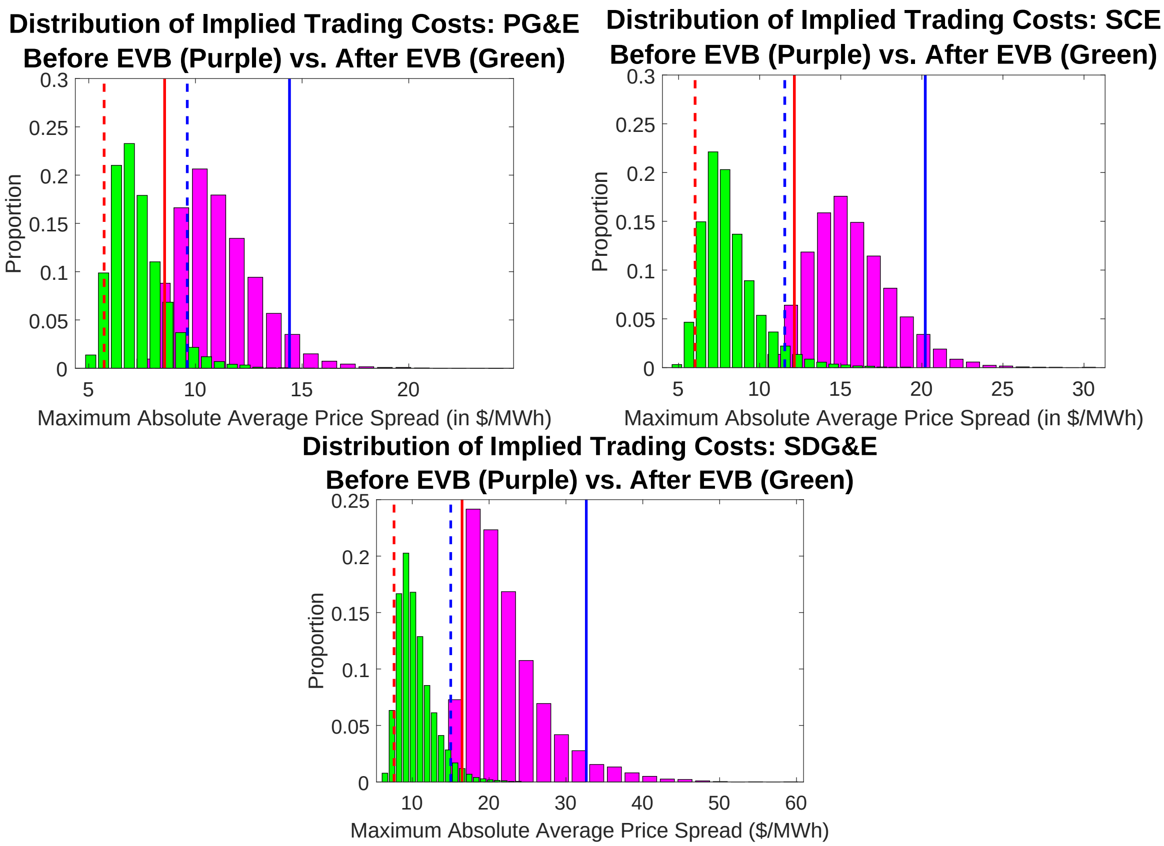
Notes: This is a map of the territories served by each of the three major investor-owned electric utilities in California. These three utilities are Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E). This map is provided by the California Energy Commission; see https://www.ferc.gov/market-oversight/mkt-electric/california.asp.

Figure A.2: Hourly Average Price Spreads 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 pre-EVB sample period spans 4/1/2009-1/31/2011 while the post-EVB sample period spans 2/1/2011-12/31/2012. 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). This figure also includes the pointwise 95% confidence interval associated with the average day-ahead/real-time price spread for each hour-of-the-day.

Figure A.3: Bootstrap Distribution of Maximum Absolute Average Price Spreads



Notes: This figure plots the bootstrap distribution of the maximum over hours-of-the-day of the absolute value of the 24×1 vector of hourly average day-ahead/real-time price spreads, which we term "implied trading costs". We plot separate distributions for the pre-EVB sample period (4/1/2009-2/1/2011) versus the post-EVB sample period (2/1/2011-12/31/2012) in purple and green respectively. The top left panel, the top right panel, and the bottom middle panel of this figure focuses on the LAP-level implied trading costs associated with PG&E, SCE, and SDG&E respectively. The solid vertical lines on each graph in this figure denote our estimated values for c_{lower} (in red) and c_{upper} (in blue) for the pre-EVB sample period while the dotted vertical lines denote our estimated values for c_{lower} and c_{upper} for the post-EVB sample. c_{lower} is the smallest value of per-unit transaction costs for which we can reject the null hypothesis that a profitable trading strategy exists, which corresponds to the 5^{th} percentile of the distribution of implied trading costs. c_{upper} is the largest value of transaction costs for which we can reject the null hypothesis that no profitable trading strategy exists, which corresponds to the 95^{th} percentile of the distribution of implied trading costs.

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 submit financial bids in California's wholesale electricity market (CAISO (2019a)). This table is split into two columns: the first column lists physical participants (i.e.: electricity suppliers and demanders) who inject or withdraw electricity from the transmission grid while the second column lists purely financial participants that don't inject or withdraw electricity from the grid.

that a profitable trading strategy exists, which corresponds to the 5^{th} percentile of the distribution of implied trading costs. c_{upper} is the largest value of transaction costs for which we can reject the null hypothesis that no profitable trading strategy exists, which corresponds to the 95^{th} percentile of the distribution of implied trading costs.

The top left panel, the top right panel, and the bottom middle panel of Appendix Fig-

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 submit financial bids to California's wholesale electricity market (CAISO (2019a)). This table is split into two columns: the first column lists physical participants (i.e.: electricity suppliers and demanders) who inject or withdraw electricity from the transmission grid while the second column lists purely financial participants that don't inject or withdraw electricity from the grid.

ure A.3 focuses on the LAP-level implied trading costs associated with PG&E, SCE, and SDG&E respectively. All three panels indicate that both c_{lower} and c_{upper} fell substantially after the introduction of explicit virtual bidding (EVB). That being said, Figure 2 presents results from a formal test of the null hypothesis that c_{lower} and c_{upper} remained the same before versus after EVB.

B Trading Fees for California's Electricity Market

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 the virtual bids they submit each day.[20] This collateral does not earn any rate of return while it is held by California's Independent System Operator (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.[21]

Purely financial participants must pay roughly 0.5 cents for each price and quantity pair associated with each virtual bid they submit. They must also pay 9 cents per MWh of virtual energy *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 \$4.55 $= (\$0.09 \times 50) + (\$0.005 \times 10)$ in transaction fees. 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.[22]

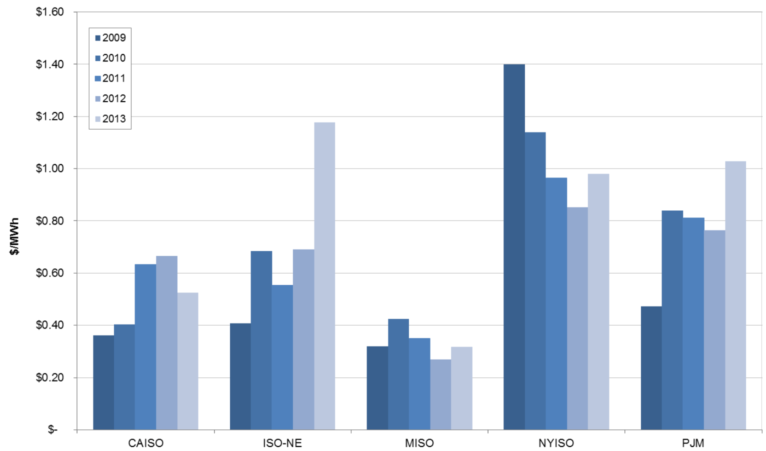
Finally, purely financial participants are required to pay "uplift" charges when system-wide virtual demand is larger than system-wide virtual supply (i.e.: net virtual demand is greater than zero). Uplift charges compensate suppliers that are forced to start up or ramp up their units in order to meet net virtual demand. System-wide uplift charges are allocated to different market participants based on their individual levels of net virtual demand cleared. Appendix Figure B.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

[20] The total value of the 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. 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.

