Testing for Market Efficiency with Transaction Costs: An Application to Financial Trading in Wholesale Electricity Markets

Akshaya Jha*

and

Frank A. Wolak†

September 6, 2015

Abstract

With risk neutral traders and zero transaction costs, the expected value of the difference between the current forward price and the spot price of a commodity at the delivery date of the forward contract should be zero. Accounting for the transaction costs associated with trading in these two markets invalidates this result. We develop a statistical framework to test whether profitable trading strategies exploiting systematic differences between spot and forward market prices exist in the presence of trading costs. We implement these tests using the day-ahead forward and real-time spot locational marginal prices from California's wholesale electricity market. We use our statistical tests to construct an estimate of the variable cost

*H. John Heinz III College, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213. Email: akshayaj@andrew.cmu.edu.

†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

of trading in this market. During our sample period, we observe the introduction of financial trading, which was aimed at reducing the costs associated with exploiting differences between forward and spot prices. Consistent with this aim, our measures of trading costs are significantly smaller after the introduction of financial trading. Prior to financial trading, day-ahead/real-time price differences could be exploited more readily at locations where generation is injected ("generation nodes"). Consistent with this, our estimated trading costs are lower for generation nodes relative to non-generation nodes before financial trading and trading costs fell more for non-generation nodes after financial trading, eliminating any difference in trading costs across the two types of nodes. We also present evidence that the introduction of financial trading reduced the total amount of input fossil fuel energy required to generate the thermal-based electricity produced in California and the total variable of costs of producing this electrical energy. Taken together, these results demonstrate that purely financial forward market trading can improve the operating efficiency of short-term commodity markets.

1 Introduction

Many commodities are traded in both forward and spot markets. With risk neutral arbitrageurs and zero transactions costs, market efficiency implies that the forward price at time t for delivery k periods in the future (F_{t+k}) is equal to the expected value of the spot price k periods in the future conditional on the information available to market participants at time t ($E_t[P_{t+k}]$). Namely, $F_{t+k} = E_t[P_{t+k}]$. After accounting for transactions costs, the existence of a profitable trading strategy implies the $|F_{t+k} - E_t[P_{t+k}]| > c$, where c is the per-unit dollar cost associated with trading in both the forward and spot markets. Specifically, the expected profits from exploiting the difference between the forward and spot price is greater than the trading costs. This paper develops a statistical framework that tests whether or not profitable trading opportunities exist in a commodity market with transaction costs. We apply this testing framework to data from California's wholesale electricity market, and derive an estimate of the per unit trading cost, c.

Wholesale electricity markets with a day-ahead forward market and real-time market are ideally suited to test for the existence of profitable trading strategies because exactly the same product—electrical energy delivered during an hour of the day—is sold in the day-ahead and real-time markets. Moreover, the time lag between the purchase/sale in the forward market and subsequent sale/purchase in the real-time market is less than one day. However, our statistical tests are complicated by the fact that trading occurs daily, with the opportunity to trade day-ahead/real-time price differences corresponding to any of the 24 hours of the day. Therefore, we base our tests on a trader with access to 24 assets corresponding to the day-ahead/real-time price spreads for different hours of the day. Using this framework, we can test both the null hypothesis that profitable trading strategies *do* exist and the null hypothesis that profitable trading strategies *do not* exist.

This analysis also has implications for the design of wholesale electricity markets because of the controversial role that purely financial traders play in these markets. Regulators have been reluctant to allow explicit financial transactions in day-ahead and real-time energy markets despite the benefits typically associated with financial markets. For example, without financial trading, it is impossible to determine whether a market participant sells (buys) a different amount of energy in the day-ahead market than their real-time production (consumption) because of new information about real-time demand or supply conditions after the close of the day-ahead market or because the market participant is attempting to profit from anticipated differences between prices in the day-ahead and real-time markets.

Exploiting anticipated differences between day-ahead and real-time prices without financial trading involves costly actions by generation unit owners and load-serving entities that can have adverse system reliability consequences. For example, if a generation unit owner expects the real-time market price to be higher than the day-ahead price, the unit owner is likely to delay selling its output until the real-time market. If enough generation unit owners share these expectations, the system operator will find that the day-ahead market clears at a level of demand below expected real-time demand. The independent system operator (ISO) must therefore purchase a substantial amount of energy in the real-time market to meet actual demand, which can be extremely challenging for the ISO to manage and can increase the total cost of serving final demand. These concerns were ultimately realized in a number of United States (U.S.) wholesale markets, which led to the introduction of explicit virtual bidding (also termed convergence bidding)—a purely financial product that is designed to allow market participants to profit from expected price differences between the day-ahead and real-time markets without these potential reliability consequences and production cost increases.

Explicit virtual bidding was implemented on February 1, 2011 in the California wholesale electricity market. It allows market participants to take purely financial positions in the day-ahead market that must be closed out in the real-time market. A trader that sells energy in the day-ahead market using an incremental or INC virtual bid has an obligation to buy back the same amount of energy as a price-taker in the real-time market. The payoff from this transaction, before accounting for trading costs, is the difference between the day-ahead and real-time prices for that hour times the number of megawatt-hours (MWhs) sold in the day-ahead market. Buying energy in the day-ahead market using a decremental or DEC virtual bid implies an obligation to sell that same amount of energy in the real-time market as a price-taker. This transaction has revenue equal to the difference between the real-time price and the day-ahead price for that hour times the number of MWhs purchased.

Explicit virtual bidding was introduced for two major reasons: (1) to reduce the cost to market participants of exploiting price differences between the day-ahead and real-time markets, and (2) to reduce the total cost of serving demand at all locations in the transmission network in real-time. We present evidence that explicit virtual bidding achieved both of these goals. Specifically, our measures of the implied per-unit cost of trading day-ahead versus real-time price differences fell for the three major pricing zones and at virtually all of the more than 4,000 price locations in the California ISO control area after the introduction of explicit virtual bidding. We also find that the variance of the difference between day-ahead and real-time prices declined and the variance of the real-time price declined after the introduction of explicit virtual bidding. Finally, we find that the total hourly input fossil fuel energy consumed fell by 6.2 percent and the total hourly variable cost of producing fossil fuel-fired electricity in California fell by 6.8 percent after the introduction of explicit virtual bidding.

The remainder of the paper proceeds as follows. The next section describes the mechanism used to set locational marginal prices and determine dispatch levels in the day-ahead and real-time markets in California and all other bid-based markets in the United States. This section also describes how the actions of generation unit owners and load serving entities influence locational marginal prices in the absence of explicit virtual bidding as well as how virtual bids influence locational marginal prices in the day-ahead and real-time markets. Section 3 describes the data used to perform our hypothesis test and presents descriptive statistics on the behavior of the average hourly differences in the day-ahead and real-time prices before versus after the implementation of explicit virtual bidding. Section 4 derives our statistical testing framework regarding the existence of a profitable trading strategy with transactions costs. We also demonstrate in this section that the empirical distribution of trading costs implied by our hypothesis test is shifted downwards after the introduction of explicit virtual bidding. We also show empirically that explicit virtual bidding actually *reduces* the variance of day-ahead/real-time prices spreads. Section 5 presents our analysis of the market efficiency consequences associated with implementing explicit virtual bidding. We conclude in Section 6 by discussing the implications of our results for the design of wholesale electricity markets.

2 Locational Marginal Pricing and Explicit Virtual Bidding in the California Market

This section first describes the important features of multi-settlement locational marginal pricing (LMP) wholesale electricity markets that currently exist throughout the United States. We explain how a market participant's actions are used to determine the electricity prices received by generation unit owners and paid by load serving entities in the day-ahead and real-time markets. We next elaborate on how suppliers and load-serving entities (read: demanders) exploited expected price differences between the day-ahead and real-time markets before the introduction of explicit virtual bidding. We then provide details on the mechanics of explicit virtual bidding, including how these purely financial transactions influence day-ahead and real-time locational marginal prices. Finally, we list the transactions costs associated with exploiting expected differences between day-ahead and real-time prices with and without explicit virtual bidding.

2.1 Locational Marginal Pricing in Multi-Settlement Markets

Short-term wholesale electricity markets differ from markets for other products because the electricity produced by a generation unit at one location and sold to a customer at another location is not actually delivered to that location in the same sense that an automobile produced in Detroit is delivered to the customer that purchased it in San Francisco. Energy injected into the transmission network flows according to Kirchhoff's laws, rather than from the seller to the buyer of the energy. The capacity of the transmission network often limits the amount that generation units at certain locations can inject and the amount that consumers at certain locations can withdraw. This circumstance is referred to as transmission congestion and it can cause a wholesale electricity market to become segmented, meaning that some generation units cannot compete to sell energy at certain locations in the transmission network. Among other reasons, segmentation can result from the configuration of the transmission network, the locations and outputs of other generation units, and the locations and levels of final demand. Due to this, a market mechanism that assumes that all generation units in the geographic region covered by the wholesale market can compete to sell energy anywhere in that geographic region will likely produce an infeasible dispatch of the available generation units, because capacity constraints in the transmission network and other operating constraints prevent the suppliers that offer the lowest prices for their output from selling all of their available energy.

For this reason, spatial pricing mechanisms that explicitly account for the configuration of the transmission network and operating constraints have become the *de facto* standard in the United States. All wholesale markets currently operating in the United States—in New England, New York, the PJM Interconnection (in Pennsylvania, New Jersey, Maryland and a number other eastern states), the Midwest, Texas, and California—use variants of the locational marginal pricing (LMP) algorithm described by Bohn, Caramanis and Schweppe (1984). This pricing mechanism sets potentially different prices at all locations (termed "nodes") in the transmission network. To compute these prices in the day-ahead market, suppliers submit generation unit-level offer curves indicating their willingness to supply energy from each generation unit they own as a function of the market-clearing electricity price. These willingness-to-supply schedules have two parts: a start-up cost offer and an energy supply curve. The start-up cost offer is a fixed dollar payment that must be paid to the generation unit owner if it is off-line at the start of the following day and the unit is accepted to produce a positive output during that day. The energy offer curve is a non-decreasing step function giving the willingness of the generation unit owner to supply additional energy as a function of the price it is paid for energy. All U.S. markets allow generation units owners to submit multiple price-quantity pairs for each generation unit for each hour of the day. For example, a supplier might be permitted to submit ten price-quantity pairs for each generation unit; an offer price-quantity step determines the minimum price a generator must be paid in order to produce the quantity associated with that step. The sum of the quantity increments is restricted to be less than the capacity of the generation unit. Offer prices are typically required to be greater than a price floor (which could be negative) and less than a price ceiling; the price floor and ceiling are approved by the Federal Energy Regulatory Commission (FERC), which is the national-level wholesale market regulator. In the day-ahead market, load-serving entities (LSEs) submit location-specific willingness-to-purchase functions that are non-increasing in the price at that location. The functions are composed of price-quantity pairs ordered from highest to lowest price; each offer quantity increment gives the amount the LSE is willing to increase its demand provided the market-clearing price at or below the corresponding offer price increment.

To compute the locational marginal prices (LMPs) at each node in the transmission network for every hour of the following day, the independent system operator (ISO) minimizes the as-offered total cost (based on the generation-unit level hourly offer curves and location-specific hourly demand curves submitted for each hour of the following day) of serving the demand for energy at all locations in the transmission network during all 24 hours of the following day subject to all relevant transmission network and other relevant operating constraints. The network constraints used to solve for the day-ahead hourly market outcomes are the ISO's best estimate of the real-time configuration of the transmission network during each hour of the following day. The solution to this as-bid cost minimization problem results in firm financial commitments for generation unit owners and load-serving entities for all 24 hours of the following day. The day-ahead generation unit and locational load schedules that solve this optimization problem are forward market sales and purchases for each hour of the following day.

For example, if a generation unit owner sells 50 MWh in the day-ahead market at a price of $40/MWh during one hour of the following day, then this supplier is guaranteed to be paid $2,000 (= 50 MWh x $40/MWh) regardless of the actual production of energy from its generation unit during that hour of the following day. Similarly, if a load-serving entity purchases 100 MWh in the day-ahead market during an hour of the following day at a price of $75/MWh, then this entity must pay $7,500 (= 100 MWh x $75/MWh) regardless of how much energy it withdraws from the network during that hour. The LMP at each node in the transmission network is equal to the increase in the minimized value of the objective function from this optimization problem as a result of increasing the amount of energy withdrawn at that location by 1 MWh. This property of the LMPs gives them their name. These LMPs for all 24 hours of the following day are computed during the afternoon of the day before the energy is scheduled to be delivered. All market participants are notified of these LMPs, their day-ahead generation unit-level energy schedules and location-specific load schedules in the afternoon of the day before their delivery date.

Starting with midnight on the delivery date, a real-time market determines the actual output of all generation units necessary to serve demand at all nodes in the transmission network. The real-time generation output and load-serving entity withdrawal levels are determined by minimizing the as-offered cost of serving the actual demand for energy at all locations in the transmission network subject to all relevant constraints in the transmission network and on generation units in the real-time market. Suppliers are allowed to change their hourly generation unit-level offer curves between the day-ahead and real-time markets.

In all U.S. ISOs, the real-time market is run every 5 minutes to determine the overall level of output from all generation units in the control area necessary to serve demand at all nodes in the transmission network. The solution to this optimization problem produces real-time locational marginal prices for each 5-minute interval within the hour. Hourly real-time prices are determined as the time-weighted average of the twelve 5-minute real-time prices during that hour. Generation unit owners that do not receive dispatch instructions within the hour receive this hourly real-time price for energy produced beyond their day-ahead forward market sales during that hour. Alternatively, if generation unit owners produce less energy in real-time than they sold for that hour in the day-ahead market, they must purchase the difference between their day-ahead forward market sales and real-time production at the hourly real-time price. Load-serving entities also only purchase or sell real-time deviations from their day-ahead schedules at the real-time price at their node in the transmission network. This combination of a day-ahead forward market and real-time spot market is called a multi-settlement market because of the property that only hourly real-time deviations from participants' hourly day-ahead schedules are settled at the hourly real-time price.

Let's return to the previous example of a generator that sold 50 MWhs of energy in the day-ahead market at a price $40/MWhs. If that generation unit only produced 40 MWhs of energy, the owner would have to purchase the remaining 10 MWhs at the real-time price in order to meet its forward market commitment. If the unit owner produced 55 MWhs, then the additional 5 MWhs beyond the unit's 50 MWhs day-ahead schedule is sold at the real-time price.

2.2 Implicit Virtual Bidding in Multi-Settlement Markets

A supplier or load serving entity that expects the real-time LMP at their node to be different from the day-ahead LMP at their node could exploit this price difference by selling or buying more or less energy than it expected to produce or consume in real-time. For example, suppose that a generation unit owner expected to ultimately produce 100 MWhs of energy from its unit. Moreover, suppose that this owner forecasts that: (1) real-time prices will be $60/MWh, and (2) real-time prices will be higher than day-ahead prices. The unit owner would simply submit price offers into the day-ahead market at or above $60/MWh, which he expects to result in selling no energy in the day-ahead market. The unit owner could then offer 100 MWhs of energy into the real-time market as a price taker to ensure that it produces its expected output of 100 MWh. This is accomplished by offering to supply this energy into the real-time market at an offer price equal to the offer price floor.

These actions by the generation unit owner are likely cause the day-ahead price to rise because less supply at or below the price of $60/MWh has been offered into the day-ahead market and the real-time price is likely to fall because more supply has been offered into the real-time market. The net impact of the supplier's actions is to increase the likelihood that the day-ahead and real-time prices are closer together than would be the case if the supplier did not submit a high offer price into the day-ahead market. For this reason, these actions by generation unit owners have been called "implicit virtual bidding or implicit convergence bidding" because the supplier is using forward market sales from its generation unit as a mechanism for exploiting expected price differences between the day-ahead and real-time markets.

Load-serving entities can also engage in implicit virtual bidding. Suppose that a load serving entity (LSE) expects real-time demand of 100 MWh. Further suppose that the LSE expects: (1) the day-ahead price to be higher than the real-time price, and (2) the real-time price to be $100/MWh. This LSE would then submit a demand bid into the day-ahead market with zero quantity demanded at prices above $100/MWh. Based on this bid, the LSE would very likely not make any purchase in the day-ahead market. Instead, its demand would be entered as a price-taker in the real-time market. As with the previous example with the generation unit owner, these actions by the LSE would reduce the difference between the day-ahead and real-time prices because demand is lower in the day-ahead market and higher in the real-time market as a result of these actions.

