Testing for Market Efficiency with Transactions Costs: An Application to Convergence Bidding in Wholesale Electricity Markets

Akshaya Jha[*]

and

Frank A. Wolak[†]

May 7, 2013

Abstract

With risk neutral traders and zero transactions 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 statistical tests of the null hypothesis that 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 and use them to construct an estimate of the cost of trading in this market. During our sample period, we observe the introduction of explicit virtual bidding, which was aimed at reducing the costs associated with exploiting differences between forward and spot prices. All our measures of trading costs are significantly smaller after the introduction of explicit virtual bidding. We also find that the mean of trading costs is lower for generation nodes relative to non-generation nodes before explicit virtual bidding. However, mean trading costs fell more for non-generation nodes after explicit virtual bidding, eliminating any difference in mean trading costs across the two types of nodes. We also present evidence that the introduction of convergence bidding reduced the total amount of input fossil fuel energy required to generate the thermal energy produced in California and the total variable of cost 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.

[*]Department of Economics, Stanford University, 579 Serra Mall, Stanford, CA 94305. e-mail: akshayaj@stanford.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

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, $E_t[P_{t+k}]$. After accounting for transactions costs, the existence of a profitable trading strategy implies the $|E_t(F_{t+k} - P_{t+k})| > c$, where c is the dollar per unit cost associated with transacting 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 tests of the null hypothesis that profitable trading opportunities exist in a commodity market with transaction costs, and implements this test using data from the California wholesale electricity market.

Wholesale electricity markets with a day-ahead forward market and real-time spot market are ideally suited to test this hypothesis because the same products—electrical energy delivered during each hour the following day—is sold in the day-ahead and real-time markets and the time lag between the purchase or sale in the forward market and subsequent sale or purchase in the spot market is less than one day. Our tests of this hypothesis are complicated by the fact that each day there are 24 hourly trading opportunities between the day-ahead price and real-time price. Therefore, we derive tests of the null hypothesis of the existence of a profitable trading strategy with transactions costs for different portfolios of the 24 hourly price differences.

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. Stakeholders and regulators have been reluctant to allow explicit financial transactions in day-ahead and real-time energy markets in spite of the fact that it is impossible to determine if the reason a market participant sells or buys more or less energy in the day-ahead market than their real-time production or consumption is because of new information about real-time demand or supply conditions after the close of 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 typically involves costly actions by generation unit owners and load-serving entities that can have adverse system reliability consequently. For example, if a generation unit owner expects the real-time market price to be higher than the day-ahead price, the unit owner will 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 (US) wholesale markets, which led to the introduction of convergence or virtual bidding—a purely financial product that is designed to allow market participants

to exploit expected price differences between the day-ahead and real-time markets without these reliability consequences or potential production cost increases.

Convergence 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 than sells energy in the day-ahead market using an incremental or INC convergence bid has an obligation to buy back the same amount of energy as a price-taker in the real-time market. The net payoff from this transaction 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 convergence bid has an obligation to sell that same amount of energy in the real-time market as a price-taker. This transaction has a net profit of the difference between the real-time price and the day-ahead price for that hour times the number of MWhs purchased.

Convergence 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) reduce the total cost of serving demand at all locations in the transmission network in real time. We present evidence that convergence bidding achieved both of these goals. Specifically, our measures of the implied cost associated with trading day-ahead versus real-time price differences fell for the three major pricing zones after the implementation of convergence bidding. We also find that the total hourly input fossil fuel energy consumed fell by 2.8 percent and the total hourly variable cost of producing fossil fuel-fired electricity in California each fell by 2.6 percent after the introduction of convergence 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. This section also describes how the actions of generation unit owners and load serving entities influence locational marginal prices in the absence of convergence bidding as well as how convergence 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 price for the 24 hours of the day before versus after the implementation of convergence bidding. Section 4 derives the three hypotheses tests of our null hypothesis of the existence of a profitable trading strategy with transactions costs. This is followed by a presentation of the pre- and post-convergence bidding implied trading costs for each of our hypothesis tests. Section 5 presents our analysis of the market efficiency consequences of implementing convergence bidding. Section 6 closes with a discussion of the implications of our results for the design of wholesale electricity markets.