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

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

Figure B.1: Annual Uplift Charges for the 5 Major ISOs: 2009-2013



Notes: This figure is taken from FERC (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 uplift charges and electricity demand for ISO-NE in 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).

significant volatility in daily uplift charges (FERC (2014)).

C Additional Empirical Results: Price Spreads

This Appendix section discusses three additional results pertaining to differences between day-ahead and real-time electricity prices (i.e.: day-ahead/real-time price spreads). The first subsection provides empirical evidence that average day-ahead/real-time price spreads are smaller in absolute value after the introduction of explicit virtual bidding (i.e.: financial trading). We next show that the volatility of both price spreads and real-time prices fell after explicit virtual bidding (EVB) was introduced. Combined, the results presented in the first two subsections suggest that day-ahead prices better reflect real-time conditions after EVB was introduced.

In the final subsection, we test whether the daily 24×1 vector of hourly price spreads is autocorrelated over days-of-sample. The results of this analysis indicate 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.

C.1 Absolute Average Price Spreads Before Versus After EVB

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) on 2/1/2011. In particular, we formulate a test of the null hypothesis that $|\mu_{pre}| > |\mu_{post}|$, where μ_{pre} (μ_{post}) is a 24×1 vector composed of the expected day-ahead/real-time price differences for each hour-of-the-day for the sample period 4/1/2009-1/31/2011 (2/1/2011-12/31/2012). We implement this statistical test separately for each pricing location (i.e.: node).

Using the methodology derived in Wolak (1989), we compute the following test statistic in order to test the null hypothesis that $|\mu_{pre}| > |\mu_{post}|$:

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

where $\overline{X}^{pre}$ ($\overline{X}^{pre}$) 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 period. We calculate the

Table C.1: LAP-Level P-values for the Absolute Difference Tests

H_0 :	$\lvert\mu_{pre}\rvert > \lvert\mu_{post}\rvert$	$\lvert\mu_{post}\rvert > \lvert\mu_{pre}\rvert$
PG&E	0.704	0.141
SCE	0.907	0.006
SDG&E	0.689	0.040

Notes: This table reports the p-values associated with the statistical test of the null hypothesis that $|\mu_{pre}| > |\mu_{post}|$ (Column 1) as well as the statistical test of the null hypothesis that $|\mu_{post}| > |\mu_{pre}|$ (Column 2). μ_{pre} (μ_{post}) is a 24×1 vector composed of the expected day-ahead/real-time price spreads 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 4/1/2009-1/31/2011 while the post-EVB sample is 2/1/2011-12/31/2012. We perform these statistical tests on the load aggregation point (LAP) level price spreads 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).

covariance matrix $\hat{V}$ as follows:

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

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 before (after) the introduction of explicit virtual bidding. $\hat{\Sigma}^{pre}$ ($\hat{\Sigma}^{post}$) is an estimate of the asymptotic covariance matrix associated with $\overline{X}^{pre}$ ($\overline{X}^{post}$); we compute these covariance matrices using the autocorrelation consistent estimator proposed by Newey and West (1987) accounting for 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. We consider tests of size $\alpha = 0.05$ in the results presented below. Finally, the test statistic and p-value associated with the null hypothesis that $|\mu_{post}| > |\mu_{pre}|$ are computed in a similar manner.

We first perform these statistical tests on the load aggregation point (LAP) level price spreads 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

Table C.2: Proportion of Nodes for which we fail to reject the Absolute Difference Test

H_0 :	$\lvert\mu_{pre}\rvert > \lvert\mu_{post}\rvert$	$\lvert\mu_{post}\rvert > \lvert\mu_{pre}\rvert$
Gen Node	0.983	0.015
Non-Gen Node	0.988	0.013

Notes: This table reports the proportion of pricing locations (i.e.: nodes) for which we fail to reject 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 expected day-ahead/real-time price spreads for each hour-of-the-day for a given node for the sample period before (after) the introduction of explicit virtual bidding (EVB). There are 653 generation nodes and 3,961 non-generation nodes that are present in both the pre-EVB sample (4/1/2009-2/1/2011) and the post-EVB sample (2/1/2011-12/31/2012).

and Electric (SDG&E). Appendix Table C.1 presents the p-values associated with these tests. We fail to reject the null hypothesis that $|\mu_{pre}| > |\mu_{post}|$ for all three LAPs, but reject the null hypothesis that $|\mu_{post}| > |\mu_{pre}|$ for SCE and SDG&E.

We also perform our statistical tests separately for each node. Specifically, Column 1 of Appendix Table C.2 lists the proportion of nodes for which we fail to reject the null hypothesis that $|\mu_{pre}| > |\mu_{post}|$, separately for generation versus non-generation nodes. We fail to reject this null hypothesis for roughly 98% of nodes for both generation and non-generation nodes. Column 2 of Appendix Table C.2 lists the proportion of nodes for which we fail to reject the null hypothesis that $|\mu_{post}| > |\mu_{pre}|$. We fail to reject this null hypothesis for only roughly 1.5% of nodes for both generation and non-generation nodes. Combined, we have strong evidence that absolute average day-ahead/real-time price spreads fell after purely financial participation was allowed. This suggests that day-ahead prices provide better information on real-time conditions after the introduction of financial trading.

C.2 Volatility in Prices Before versus After EVB

As discussed in Section 2, we expect financial participation in wholesale electricity markets to reduce day-ahead uncertainty regarding real-time conditions. As a consequence, we should expect both the variance of day-ahead/real-time price spreads and the variance of real-time prices to fall after the introduction of explicit virtual bidding (EVB). To test this hypothesis, we compare estimates of the covariance matrices associated with each of these magnitudes before versus after EVB was introduced. For ease of exposition, we discuss the methodology underlying this volatility test referring to price spreads; we compute test statistics and p-values for the volatility of real-times prices using the same method.

Setting up notation, let $X_{d,h}$ be the realized price spread in hour-of-the-day h in day-of-sample d; let $\mathbf{X}_d$ be the 24×1 vector composed of realized price spreads for each of the 24 hours of day d. This vector is drawn independently and identically from a distribution with covariance matrix Λ^{pre} versus Λ^{post} depending on whether we're considering the sample period 4/1/2009-1/31/2011 versus 2/1/2011-12/31/2012.

The variance of price spreads is larger pre-EVB relative to post-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 $\{\hat{\omega}_j\}_{j=1}^{24}$ associated with $\hat{\Lambda}^{\text{diff}} \equiv \hat{\Lambda}^{pre} - \hat{\Lambda}^{post}$. We test the joint null hypothesis that all of these eigenvalues are greater than or equal to zero using the methodology developed by Wolak (1989).