Implicit virtual bidding can have severe system reliability consequences. The combination of the example of a supplier that submits high offer prices in the day-ahead market because of a desire to sell at a higher price in the real-time market and the example of a load-serving entity wishing to purchase at a lower price in the real-time market can result in aggregate day-ahead forward market generation and load schedules that are below actual real-time demand levels. In this case, the system operator may be forced to find large amounts of additional energy after the close of the day-ahead market to ensure that actual demand is met. Wolak (2003) notes that this day-ahead/real-time imbalance is precisely what happened during the summer of 2000 in California's electricity market, exacerbated by the fact that the offer price cap for the day-ahead market was substantially higher than the offer price cap for the real-time market. Load-serving entities submitted demand bids into the day-ahead with zero quantity demanded at offer prices above the offer cap on the real-time market. Suppliers submitted offer prices into the day-ahead market at or above the offer cap on the real-time market for much of their anticipated real-time output, which resulted in the day-ahead market clearing at a quantity far below the anticipated real-time demand. This left the California ISO scrambling to find additional energy, often over 1/4 of the anticipated real-time demand, to ensure that real-time system demand would be met.

Implicit virtual bidding can also increase the cost of serving system demand. All wholesale electricity markets have generation units that take a number of hours to start, but can produce at a low variable cost once started. Implicit virtual bidding by both generation unit owners and load-serving entities can result in these long-start, low-operating-cost units not producing output. Although it may be unilaterally expected profit-maximizing for the owner of a portfolio of long-start, low-cost units and short-start, high-cost units to submit bids that cause some of these low-cost units not to operate, these actions increase the total cost of serving system demand. Additionally, one of the key potential benefits associated with the introduction of financial trading (termed explicit virtual bidding) is more and better information aggregated across a larger number of participants. If this results in the day-ahead price that more accurately reflects real-time conditions, purely financial market for wholesale electricity can decrease the cost of serving system demand even in the absence of implicit virtual bidding.

2.3 Explicit Virtual Bidding versus Implicit Virtual Bidding

The two major motivations for introducing explicit virtual bidding are: (1) to eliminate the adverse reliability consequences of market participants attempting to exploit expected price differences between the day-ahead and real-time markets and (2) to reduce the total cost of serving final demand. Explicit virtual bidding introduces a purely financial instrument that allows generation unit owners, load-serving entities and energy traders to exploit LMP differences between the day-ahead and real-time markets so that generation unit owners and load-serving entities do not distort their bidding and offer behavior in the day-ahead market in ways that increase their costs and potentially harm system reliability. Additionally, California's ISO allowed purely financial participants to trade these financial instruments; previously, only generation unit owners and load-serving entities were allowed to participate in California's wholesale electricity market.

Virtual (or convergence) bids are classified as either incremental (INC) or decremental (DEC) bids and are explicitly identified as such to the system operator. Market participants can submit either type of bid at any node in the transmission network. An INC bid at a node is treated just like a generation bid at the node. It is a step-function offer curve to supply additional energy in the day-ahead market. The only difference between an accepted virtual bid and an accepted bid from a generation unit owner is that the ISO knows that the energy sold in the day-ahead market from a virtual bid will be purchased in the real-time market as a price-taker. A DEC virtual bid is treated just like a physical demand bid in the day-ahead market. It is a step function bid curve to purchase additional energy in the day-ahead market. An accepted DEC virtual bid implies an obligation to sell this energy in the real-time market as a price-taker.

As should be clear from the above description, an INC virtual bid has a revenue stream equal to the difference between the day-ahead and real-time LMPs at that node times the amount of MWhs sold in the day-ahead market and a DEC virtual bid has a revenue stream equal to the difference between the real-time and day-ahead LMPs at that node times the amount of MWhs purchased in the day-ahead market. An INC virtual bid earns positive revenues if the day-ahead price is higher than the real-time price. However, the actions of INC virtual bidders make earning these profits less likely because supply is higher in the day-ahead market and demand is higher in the real-time market as a result of the INC bids. A DEC virtual bid earns positive revenues if the real-time price is higher than the day-ahead price. Again, the actions of DEC virtual bidders make this outcome less likely because demand in the day-ahead market is higher and supply in the real-time market is higher as a result of the DEC bids.

There are a number of reasons to believe that the introduction of explicit virtual bidding will lead to smaller realized nodal price differences between the day-ahead and real-time markets. First, relative to implicit virtual bidding, submitting an explicit virtual bid is a lower cost way for a market participant to take a financial position designed to profit from expected price differences between day-ahead and real-time markets. By submitting an INC virtual bid with an offer price below the price it expects in the real-time market, a market participant can earn the difference between day-ahead and real-time market prices. The availability of this financial instrument makes it unnecessary for a supplier or load-serving entity to employ more costly distortions in their day-ahead energy purchases or sales in order to exploit expected day-ahead versus real-time price differences. Instead, the supplier can offer their generation unit into the day-ahead market at its variable cost and submit decremental virtual bids with offer prices equal to the generation unit owner's expected real-time market price. In this way, the generation unit owner does not distort the *physical* bids associated with its generation units in order to exploit expected price differences between the day-ahead and real-time markets.

A second reason that nodal-level day-ahead versus real-time price differences are likely to be smaller after the introduction of explicit virtual bidding is because it gives market participants greater flexibility to exploit locational price differences. A generation unit owner can only implicitly virtual bid total MWhs less than or equal to the capacity of their generation unit at a given node. An implicit virtual bidding supplier has no recourse if withholding output equal to the capacity of this unit from the day-ahead market is insufficient to increase the day-ahead price enough to cause it to equal the expected real-time price at that location. However, with (explicit) virtual bidding, the supplier can submit an almost unlimited quantity of DEC bids at that location to raise the price at that node in the day-ahead market. The same logic goes for a load-serving entity (LSE) engaging in implicit virtual bidding. The actual demand than an LSE must satisfy limits the amount of demand it can bid into the day-ahead market. For example, without explicit virtual bidding, if bidding zero demand into the day-ahead market still does not reduce the LMP at that node to the level the load-serving entity expects in the real-time market, that supplier has no other way to reduce the day-ahead price at that node. However, with a sufficient volume of INC bids, the load-serving entity can reduce the price at that node to any level it expects to prevail in the real-time market.

Before nodal-level (explicit) virtual bidding was introduced in California, only physical players (i.e: generation unit owners and load-serving entities) were allowed to participate in the wholesale electricity market. Due to this, the opportunities to implicit virtual bid at the nodal level were limited to locations with generation units. Load-serving entities cannot place physical bids at a nodal-level; the California market requires the three large load-serving entities—Southern California Edison (SCE), Pacific Gas and Electric (PG&E), and San Diego Gas and Electric (SDG&E)—to bid their service area-level demand into the day-ahead market. The California ISO then allocates this demand to all nodes in the load-serving entity's service territory using load-distribution factors (LDFs) that the ISO produces. For example, if a load-serving entity has 100 MWhs of load and the ISO computes equal LDFs for the ten nodes in its service area, then the load-serving entity's LDFs are equal to 1/10 for each node. This implies that it is very costly for a load-serving entity (LSE) to implicitly virtual bid 1 MWh at one node, because this would effectively require 1 MWh of implicit virtual bids at all nodes within the LSE's service area. With the introduction of explicit nodal-level virtual bidding, load-serving entities and generation unit owners can exploit day-ahead and real-time price differences at any node, even those with no generation units, by submitting a virtual bid at that node. Moreover, purely financial traders can now also enter the market and submitting a virtual bid at a given node in order to exploit day-ahead/real-time price spreads at that node.

A final market efficiency benefit of introducing explicit virtual bidding is that it makes it much easier for market monitors and regulatory authorities to identify implicit virtual bidding. Before the introduction of explicit virtual bidding, a generation unit owner or load-serving entity could always claim that the reason their day-ahead sales or purchases were substantially less than their real-time production or consumption is because of the expectation of more favorable prices in the real-time versus day-ahead market. With the introduction of explicit virtual bidding, regulators can argue that suppliers and load-serving entities should sell and purchase their best estimate of their expected real-time production and consumption in the day-ahead market, because they can use explicit virtual bidding to exploit any expected differences between day-ahead and real-time prices. The existence of this additional product to exploit expected price differences allows the regulator to be tougher on actions that might be unilaterally profit-maximizing for suppliers and load-serving entities but also reduce system reliability and overall market efficiency.

3 Descriptive Statistics for California Market

This section summarizes our evidence on hourly price convergence between the day-ahead and real-time markets for the three large load-serving entities in California before and after the implementation of explicit virtual bidding (EVB). We first present the results of a test of the null hypothesis that the 24x1 vector of day-ahead/real-time price spread means, averaged over days-of-sample for each hour-of-the-day, is equal to zero for these three load-serving entities. We find that we overwhelmingly reject this null hypothesis for all three load-serving entities. However, these naive tests do not account for the transactions costs associated with actually taking action (such as submitting a virtual bid) in order to exploit these mean price differences. This motivates the development of our statistical testing procedure, which does account for transactions costs.

Our hypothesis tests are implemented using hourly data from April 1, 2009[1] to December 31, 2012 on day-ahead and real-time wholesale electricity prices at all nodes (read: locations) in California's ISO area. There are over 5,000 nodes, all with potentially different prices. However, each of the three large load-serving entities faces a single load aggregation point (LAP) day-ahead price and a single LAP-level real-time price each hour of the day. These LAP-level prices are computed as the nodal quantity-weighted average price (either day-ahead or real-time) for that load-serving entity, summed over all nodes in the load-serving entity's service area with a positive amount of energy withdrawn from the transmission network during that hour. Each of the three large load-serving entities has its own day-ahead and real-time LAP price determined by the California ISO.

Figure 1 presents a comparison by hour-of-the-day of the average difference between the day-ahead and real-time prices for the Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E) LAPs both before and after the introduction of explicit virtual bidding. This figure provides descriptive evidence that the day-ahead/real-time price spread is more pronounced prior to the introduction of explicit virtual bidding relative to after its introduction for each of the three load-serving entities. For example, the average day-ahead price for PG&E is much

[1]California introduced nodal pricing on this date.

lower than the average real-time price during the hours of 8PM-12AM. These results immediately raise the question: do these mean price differences reflect the existence of profitable trading strategies or are they simply due to the existence of non-zero trading costs that allow non-zero mean price differences?

To further motivate our subsequent analysis, we present a version of an arbitrage test that holds only if there are *zero* transactions costs. Namely, we plot in Figure 2 the average day-ahead/real-time spread along with point-wise 95% confidence intervals around these means for the PG&E, SCE, and SDG&E LAPs after the introduction of (explicit) virtual bidding. For all three load-serving entities for some hours of the day, we can reject at a 5% significance level that the price spread is zero. Along these same lines, we can also simply perform a joint test that average day-ahead and real-time price differences are zero for all hours of the day. We use the Newey and West (1987) autocorrelation consistent asymptotic covariance matrix estimate, $\hat{\Sigma} = \hat{\Lambda}_0 + \sum_{j=1}^{m} w(j,m)(\hat{\Lambda}_j + \hat{\Lambda}_j{}')$, where $\hat{\Lambda}_j = \frac{1}{T}\sum_{t=j+1}^{T}(X_t - \overline{X})(X_{t-j} - \overline{X})'$, $\overline{X} = \frac{1}{T}\sum_{t=1}^{T} X_t$, $w(j,m) = 1 - \frac{j}{m+1}$ for $m = 14$ to construct the chi-squared test statistics. These test statistics are presented for each LAP before and after the introduction of virtual bidding in Table 1. Note that these test statistics are quite large. We would reject the null hypothesis that all of the hour-of-day price difference means are zero in all cases.[2] However, the statistical tests described in this paragraph fail to account for the potentially sizable transaction costs present in nearly every commodities market. In the next section, we present a hypothesis test for the existence of arbitrage that accounts for the fact that day-ahead/real-time price spreads can differ from zero simply due to positive transaction costs.

4 Tests for the Existence of a Profitable Trading Strategy

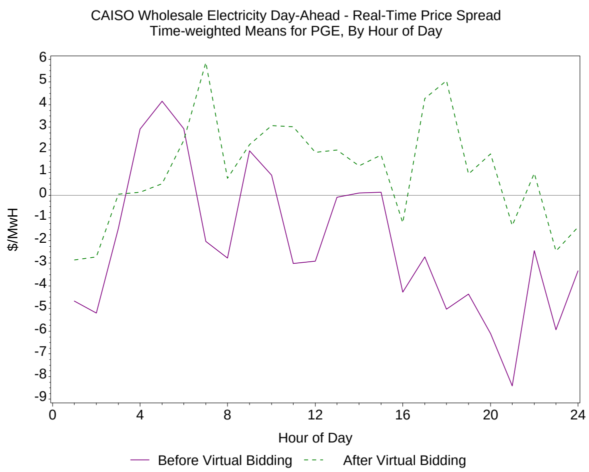
4.1 Introduction

In this section, we develop a statistical framework in order to test whether or not a profitable trading strategy exists when accounting for the presence of transactions costs. For simplicity, we restrict attention to trading strategies that only condition on the value of

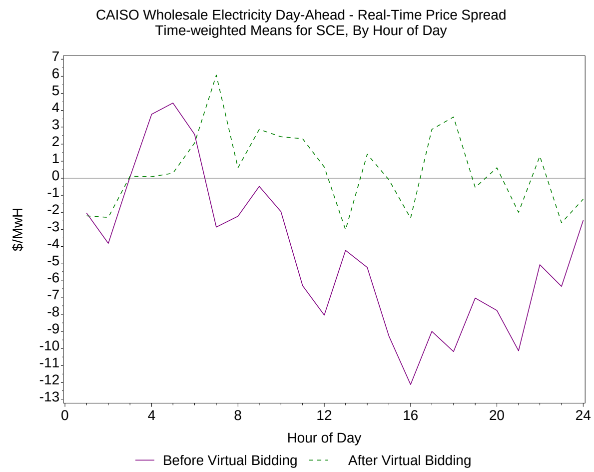
[2]The upper $\alpha = 0.05$ critical value for the $\chi^2(24)$ distribution is 36.415.

Figure 1: Hourly Graphs of Day-Ahead/Real-Time Price Differences: Before and After EVB

(a)



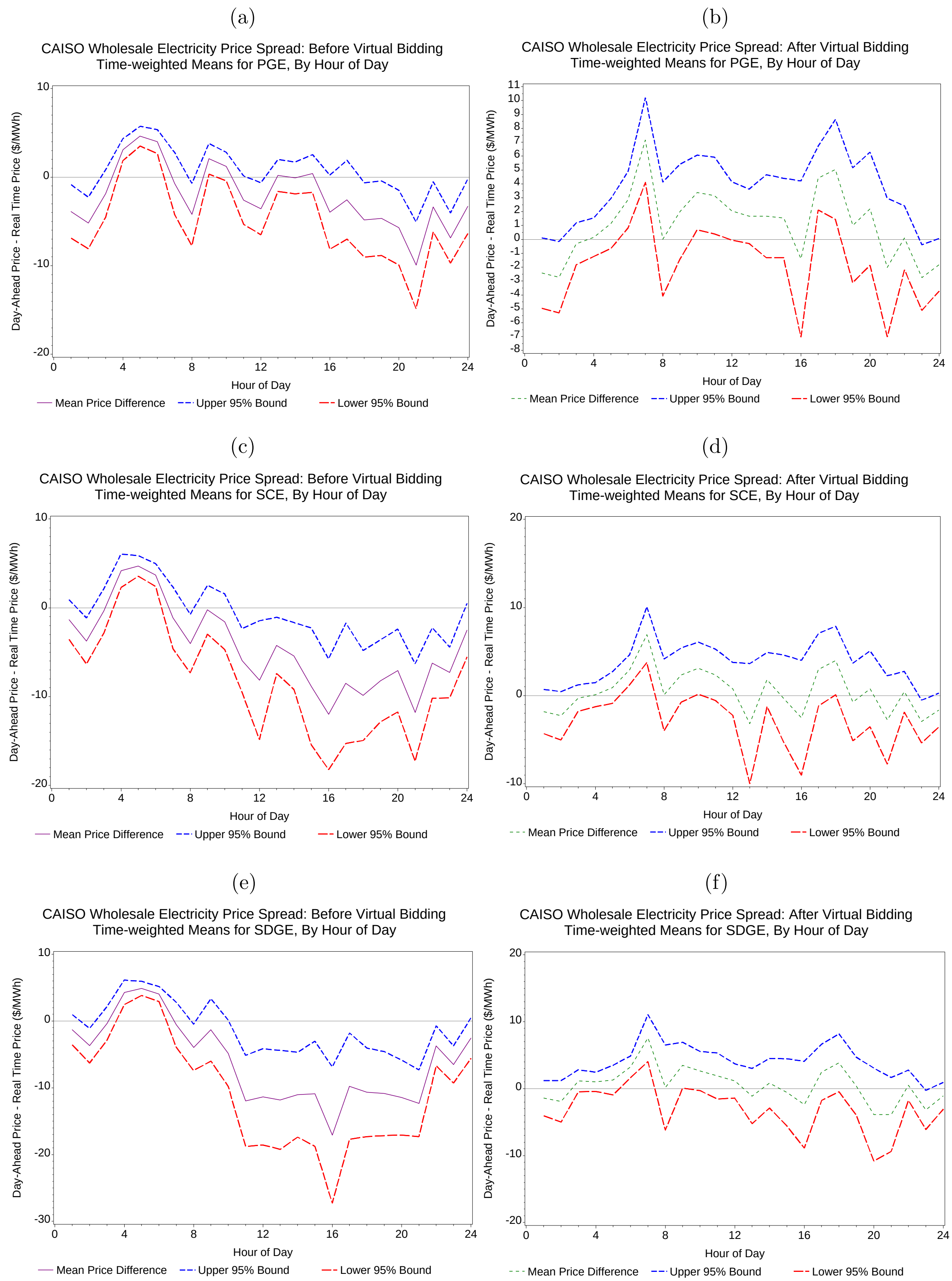
(b)



(c)

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. The three LAPs correspond to Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E).