2 Locational Marginal Pricing and Convergence Bidding in the California Market

This section first describes the important features of multi-settlement locational marginal pricing wholesale electricity markets that currently exist throughout the United States. In the process we describe how a market participant's actions are used to determine the prices received by generation unit owners and paid by load serving entities in the day-ahead and real-time markets. We then describe how suppliers and load-serving entities exploit expected price differences between the day-ahead and real-time markets before the introduction of explicit convergence bidding. We then explain the mechanics of convergence bidding, including how these purely financial transactions influence day-ahead and real-time locational marginal prices. Finally, the transactions costs associated with exploiting expected differences between day-ahead and real-time prices with and without convergence bidding are discussed.

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 because the configuration of the transmission network, the locations and outputs of other generation units, and the locations and levels of final demand do not allow it. Under these circumstances, 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 on the transmission network and generation units 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 or nodes in the transmission network. To compute

these prices in the day-ahead market, generation unit-owners submit unit-level offer curves giving their willingness to supply energy as a function of the price at the location for each generation unit they own. These willingness-to-supply schedules have two parts-a start-up cost offer and 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 next day and the unit is accepted to produce a positive output during that day. The energy offer curves are 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 US markets allow generation units owners to submit multiple price and quantity pairs for each generation unit each hour of the day. For example, a supplier might be permitted to submit ten price and quantity pairs for each generation unit, with the offer price giving the minimum price at which the supplier is willing to supply the incremental amount of output in the quantity offer associated with that offer price. The sum of the quantity increments is restricted to be less than the capacity of the generation unit and offer prices are typically required to be greater than a price floor (which could be negative) and less than a price ceiling, both which are approved by the Federal Energy Regulatory Commission (FERC), the US wholesale market regulator. In the day-ahead market, load-serving entities submit location-specific willingness-to-purchase functions that are decreasing functions of the price at that location. The functions are composed of price quantity pairs ordered from highest to lowest price where each quantity increment gives the amount the LSE is willing reduce its demand if the price is at or below that price level. All LSEs also submit an inelastic demand that they are willing to purchase at the price floor.

In all US markets simultaneously operate the ancillary services market along with the energy market. Generation unit owners and submit non-decreasing step functions giving their willingness-to-supply each ancillary service. These offer curves are generation unit-specific and unit owners are only allowed to submit offers to supply an ancillary service from their generation unit that the ISO has certified that their unit is able to provide. All ISOs operate markets for spinning reserve, non-spinning reserves and regulation reserve (automatic generation control). In the day-ahead market, the amounts of each operating reserve accepted from each generation unit and the price paid for that operating reserve is determined simultaneously with the generation schedules and LMPs for energy.

To compute the locational marginal prices or LMPs at each node in the transmission network and prices for each ancillary service 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 and ancillary services 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. Although the locational demands for energy are determined by the offer curves submitted by the LSEs, the locational demand for each ancillary service is determined by the ISO. The network constraints used to solve for the day-ahead market outcomes are the ISO's best estimate of real-time configuration of the transmission network during each hour of the following day. The solution to this as-bid cost minimization problem determines firm financial commitments for generation unit owners and load-serving entities for all 24 hours of the following day. The day-ahead

4

generation unit and locational load schedules and ancillary service 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 following day. Similarly, if a load-serving entity purchases 100 MWh in the day-ahead market during 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 in real-time. 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 locational prices gives them their name. For ancillary services, the locational price is also the increase in the minimized value of the objective function associated with increasing the locational demand for that ancillary service by 1 MW. These prices 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 prices and their day-ahead generation unit-level energy and ancillary services schedules and location-specific load schedules in the afternoon of the day-ahead before they are valid.