Specifically, our test statistic is:

$$TS = \min_{z \geq 0} (\hat{\Omega} - z)' [\widehat{Var(\hat{\Omega})}]^{-1} (\hat{\Omega} - z)$$

where $\hat{\Omega}$ is the 24×1 vector containing the eigenvalues $\{\hat{\omega}_j\}_{j=1}^{24}$. The covariance matrix $Var(\hat{\Omega})$ is estimated using a moving-block bootstrap procedure.[23] As described in the previous subsection, the asymptotic distribution of test statistic TS is the weighted sum of chi-squared random variables under the null hypothesis.

We also calculate the test statistic associated with the null hypothesis that price spreads are more volatile post-EVB relative to pre-EVB (i.e.: $\Lambda^{post} - \Lambda^{pre}$ is a positive semi-definite matrix). In addition, we use the same testing procedure considering real-time prices instead of price spreads. The p-values associated with each of these statistical tests, conducted separately for each load aggregation point (LAP), are documented in Appendix Table C.3.

The results listed in Column 1 of the top panel of this table indicate that we fail to

[23] We construct L moving-block re-samples separately for the sample periods before versus after EVB. For each re-sample $b \in \{1, 2, ..., L\}$, we estimate the contemporaneous covariance matrices associated with day-ahead/real-time prices spreads in each sample period (i.e.: $\hat{\Lambda}_b^{pre}$ and $\hat{\Lambda}_b^{post}$). This allows us to compute $\hat{\Lambda}_b^{\text{diff}} \equiv \hat{\Lambda}_b^{pre} - \hat{\Lambda}_b^{post}$ as well as the 24 eigenvalues associated with $\hat{\Lambda}_b^{\text{diff}}$. We denote the 24×1 vector of these eigenvalues $\hat{\Omega}_b$. Finally, we calculate the empirical covariance of $\hat{\Omega}_b$ across the L bootstrap re-samples in order to obtain our estimate of the covariance matrix associated with $\hat{\Omega}$.

Table C.3: LAP-Level P-values for Volatility Tests

		(1) Price Spread	(2) Real-Time Price
	PG&E	0.284	0.516
Pre - Post	SCE	0.509	0.697
	SDG&E	0.476	0.647
	PG&E	0.001	0.016
Post - Pre	SCE	0.001	0.034
	SDG&E	0.028	0.165

Notes: This table reports the p-values associated with our statistical test for reductions in volatility. The first (second) column of the top panel focuses on the null hypothesis that the volatility of day-ahead/real-time price spreads (real-time prices) falls after the introduction of explicit virtual bidding (EVB). The first (second) column of the bottom panel considers the null hypothesis that the volatility of price spreads (real-time prices) is higher post-EVB relative to pre-EVB. We perform these statistical tests using the load aggregation point (LAP) level day-ahead and real-times 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). To implement our test, we estimate the contemporaneous covariance matrix associated with the 24×1 vector of price spreads as well as the covariance matrix associated with 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). We then test the null hypothesis that the pre-EVB covariance matrix minus the post-EVB covariance matrix is positive semi-definite (top panel) as well as the null hypothesis that the post-EVB covariance matrix minus the pre-EVB covariance matrix is positive semi-definite (bottom panel).

reject the null hypothesis that day-ahead/real-time price spreads are more volatile before the introduction of explicit virtual bidding (EVB) for all three LAPs. Similarly, we fail to reject the null hypothesis that real-time prices are more volatile before EVB for all three LAPs (see Column 2 of the top panel of Appendix Table C.3). Focusing on the bottom panel of this table, we can reject the null hypothesis that the volatility of price spreads is higher after EVB was introduced for all three LAPs. We can reject the null hypothesis that real-time prices are more volatile after EVB for PG&E and SCE but not SDG&E.

We also conduct these statistical tests at the nodal level. Specifically, Appendix Table C.4 presents the proportion of nodes for which we *fail to* reject each of our aforementioned null hypotheses, separately for nodes associated with generation units ("Gen Nodes") versus nodes not associated with generation units ("Non-Gen Nodes"). The results of this table indicate that there aren't marked differences across generation versus non-generation locations in the proportion of nodes for which we reject any specific null hypothesis. Thus, focusing on the rows corresponding to overall proportions (labeled "Overall"), we fail to reject the null hypothesis that the volatility of price spreads is lower (higher) after EVB for 74% (17%) of locations. Similarly, the proportion of nodes for which we fail to reject the null hypothesis that the volatility of real-time prices is lower (higher) after EVB is 0.752 (0.128).

Combined, our statistical evidence indicates that the volatility of day-ahead/real-time price spreads fell after the introduction of explicit virtual bidding. This is consistent with the intuition that allowing purely financial participation resulted in day-ahead market outcomes that more closely resemble real-time market conditions. Appendix Tables C.3 and C.4 also provide statistical evidence that the volatility of real-time prices decreased after EVB was introduced. This result runs counter to the claim that financial trading increases the volatility of real-time market outcomes, causing undue stress to physical operating constraints such as transmission congestion.

C.3 Testing for Autocorrelation in Price Spreads

The statistical test of arbitrage opportunities accounting for transaction costs discussed in Section 5 considers trading strategies that vary only by hour-of-the-day. Specifically,

Table C.4: Proportion of Nodes for which we fail to reject the Volatility Test

		Price Spread	Real-Time Price
Pre-Post	Gen Node	0.736	0.751
	Non-Gen Node	0.744	0.758
	Overall	0.737	0.752
Post-Pre	Gen Node	0.168	0.132
	Non-Gen Node	0.156	0.104
	Overall	0.166	0.128

Notes: This table reports the proportion of nodes (i.e.: 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 associated with generation units ("Gen Nodes") versus nodes not associated with generation units ("Non-Gen Nodes"). We also report the overall proportions aggregated across all pricing locations ("Overall"). To implement our tests, we estimate the contemporaneous covariance matrix associated with the 24×1 vector of hourly average price spreads as well as the covariance matrix associated with the 24×1 vector of hourly average 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 price spreads and real-time prices, we test the null hypothesis that the pre-EVB covariance matrix minus the post-EVB covariance matrix is positive semi-definite (top panel) as well as the null hypothesis that the post-EVB covariance matrix minus the pre-EVB covariance matrix is positive semi-definite (bottom panel).

we do not allow our hypothetical trader to update her strategy based on information from past days. We justify this restriction on trading strategies in this subsection.

Traders submit virtual bids to buy (sell) one MWh of virtual electricity in the day-ahead market at a given location for a given hour with the obligation to sell (buy) this virtual electricity back in the real-time market at the same location for the same hour. Traders simultaneously submit virtual bids corresponding to all 24 hours of the next day to the day-ahead market. Trading strategies are thus based on the conditional expectation of the 24×1 vector of realized day-ahead/real-time price spreads for each hour-of-the-day. However, trading strategies in practice cannot be a function of information from the prior day because the values of the 24×1 vector of hourly real-time prices for day $d-1$ are not known before virtual bids are submitted to the day-ahead market for day d. Thus, restricting to trading strategies that condition only on hour-of-the-day is valid if and only all of the autocorrelation matrices associated with the time series process governing the daily vector of price spreads are zero excepting the autocorrelation matrix associated with the first lag.

We denote the τ^{th} auto-covariance matrix associated with the 24×1 vector of price spreads $\Gamma(\tau) = E[(X_t - \mu)(X_{t-\tau} - \mu)']$. Consistent with our above discussion, we expect $\Gamma(1)$ to be non-zero but restrict $\Gamma(\tau) = 0$ for all $\tau > 1$. We thus formulate a statistical test of the following null hypothesis:

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

for a fixed value of R. Empirically, we set $R = 10$.