Figure 2: Hourly Graphs of Price Differences with 95% C.I: Before and After EVB



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

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

	Before EVB	After EVB
PG&E	141.738	88.158
SCE	140.140	105.127
SDG&E	157.742	86.084

Notes: This table presents the chi-squared test statistics for each load aggregation point before and after the introduction of explicit virtual bidding (EVB) of the null hypothesis that all 24 hour-of-the-day average day-ahead/real-time price spreads are equal to zero. The three LAPs correspond to Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E). We use the Newey and West (1987) autocorrelation consistent asymptotic covariance matrix estimate, $\hat{\Sigma} = \hat{\Lambda}_0 + \sum_{j=1}^{m} w(j,m)(\hat{\Lambda}_j + \hat{\Lambda}_j{}')$, where $\hat{\Lambda}_j = \frac{1}{T}\sum_{t=j+1}^{T}(X_t - \overline{X})(X_{t-j} - \overline{X})'$, $\overline{X} = \frac{1}{T}\sum_{t=1}^{T} X_t$, $w(j,m) = 1 - \frac{j}{m+1}$ for $m = 14$ days of lag in order to construct the chi-squared test statistics. We reject the null hypothesis that all of the hour-of-day price difference means are zero at an $\alpha = 0.05$ critical value if the test statistic is larger than $\chi^2_{0.95}(24) = 36.415$.

the (24x1) vector of hour-of-day day-ahead minus real-time mean price differences. We denote this 24x1 vector of means as μ. Using our framework, we can assess whether the data provide evidence against: 1) the null hypothesis that a profitable trading strategy exists based on 24 assets with (unconditional) means μ and covariance matrix Σ, and 2) the null hypothesis that *no* profitable trading strategy exists based on 24 assets with (unconditional) means μ and covariance matrix Σ. The timing of the day-ahead and real-time markets precludes trading strategies that condition of the first lag of the price difference vector because market participants submit their offers into the day-ahead market for date t *before* knowing the real-time prices for *any* of hour-of-the-day for date $t-1$.[3] Thus, we formulate a statistical test of the null hypothesis that all autocorrelations in the daily day-ahead/real-time price differences vector *beyond the first lag* are jointly zero. We find no empirical evidence against this null hypothesis, justifying our consideration of trading strategies based only on the unconditional means μ and covariance matrix Σ of hour-of-the-day day-ahead/real-time price spreads.

We motivate our statistical test by considering the problem facing a market participant maximizing expected profits from trading day-ahead versus real-time price differences. In particular, the trader chooses a portfolio based on 24 assets consisting of day-ahead/real-time price differences for each hour of the day. California's independent

[3]Offers to the day-ahead market must be submitted by noon the day before actual system operation, so it is not possible to base a daily trading strategy on knowledge of the first-order autocorrelation of the price differences vector.

system operator (CAISO) sets explicit trading costs associated with the financial trading of day-ahead/real-time price differences; however, these explicit trading costs are only one component of the overall costs of exploiting these price differences. Due to this, we use our statistical test to recover an estimate of the *lowest* trading costs for which we can reject the null hypothesis that arbitrage trading opportunities *exist*. Similarly, we also recover an estimate of the *highest* trading costs for which we can *reject* the null hypothesis that arbitrage trading opportunities *do not exist*. Fang and Santos (2014) develops theoretically how to construct confidence intervals around directionally differentiable functions of a regular parameter estimate; we implement this confidence interval using a bootstrap procedure described by Hong and Li (2015). In short, we are also able to compute an estimate of the *distribution* of the implied trading cost estimate.

We estimate these implied trading costs separately for each location, both at the aggregated LAP-level as well as the nodal level. Importantly, we also estimate these costs separately for the sample periods before versus after the introduction of explicit virtual bidding. Comparing these estimated trading cost distributions before versus after the introduction of explicit virtual bidding allows us to assess whether the point estimates of our implied trading costs are statistically significantly different before versus after the introduction of explicit virtual bidding. As an alternative approach, we also perform a test of the null hypothesis that the expected profits traders can earn by buying/selling portfolios based on the 24 hourly day-ahead/real-time price spreads fell after the implementation of explicit virtual bidding. For this hypothesis test, we use the multivariate inequality constraints testing procedure described in Wolak (1989).

4.2 The Trader's Problem

Consider a trader with access to 24 assets, where asset X_h for $h \in \{1, ..., 24\}$ is equal to the difference between day-ahead and real-time electricity prices for hour h of the day. In math, $X_h \equiv P_h^{DA} - P_h^{RT}$, where P_h^{DA} is the day-ahead price for hour h and P_h^{RT} is the real-time price for hour h. Purchasing this security requires the trader to sell 1 MWh more energy in the day-ahead market than it produces in real-time. Selling this security requires that the trader buy 1 MWh more energy in the day-ahead market than it consumes in real-time. Let $\mu_h = E(X_h) = E(P_h^{DA}) - E(P_h^{RT})$ for $h = 1, 2., ..., 24$. Define μ as the 24 x 1 vector composed of $(\mu_1, \mu_2, ..., \mu_{24})'$ and X_d as the 24 x 1 vector composed of $(X_{d,1}, X_{d,2}, ..., X_{d,24})'$ for day-of-sample d. Let Λ_0 be the 24 x 24 contemporaneous covariance matrix of X_d. Finally, let the per-unit trading cost of buying or selling this security be c. Then, the expected profit-maximization problem of the trader is:

$$\begin{matrix} max \\ a \in R^{24} \end{matrix} \; a'\mu - c\sum_{i=1}^{24} |a_i| \;\; subject \;\; to \;\; \sum_{i=1}^{24} |a_i| \; = \; 1. \tag{1}$$

where the trader optimizes over asset weights $a = (a_1, a_2, ..., a_{24})'$. Note that each a_i can be positive or negative. We denote $a^*(\mu) \in R^{24}$ as the weights that maximize the trader's expected profits (i.e: solve the optimization problem described in Equation 1).

We consider both the null hypothesis that $a^*(\mu)'\mu \; - \; c > 0$ ("profitable trading strategies exist") and the null hypothesis that $a^*(\mu)'\mu \; - \; c \leq 0$ ("no profitable trading strategies exist"). The trader pays a per-unit trading cost c regardless of whether they buy or sell a unit of the asset; this is why the overall trading costs are calculated based on the sum of the absolute value of the portfolio weights: $(c\sum_{i=1}^{24}|a_i|)$. Moreover, it is due to these per-unit trading costs that we impose the normalization that $\sum_{i=1}^{24}|a_i| \; = \; 1$ rather than the "traditional" normalization that $\sum_{i=1}^{24} a_i \; = \; 1$.[4] The optimized value of the objective function reduces to

$$a^*(\mu)'\mu \; = \; \begin{matrix} max \\ i \in \{1, ..., 24\} \end{matrix} |\mu_i| \tag{2}$$

This setting falls into the framework developed by Fang and Santos (2014) for testing hypotheses involving directionally differentiable functions of a regular parameter estimate. First, the function $\phi(\mu) \; \equiv \; a^*(\mu)'\mu$ is a directionally differentiable function of the parameter vector μ. Also, our estimate of the true parameter vector μ_0 is simply the sample average $\overline{X} \; = \; \frac{1}{N}\sum_{d=1}^{N} X_d$ where N is the number of days in the sample. Because $\sqrt{N}(\overline{X} \; - \; \mu)$ possesses an asymptotic normal distribution, the sample mean of the X_d is regular estimate of μ_0.

[4]For example, the "net position" normalization $(\sum_{i=1}^{24} a_i \; = \; 1)$ is imposed in the canonical portfolio choice model formulated by Markowitz (1952).

4.3 Tests Regarding Existence of a Profitable Trading Strategy

To implement the hypothesis test, we compute $\phi(\overline{X})$, which is defined to be the element of $\overline{X}$ that is largest in absolute value. The difference between $\phi(\overline{X})$ and the trading cost c is our test statistic. Fang and Santos (2014) present a modified bootstrap estimator of the asymptotic distribution of a directionally differentiable function of $\overline{X}$; applied to our context, we compute an estimate of the asymptotic distribution of $\sqrt{N}(\phi(\overline{X}) - \phi(\mu))$. For this estimation, we utilize a numerical derivative-based procedure developed by Hong and Li (2015) for simulating the distribution of $\phi(\overline{X})$. For this procedure, we first compute a moving blocks bootstrap re-sample of $\overline{X}$ with block size equal to the largest integer less than or equal to $N^{1/3}$. We denote the b^{th} bootstrap re-sample of $\overline{X}$ as $\overline{X}^b$. 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 (2015) 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 ϵ tends to zero as N goes to infinity but $\sqrt{N}\epsilon$ tends to infinity as N goes to infinity. To satisfy these conditions, we set $\epsilon = N^{-1/3}$, which is the value recommended by Hong and Li (2015).

Although we know the explicit trading charges associated with purchasing or selling elements of X_d set by the California ISO market, this is just one component of overall per-unit trading costs. Setting our parameter for trading costs, c, equal to this magnitude and performing our hypothesis test implies that there is no opportunity cost associated with the time of the individual undertaking the trades, no up-front costs of participating in the ISO markets, and no other cost associated with preparing or updating a strategy for trading day-ahead and real-time price differences. For this reason, we instead use our hypothesis testing procedure to compute *implied* trading costs. We compare these implied trading costs to the actual cost of purchasing and selling the 24 elements of X in the ISO market, including both explicit trading charges and conservative estimates of other transactions costs.

The bootstrap distribution of Z^b is an estimate of the distribution of $\phi(\overline{X})$. We compute each bootstrap re-sample of $\phi(\overline{X})$ as:

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

We then use this bootstrap distribution to compute two values associated with trading costs c:

1. c_{lower}: the smallest value of c that results in rejection of the $\alpha = 0.05$ size test of the null hypothesis that $a^*(\mu)'\mu - c > 0$

2. c_{upper}: the largest value of c that causes rejection of the $\alpha = 0.05$ size test of the null hypothesis that $a^*(\mu)'\mu - c \leq 0$.

The first value, c_{lower}, is the smallest value of the dollar per MWh trading cost that would cause rejection of the null hypothesis that a profitable trading strategy exists. It is computed as the lower 5^{th} percentile of the distribution of $\phi(\overline{X})$. The second magnitude, c_{upper}, is the largest value of the trading charge that causes rejection of the null hypothesis that no profitable trading strategy exists. It is computed as the 95^{th} percentile of the distribution of $\phi(\overline{X})$.

4.4 Test for Difference in the Absolute Value of Means Before versus After Explicit Virtual Bidding

We also test whether expected trading profits fall after the introduction of explicit virtual bidding using a multivariate inequality constraints test. If we let the trading costs prior to explicit virtual bidding be c^{pre} and the trading costs after explicit virtual bidding be c^{post}, then a test of the null hypothesis that trading profits fell after the introduction of explicit virtual bidding can be formulated as $|\mu_{pre}| - \mathbf{1}c^{pre} > |\mu_{post}| - \mathbf{1}c^{post}$. $|\mu_J|$ for $J \in \{pre,\ post\}$ is a 24 x 1 vector composed of the absolute value of the average day-ahead/real-time price differences for hours-of-the-day $h \in \{1, 2, ..., 24\}$, where averages are taken separately for samples before versus after explicit virtual bidding; $\mathbf{1}$ is 24 x 1 vector of 1's. The difference $|\mu_{pre}| - \mathbf{1}c^{pre}$ is the expected profits associated with buying (selling) one unit of the day-ahead/real-time price difference for each hour-of-the-day $h \in \{1, 2, ..., 24\}$ if it is positive (negative). Re-arranging this inequality, we see that the null hypothesis becomes: $|\mu_{pre}| - |\mu_{post}| > \mathbf{1}(c^{pre} - c^{post})$. If we assume that $c^{pre} > c^{post}$, which is consistent with the results presented in Section 5, then testing the null hypothesis that $|\mu_{pre}| - |\mu_{post}| > 0$ is a conservative test with respect to our original null hypothesis that $|\mu_{pre}| - |\mu_{post}| > \mathbf{1}(c^{pre} - c^{post})$. As we don't observe trading charges c^{pre} and c^{post}, we focus empirically on the statistical test of the null hypothesis that $|\mu_{pre}| - |\mu_{post}| > 0$. Conversely, if we want to test whether expected trading profits rose after the introduction of explicit virtual bidding, our null hypothesis would be that $|\mu_{post}| - |\mu_{pre}| > \mathbf{1}c^{post} - \mathbf{1}c^{pre}$. Thus, if we fail to reject the null hypothesis that $|\mu_{post}| - |\mu_{pre}| > 0$, we can conclude that we would also fail to reject the null hypothesis that trading profits were higher after the introduction of explicit virtual bidding can be rejected. If we fail to reject the null hypothesis that $|\mu_{pre}| > |\mu_{post}|$ but reject the null hypothesis that $|\mu_{post}| > |\mu_{pre}|$, then we have evidence that trading profits fell after the introduction of explicit virtual bidding.

We implement these two multivariate nonlinear inequality constraints tests using the methodology derived in Wolak (1989). We present the procedure for $|\mu_{pre}| > |\mu_{post}|$ below:

Proposition 1 ***Direct Test of Null Hypothesis that*** $|\mu_{pre}| > |\mu_{post}|$
Let:
$\hat{V} = \frac{1}{N^{pre}} diag[SIGN(\overline{X}^{pre})]'\hat{\Sigma}^{pre} diag[SIGN(\overline{X}^{pre})] + \frac{1}{N^{post}} diag[SIGN(\overline{X}^{post})]'\hat{\Sigma}^{post} diag[SIGN(\overline{X}^{post})]$ *Calculate the test statistic:*

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

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.*

Cataloging notation, the $diag[Z]$ operator takes a vector Z, and returns a diagonal matrix with elements of Z on the diagonal. All objects with a "pre" superscript are based on the 4/1/2009-2/1/2011 period prior to the introduction of explicit virtual bidding (EVB); the "post" superscript refers to the 2/1/2011-12/31/2012 period after the introduction of explicit virtual bidding. $N^{pre} = 657$ is the number of days in the sample prior to the start of explicit virtual bidding; $N^{post} = 410$ is the number of days in the sample after explicit virtual bidding was introduced. $\overline{X}^{pre}$ is a 24×1 vector of the average day-ahead/real-time price differences, averaged for the pre-EVB sample separately for each hour-of-the-day. $\hat{\Sigma}^{pre}$ is a 24×24 estimate of the asymptotic covariance matrix corresponding to $\overline{X}^{pre}$; we compute $\hat{\Sigma}^{pre}$ using the autocorrelation consistent estimator proposed by Newey and West (1987) with $m = 14$ days of lagged data. $\overline{X}^{post}$ and $\hat{\Sigma}^{post}$ are calculated similarly using data from after explicit virtual bidding was introduced. Finally, we calculate $w(24, 24-h, \hat{V})$ using the simulation method described in Wolak (1989).

4.5 Test for the Difference Between Variance Matrices Before versus After Explicit Virtual Bidding

We also expect that the introduction of explicit virtual bidding will reduce the day-ahead uncertainty about real time prices. We would therefore expect both the variance of day-ahead/real-time price differences and the variance of real-time prices to fall after the introduction of explicit virtual bidding.