Starting with midnight the following day, 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 and ancillary services 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. In all of the markets, suppliers are allowed to change their real-time hourly generation unit-level offer curves between the day-ahead and real-time markets.

The real-time market is run every 5 minutes to determine the level of output of 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 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, they must purchase any energy sold in the day-ahead market during that hour that their unit does not produce 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 deviations from hourly day-ahead schedules are settled at the real-time price.

Returning to the above example of the generator that sold 50 MWh of energy in the day-ahead market at a price \$40/MWh, if that generation unit only produced 40 MWh of

energy, the owner would have to purchase the remaining 10 MWh at the real-time price to meet it forward market commitment. If the unit owner produced 55 MWh, then the additional 5 MWh beyond the unit's 50 MWh 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 owners expected to ultimately produce 100 MWh of energy from its unit and forecast a \$60/MWh real-time price that it expected to be higher than the day-ahead price. The unit owner would simply submit price offers into the day-ahead market at or above \$60/MWh, which could cause it to sell no energy in the day-ahead market. The supplier could then offer 100 MWh 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 an offer price of \$60/MWh has been offered into this market and the real-time price is likely to fall because more supply has been offered into this 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 convergence or virtual bidding" because the supplier is using forward market sales from its generation unit as mechanism for exploiting expected price differences between the day-ahead and real-time markets.

Load-serving entities can also engage in implicit convergence or virtual bidding. Suppose that a load serving entity with a 100 MWh real-time demand expects the day-ahead price to be higher than the real-time price, which it expects to be \$100/MWh. This load-serving entity would then submit a demand bid into the day-ahead market with zero quantity demanded at prices above \$100/MWh. The load-serving entity would very likely not make any purchase in the day-ahead market and instead its demand would be entered as a price-taker in the real-time market. These actions by the load-serving entity would reduce the difference between the day-ahead and real-time price, because demand is lower in the day-ahead market and higher in the real-time market as a result of these actions.

Implicit convergence bidding can have severe system reliability consequences and increase the cost of serving system demand. 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 desire of a load-serving entity 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. This can make it necessary for the system operator to have to find large amounts of additional energy between the close of the day-ahead market to ensure that actual demand is met. Wolak (2003) notes that during the

summer of 2000 in the California electricity market this is precisely what happened in part because the offer cap on the day-ahead market was substantially higher than the offer cap on 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 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 this demand would be met.

Besides the reliability consequences of implicit virtual bidding, there are also total variable cost consequences of these actions. All wholesale electricity markets have generation units that take a a number of hours to start up, but once started they are able to produce at a very low variable cost. The implicit virtual bidding by both generation unit owners and load-serving entities can result in long-start, low-operating-cost units to be excluded from producing. Although it may be unilateral profit-maximizing for the owner of a portfolio of long-start, low-cost units and short-start, high-cost units to allow implicit virtual demand bids to cause some of these low-cost units not to operate, these actions increase the total cost of serving system demand.

2.3   Explicit Convergence Bidding versus Implicit Convergence Bidding

The major motivations for introducing explicit virtual bidding are to eliminate the adverse reliability consequences of market participants attempting to exploit expected price differences between the day-ahead and real-time markets and reduce the total cost of serving final demand because market participants have lower cost options besides withholding long-start, low variable cost generation units to exploit day-ahead and real-time price differences. Convergence 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 will not need to distort their bidding and offer behavior in the day-ahead market in ways that increase their costs and potentially harm system reliability.