To implement this hypothesis test, we first define:

$$\xi \equiv [vec(\Gamma(2))', vec(\Gamma(3))', ..., vec(\Gamma(R))']'$$

where the $vec(.)$ operator takes each 24×24 auto-covariance matrix and stacks it columnwise to create a 576×1 vector. Therefore, ξ has $5{,}184 = 576 \times 9$ elements, all of which must equal zero under the null hypothesis. We use the moving block bootstrap discussed in the previous subsection in order to estimate the $5{,}184 \times 5{,}184$ covariance matrix associated with $\hat{\xi}$. Our Wald statistic $TS = \hat{\xi}'\hat{\Sigma}^{-1}_{\xi,boot}\hat{\xi}$ is asymptotically chi-squared distributed with $576 \times (R-1)$ degrees of freedom under the null hypothesis.

We first conduct this statistical test separately for each of three load aggregation points (LAPs) both before and after EVB was introduced. We consider the LAP-level price spreads faced by each of California's three major investor-owned utilities: Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E). Appendix Table C.5 reports the resulting test statistics; the upper $\alpha = 0.05$ critical value for these test statistics is $\chi^2_{(5{,}184)} = 5{,}352.6$. We fail to reject the null hypothesis that the second through tenth auto-covariance matrices are zero for all three LAPs both before and after the introduction of EVB.

We also conduct these autocorrelation tests at the nodal level, reporting the results in Appendix Table C.6. Prior to the introduction of EVB, we can reject the null hypothesis of no second through tenth degree autocorrelation at approximately 70-75 percent of nodes; these rejections occur less frequently at locations where electricity generation is injected (i.e: "generation nodes"). This is consistent with the fact that, before EVB, only suppliers could arbitrage price spreads and they could do so only by adjusting their physical bids at the locations corresponding to their generation units.

In contrast, after EVB is introduced, we reject the null hypothesis of no second

Table C.5: Test Statistics for Autocorrelation ($1 < L \leq 10$) in Daily Price Spreads

	Before EVB	After EVB
PG&E	2,862.2	2,767.0
SCE	2,789.2	2,842.6
SDG&E	3,082.1	2,700.7

Notes: This table presents chi-squared test statistics corresponding to the null hypothesis that the second through tenth auto-covariance matrices associated with the 24×1 vector of day-ahead/real-time price spreads for each hour-of-the-day are zero; formally, we are testing the null hypothesis that $\Gamma(2) = \Gamma(3) = ... = \Gamma(10) = 0$. We perform this test separately for each load aggregation point (LAP) before versus after the introduction of explicit virtual bidding (EVB). We consider the LAP-level price spreads faced by each of California's three major investor-owned electricity distribution utilities: Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E). To implement this statistical test, we first estimate auto-covariance matrices $(\Gamma(2), \Gamma(3), ..., \Gamma(10))$ and stack the elements columnwise. This results in a $5{,}184 = 24^2 \times 9$ element vector. We use a moving block bootstrap procedure in order to estimate the covariance matrix associated with this vector. The upper $\alpha = 0.05$ critical value for these test statistics is $\chi^2_{(5,184)} =$ 5,352.6.

through tenth degree autocorrelation at only approximately 7-9% percent of both generation and non-generation nodes. This is consistent with the intuition that financial traders can quickly take advantage of any systematic auto-correlation in price spreads after EVB is introduced. Summarizing, the results from this subsection provide evidence that traders cannot earn significantly more profits by conditioning on previous realizations of price spreads. This helps to justify our focus in Section 5 on trading strategies that do not condition on past lags of price spreads.

Table C.6: Proportion of Nodes for which we fail to reject the Autocorrelation Test

	Before EVB	After EVB
Non-Generation Nodes	0.299	0.912
Generation Nodes	0.265	0.932

Notes: This table presents the proportion of nodes (i.e.: pricing locations) for which we fail to reject a size $\alpha = 0.05$ test of the null hypothesis that the second through tenth auto-covariance matrices of the 24×1 vector of day-ahead/real-time price spreads for each hour-of-the-day are zero; formally, 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). To implement this test, we estimate auto-covariance matrices $(\Gamma(2), \Gamma(3), ..., \Gamma(10))$ and stack the elements column-wise; this results in a 5,184 $= 24^2 \times 9$ element vector. We use a moving block bootstrap procedure in order to estimate the covariance matrix associated with this vector. The upper $\alpha = 0.05$ critical value for these test statistics is $\chi^2_{(5{,}184)} = 5{,}352.6$. Generation units inject electricity at some locations on the transmission grid; these locations are called "generation nodes". The remaining locations are categorized as "non-generation" nodes. The number of generation nodes (non-generation nodes) in the pre-EVB sample is 669 (4,031) and the number of generation nodes (non-generation nodes) in the post-EVB sample is 673 (4,386).

D Difference-in-Differences Robustness Checks

This Appendix section describes robustness checks pertaining to the difference-in-differences results presented in Table 4.

D.1 Pre-Trends Analysis

In this subsection, we argue that the results presented in Table 4 are not driven by pre-existing differences in the trends in outcomes across high versus low demand hours. Specifically, in order to interpret the findings from our difference-in-differences analysis as causal, average outcomes in low versus high demand hours should have the same trends over time prior to the introduction of explicit virtual bidding (i.e.: financial trading) on 2/1/2011. We first provide descriptive evidence suggesting that trends in monthly average residualized outcomes do not seem to be different across high versus low demand hours before explicit virtual bidding (EVB) was introduced. As a more formal test of this common trends assumption, we next demonstrate that the first difference in outcomes is not statistically different across high versus low demand hours before 2/1/2011 even after controlling for the same variables as in our primary specification.

D.1.1 Trends in Residualized Outcomes

This subsubsection provides descriptive evidence that residualized trends over time in our outcome variables are not different across high versus low demand hours prior to the introduction of EVB. We proceed in three steps. First, we residualize each of our outcome variables by estimating the following regression specification:

$$Y_t = \alpha_m + \gamma_h + \theta_w + X_t\phi + u_t \tag{6}$$

where 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 control variables X_t: the log of total electricity demand, the log of total electricity imports, two separate controls for the logs of the natural gas prices paid at the PG&E and SCG citygates, 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.

We next take the monthly average of the residuals $\hat{u}_t$ generated by estimating Equation (6). We do so separately for high versus low demand hours. Denote the time series of monthly average residuals for high and low demand hours $\{\hat{\bar{u}}_m^H\}_{m=1}^M$ and $\{\hat{\bar{u}}_m^L\}_{m=1}^M$ respectively. As an example:

$$\hat{\bar{u}}_m^H = \frac{\sum_{h=0}^{K_m-1} \hat{u}_{t+h} 1(D_{t+h} > 90^{th} \text{ Percentile})}{\sum_{h=0}^{K_m-1} 1(D_{t+h} > 90^{th} \text{ Percentile})}$$

where t is the first hour-of-sample in month-of-sample m and the indicator variable $1(D_{t+h} > 90^{th}$ Percentile) is equal to one if and only if hourly demand D_{t+h} is greater than the 90^{th} percentile of the distribution of hourly demands across our 4/1/2009-3/31/2012 sample period. K_m is the number of hours in month m.