With nodal-level explicit virtual bidding, market participants can profit from their ability to forecast real-time system conditions at any location in the transmission network. A market participant who believes that the real-time price will be higher than the day-ahead price at a given location will submit a DEC bid to purchase energy at that location in the day-ahead market that is subsequently sold at the real-time price. If this market participant is correct, she will be rewarded with positive trading profits. However, these actions will also cause the day-ahead price to rise (because of the higher day-ahead demand associated with the DEC bid) and the real-time price to fall (because of the lower real-time demand due to the sale of the accepted DEC bid in the real-time market); this reduces the market participant's trading profits. However, profits will not go to zero unless the total (across all market participants) amount of day-ahead DEC bids at that location is large enough to close the day-ahead/real-time price gap. Conversely, market participants that believe that the real-time price will be lower than the day-ahead price at a given location, perhaps because they believe the real-time demand at that location will be lower than expected, will submit INC bids in the day-ahead market and subsequently purchase the energy sold in the day-ahead market from the real-time market. If they end up being incorrect, they will lose money from these actions.

Given that physical players can use the financial commodity rather than their physical bids in order to arbitrage expected day-ahead/real-time price spreads and financial players can participate in the market after the introduction of explicit virtual bidding, we should expect final day-ahead generation schedules to be closer to the real-time output of these generation units. For this reason, we should see a decrease in the volatility of day-ahead/real-time price spreads after the introduction of explicit virtual bidding. Additionally, as final day-ahead generation schedules are a better predictor of the real-time output of these generation units after the introduction of explicit virtual bidding, California's ISO will have to make substantial purchases or sales in the real-time market less often after explicit virtual bidding. Thus, we should expect the volatility of real-time prices to be lower after the introduction of explicit virtual bidding (EVB) as well.

Formally, we consider the Null hypothesis H_1 that $\Lambda^{pre} - \Lambda^{post}$ is a positive semi-definite matrix, where Λ^{pre} (Λ^{post}) is the 24 x 24 contemporaneous covariance matrix for the time period before (after) explicit virtual bidding. In order to implement this test, we find the eigenvalues $\hat{\omega}_j$ ($j = 1, 2, ..., 24$) of $\hat{\Lambda}^{diff} \equiv \hat{\Lambda}^{pre} - \hat{\Lambda}^{post}$ and test the joint null hypothesis that all of these eigenvalues are greater than or equal to zero. We use the same multivariate inequality constraints test employed in the previous section. We obtain the covariance matrix for our estimated eigenvalues $\hat{\omega}_j$ ($j = 1, 2, ..., 24$) using a moving-block bootstrap procedure. Briefly, this moving block procedure accounts for fact that day-ahead/real-time price spreads (X_d) may be autocorrelated across days.

More formally, let $\hat{\omega}$ equal the 24×1 vector of eigenvalues of $\hat{\Lambda}^{diff}$. We first re-sample contiguous blocks of length $B^{pre} = (N^{pre})^{1/3}$ (where N^{pre} is the pre-EVB sample size) from the time series of daily price difference vectors. We repeat this process L times, estimating $\hat{\Lambda}_b^{pre}$ for each re-sample $b \in \{1, 2, ..., L\}$. We similarly re-sample $B^{post} = (N^{post})^{1/3}$ times, obtaining $\hat{\Lambda}_b^{post}$ for each re-sample $b \in \{1, 2, ..., L\}$ (where "post" refers to the post-EVB sample period). For each re-sample $b \in \{1, 2, ..., L\}$, we next compute $\hat{\Lambda}_b^{diff} \equiv \hat{\Lambda}_b^{pre} - \hat{\Lambda}_b^{post}$ as well as the eigenvalues associated with $\hat{\Lambda}_b^{diff}$. We denote the eigenvalues associated with re-sample b $\hat{\omega}_b$ and find the empirical co-variance of these eigenvalues across our L re-samples in order to get an estimate of the covariance matrix for $\hat{\omega}$. We call this moving-block bootstrap estimated covariance matrix $Var(\hat{\Lambda}_b)$. Our test statistic is $TS = \min_{z \geq 0} N(\hat{\Lambda}_c - z)'[Var(\hat{\Lambda}_c)]^{-1}(\hat{\Lambda}_c - z)$; TS is asymptotically distributed as the weighted sum of chi-squared random variables given in the previous section under the null hypothesis.

We can also perform this test for the null hypothesis H_2 that $\Lambda^{post} - \Lambda^{pre}$ is a positive semi-definite matrix. Failing to reject H_1 and rejecting H_2 (for both the vector of price differences and the vector of real time prices) would give us strong evidence consistent with our prediction that the introduction of explicit virtual bidding reduced the variance in day-ahead/ real-time price spreads as well as the variance of real-time prices.

4.6 Why not condition on past values of X_d?

The values of the 24×1 vector of real-time prices for day $d-1$ are not known before offers are submitted to the day-ahead market for day d; thus, any first-order autocorrelation between realizations of X_d *cannot* feasibly be exploited in a trading strategy. Specifically, any trading strategies involving portfolios of the 24×1 price differences that condition on X_{d-k}, for $k > 0$, would have to condition on values from at least $k = 2$ days ago, because only realizations of X_{d-k} for $k > 1$ are known when a market participant submits bids or offers into the day-ahead market for day d. Our analysis is restricted to trading strategies that do not condition on past values of X_{d-k}; noting that the realization of X_{d-1} cannot be feasibly traded upon, X_d must follow a vector MA(1) process in order for our restriction to be justified. To investigate this hypothesis, we would ideally like to estimate a vector MA(1) process for X_d and test the null hypothesis that the errors from this model are multivariate white noise. However, estimating the 24 x 1 vector MA(1) model necessary to test this hypothesis has proven extremely difficult to compute in finite time.

Due to this, we formulate a different approach that does not rely on estimating a vector MA(1) model for the daily price difference vector. Denote the τ^{th} (24×24) auto-correlation matrix: $\Gamma(\tau) = E(X_t - \mu)(X_{t-\tau} - \mu)'$. Consistent with our above discussion,we expect $\Gamma(1)$ to be non-zero, but $\Gamma(\tau) = 0$ for all $\tau > 1$. Thus, we consider the Null hypothesis:

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

for a fixed value of R. For our application, we test using $R = 10$. This hypothesis test is implemented by first defining $\xi \equiv [vec(\Gamma(2))', vec(\Gamma(3))', ..., vec(\Gamma(L))']'$, where the $vec(.)$ operator takes a (24 x 24) matrix and stacks it column-wise to create a (576 x 1) vector. Therefore, ξ has 5760 = 576 * 10 elements, which all must equal zero under the Null hypothesis. We create a simple Wald Statistic, using the moving block bootstrap (described more fully in the previous subsection) in order to estimate the 5760 x 5760 covariance matrix associated with $\hat{\xi}$. Our Wald statistic $TS = N\hat{\xi}'\hat{\Sigma}^{-1}_{\xi,boot}\hat{\xi}$ is asymptotically distributed as a chi-squared with $24^2*(R-1)$ degrees of freedom under the null hypothesis. We perform this test separately for before versus after the introduction of explicit virtual bidding.

5 Empirical Results

This section presents our estimation of the smallest trading charge for which we can reject the null hypothesis of arbitrage (c_{lower}) and largest trading charge for which we can reject the null hypothesis of no arbitrage(c_{upper}). We also provide the results of our tests that expected trading profits fell after the introduction of explicit virtual bidding. Finally, we demonstrate that the volatility of both day-ahead/real-time price spreads and real-time prices themselves fell after the introduction of explicit virtual bidding. Before we discuss these results, we show that more complex trading strategies based on lagged values of price differences are unlikely to yield significant profit improvements relative to a strategy based only on the unconditional average price differences for each hour-of-the-day (μ).

5.1 Is there autocorrelation in daily price differences beyond the first lag?

Recall from the previous section that we want to test the null hypothesis that the second through tenth autocorrelation matrices for X_d are zero: $\Gamma(2) = \Gamma(3) = ... = \Gamma(10) = 0$. We test separately for each LAP (i.e: PG&E, SCE, SDG&E), both before and after the introduction of explicit virtual bidding. The test statistics are recorded in Table 2. The upper $\alpha = 0.05$ critical value for these test statistics is $\chi^2(5184) = 5352.6$.

We fail to reject the null hypothesis that the second through tenth autocorrelation

Table 2: Test Statistics for Autocorrelation ($1 < L \leq 10$) in Daily Price Differences

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

Notes: This table presents the chi-squared test statistics for each load aggregation point (LAP) before and after the introduction of explicit virtual bidding (EVB) of the null hypothesis that the second through tenth autocorrelation matrices for daily day-ahead/real-time price difference vector X_d (which is 24 x 1) are zero: $\Gamma(2) = \Gamma(3) = ... = \Gamma(10) = 0$. The three LAPs correspond to Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E). We estimate $(\Gamma(2), \Gamma(3), ..., \Gamma(10))$ (which are each 24 x 24) pre-EVB and post-EVB for each LAP and stack the elements column-wise; this results in a $5760 = 24^2 * 10$ element vector. We use the moving block bootstrap in order to estimate the covariance matrix associated with this vector. The upper $\alpha = 0.05$ critical value for these test statistics is $\chi^2(5184) = 5352.6$.

matrices are zero for any LAP at the 5% level, either before or after the introduction of explicit virtual bidding. This lends strong evidence in favor of our assertion that daily price differences follow an MA(1) process. As traders cannot condition on the previous day's price realizations when submitting into the day-ahead market, this test helps to justify our focus on trading strategies that do not condition on past lags of the daily price difference vector (X_{d-k} for $k > 0$).

We repeat these same autocorrelation tests at the nodal level and found that before the implementation of explicit virtual bidding, particularly at non-generation nodes, the null hypothesis that $\Gamma(2) = \Gamma(3) = ... = \Gamma(10) = 0$ could be rejected at approximately 70 percent of the nodes. However, after the implementation of explicit virtual bidding this null hypothesis was rejected at approximately five percent of the generation and non-generation nodes which is consistent with this null hypothesis being true for all nodes after the implementation of explicit virtual bidding, because the size of each individual nodal-level test was $\alpha = 0.05$.

5.2 Results from Trading Costs Hypothesis Tests

We first implement our trading cost hypothesis tests at the load aggregation point (LAP) level. These results are presented in Table 5. For each LAP, we report the values of c_{lower} and c_{upper} both before and after the introduction of explicit virtual bidding. Recall that

Table 3: Percentage of Autocorrelation Tests that Fail to Reject ($\alpha = 0.05$)

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

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

Table 4: Sample Counts of Nodes By Cell

	Before EVB	After EVB
Non-Generation Node	4,031	4,386
Generation Node	669	673

Notes: This table presents the number of nodes (read: locations) in each category (generation versus non-generation) before and after the introduction of explicit virtual bidding. Generators inject electricity at some nodes; these nodes are called "generation nodes". The remaining nodes are termed "non-generation" nodes.

Table 5: LAP level Implied Trading Costs–c_{lower} and c_{upper}

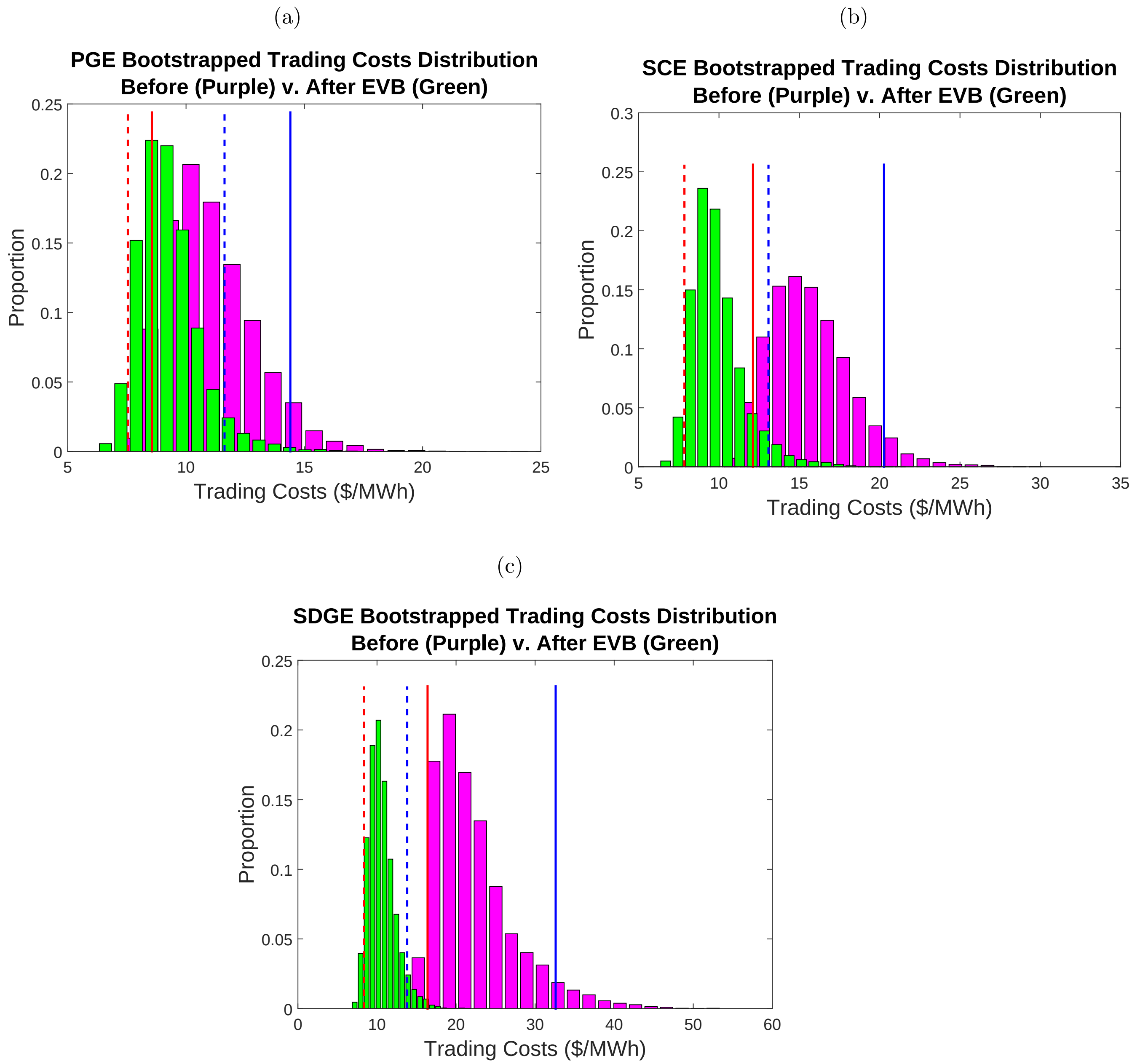
		Before EVB	After EVB
Lower 5% C.I	PG&E	8.591	7.531
	SCE	12.112	7.845
	SDG&E	16.453	8.393
Upper 95% C.I	PG&E	14.385	11.684
	SCE	20.185	13.209
	SDG&E	32.391	13.825

Notes: This table We first implement our trading cost hypothesis tests at the load aggregation point (LAP) level. These results are presented in Table 5. For each LAP, we report the values of c_{lower} and c_{upper} both before and after the introduction of explicit virtual bidding. Recall that c_{lower} is the smallest value of per-unit trading costs c such that the null hypothesis of the existence of an expected profit-maximizing trading strategy can be rejected and that c_{upper} is the largest value of c such that the null hypothesis that no profitable trading strategy exists can be rejected.

c_{lower} is the smallest value of per-unit trading costs c such that the null hypothesis of the existence of an expected profit-maximizing trading strategy can be rejected and that c_{upper} is the largest value of c such that the null hypothesis that no profitable trading strategy exists can be rejected. Table 5 demonstrates that the values of both c_{lower} and c_{upper} fall after the introduction of explicit virtual bidding for all LAPs. This is consistent with the logic outlined in Section 2 that the costs of trading day-ahead/real-time price differences decrease after the introduction of explicit virtual bidding (EVB). Figure 3 plots the bootstrap distributions of implied trading costs $\phi(\overline{X})$ for the pre-EVB and post-EVB sample periods for each of the three LAPs. The solid vertical lines on each graph are the values of c_{lower} and c_{upper} for the pre-EVB sample period and the dotted vertical lines on each graph are the values of c_{lower} and c_{upper} for the post-EVB sample.