Convergence or virtual bids are classified as either decremental (DEC) or incremental (INC) 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 convergence bid and a bid from a generation unit owner is that the ISO knows that the energy sold in the day-ahead market from a convergence bid will be purchased out the real-time market as a price-taker. A DEC convergence 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 convergence 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 convergence 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 convergence bid has the 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 convergence bid earns positive revenues if the day-ahead price is higher than the real-time price, but the actions of INC convergence bidders made earning these profits less likely because the 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 convergence bid earns positive revenues if the real-time price is higher than the day-ahead price. The actions of DEC convergence bidders makes 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 convergence bidding will lead to smaller realized nodal price differences between the day-ahead and real-time markets. First, submitting a convergence bid is a lower cost way for a market participant to take a financial position designed to profit from expected price differences between the day-ahead and real-time markets. By submitting an INC convergence 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 convergence bids with offer prices equal the generation unit owner's expected real-time market price. In this way, the generation unit owner does not distorts its offer prices for its generation units in order to exploit expected price differences between the day-ahead and real-time markets.

A second reason that node-level day-ahead versus real-time price differences are likely to be smaller is because explicit virtual bidding 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 this generation unit from the day-ahead market cannot raise 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 amount 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 engaging in implicit virtual bidding. The actual demand of a load-serving entity limits the amount of demand it can bid into the day-ahead market. For example, without explicit virtual bidding, if bidding no 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 node-level convergence bidding was introduced in California, the opportunities to

implicit virtual bid at the node level was limited to locations with generation units. Implicit virtual bidding at nodes with no generation units was not in general possible. The California market requires the three large load-serving entities in California—Southern California Edison, Pacific Gas and Electric, and San Diego Gas and Electric—to bid their service area-level demand into the day-ahead market and the California ISO allocates this demand to all nodes in the load-serving entity's service territory using load-distribution factors (LDFs) that it produces. For example, if a load-serving entity has 100 MW 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 the load-serving entity to implicitly virtual bid 1 MWh at one node, because this would effectively require 1 MWh of implicit virtual bids at all nodes. With the introduction of explicit node-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.

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 was 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 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 Markets

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—Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E) before and after the implementation of explicit convergence bidding. We also present the results of a test of the null hypothesis that the mean price difference vector for the 24 day-ahead and real-time hourly prices is equal to zero for these three load-serving entities and find that we overwhelmingly reject this null hypothesis in all cases. However, these tests do not account for the transactions costs associated with attempting to exploit these mean price differences which out testing procedures account for.

These hypothesis tests are implemented using daily price data from April 1, 2009 when nodal pricing was implemented in California, to December 31, 2012 for the 24 hourly real time and day-ahead wholesale electricity prices. These prices are set at the node level and

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) price each hour of the day which is computed as a nodal quantity-weighted average price for that load-serving entity summed over all nodes in the load-serving entity's service area with 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. For each of these LAPs, we compute the hour-of-day average price difference for all hours of the day.

Figure 1 presents a comparison by hour of day of the average difference between the day-ahead and real-time prices for the PG&E, SCE, and SDG&E LAPs both before and after the introduction of explicit virtual bidding. This figure provides descriptive evidence that the day-ahead/real-time spread is more pronounced prior to the introduction of virtual bidding than afterwards for each of the load-serving entities. For example, for PG&E, the average day-ahead price is much lower than the average real-time price during the hours of 8PM–12AM. These results immediately raise the question of whether these mean price differences reflect the existence of profitable trading strategies or are simply due to the existence of non-zero trading costs that allow non-zero mean price differences.

[Figure 1 about here.]

To further motivate our subsequent analysis, we present a zero transaction cost version of an arbitrage test for the PG&E, SCE, and SDG&E LAPs after the introduction of explicit virtual bidding in Figure 2. Namely, we plot the average day-ahead/real-time spread along with pointwise 95% confidence intervals around these means. 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 the daily mean of the vector of day-ahead and real-time price differences is 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 $\hat{\Lambda}_j = \sum_{t=j+1}^{T}(X_t - \overline{X})(X_{t-j} - \overline{X})'/T$, $\overline{X} = \sum_{t=1}^{T} X_t/T$, $w(j,m) = 1 - [j/(m+1)]$ for $m = 14$ to construct the chi-squared test statistics. These test statistics are presented for each LAP before and after the introduction of explicit virtual bidding in Table 1. Note that these test statistics are quite large. We would reject the null hypothesis that the mean of the price difference vector is zero in all cases.[1] However, these two tests fail to account for the potentially sizable transaction costs present in nearly every commodities market. In the next section, we present hypothesis testing procedures that account for the fact that the day-ahead/real-time price spread can differ from zero simply due to positive transaction costs.