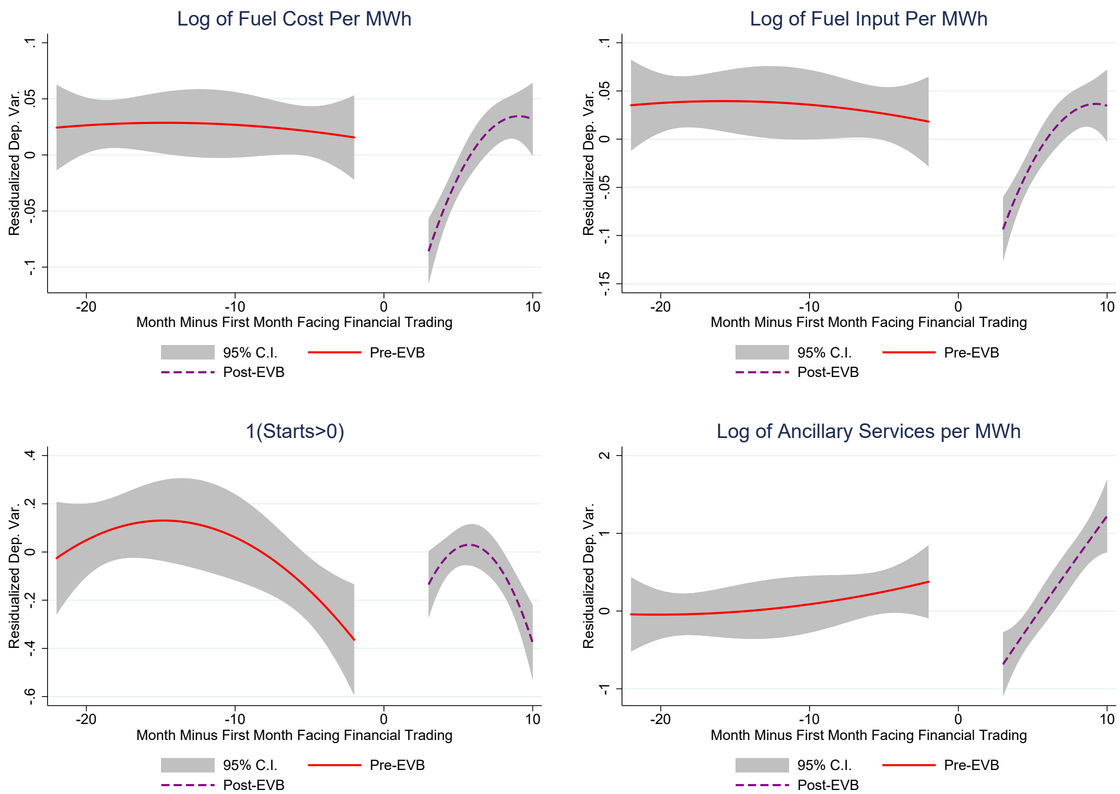
Finally, we plot the quadratic best fit of the difference between the monthly average residuals in high versus low demand hours. This quadratic fit is estimated separately for months-of-sample before versus after the introduction of EVB. Specifically, we plot the predictions from the following regression:

$$\hat{\bar{u}}_{m,p}^H - \hat{\bar{u}}_{m,p}^L = \alpha_{0,p} + \alpha_{1,p}(m - m_0) + \alpha_{2,p}(m - m_0)^2 + \epsilon_{m,p}$$

where m indexes month-of-sample in sample period p (i.e.: either 4/2009-1/2011 or 2/2011-3/2012). We center at the month-of-sample that EVB was introduced (i.e.: m_0 is 2/2011). Standard errors are not adjusted for any form of autocorrelation. This is by design; by considering anti-conservative 95% confidence intervals, we're more likely to reject the null hypothesis of common pre-existing trends. That being said, the figures presented below do not constitute a formal statistical test of the common trends assumption, but rather, a descriptive exploration of whether our findings are driven by pre-existing trends.

Appendix Figures D.1 and D.2 plot the best quadratic fit of the difference in monthly average residuals across high versus low demand hours for each of the seven outcomes considered in Table 4. Prior to the introduction of EVB on 2/1/2011, the residualized outcome is higher on average for high demand hours relative to low demand hours for all of our outcomes. This is consistent with intuition: we should expect fuel costs per MWh, input fuel per MWh, incidence of start-ups, and absolute day-ahead/real-time price spreads to be higher in high demand hours relative to low demand hours.

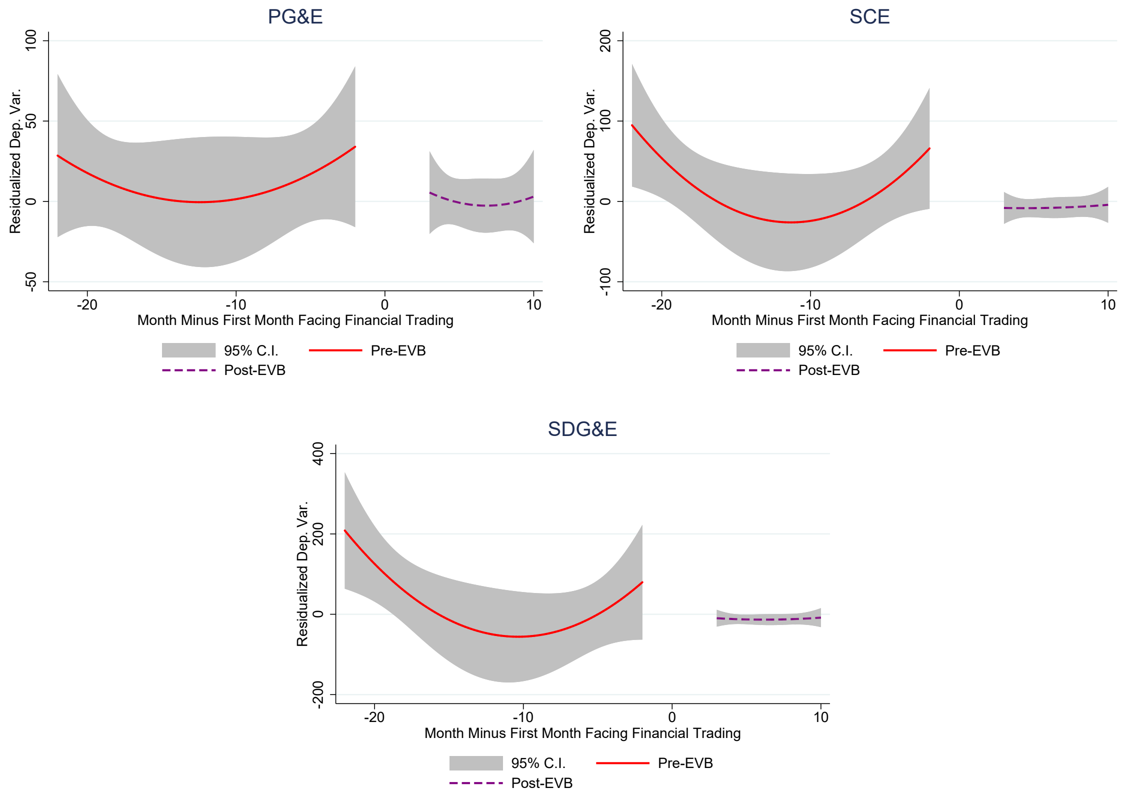
Figure D.1: Monthly Average Residualized Market Outcomes



Notes: This figure presents the quadratic fit of the difference in the monthly residualized averages across high versus low demand hours for four different outcome variables. This quadratic fit is estimated separately for months-of-sample before versus after the introduction of financial trading on 2/1/2011. The dependent variables considered in this figure are the log of fuel costs per MWh (top left panel), the log of input energy per MWh (top right panel), an indicator that's equal to one if and only if any fossil-fuel-fired units started up (bottom left panel), and the log of ancillary service costs per MWh (bottom right panel). The gray-shaded area is the 95% confidence interval associated with this quadratic fit, based on unadjusted standard errors. The x-axis plots the number of months away from 2/2011. To construct the monthly average residualized outcome in high versus low demand hours, we first regress each outcome on 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, the log of net electricity imports, two separate controls for the logs of the natural gas prices paid at the PG&E and SCG citygates 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. Next, we take the monthly average of the residuals from this regression, separately for high versus low demand hours.