To obtain a more formal comparison of the implied trading costs before versus after the introduction of explicit virtual bidding, we compute the bootstrap distribution of the difference in implied trading costs for each LAP before versus after the implementation of explicit virtual bidding. Figure 5 plots the bootstrap distribution of the difference in trading costs for each of the three LAPs. The left vertical line on the graph 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 5^{th} percentile of the distribution of $c_{pre} - c_{post}$ is greater than zero, then we can reject the $\alpha = 0.05$ test of the the null hypothesis that $c_{pre}^{true} \leq c_{post}^{true}$. If 95^{th} percentile of the distribution of $c_{pre} - c_{post}$ is less than zero, then we can reject the

Figure 3: Bootstrap Distribution of $\phi(\overline{X})$ with 95% C.I: Before and After EVB



Notes: This figure plots the bootstrap distributions of implied trading costs $\phi(\overline{X})$ for the sample periods before and after the introduction of explicit virtual bidding (EVB) for each of the three load aggregation points (LAPs). The three LAPs correspond to Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E). The solid vertical lines on each graph are the values of c_{lower} and c_{upper} for the pre-EVB sample period and the dotted vertical lines on each graph are the values of c_{lower} and c_{upper} for the post-EVB sample.

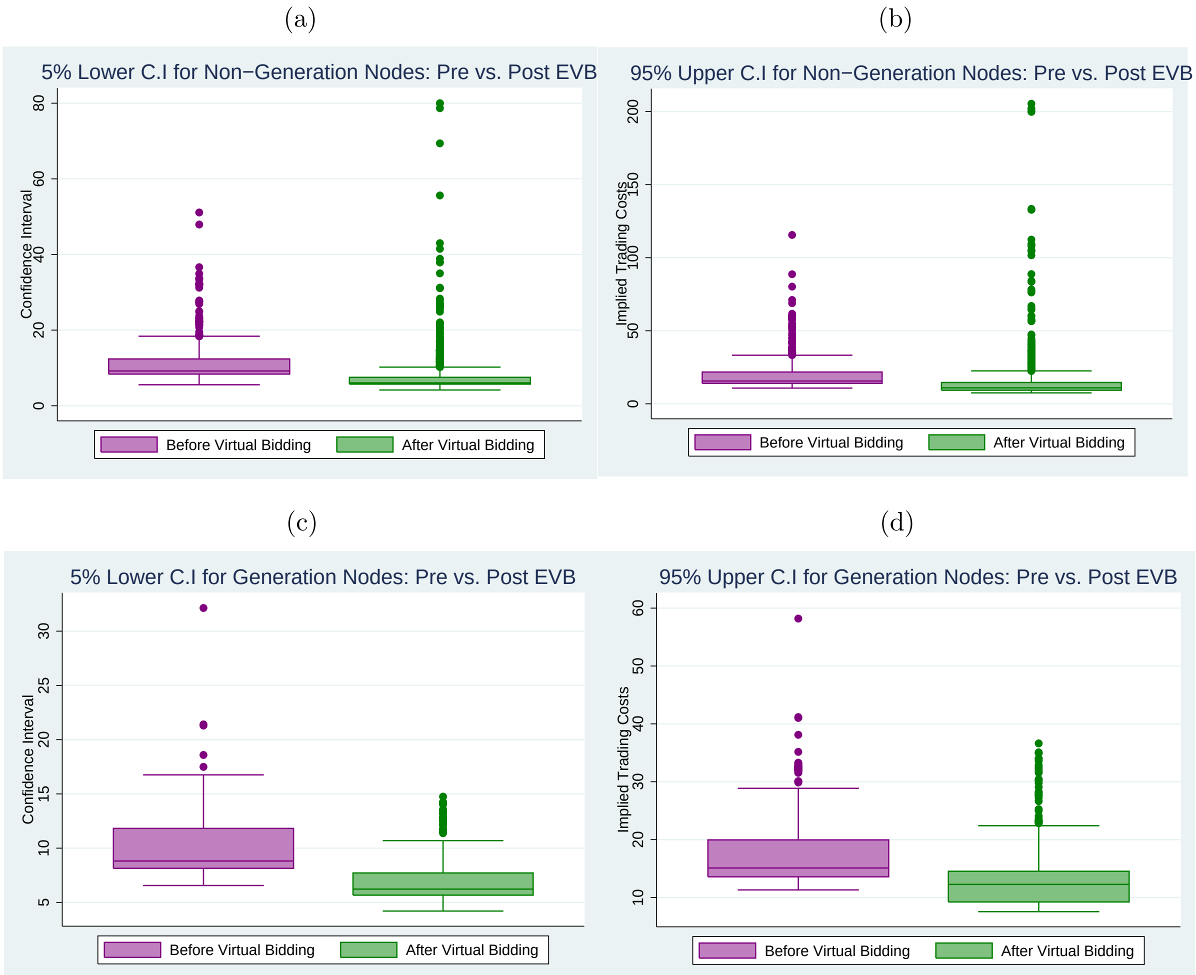
$\alpha = 0.05$ test of the the null hypothesis that $c_{pre} \geq c_{post}$. For both SCE and SDG&E, the null hypothesis that the difference in trading costs pre- versus post-EVB is less than or equal to zero can be rejected and the null hypothesis that the trading charges pre- versus post-EVB is greater than zero cannot be rejected. For PG&E, the null hypotheses cannot be rejected for both tests.

We can also compute the values of c_{lower} and c_{upper} for each node (read: location) in the California ISO control area. Figure 4 plots the values of c_{lower} and c_{upper} for each node before and after the introduction of EVB.[5] We plot the across-node distributions of c_{lower} and c_{upper} separately for nodes associated with generation units and nodes not associated with generation units. The inter-quartile range of the distribution these two measures of implied trading costs across nodes markedly decreases when we calculate it using data after the introduction of explicit virtual bidding. Recall that we reject the null hypothesis that a profitable trading strategy exists if actual trading costs are larger than than the values plotted for c_{lower}. Therefore, for any value of actual trading costs, we reject the null hypothesis that "a profitable trading strategy exist at a given node" for more nodes after the introduction of explicit virtual bidding.

We repeat the bootstrap estimation of the distribution of $c_{pre} - c_{post}$ for each of the more than 4,000 nodes in the California ISO control area. The first line of Table 6 reports the fraction of nodes for which the null hypothesis that $c_{pre}^{true} \leq c_{post}^{true}$ can be rejected, separately for generation nodes (Gen Node) and for non-generation nodes (Non-Gen Node). The second line of Table 6 reports fraction of nodes that the null hypothesis that $c_{pre}^{true} \geq c_{post}^{true}$ can be rejected for generation nodes (Gen Node) and for non-generation nodes (Non-Gen Node). The null hypothesis that the implicit trading charge increased after the introduction of EVB can be rejected for more than 70 percent of the nodes. The percentage of rejections at non-generation nodes is slightly higher than at generation nodes. For less than 5 percent of the nodes, the null hypothesis that the trading charge fell after the introduction of EVB can be rejected. This rejection frequency is consistent with this null hypothesis being true for all nodes, because the size of each hypothesis test is 0.05.

[5]Note that the box portion of box and whiskers plot corresponds to the 25% through 75% of the distribution of trading costs over nodes. The upper (lower) whisker corresponds to data points within 1.5(IQR) of the 75% (25%) quantile point, where IQR is the inter-quartile range defined by the distance between the 25% and 75% quartiles. Finally, the remaining points are outliers outside of the aforementioned range.

Figure 4: Nodal-Level Distribution of Confidence Intervals: Before and After EVB



Notes: This table We can also compute the values of c_{lower} and c_{upper} for each node (read: location) in the California ISO control area. Figure 4 plots the values of c_{lower} and c_{upper} for each node before and after the introduction of EVB.[6] We plot the across-node distributions of c_{lower} and c_{upper} separately for nodes associated with generation units and nodes not associated with generation units. The inter-quartile range of the distribution these two measures of implied trading costs across nodes markedly decreases when we calculate it using data after the introduction of explicit virtual bidding. Recall that we reject the null hypothesis that a profitable trading strategy exists if actual trading costs are larger than than the values plotted for c_{lower}. Therefore, for any value of actual trading costs, we reject the null hypothesis that "a profitable trading strategy exist at a given node" for more nodes after the introduction of explicit virtual bidding.

Table 6: Proportion of Nodes that Reject the Two Null Hypotheses

	Total	1(Gen Node)	1(Non-Gen Node)
1(5% Lower Bound>0)	0.707	0.659	0.711
1(95% Upper Bound<0)	0.042	0.076	0.039
Number of Observations	4316	355	3961

Notes: This table The first line of Table 6 reports the fraction of nodes for which the null hypothesis that $c_{pre}^{true} \leq c_{post}^{true}$ can be rejected, separately for generation nodes (Gen Node) and for non-generation nodes (Non-Gen Node). The second line of Table 6 reports fraction of nodes that the null hypothesis that $c_{pre}^{true} \geq c_{post}^{true}$ can be rejected for generation nodes (Gen Node) and for non-generation nodes (Non-Gen Node).

From the logic discussed in sections 2.2 and 2.3, we expect the following two relationships to hold between the true values of the implied trading costs across generation versus non-generation nodes before versus after the introduction of explicit virtual bidding. First, suppliers could only implicitly virtual bid at the nodal level before the implementation of the explicit virtual bidding through how they operated their generation units; therefore, load-serving entities (read: demanders) could only bid at the LAP level before explicit virtual bidding. Therefore, we expect the implied trading costs to be higher at non-generation nodes relative to generation nodes before the implementation of explicit virtual bidding, as neither suppliers nor demanders could implicitly virtual bid at non-generation nodes. The introduction of explicit virtual bidding allowed any market participant to place virtual bids non-generation nodes; based on this, we expect the mean reduction in implied trading costs after EVB to be larger for non-generation nodes relative to generation nodes. To test these two hypotheses, we regressed the value of c_{lower} at each node both before and after explicit virtual bidding on a constant, an indicator variable for whether the node was a generation node, an indicator variable for whether the implied trading cost was from the post-EVB period, and an indicator variable for whether the observation was from a generation node during the post-EVB period (the interaction term between "generation node" and "post-EVB"). We also run the same node/pre-vs-post EVB level regression with c_{upper} as the dependent variable.

Table 7 reports the results of estimating these difference-in-differences style regressions for c_{lower} and c_{upper}. For both percentiles of the distribution of nodal-level implicit trading costs, we find strong evidence consistent with both of our hypotheses. The best linear prediction of both c_{lower} and c_{upper} before the introduction of explicit virtual bidding is significantly lower for generation nodes and this difference is essentially eliminated after the introduction of explicit virtual bidding. Specifically, for both c_{lower} and c_{upper}, we fail to reject the null hypothesis that the sum of the coefficient on "Generation Node Indicator" and the coefficient on "Interaction Between Generation Node and Post EVB Indicator" is zero. Therefore, as we expected, the difference in implied trading cost before versus after explicit virtual bidding (i.e: $c^{pre} - c^{post}$ fell more for non-generation nodes than for generation nodes.

Figure 6 contains monthly average hourly virtual supply offered and cleared and virtual demand offered and cleared for October 2011 to December 2012.; this figure is taken directly from the California ISO Department of Market Monitoring's Q4 Report on Market Issues and Performance (issued February 13, 2013). This graph shows that slightly less than 1,000 MWh of virtual supply clears each hour and approximately the same level of virtual demand clears each hour. Roughly half of the virtual supply and virtual demand offers that are submitted each hour clear. Because there are over 4,000 nodes in the ISO system and the minimum virtual bid offer is 1 MWh, there are many nodes each hour that do not receive nodal-level virtual bids. Figure 7 shows the average offered and cleared virtual demand and supply bids by hour of the day for October to December of 2012. For demand bids, there are significantly higher levels of offered and cleared bids during the peak demand hours of the day; in contrast, the patterns of both virtual supply offered and cleared bids is fairly constant throughout the day.

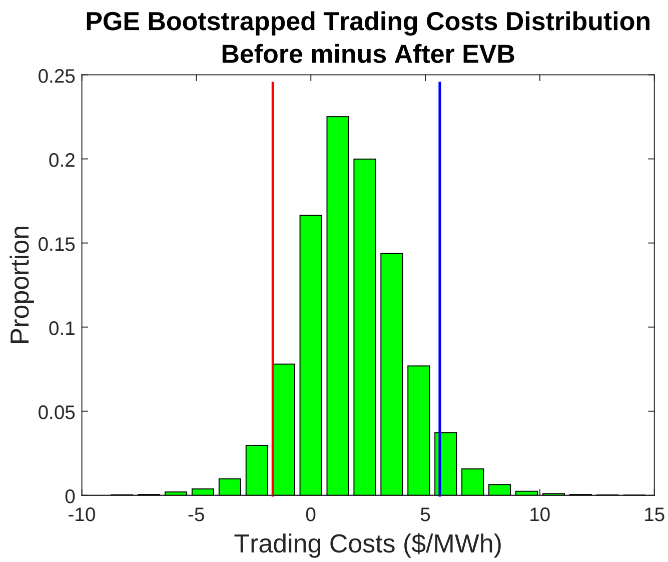
5.3 Results from Test for a Fall in Trading Profits

In this section, we implement the direct tests that $|\mu_{pre}| > |\mu_{post}|$ and $|\mu_{post}| > |\mu_{pre}|$. If we assume that $c^{pre} > c^{post}$ as appears to be the case from the implied trading cost results presented in the previous section, then $|\mu_{pre}| > |\mu_{post}|$ is a test of the null hypothesis that expected trading profits declined as a result of the introduction of explicit virtual bidding. The p-values corresponding to these tests for each LAP are presented below in Table 8:

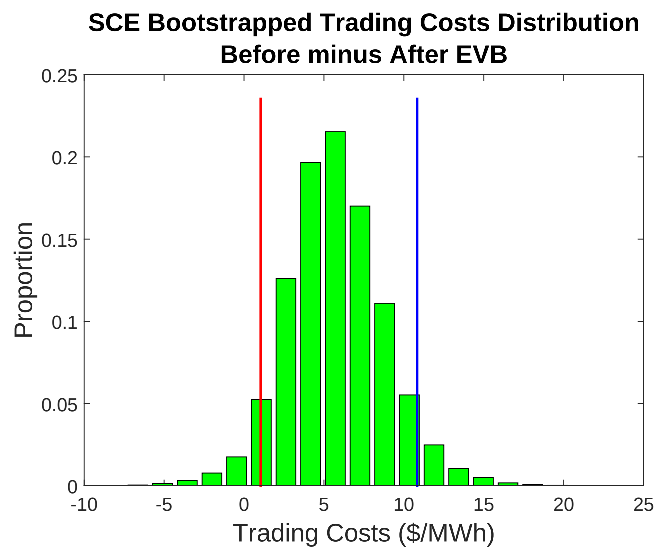
We cannot reject the null hypothesis that $|\mu_{pre}| > |\mu_{post}|$ for any of the three LAPs, while we can reject the null hypothesis that $|\mu_{post}| > |\mu_{pre}|$ at the 5% level for two of the three LAPs. If $c^{pre} > c^{post}$, as is implied by the results in Figure3 and Table5, these

Figure 5: Bootstrap Distribution of the Difference in Trading Costs

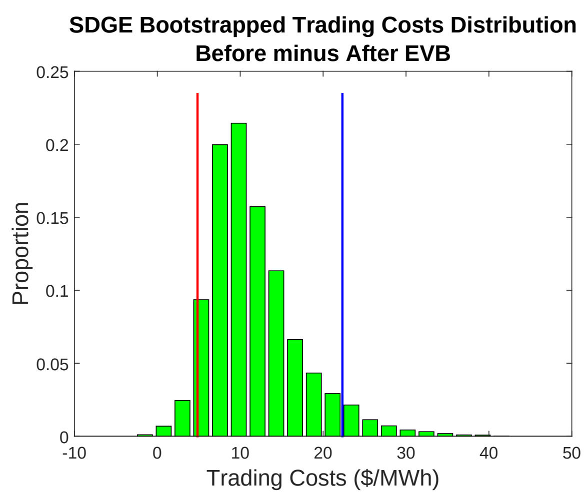
(a)



(b)