[Figure 2 about here.]

[Table 1 about here.]

---

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

4   Testing the Null Hypothesis of the Existence of a Profitable Trading Strategy

4.1   Introduction

In this section, we first develop three tests of the null hypothesis that a profitable trading strategy exists. For simplicity, we restrict attention to the set of trading strategies that only condition on the value of μ. Our null hypothesis is that profitable trading strategies exist. Rejection of this null hypothesis implies that the data provides evidence against the existence of a profitable trading strategy based on the unconditional mean of the (24 x 1) vector of daily price differences. We then present empirical evidence that strategies that condition on past price difference vector realizations are unlikely to be of practical importance, because past daily price differences beyond the first lag are poor predictors of future daily price differences. The market rules prohibit strategies that condition of the first lag of the price difference vector because the all real-time prices for the current day are not known when market participants submit their offers into the day-ahead market for the following day.[2]

It is often the case when analyzing the performance of a new drug relative to an existing drug that the researcher would like to conclude the two drugs are bioequivalent in terms of their efficacy. In this general case, the researcher formulates the null hypothesis as some nonlinear function $g(\theta)$ of the parameter vector of interest lies outside of the set (a,b), versus the alternative that it lies in the set (a,b). If the interval (a,b) contains zero, then rejection of the null hypothesis implies the two drug are bioequivalent, because the difference in their efficacy does not lie outside the interval (a,b). For this reason, this class of hypotheses are called equivalence hypotheses. See Romano (2005) for a discussion of optimal equivalence tests.[3] Note than the typical approach to testing market efficiency as the null hypothesis is to find no evidence against market efficiency by failing to reject the null hypothesis. By formulating the test as an equivalence hypothesis, failing to reject the null hypothesis says that we have no evidence against the null hypothesis that a profitable trading strategy exists. On the other hand, rejection of the hypothesis implies that the data is inconsistent with the existence of a profitable trading strategy based on the unconditional mean of the price differences.

We motivate three statistical tests by considering the problem of a trader. The different statistical tests are derived from different trading strategies involving the 24 assets (one for each hour of the day). For example, a trader can buy (or sell) one unit of the asset with the maximum, across all hours of the day, absolute value of the expected difference between the day-ahead and real-time prices. The trader can also buy one unit of all assets with positive expected price differences and sell one unit of all assets with negative expected

---

[2]Offers to the day-ahead market must be submitted by noon the day before actual system operation.

[3]Testing equivalence hypotheses has a rich tradition in the statistics literature. See Berger and Hsu (1996), Perlman and Wu (1999), and Munk and Pflüger (1999).

price difference. Each trading strategy results in a statistical test with higher power against different alternatives.

Because the explicit costs of buying and selling these assets is only one component of the cost of exploiting these price differences, we use each of these statistical tests to recover an implied trading cost, which is the smallest value of the trading cost that causes us to reject the null hypothesis that a profitable trading strategy exists. We do this both at the LAP and nodal level, for before and after the introduction of explicit virtual bidding. Using the bootstrap, we compute an estimate of the distribution these trading cost estimates. Comparing these estimated trading cost distributions before versus after the introduction of convergence 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 convergence bidding. We also perform a test of the null hypothesis that the profits traders expected to earn from buying and selling differences between the 24 day-ahead and real-time prices fell after the implementation of convergence bidding using the multivariate inequality constraints testing procedure of Wolak (1989).