Moreover, for every outcome except the indicator for at least one unit starting up in the hour, we see that the residualized trend in outcome is either flat or moving slightly in the opposite direction of the estimated effect prior to 2/2011. For example, the monthly average of residualized absolute day-ahead/real-time price spreads for each of California's three major service territories (PG&E, SCE, SDG&E) is trending slightly upward as 2/2011 approaches, but drops after EVB is introduced on 2/1/2011. Thus, if anything, the pre-existing trends in these outcomes would bias us **against** finding that financial

Figure D.2: Monthly Average Residualized Absolute Price Spreads



Notes: This figure presents the quadratic fit of the difference in the monthly residualized averages across high versus low demand hours for four different outcome variables. This quadratic fit is estimated separately for months-of-sample before versus after the introduction of financial trading on 2/1/2011. The dependent variables considered in this figure are the absolute day-ahead/real-time price spreads faced by each of California's three major electricity distribution utilities: PG&E (top left panel), SCE (top right panel), and SDG&E (bottom middle panel). The gray-shaded area is the 95% confidence interval associated with this quadratic fit, based on unadjusted standard errors. The x-axis plots the number of months away from 2/2011. To construct the monthly average residualized outcome in high versus low demand hours, we first regress each outcome on 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, the log of net electricity imports, two separate controls for the logs of the natural gas prices paid at the PG&E and SCG citygates 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. Next, we take the monthly average of the residuals from this regression, separately for high versus low demand hours.

trading reduced absolute price spreads in high demand hours.

D.1.2 Statistical Test of Common Trends Using 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 1^{st} 2011:

$$Y_t - Y_{t-1} = \alpha_m + \gamma_h + \theta_w + X_t\phi + \beta_0 \mathrm{HIGH}_t + u_t \tag{7}$$

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, the log of net electricity imports, two separate controls for the logs of the natural gas prices paid at the PG&E and SCG citygates, 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. As before, HIGH_t is an indicator variable that's equal to one if and only if system-wide demand in hour-of-sample t is greater than the 90^{th} percentile of the distribution of hourly demands across our 4/1/2009-3/31/2012 sample period. Finally, we cluster standard errors at the day-of-sample.

Appendix Table D.1 presents the results from estimating Equation (7). These results indicate that the first difference of the outcome variable is not statistically different across high versus low demand hours prior to 2/1/2011 for all of our outcome variables. This provides formal statistical evidence that the findings from our difference-in-differences framework are not driven by pre-existing differences in the time trend of our outcomes in high versus low demand hours.

D.2 Additional Robustness Checks

This subsection considers additional sensitivity analyses pertaining to the difference-in-differences results presented in Table 4. First, Appendix Table D.2 shows that our empirical results remain similar if we estimate Equation (5) in Section 7 dropping the 28 days before and after the introduction of financial trading on 2/1/2011. Thus, our findings are unlikely to be driven by short-run adjustments to this policy change.

Next, Appendix Table D.3 presents the results from estimating Equation (5) using only data from the 6 months before and after 2/2011. The empirical results are quantitatively quite similar to our primary findings from Table 4. Further, Appendix Table D.4 presents estimates from our primary specification along with standard errors calculated

Table D.1: Checking For Common Pre-Existing Trends Using First-Differences

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1(Demand > 90%)	0.001	0.001	-0.008	-0.000	0.620	-0.422	-1.501
	(0.001)	(0.001)	(0.009)	(0.008)	(0.858)	(1.807)	(1.724)
R^2	0.030	0.018	0.018	0.112	0.015	0.013	0.010
Mean of Dep. Var.	0	-0.00001	-0.00001	0	-0.00721	-0.00684	-0.00672
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 day-of-sample and are reported in parentheses. The dependent variables considered in this table are the first differences of: (1) the log of fuel costs per MWh, (2) the log of input energy per MWh, (3) an indicator variable that's equal to one if and only if at least one fossil-fuel-fired unit started up, (4) the log of ancillary service costs per MWh, (5) the absolute value of the day-ahead/real-time price spread in PG&E, (6) the absolute value of the day-ahead/real-time price spread in SCE, and (7) the absolute value of the day-ahead/real-time price spread in SDG&E. The row titled "Mean of Dep. Var." reports the mean of the relevant dependent variable. All of the regressions listed in this table 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, the log of total net electricity imports, two separate controls for the logs of the natural gas prices paid at the PG&E and SCG citygates 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 4. Specifically, we estimate the specification documented in Equation (7) in Appendix Section D.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. The indicator variable 1(Demand > 90%) is equal to one if and only if system-wide electricity demand in the hour is greater than the 90^{th} percentile of the distribution of hourly demands across our 4/1/2009-3/31/2012 sample period.

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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1(Demand > 90%)	-0.033***	-0.042***	-0.042*	-0.062	-5.833**	-14.090***	-21.664***
× 1(Post EVB)	(0.006)	(0.007)	(0.022)	(0.072)	(2.699)	(4.668)	(7.433)
1(Demand > 90%)	0.027***	0.035***	0.003	0.142***	3.205*	6.831**	8.395
	(0.003)	(0.003)	(0.017)	(0.035)	(1.860)	(3.407)	(5.247)
R^2	0.902	0.366	0.127	0.610	0.041	0.039	0.050
Mean of Dep. Var.	3.579	2.081	0.684	0.026	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 day-of-sample and are reported in parentheses. The dependent variables considered in this table are: (1) the log of fuel costs per MWh, (2) the log of input energy per MWh, (3) an indicator variable that's equal to one if and only if at least one fossil-fuel-fired unit started up, (4) the log of ancillary service costs per MWh, (5) the absolute value of the day-ahead/real-time price spread in PG&E, (6) the absolute value of the day-ahead/real-time price spread in SCE, and (7) the absolute value of the day-ahead/real-time price spread in SDG&E. The row titled "Mean of Dep. Var." reports the mean of the dependent variable. All of the regressions listed in this table 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, the log of total electricity imports, two separate controls for the logs of the natural gas prices paid at the PG&E and SCG citygates 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 4, 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 equal to one if and only if the day-of-sample is on or after 2/1/2011. The indicator variable 1(Demand > 90%) is equal to one if and only if system-wide electricity demand in the hour is greater than the 90^{th} percentile of the distribution of hourly demands across our 4/1/2009-3/31/2012 sample period.