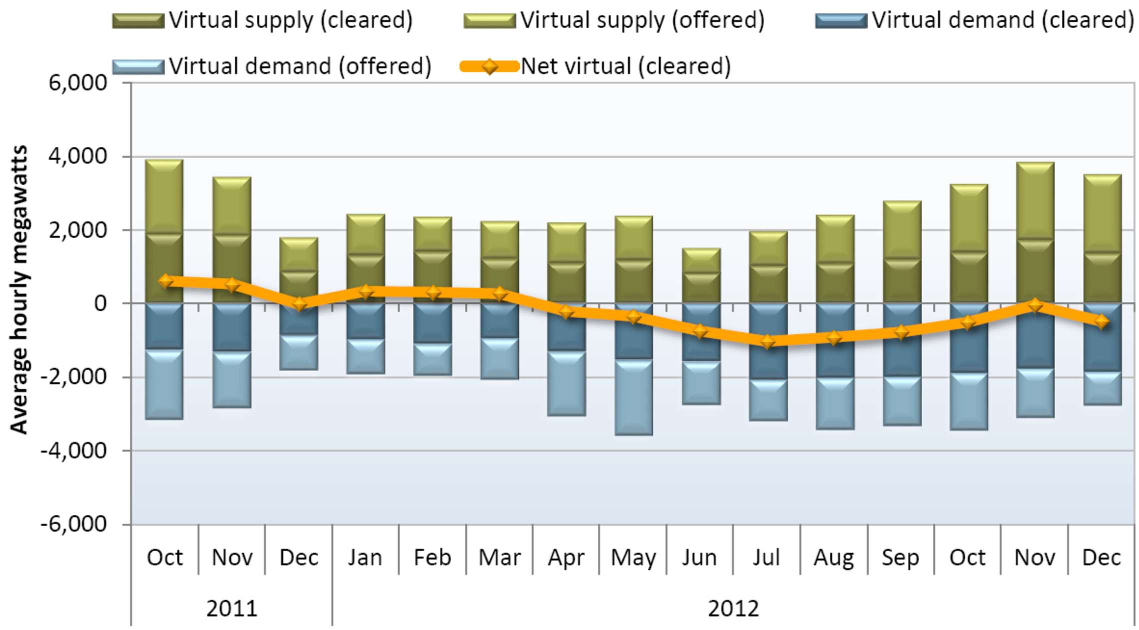


(c)



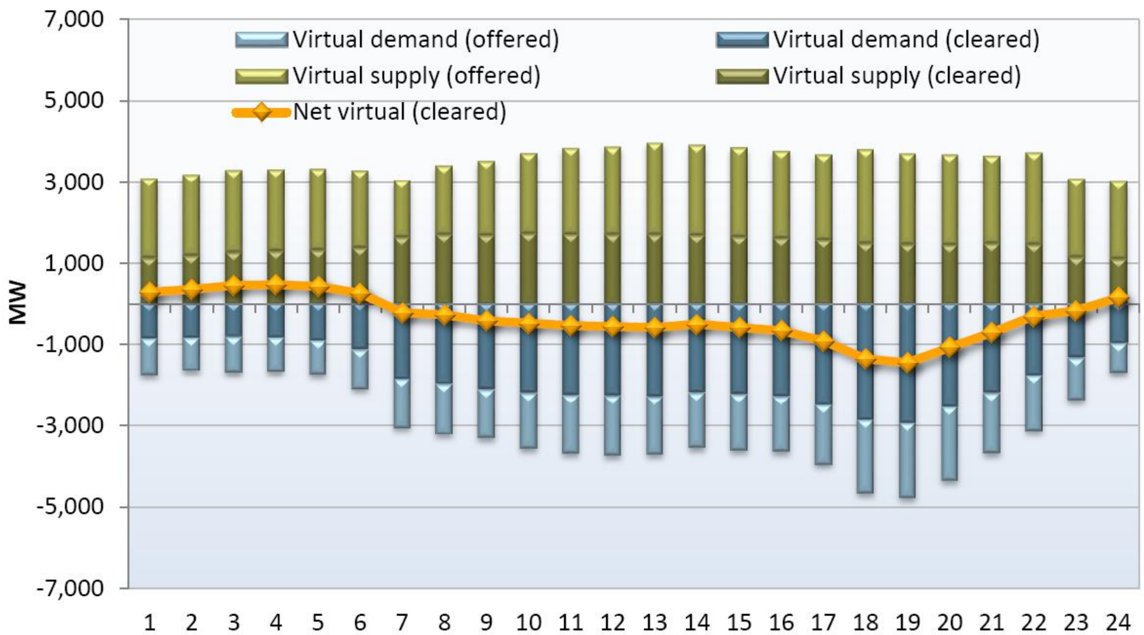
Notes: This table

Figure 6: Average Hourly MW Virtual Supply and Demand Offered and Cleared: Monthly



Notes: This table

Figure 7: Average Hourly MW Virtual Supply and Demand Offered and Cleared: Hourly



Notes: This table

Table 7: Regression Results Associated with Implied Trading Costs–c_{lower} and c_{upper}

VARIABLES	(1) 5% Lower Bound	(2) 95% Upper Bound
1(Post EVB)*1(Gen Node)	0.532	1.421
	(0.174)	(0.431)
1(Post EVB)	-3.527	-5.404
	(0.0752)	(0.193)
1(Gen Node)	-0.654	-1.765
	(0.119)	(0.250)
Constant	10.72	19.16
	(0.0538)	(0.118)
Observations	9,791	9,791
R-squared	0.202	0.080

Robust standard errors in parentheses

Notes: This table

Table 8: P-values associated with the Absolute Difference Tests

	$\lvert\mu_{pre}\rvert > \lvert\mu_{post}\rvert$	$\lvert\mu_{post}\rvert > \lvert\mu_{pre}\rvert$
PG&E	0.705	0.144
SCE	0.908	0.006
SDG&E	0.687	0.040

Notes: This table

Table 9: P-values associated with Volatility Tests

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

Notes: This table

hypothesis testing results provide evidence in favor of the view that trading profits fell after the introduction of explicit virtual bidding.

5.4 Results from Test for a Fall in Volatility

As outlined in Section 4.6, we expect that the introduction of explicit virtual bidding results in a fall in the volatility of the day-ahead/real-time price difference, as well as the volatility of the real-time price itself. Formally, we compare the covariance matrices associated with the price differences (and real-time prices) prior to versus after explicit virtual bidding, testing whether the difference between the covariance matrices is a positive semi-definite matrix. Formally, this is a (24x1) multivariate nonlinear inequality constraints test on the eigenvalues of the difference between the two covariance matrices. These results are documented in Table 9, where we report the probability of obtaining a value from the distribution of the test statistic under the null hypothesis greater than the actual statistic (i.e: the p-value). We reject a size $\alpha = 0.05$ test if this probability is less than 0.05. We fail to reject the null hypothesis if it is greater than 0.05.

We fail to reject the null hypothesis that the daily price differences and real time prices prior to explicit virtual bidding are more volatile relative to the differences and real-time prices after explicit virtual bidding. Moreover, we reject the opposite null hypothesis that the volatility of price differences and real-time prices are lower pre-EVB versus post-EVB for all cases but the real-time price results for SDG&E. These results are consistent with the claim that explicit virtual bidding resulted in the day-ahead market producing generation and load schedules closer to actual physical conditions in the real-time market, leading to less "residual" deviations between day-ahead schedules

Table 10: Proportion of Nodes for which we fail to reject the $\alpha = 0.05$ sized Volatility Test

	Node Type	Price Difference	Day-Ahead Price	Real-Time Price
Pre-Post	1(Gen Node)	0.736	0.049	0.751
	1(Non-Gen Node)	0.744	0.061	0.758
	Total	0.737	0.051	0.752
Post-Pre	1(Gen Node)	0.168	0.902	0.132
	1(Non-Gen Node)	0.156	0.907	0.104
	Total	0.166	0.902	0.128

Notes: This table

and real-time market outcomes. This result is consistent with generation unit owners and load-serving entities taking costly actions to attempt to profit from differences between day-ahead and real-time prices prior to explicit virtual bidding leading to a more frequent need for the ISO to make significant adjustments to day-ahead generation schedules to meet real-time demand at all locations in the transmission network. Therefore, prior to explicit virtual bidding, large differences between day-ahead and real-time prices reflected both genuine shocks to the electricity production process as well as financially motivated distortions in bid and offer behavior motivated by divergent expectations over day-ahead versus real-time prices. The logic underlying the cause of these variance reduction results is consistent with the market efficiency results presented in the next section.

6 Measuring Market Efficiency Implications of Explicit Virtual Bidding

This section describes the data used and analysis performed to assess the market efficiency consequences of the introduction of explicit virtual bidding. The three market outcome measures we compare before versus after the introduction of explicit virtual bidding are: (1) the total millions of British Thermal Units (MMBTUs) used each hour to produce the fossil fuel electricity generated during that hour, (2) the total variable cost of producing the fossil fuel electricity generated during that hour, and (3) the total number of thermal generation units started during that hour.[7] Explicit virtual bidding was introduced in

[7]The vast majority of thermal-based generation in California comes from natural gas; there is a small amount of generation from oil-fired sources.

2/1/2011; for this section, we consider a 2/1/2010-2/1/2012 sample period that spans exactly one year both before and after the introduction of this policy.

We employ the Robinson (1988) partially linear model to estimate the effects of explicit virtual bidding on the conditional mean function for each of our three market performance measures. We non-parametrically control for differences in California's total electricity generation from all sources, California's total net imports of electricity, total output from all intermittent renewable resources[8] as a percentage of overall electricity generation and daily wholesale prices of natural gas delivered to northern California and to southern California. Finally, all specifications additionally control for differences across hours of the day and months of the year.

6.1 Data Description

For this section, we utilize hourly data on generator-level output from California's independent system operator (CAISO), daily natural gas prices for each of Northern and Southern California from SNL Financial[9], and daily oil prices from the Energy Information Administration. We also use generator-level data on capacity, fuel type (i.e: natural gas, oil, wind, etc.), and heat rate curve; these variables are provided by CAISO.

We construct the total hourly MMBTUs of energy consumed by all thermal-based generation units, our first market outcome measure, as follows. First, we combine the hourly metered output of each thermal generation unit, obtained from CAISO's settlement system, with the generation unit-level heat rate curve that all generation unit owners are required to submit as part of CAISO's local market power mitigation mechanism. This curve is a piecewise linear function that can have up to ten pairs of heat rate level (in MMBTU per MWh) and output quantity (in MWh) steps, where the sum of the output quantity steps equals the generator unit's capacity. Put another way, a heat rate curve tells us: "if a generation unit is currently utilizing X% of their generation capacity, we need Y mmBTU of thermal input in order to produce one more unit of generation (in MWh)". The heat rate value on this piecewise linear curve times the generation unit's metered output for that hour is the first component of the total MMBTUs of energy

[8]This category includes wind, solar, biomass, and biogas based sources, as well as hydroelectric-based sources with capacity less than 30MW.

[9]We use natural gas prices at the Pacific Gas and Electric Citygate delivery point and the Southern California Gas Citygate delivery point respectively.

consumed by that generation unit during the hour.

However, thermal energy is also required to start up a generation unit; from our data, a unit is defined as starting in hour t if its output in hour t-1 is zero and its output in hour t is greater than zero. Generation unit owners are also required to file information on the total amount of MMBTUs required to start each generation unit with CAISO as part of its local market power mitigation mechanism. Thus, we sum the thermal energy required to start up the thermal units that started up in hour-of-sample t and the thermal energy required for all thermal generation in hour-of-sample t to calculate our first market outcome measure TOTAL ENERGY$_t$. We also consider as a market outcome measure the total number of thermal-based generation units started in an hour t, $STARTS_t$.

The final market performance measure, TOTAL VC$_t$, is the total variable cost of all thermal-based generation units in hour t. The marginal cost for each generation unit is computed by multiplying the heat rate associated the unit's metered output for that hour (computed from the piecewise linear heat-rate curve) times the daily fuel price (either natural gas or oil) for that unit plus the variable operating and maintenance cost that the unit's owner submits to CAISO's local market power mitigation mechanism. The start-up cost associated with a unit that starts up in hour-of-sample t is calculated by multiplying the MMBTUs of energy consumed to start the unit by the daily fuel price. The total variable cost for a given unit in a given hour is simply the unit's marginal cost multiplied by its metered output, which is then added to the unit's start-up costs in that hour (if any). Summing these unit-level variable costs over all generation units operating in hour t yields the value of TOTAL VC$_t$.

We specify semi-parametric functions for each of the three market performance measures in order to estimate the difference in the mean of each of the three hourly market performance measures before versus after the implementation of explicit virtual bidding. The conditional mean functions can be written as:

$$y_t = W_t'\alpha + X_t'\beta + \theta(Z_t) + \epsilon_t \tag{5}$$

where y_t is one of our three market performance measures. The function $\theta(Z_t)$ is an unknown function of the vector Z_t, W_t is a (24x1) vector of hour-of-day dummy variables, and α and β are unknown parameter vectors. X_t is a single dummy variable that takes on the value 1 for all hours after midnight January 31, 2011 and zero otherwise; X_t is an indicator for whether hour-of-sample t is before versus after explicit virtual bidding. Finally, Z_t is five dimensional vector composed of the (log of) total output in MWhs of all generation units in California during hour t, the (log of) total imports minus total exports in MWhs for California during hour t, the total output in MWhs of all renewable generation units in California during hour t divided by the total output in MWhs of all generation units in California during hour t, the (logged) price of natural gas in northern California during hour t, and the (logged) price of natural gas in Southern California during hour t. We consider three dependent variables: the natural logarithm of total variable costs (including start-up costs) of thermal generation in hour t, $log(\text{TOTAL VC})_t$, the natural logarithm of total input energy used (in MMBTU) in hour t, $log(\text{TOTAL ENERGY})_t$, and the total number of thermal units starting up in hour t, $STARTS_t$. We also estimate specifications allowing the mean effects of explicit virtual bidding to vary by the hour of the day. In this case, X_t is a (24x1) vector with k^{th} element $X_{t,k}$; $X_{t,k}$ equals one during hour-of-the-day k for all days from February 1, 2011 until the end of the sample period. For all specifications, we assume that $E[\epsilon_t|(W_t, X_t, Z_t)] = 0$.

Our goal in this section is to answer: "for the same level of output, how are total variable costs, input levels, and total starts (our market outcome measures) affected by explicit virtual bidding?". Due to this, we first control for California's hourly total in-state generation; Figures 8 and 9 show that there are slight differences in the averages and standard deviations over days-of-sample by hour-of-the-day in total in-state generation before versus after explicit virtual bidding. However, to isolate the effects of explicit virtual bidding on our market outcome measures, we control for other factors as well. First, we flexibly account for differences in the hourly percentage of total generation coming from renewable sources; this control is necessary because the share of electricity generated from renewable resources grew significantly over our sample period. This growth is the result of California's Renewables Portfolio Standard (RPS), which requires all California load-serving entities to procure 33 percent of their electricity generation from qualified renewable sources by 2020. Figure 10 plots the average over days-of-sample for each hour-of-the-day of the percentage of in-state generation from renewable resources during the year before virtual bidding and year after virtual bidding; we see a roughly 2% increase in the average percentage of generation from renewable sources in the 2/1/2011-2/1/2012 sample period relative to the 2/1/2010-2/1/2011 sample period. Moreover, Figure **??** demonstrates that the standard deviation of this percentage also increases substantially after explicit virtual bidding, speaking to the importance of the flexibility provided by *non-parametrically* controlling for differences in percentage of generation from renewables. In practice, this increased volatility of renewable resources (termed "intermittancy") implies that more thermal resources must be held as operating reserves and stand ready to supply additional energy if the renewable resources disappear suddenly. We also control flexibly for California's overall hourly net imports of electricity, as both the average and standard deviation of net imports increases substantially in the sample period after explicit virtual bidding was introduced; see Figures 12 and 13 for these summary statistics taken over days-of-sample by hour-of-the-day, separately for the sample periods before versus after explicit virtual bidding. Finally, we also flexibly control for differences in the natural gas prices in Northern California and Southern California, as these two price series obviously affect the thermal resources firms decide to utilize in a given hour-of-sample.