4.2   Motivation: The Trader's Problem

Consider a trader with access to 24 assets, where asset X_h for $h \in \{1, ..., 24\}$ is equal the difference between the day-ahead and real-time price. This implies $X_h = P_h^{DA} - P_h^{RT}$, where P_h^{DA} is the day-ahead price during hour h and P_h^{RT} is the real-time price during 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 $\mu_h = E(X_h) = E(P_h^{DA}) - E(P_h^{RT})$. Define μ as the 24 x 1 vector composed of $(\mu_1, \mu_2, ..., \mu_{24})'$ for $h = 1, 2., ..., 24$ and X equal the 24 x 1 vector composed of $(X_1, X_2, ..., X_{24})'$. Let Λ_0 equal the 24 x 24 contemporaneous covariance matrix of X. Suppose the per-unit trading cost of buying or selling this security is c. The expected profit-maximization problem of a trader holding the net portfolio with weights vector, $a = (a_1, a_2, ..., a_{24})'$, where each a_i can be positive or negative, and paying a per unit trading cost c is:

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

subject to different constraints on the elements of a. Then, our null hypothesis is that this optimization problem results in a positive expected profit; for some $a^* \in R^{24}$, so that our equivalence null hypothesis is that $a^{*'}\mu - c \sum_{i=1}^{24} |a_i^*| > 0$ for this value of a^*. Note the trader pays the same trading charge or sales and purchases of this asset, which is why the trading charge is assessed on the sum of the absolute values of the individual portfolio weights, a_i. Define the function $sign(x)$ which equals 1 if $x > 0$ is equal to -1 if $x < 0$ and zero if $x = 0$. Our three hypothesis tests correspond to different choices of a^*:

- "Intersection-Union (IU)": $a = MAXE(\mu)$,

- "Square": $a = SIGN(\mu)$.

- "Ellipsoid": $a = \Sigma^{-1}\mu$,

where $MAXE(\mu) = (I_1 * |\mu_1|, I_2 * |\mu_2|, ..., I_{24} * |\mu_{24}|)'$, $I_j = 1$ if the jth element of μ is largest in absolute value element of the vector and $I_j = 0$ if that is not the case, and $SIGN(\mu) = (sign(\mu_1), sign(\mu_2), ..., sign(\mu_{24}))'$. Figure 3 presents this same intuition in graphical form. We see that the IU and Square tests correspond to a square rejection region, while the ellipsoid test corresponds to an ellipsoid rejection region. Note that we reject the null hypothesis of the existence of a profitable trading strategy if the sample mean of the price differences lies inside the shaded area, and fail to reject if the sample mean of the price differences lies outside the shaded area.

[Figure 3 about here.]

4.3 Three Tests for the Existence of a Profitable Trading Strategy

To implement our hypothesis tests, we compute the sample mean and an autocorrelation consistent estimate of the asymptotic variance of this sample mean. Let N denote the number of days in our sample and $\overline{X}$ denote the sample mean of X, and the estimate of the variance of the asymptotic distribution of $\sqrt{N}(\overline{X} - \mu)$, denoted $\hat{\Sigma}$, is calculated using Newey and West (1987) autocorrelation consistent covariance matrix with $m = 14$. Assume that the trading cost c is known. In the next subsection, we discuss how to "invert" our hypothesis tests to recover the trading costs implied by just rejecting our null hypothesis.

We first consider the Intersection-Union (IU) Test, which intersects the rejection region of 24 individual equivalence tests that $|\mu_h| > c$ for $h = 1, 2, ..., 24$. This test can be stated as follows:

Proposition 1 *Intersection-Union Test*
Consider the hypothesis test $H_h : |\mu_h| > c$. For each h, perform the size $\alpha_h = 0.05$ test of H_h as follows: reject if and only if $|\overline{X_h}| + N^{-\frac{1}{2}}\hat{\Sigma}_{hh}z_{1-\alpha_h} < c$, where $\hat{\Sigma}_{hh}$ is the hth diagonal element of $\hat{\Sigma}$.[4] Then, we reject the Null hypothesis that a profitable trading strategy exists if and only if we reject H_h for all $h \in \{1, ..., 24\}$. This is an overall level $\alpha = 0.05$ test of this multivariate hypothesis.