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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1(Demand > 90%)	-0.051***	-0.062***	-0.006	-0.463***	-9.740**	-14.367***	-12.471**
× 1(Post EVB)	(0.009)	(0.010)	(0.033)	(0.094)	(4.725)	(5.477)	(5.370)
1(Demand > 90%)	0.056***	0.064***	-0.032	0.303***	0.092	2.245	-1.367
	(0.007)	(0.008)	(0.033)	(0.065)	(4.260)	(4.932)	(5.062)
R^2	0.668	0.375	0.120	0.665	0.038	0.040	0.038
Mean of Dep. Var.	3.646	2.102	0.706	0.332	15.298	17.023	18.270
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 day-of-sample and are reported in parentheses. The dependent variables considered in this table are: (1) the log of fuel costs per MWh, (2) the log of input energy per MWh, (3) an indicator variable that's equal to one if and only if at least one fossil-fuel-fired unit started up, (4) the log of ancillary service costs per MWh, (5) the absolute value of the day-ahead/real-time price spread in PG&E, (6) the absolute value of the day-ahead/real-time price spread in SCE, and (7) the absolute value of the day-ahead/real-time price spread in SDG&E. The row titled "Mean of Dep. Var." reports the mean of the dependent variable. All of the regressions listed in this table 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, the log of total electricity imports, two separate controls for the logs of the natural gas prices paid at the PG&E and SCG citygates 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 4, the data used for this table span the sample period 8/1/2010-8/31/2011 (i.e.: the 6 months before and after 2/2011). California introduced explicit virtual bidding (EVB) on 2/1/2011; the "Post EVB" indicator is thus equal to one if and only if the day-of-sample is on or after 2/1/2011. The indicator variable 1(Demand > 90%) is equal to one if and only if system-wide electricity demand in the hour is greater than the 90^{th} percentile of the distribution of hourly demands across our 4/1/2009-3/31/2012 sample period.

using the Newey-West formula accounting for 168 hours (i.e.: one week) of autocorrelation in the outcome variables (Newey and West, 1987). Our conclusions remain the same when assessing the results from this table: financial trading reduces fuel costs per MWh and absolute price spreads in high demand hours without exacerbating physical constraints such as generation unit start-ups or the need for ancillary services. Finally, Appendix Table D.5 demonstrates that our estimates remain economically and statistically significant if we consider day-of-sample fixed effects rather than month-of-sample fixed effects.

Finally, Appendix Table D.6 presents our specifications estimated on data aggregated to the daily level. In particular, we take the daily sum over hours-of-the-day of electricity demand, electricity imports, fuel costs, thermal input energy, electricity production by

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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1(Demand > 90%)	-0.031***	-0.040***	-0.036	-0.036	-5.742*	-14.024**	-21.292**
× 1(Post EVB)	(0.011)	(0.012)	(0.028)	(0.133)	(3.049)	(5.642)	(9.251)
1(Demand > 90%)	0.033***	0.041***	0.022	0.211***	2.428	5.925	8.575
	(0.005)	(0.006)	(0.023)	(0.054)	(2.137)	(4.439)	(6.202)
R^2	0.900	0.370	0.127	0.605	0.041	0.039	0.049
Mean of Dep. Var.	3.583	2.082	0.686	0.014	13.199	15.055	16.426
Number of Obs.	26,277	26,277	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 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 costs per MWh, (2) the log of input energy per MWh, (3) an indicator variable that's equal to one if and only if at least one fossil-fuel-fired unit started up, (4) the log of ancillary service costs per MWh, (5) the absolute value of the day-ahead/real-time price spread in PG&E, (6) the absolute value of the day-ahead/real-time price spread in SCE, and (7) the absolute value of the day-ahead/real-time price spread in SDG&E. The row titled "Mean of Dep. Var." reports the mean of the dependent variable. All of the regressions listed in this table 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, the log of total electricity imports, two separate controls for the logs of the natural gas prices paid at the PG&E and SCG citygates 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 equal to one if and only if the day-of-sample is on or after 2/1/2011. The indicator variable 1(Demand > 90%) is equal to one if and only if system-wide electricity demand in the hour is greater than the 90^{th} percentile of the distribution of hourly demands across our 4/1/2009-3/31/2012 sample period.

Table D.5: Diff-in-Diff Robustness Check: Day-of-Sample Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1(Demand > 90%)	-0.026***	-0.031***	-0.021	-0.118*	-7.880***	-14.939***	-21.595***
× 1(Post EVB)	(0.006)	(0.007)	(0.025)	(0.071)	(2.688)	(4.504)	(6.525)
1(Demand > 90%)	0.034***	0.040***	0.023	0.217***	1.176	2.222	3.661
	(0.003)	(0.003)	(0.019)	(0.035)	(2.082)	(2.993)	(4.640)
R^2	0.918	0.470	0.175	0.701	0.126	0.133	0.148
Mean of Dep. Var.	3.583	2.082	0.686	0.014	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 day-of-sample and are reported in parentheses. The dependent variables considered in this table are: (1) the log of fuel costs per MWh, (2) the log of input energy per MWh, (3) an indicator variable that's equal to one if and only if at least one fossil-fuel-fired unit started up, (4) the log of ancillary service costs per MWh, (5) the absolute value of the day-ahead/real-time price spread in PG&E, (6) the absolute value of the day-ahead/real-time price spread in SCE, and (7) the absolute value of the day-ahead/real-time price spread in SDG&E. The row titled "Mean of Dep. Var." reports the mean of the dependent variable. All of the regressions listed in this table control for day-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, the log of total electricity imports, two separate controls for the logs of the natural gas prices paid at the PG&E and SCG citygates 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 equal to one if and only if the day-of-sample is on or after 2/1/2011. The indicator variable 1(Demand > 90%) is equal to one if and only if system-wide electricity demand in the hour is greater than the 90^{th} percentile of the distribution of hourly demands across our 4/1/2009-3/31/2012 sample period.

Table D.6: Diff-in-Diff Robustness Check: Daily-Level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1(Demand > 90%)	-0.020**	-0.032**	-1.225*	-0.301*	-2.827	-2.500	-2.755
× 1(Post EVB)	(0.009)	(0.012)	(0.713)	(0.150)	(2.872)	(3.485)	(3.890)
1(Demand > 90%)	0.035***	0.046***	0.997*	0.297***	3.270	-0.930	-3.776
	(0.006)	(0.007)	(0.518)	(0.079)	(2.325)	(2.476)	(5.362)
R^2	0.983	0.759	0.382	0.851	0.125	0.104	0.179
Mean of Dep. Var.	3.580	2.079	16.455	0.170	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 month-of-sample and are reported in parentheses. The dependent variables considered in this table are: (1) the log of fuel costs per MWh, (2) the log of input energy per MWh, (3) the log of the total number of fossil-fuel-fired units that started up, (4) the log of ancillary service costs per MWh, (5) the absolute value of the day-ahead/real-time price spread in PG&E, (6) the absolute value of the day-ahead/real-time price spread in SCE, and (7) the absolute value of the day-ahead/real-time price spread in SDG&E. The row titled "Mean of Dep. Var." reports the mean of the dependent variable. All of the regressions listed in this table 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, the log of total electricity imports, two separate controls for the logs of the natural gas prices paid at the PG&E and SCG citygates 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 equal to one if and only if the day-of-sample is on or after 2/1/2011. The indicator variable 1(Demand > 90%) is equal to one if and only if system-wide total electricity demand for the day is greater than the 90^{th} percentile of the distribution of daily demands across our 4/1/2009-3/31/2012 sample period.

type (thermal, nuclear, hydro, and renewables), number of generation unit start-ups and ancillary service costs. We then divide daily total fuel costs, thermal input energy and ancillary services costs by daily total electricity production from fossil-fuel-fired units in order to construct four of the outcome variables considered in Appendix Table D.6. The outcome variable considered in Column 3 of this table is the log of the total number of generation unit start-ups in the day.[24] Finally, we average hourly day-ahead and real-time electricity prices for each service area to the daily-level. Specifically, our last three outcome measures are the absolute differences between the daily average day-ahead and real-time prices faced by PG&E, SCE, and SDG&E.