6.2 Empirical Methodology

Examining Equation 5, if we assume that $E[\epsilon_t|X_t, Z_t, W_t] = 0$, we can estimate parameters (α, β) using the procedure described in Robinson (1988). This procedure has the following two steps:

1. We non-parametrically estimate $E[\hat{y_t}|Z_t]$, $E[\hat{W_t}|Z_t]$ and $E[\hat{X_t}|Z_t]$.[10]

2. From Equation 5, we can see that:

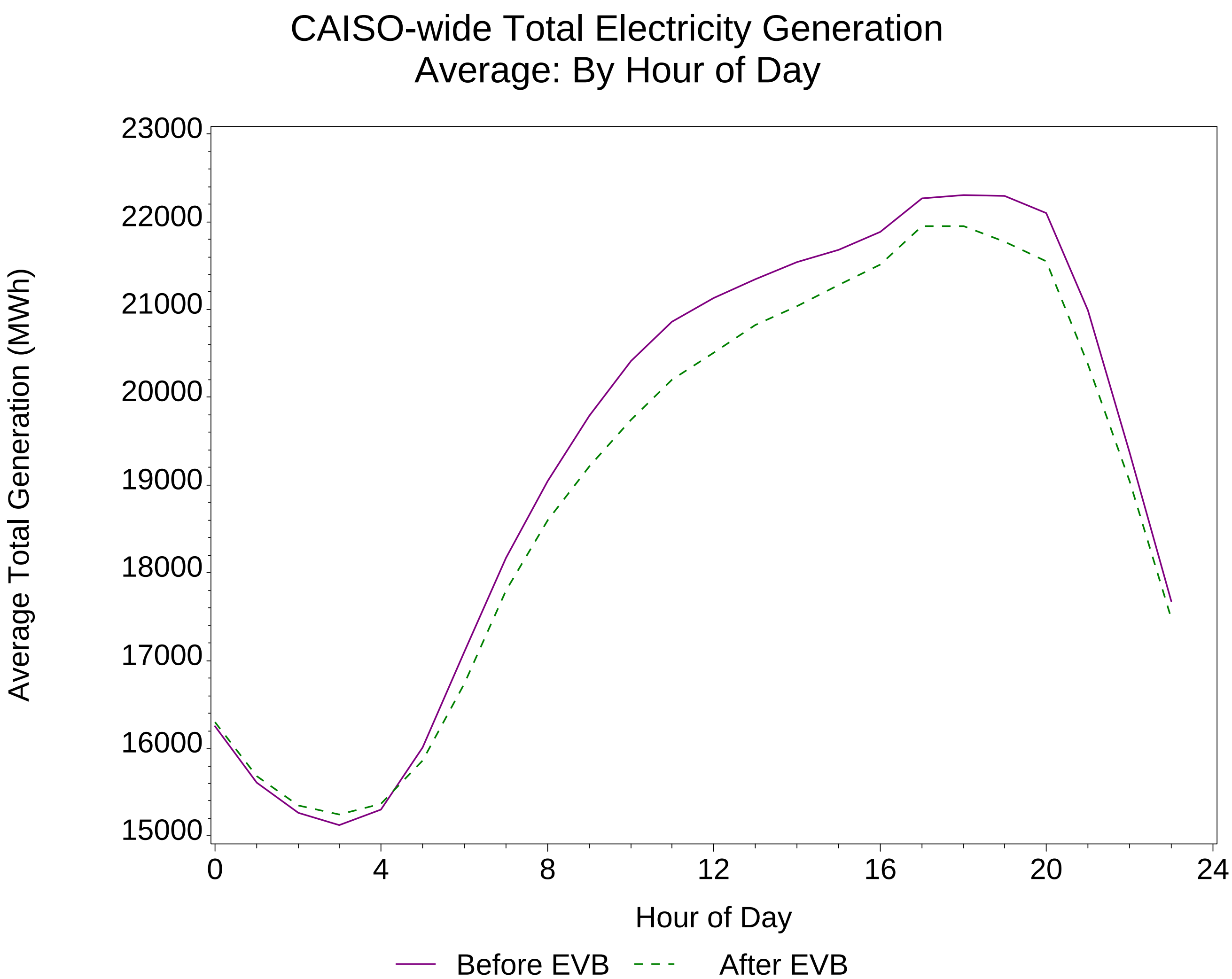
$$E[y_t|Z_t] = E[W_t|Z_t]'\alpha + E[X_t|Z_t]'\beta + \theta(Z_t)$$

   Solving for $\theta(Z_t)$ and plugging back into Equation 5:

$$y_t = W_t'\alpha + X_t'\beta + E[y_t|Z_t] - E[W_t|Z_t]'\alpha - E[X_t|Z_t]'\beta + \epsilon_t$$

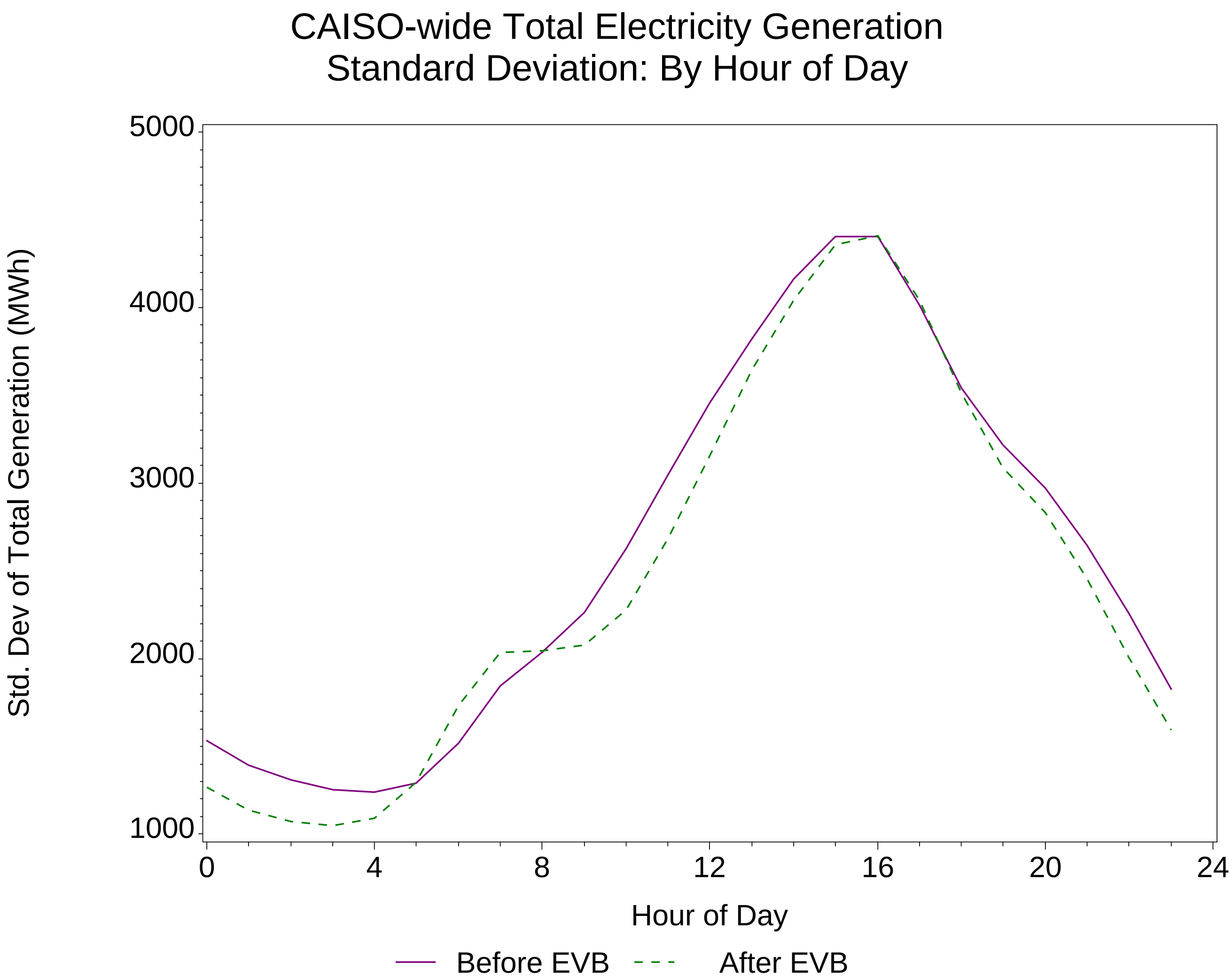
[10] We use the Nadaraya-Watson kernel regression estimator with a Gaussian kernel. We find optimal smoothing parameters separately for each covariate using leave-one-out cross validation. Finally, we report results based on jointly standardizing the independent variables using the unconditional averages and empirical covariance matrix estimated with our whole 2/1/2010-2/1/2012 sample period (i.e: we apply the Mahalanobis transformation $\tilde{X}_i \equiv (X_i - \overline{X})\hat{\Sigma}^{\frac{-1}{2}}$). As expected, the decision to standardize jointly, standardize each covariate individually, or not standardize the data at all does not significantly affect our results.