Note that each individual test H_h is performed at size $\alpha_h = 0.05$. We do *not* need to appeal Bonferroni's inequality and test each H_h at size $\alpha_h = \frac{0.05}{24}$. However, the IU Test is known to be very conservative. Intuitively, we can think of the trader transacting $\frac{1}{24}$ of each asset (buying if the expected price difference is positive, and selling otherwise). This trading strategy has maximum probability of rejection, which we know to be $\alpha = 0.05$, at one of the vertices of the square presented in Figure 1, and must therefore be conservative at other points in the rejection region. This same intuition underlies the square-based test presented below:

---

[4]$z_{1-\alpha_h}$ satisfies $\Phi(z_{1-\alpha_h}) = 1 - \alpha_h$, where $\Phi(t)$ is the cumulative distribution function of a standard normal random variable.

Proposition 2 *Square-Based Test*

Define $\hat{V} \equiv SIGN(\overline{X})'\hat{\Sigma} SIGN(\overline{X})$. Then, we have test statistic $TS = \sqrt{N}(\overline{X}'SIGN(\overline{X}) - c)\hat{V}^{-\frac{1}{2}} \to^d N(0,1)$ for μ at the boundary of the set defined by our null hypothesis. Therefore, we reject a size $\alpha = 0.05$ test of our null hypothesis if and only if $\Phi(\frac{\sqrt{N}(\overline{X}'SIGN(\overline{X})-24c)}{\sqrt{V}}) - \Phi(\frac{\sqrt{N}(-\overline{X}'SIGN(\overline{X})-24c)}{\sqrt{V}}) \leq 0.05$.[5]

As before, this test has size $\alpha = 0.05$ at one of the vertices of the square rejection region. The $24c$ term comes from trading 1 MWh of each of the 24 assets based on the sign of the expected price difference. However, as foreshadowed in Figure 1, we can also consider an ellipsoid rejection region which has greater power against a different set of alternatives than the Square test. See Munk and Pflüger (1999) for a discussion on the advantages of testing equivalence using ellipsoidal rather than rectangular rejection regions. These authors note that the ellipsoid test is also likely to be more powerful than the Intersection-Union Test, an intuition that is borne out in our empirical results.

Proposition 3 *Ellipsoidal Test*

If we define test statistic $TS = N\overline{X}'\hat{\Sigma}^{-1}\overline{X} \to^d \chi^2_{24}(Nc[\sum_{i=1}^{24}|(\Sigma^{-1}\mu)_i|])$, where $(\Sigma^{-1}\mu)_i$ is the ith element of the vector $\Sigma^{-1}\mu$, for μ at the boundary of the set defining our null hypothesis, then we reject the Null hypothesis if and only if $Pr[\chi^2_{24}(Nc\sum_{i=1}^{24}|(\hat{\Sigma}^{-1}\overline{X})_i|) \leq TS] = 0.05$, where $\chi^2_k(\lambda)$ is a non-central chi-squared random variable with k degrees of freedom and non-centrality parameter λ.

4.4 Deriving Trading Costs Implied by Rejection of the Null

Although we can compute the cost of purchasing or selling elements of X in the California ISO market, this is just one component of the trading cost. Setting the trading cost, c, equal to this magnitude implies that there is no opportunity cost of 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 use our hypothesis testing results to compute implied trading costs. We can then compare these implied trading costs to the actual cost of purchasing and selling the 24 elements of X in the ISO market, including conservative estimates of other transactions costs. We take each of the three tests described above, and find the value of c that would just reject the Null hypothesis at a 5% significance level. We denote this value by c^I. Then, if the true level of trading costs is above c^I, we would reject the null hypothesis that profitable trading strategies exist. Otherwise, we would fail to reject the existence of a profitable trading strategy based on μ.