This difference-in-differences specification thus considers days-of-sample above versus below the 90^{th} percentile of the distribution of daily total electricity demand before versus after the introduction of financial trading. Standard errors are clustered by month-of-sample. The results of this analysis, presented in Appendix Table D.6, are broadly consistent with the results presented in Table 4 with one exception: the estimated effects for absolute day-ahead/real-time price spreads are no longer statistically significant. This is unsurprising given that the reductions in price spreads due to financial trading are likely to be largest for the hours of the day with the highest average demand.

D.3 Difference-in-Differences Using Daily Number of Starts

This subsection compares market outcomes before versus after EVB on days with more versus less generation unit start-ups. To see the intuition behind this empirical approach, consider the following example. Suppose that a generation unit must start up at some point during the day in order to satisfy the maximum level of hourly electricity demand across the 24 hours of the day (i.e.: "peak demand"). Suppose further that the unit's owner has better day-ahead information regarding when real-time demand will be the highest or the level of this peak demand as a consequence of purely financial participation. Based on this information, the supplier may be able to commit a unit with lower marginal costs but higher start-up times to produce during the peak hours of the day rather than a higher marginal cost unit that can start up more quickly.

[24]We consider the log of total starts rather than an indicator variable that's equal to one if at least one thermal unit started up during the day because at least one fossil-fuel-fired unit started up in every day-of-sample.

We estimate the following specification in order to quantify how outcomes change before versus after EVB on days with a relatively large number of starts:

$$Y_t = \alpha_m + \gamma_h + \theta_w + X_t\phi + \beta_0 \mathrm{HIGH}_d + \delta_{DD}(\mathrm{HIGH}_d \times \mathrm{POSTEVB}_t) + u_t \tag{8}$$

where t indexes each hour-of-sample in day-of-sample d in month-of-sample m. As before, 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, the log of net electricity imports, two separate controls for the logs of the natural gas prices paid at the PG&E and SCG citygates, 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 HIGH_d is equal to one if and only if the total number of generation unit start-ups in day-of-sample d is larger than the 90^{th} percentile of the distribution of daily total number of starts. The indicator variable $\mathrm{POSTEVB}_t$ is equal to one if and only if the day-of-sample is on or after 2/1/2011. Finally, standard errors are clustered by day-of-sample.

Appendix Table D.7 presents the results from estimating Equation (8). Column 1 of this table indicates that average fuel costs per MWh fell after EVB on days with a large number of starts relative to days with a small number of starts. This reduction is consistent with the hypothesis that peak demand is more likely to be met by low-cost, long-start units rather than high cost, quick-start units as a consequence of the day-ahead information on real-time conditions provided by purely financial participation.

The results for the other outcome variables tell a similar story. Specifically, Columns 2 and 3 document reductions in average heat energy per MWh and ancillary service costs per MWh after EVB on days with a relatively large number of starts. This suggests that the costs associated with balancing supply and demand at every instant (i.e.: ancillary service costs) are lower as a consequence of financial participation. Finally, Columns 4-6 document reductions in the absolute differences between the day-ahead and real-time electricity prices faced by Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E) respectively, though only the effect for SDG&E is statistically significant at conventional levels.

Finally, Appendix Table D.8 demonstrates that our estimates remain similar when

Table D.7: Diff-in-Diff: High vs. Low Daily Starts Before vs. After EVB

	(1)	(2)	(3)	(4)	(5)	(6)
1(Daily Starts > 90%)	-0.014**	-0.017**	-0.202***	-3.349	-5.533	-11.328**
× 1(Post EVB)	(0.006)	(0.007)	(0.064)	(3.520)	(3.865)	(5.749)
1(Daily Starts > 90%)	0.029***	0.034***	0.016	10.230***	10.081***	14.630***
	(0.004)	(0.004)	(0.040)	(2.816)	(3.026)	(4.817)
R^2	0.900	0.371	0.605	0.045	0.040	0.050
Mean of Dep. Var.	3.583	2.082	0.014	13.199	15.055	16.426
Number of Obs.	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 day-of-sample and are reported in parentheses. The dependent variables considered in this table are: (1) the log of fuel costs per MWh, (2) the log of input energy per MWh, (3) the log of ancillary service costs per MWh, (4) the absolute value of the day-ahead/real-time price spread in PG&E, (5) the absolute value of the day-ahead/real-time price spread in SCE, and (6) the absolute value of the day-ahead/real-time price spread in SDG&E. The row titled "Mean of Dep. Var." reports the mean of the dependent variable. All of the regressions listed in this table 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, the log of total electricity imports, two separate controls for the logs of the natural gas prices paid at the PG&E and SCG citygates 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 days with more versus less generation unit start-ups. 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 equal to one if and only if the day-of-sample is on or after 2/1/2011. The indicator variable 1(Daily Starts > 90%) is equal to one for days-of-sample in which the number of instances that electricity generating units started up is greater than the 90^{th} percentile of the distribution of daily total number of starts across our 4/1/2009-3/31/2012 sample period.

Table D.8: Diff-in-Diff Robustness Check: By Percentage of Daily Starts

	Log of Average Fuel Cost Per MWh				
	(1)	(2)	(3)	(4)	(5)
1(Starts > Cut-off)	-0.007**	-0.014***	-0.014**	-0.023**	0.000
× 1(Post EVB)	(0.004)	(0.004)	(0.006)	(0.009)	(.)
1(Starts > Cut-off)	0.015***	0.022***	0.029***	0.036***	0.044***
	(0.002)	(0.002)	(0.004)	(0.005)	(0.008)
Daily Starts Cut-Off	50%	75%	90%	95%	99%
Mean of Dep. Var.	3.583	3.583	3.583	3.583	3.583
R^2	0.900	0.900	0.900	0.900	0.900
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 day-of-sample and are reported in parentheses. The row titled "Mean of Dep. Var." reports the mean of the log of fuel costs per MWh. All of the regressions listed in this table 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, the log of total electricity imports, two separate controls for the logs of the natural gas prices paid at the PG&E and SCG citygates 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 days with more versus less generation unit start-ups. 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 equal to one if and only if the day-of-sample is on or after 2/1/2011. The indicator variable 1(Starts > Cut-off) is equal to one for days-of-sample in which the number of instances that electricity generating units started up is greater than the X^{th} percentile of the distribution of daily total number of starts across our 4/1/2009-3/31/2012 sample period; X is equal to the 50^{th}, 75^{th}, 90^{th}, 95^{th}, or 99^{th} percentile depending on whether we're considering the specification estimated in Columns 1, 2, 3, 4, or 5 respectively.

considering days-of-sample with total number of starts above the 50^{th}, 75^{th}, or 95^{th} percentiles of the distribution of daily total starts rather than 90^{th} percentile. In particular, the estimated reduction in average fuel costs per MWh on days with a relatively large number of starts due to EVB is roughly 1%-2% across specifications. This magnitude is a bit smaller than the corresponding effect for high demand hours estimated in Table 4. This is intuitive: financial participation likely reduces production costs through channels other than changes in which units start up. For example, the information provided by financial traders may help market operators better manage transmission congestion, allowing the electricity produced by lower-cost units to be transmitted to satisfy demand at potentially distant locations.