Figure 8: Average Total Generation from All Sources: By Hour of Day



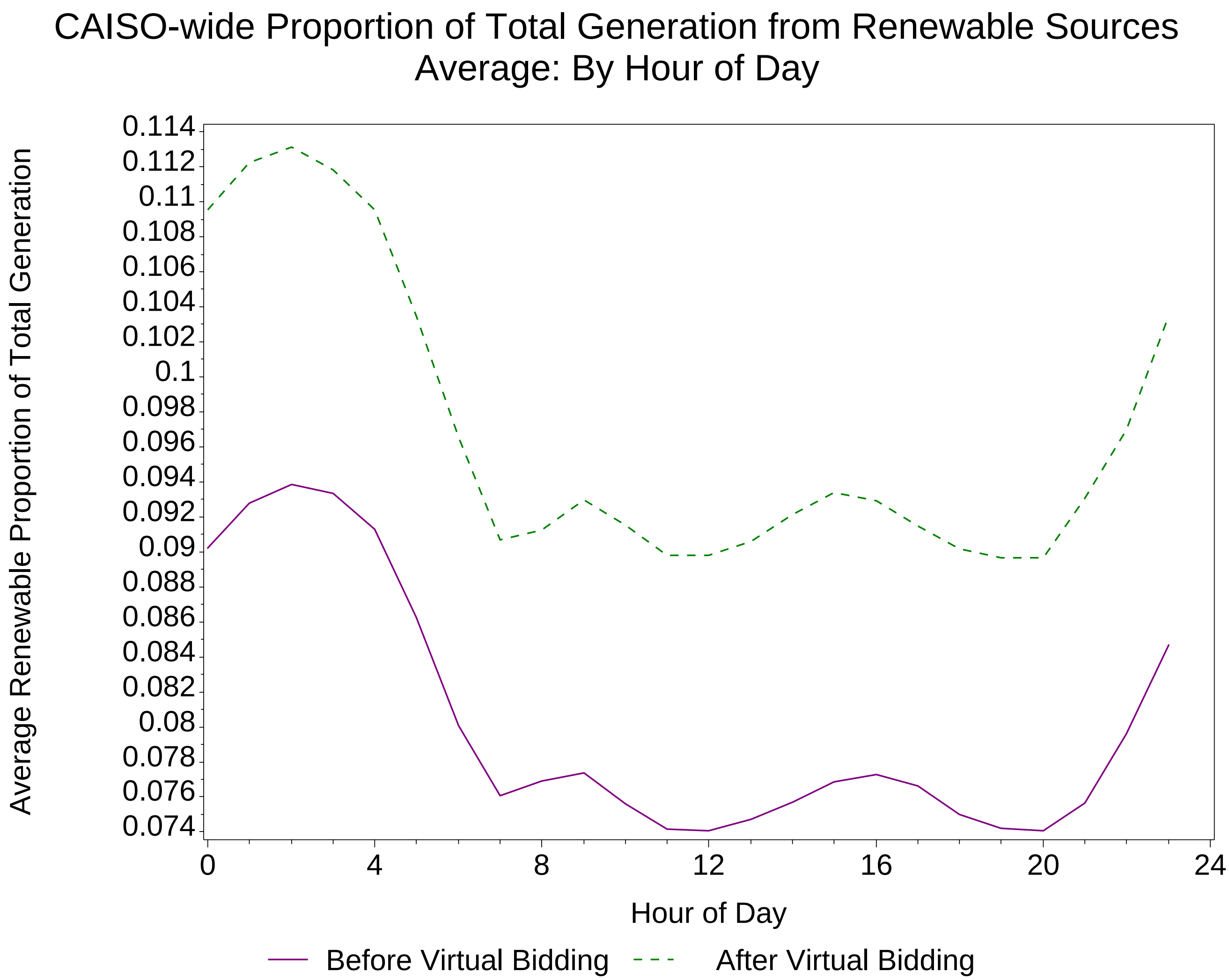
Notes: This table

Figure 9: Std. Dev. of Total Generation from All Sources: By Hour of Day



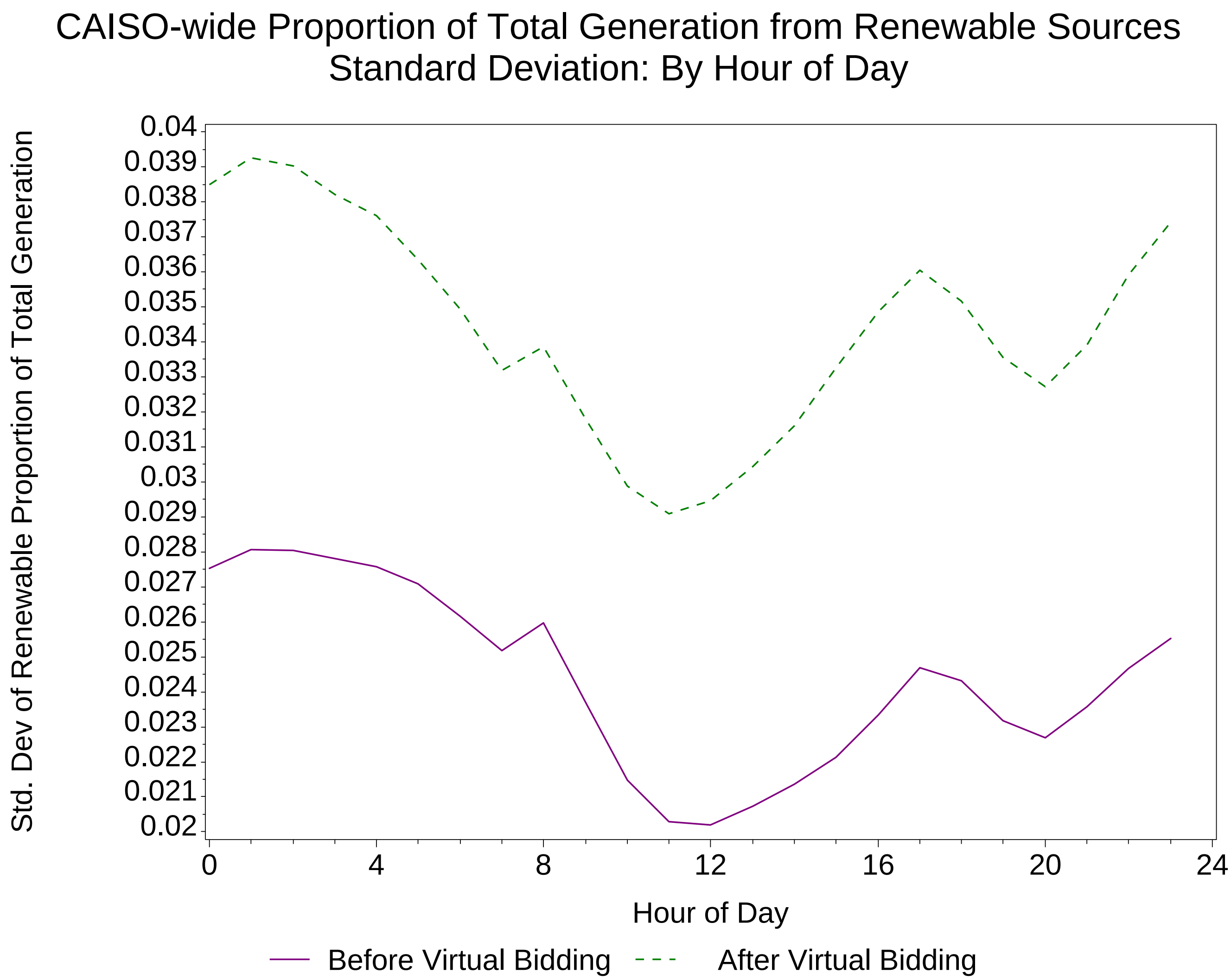
Notes: This table

Figure 10: Average Renewable Proportion of Total Generation: By Hour of Day



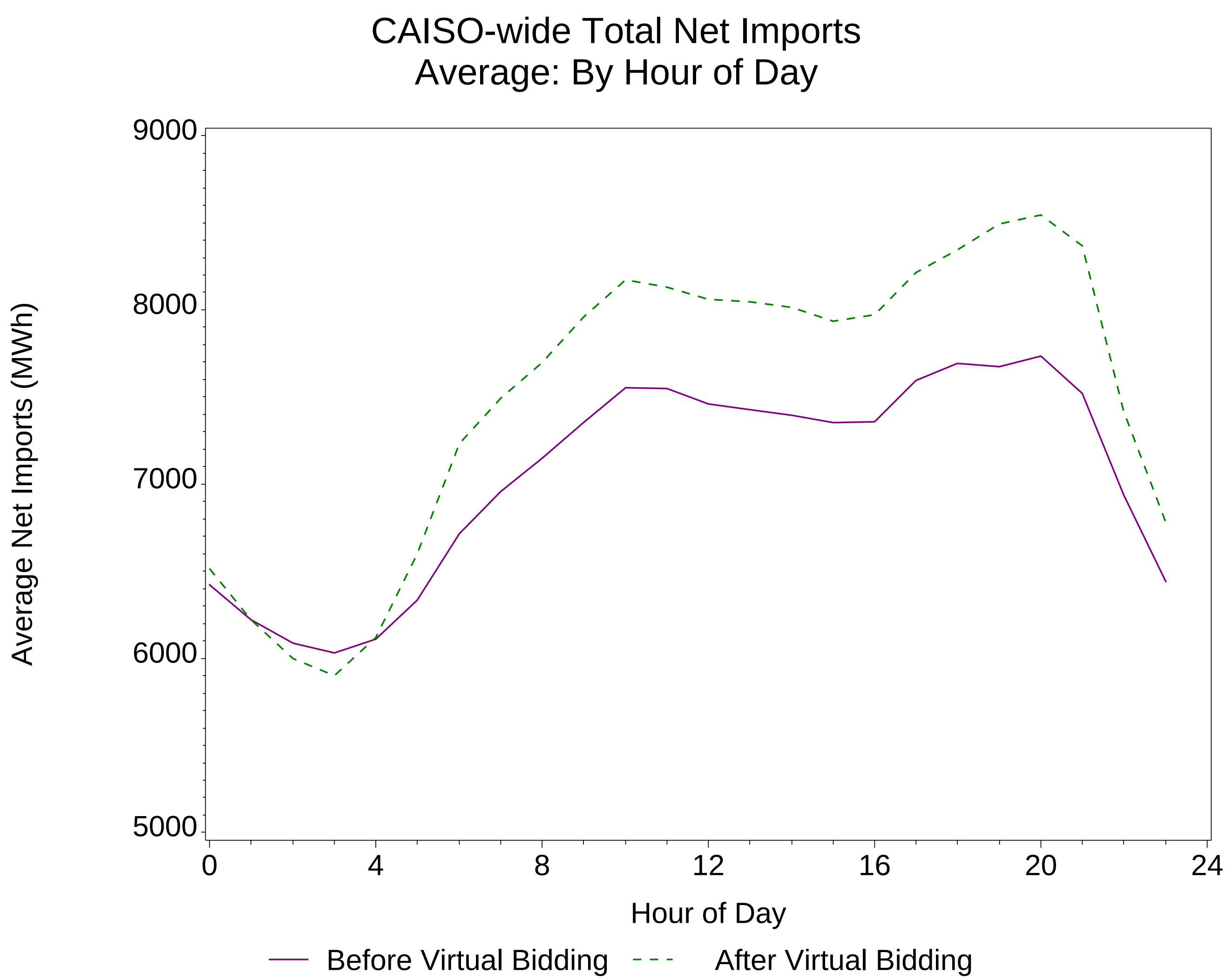
Notes: This table

Figure 11: Std. Dev. of Renewable Proportion of Total Generation: By Hour of Day



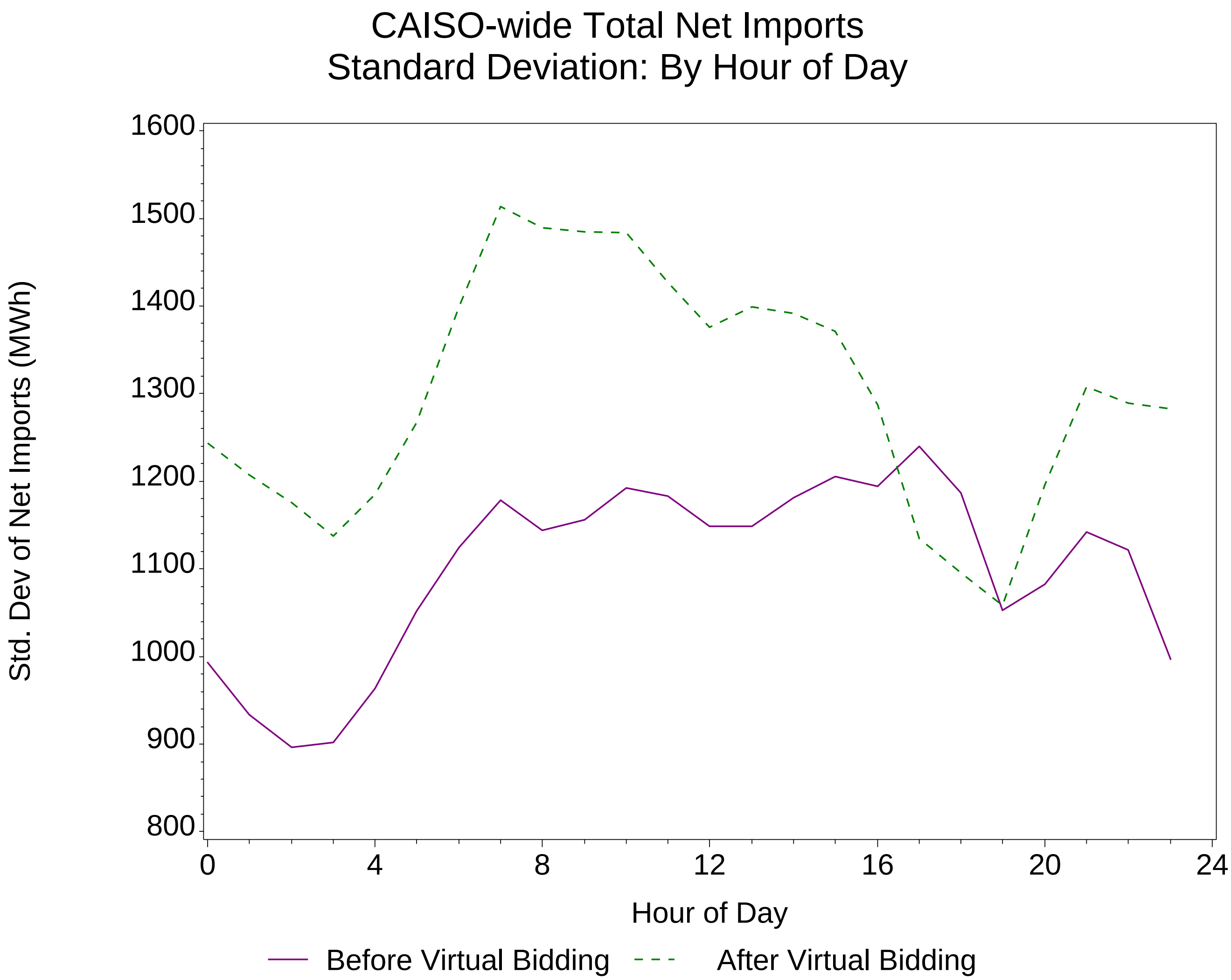
Notes: This table

Figure 12: Average Total Net Imports: By Hour of Day



Notes: This table

Figure 13: Std. Dev. of Total Net Imports: By Hour of Day



Notes: This table

Finally, we can plug in our first stage estimates and consistently estimate (α, β) by running an ordinary least squares regression of y_t on $(E[\hat{y_t}|Z_t], W_t - E[\hat{W_t}|Z_t], X_t - E[\hat{X_t}|Z_t])$.

A key innovation from Robinson (1988) is that the standard errors for $(\hat{\alpha}, \hat{\beta})$ reported from the second stage ordinary least squares regression are asymptotically valid; we report robust standard errors for all specifications.

6.3 Empirical Results

Table 11 reports the coefficient estimate for X_t for each measure of market performance, where X_t is a single dummy variable that takes on the value 1 for all hours after midnight on January 31, 2011 and zero otherwise. These estimates imply that the conditional mean of total variable costs (controlling for the total hourly output from all generation units in California, California's total hourly net imports, the total hourly output from renewable resources divided by total hourly generation, the prices of natural gas in northern and southern California and the hour of the day) is 6.8 percent lower after February 1, 2011. The conditional mean of total hourly energy is 6.2 percent lower after February 1, 2011. Finally, the conditional mean of total hourly starts (controlling for the same variables) is 0.34 starts higher after February 1, 2011. Interpreting these findings, we see that the "right" (read: lower cost) thermal units started up more often after the introduction of explicit virtual bidding; this is suggestive evidence that day-ahead prices provided a better signal of real-time conditions after financial trading was introduced.

Table 11: Semiparametric Coefficient Results

Dependent variable	$log(\text{TOTAL COSTS})_t$	$log(\text{TOTAL INPUT ENERGY})_t$	$STARTS_t$
β	-0.0678	-0.0615	0.3434
Standard error	0.0100	0.0101	0.0672

Notes: This table

Figures 14 plots the estimates of hour-of-the-day change in the conditional mean of the three hourly market performance measures after the implementation of explicit virtual bidding along with the point-wise upper and lower 95% confidence intervals for each hour-of-the-day estimate. For the case of total hourly energy, the largest in absolute value reduction occurs in the early morning hours beginning at 12 am and ending at 3 am. The hourly mean reductions are the smallest in absolute value during the hours beginning 5 am and ending at 8 am, with the remaining hours of the day slightly higher in absolute value. For total starts, the largest increase is during the hour starting at 3 pm and ending at 5 pm. Starts also increase after the implementation of explicit virtual bidding in hours beginning with 4 am and ending at 7 am. For total variable costs, the pattern of the absolute values of the hour-of-the-day reductions is similar to that for total hourly energy. The largest in absolute value reductions occur in morning hours from 12 am to 3 am.

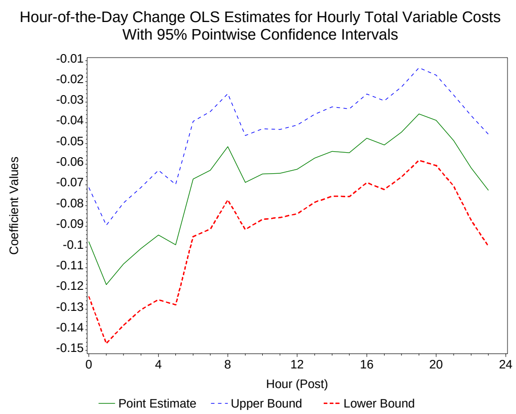
Although the percent hourly total energy and cost reductions are small, on an annual basis the implied cost savings and carbon dioxide emissions reductions can be substantial. The annual total cost of fossil fuel energy is $2.8 billion the year before explicit virtual bidding and $2.2 billion the year after explicit virtual bidding. Applying the 2.6 percent reduction to these figures implies an annual cost savings for the variable cost of fossil fuel energy of roughly 70 million dollars per year. Applying the total MMBTU figures, implies that the introduction of explicit virtual bidding reduced the greenhouse gas emissions from fossil fuel generation in California by 2.8 percent. The average heat rate of fossil fuel units in California is approximately 9 MMBTU/MWh and the typical natural gas-fired generation unit produces approximately a half of a ton of carbon dioxide per MWh of energy produced. In the year before explicit virtual bidding, 585 million MMBTUs were consumed to produce electricity and the year after 484 million MMBTUs were consumed. Applying our 2.8 percent reduction figure to these two numbers implies that the introduction of explicit virtual bidding reduced carbon dioxide emissions by between 650,000 and 537,000 tons annually. Both of these results point to sizable economic and environmental benefits from the introduction of explicit virtual bidding in California.

7 Implications of Results for Design of Electricity Markets

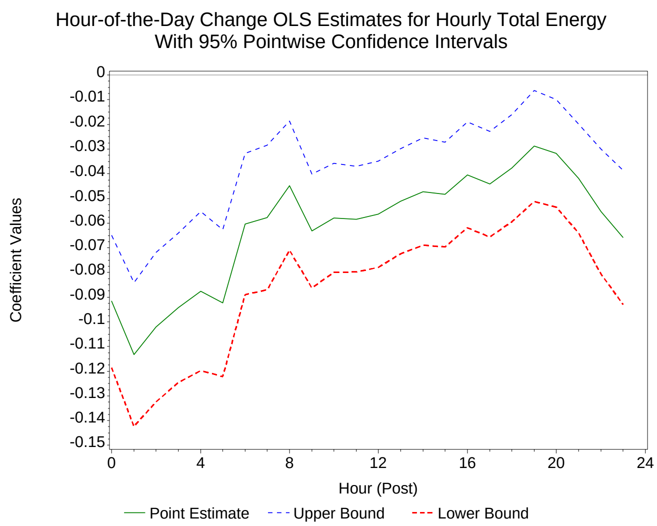
The results in the previous sections provide evidence that the introduction of explicit virtual bidding significantly reduced the transactions costs associated with attempting to profit from differences between the day-ahead and real-time market prices at the same

Figure 14: Hour-of-the-Day Percent Change Estimates from Semi-Parametric Regressions

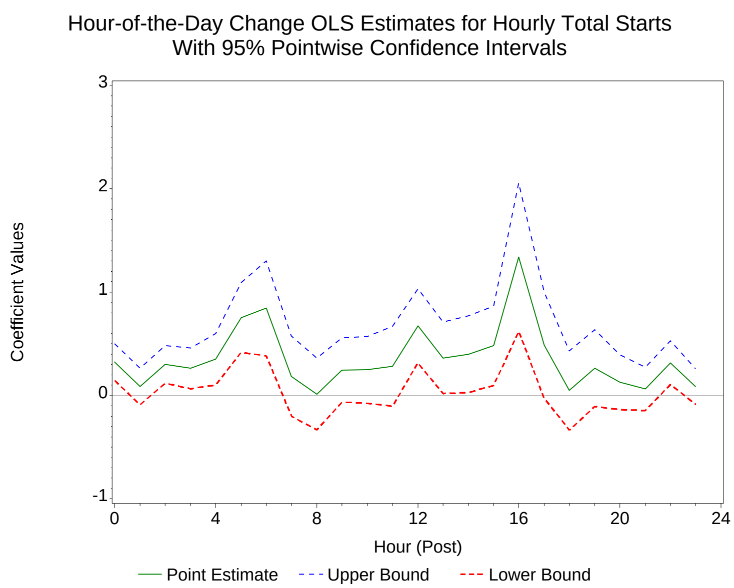
(a)



(b)



(c)



Notes: This table

location in the transmission network. In addition, these results demonstrate economically significant economic and global environmental benefits associated with the introduction of explicit virtual bidding. Although it was possible to implicit virtual bid before the introduction of explicit virtual bidding, the evidence from our analysis is that the introduction of this product significantly improved the degree of price convergence between the day-ahead and real-time markets and reduced the cost of serving load in the California ISO control area.

These results emphasize an important role for forward financial markets in improving the performance of short-term commodity markets. The financial commitments that producers and consumers make in forward markets can provide important information and feedback to market participants that improves the subsequent performance of short-term physical markets. Although explicit virtual bids are purely financial transactions, they reduce the incentive of both generation unit owners and load-serving entities to take forward market positions designed to raise prices in the short-term market. These results argue in favor of recognizing the fundamentally financial nature of day-ahead wholesale electricity markets. If explicitly financial products are not available, markets participants will still attempt to engage in profitable financial transactions, even though these transactions may require costly deviations from what the generation unit owner would do if explicit virtual bidding was possible. This appears to be the case before virtual bidding was implemented in the California market. Therefore, rather than resisting the desire of many market participants to allow purely financial transactions, these actions should be allowed and encouraged through explicit virtual bidding as a way to improve the performance of the wholesale electricity market.

References

Berger, Roger L, and Jason C Hsu. 1996. "Bioequivalence trials, intersection-union tests and equivalence confidence sets." *Statistical Science*, 11(4): 283–319.

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.

Delgado, Michael S, and Christopher F Parmeter. 2014. "A simple estimator for partial linear regression with endogenous nonparametric variables." *Economics Letters*, 124(1): 100–103.

Fang, Zheng, and Andres Santos. 2014. "Inference on directionally differentiable functions." *arXiv preprint arXiv:1404.3763*.

Hong, Han, and Jessie Li. 2015. "The Numerical Delta Method and Bootstrap."

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

Munk, Axel, and Rafael Pflüger. 1999. "1α Equivariant Confidence Rules for Convex Alternatives are $\alpha/2$–Level TestsWith Applications to the Multivariate Assessment of Bioequivalence." *Journal of the American Statistical Association*, 94(448): 1311–1319.

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.

Perlman, Michael D, and Lang Wu. 1999. "The emperors new tests." *Statistical Science*, 14(4): 355–369.

Robinson, Peter M. 1988. "Root-N-consistent semiparametric regression." *Econometrica: Journal of the Econometric Society*, 931–954.

Romano, Joseph P. 2005. "Optimal testing of equivalence hypotheses." *Annals of statistics*, 1036–1047.

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. 2003. "Diagnosing the California electricity crisis." *The Electricity Journal*, 16(7): 11–37.