For example, recall that for the Intersection Union test, we reject our null hypothesis if and only if all individual H_h are rejected. That is, we reject if and only if for all $h \in \{1, ..., 24\}$, $|\overline{X_h}| + N^{-\frac{1}{2}}\hat{\Sigma}_{hh}z_{1-\alpha_h} < c$. In this case, $c^I = \underset{h \in \{1,...,24\}}{max} |\overline{X_h}| + N^{-\frac{1}{2}}\hat{\Sigma}_{hh}z_{0.95}$. In words, we choose the largest of all implied trading cost over all of the individual equivalence hypotheses tests.

---

[5]As before, $\Phi(t)$ is the cumulative distribution function of a standard Normal random variable.

4.5 A Direct Test for Difference in Means Before and After Virtual Bidding

We also directly test whether expected trading profits fell after the introduction of convergence 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 $|\mu_{pre}| - \mathbf{1}c^{pre} > |\mu_{post}| - \mathbf{1}c^{post}$, where $|\mu|$ is the vector composed of the absolute value of the individual elements of the vector μ and $\mathbf{1}$ is 24 x 1 vector of 1's. The difference $|\mu_{pre}| - \mathbf{1}c^{pre}$ is the expected profits associated with buying one unit of μ_h if it is positive and selling one unit μ_h if it is negative for $h = 1, 2, ..., 24$. Consequently, re-arranging this inequality we see that it implies $|\mu_{pre}| - |\mu_{post}| > \mathbf{1}(c^{pre} - c^{post})$. If we assume that $c^{pre} > c^{post}$, which is consistent with the results presented in Section 5, then the null hypothesis that expected trading profits fell after the introduction of convergence bidding is that $|\mu_{pre}| - |\mu_{post}| > 0$. Therefore, testing $|\mu_{pre}| - |\mu_{post}| > 0$ is a test of this null hypothesis. Conversely, by rejecting the null hypothesis $|\mu_{post}| > |\mu_{pre}|$, we can conclude that the null hypothesis that trading profits were higher after the introduction of convergence 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 convergence bidding.

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

Proposition 4 *Direct Test of Null Hypothesis that* $|\mu_{pre}| > |\mu_{post}|$
Let $\hat{V} = diag[SIGN(\overline{X}^{pre})]'\frac{\hat{\Sigma}^{pre}}{N^{pre}}diag[SIGN(\overline{X}^{pre})]+diag[SIGN(\overline{X}^{post})]'\frac{\hat{\Sigma}^{post}}{N^{post}}diag[SIGN(\overline{X}^{post})]$
and calculate the test statistic:

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

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* χ^2_h *is a chi-squared random variable with h degrees of freedom and* $w(24, 24-h, \hat{V})$ *are the weights defined in Wolak (1989) and* α *ia 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. $\hat{\Sigma}^{pre}$ is the estimated autocorrelation consistent asymptotic covariance matrix with $m = 14$ of the vector of means prior to the introduction of explicit virtual bidding and N^{pre} is the number of days in the sample prior to explicit virtual bidding. $\hat{\Sigma}^{post}$ is the estimated autocorrelation consistent covariance matrix with $m = 14$ of the vector of means after to the introduction of explicit virtual bidding and N^{post} is the number of days in the sample after explicit virtual bidding. These are the same estimates used in the prior subsection describing the trading costs-based approach. Note that we calculate $w(24, 24 - h, \hat{V})$ using the simulation method outlined in Wolak (1989).

5   Results from Our Hypothesis Tests

This section presents the results of implicit trading cost calculation and our tests that expected trading profits fells after the introduction explicit virtual bidding. Before we present these results, we provide some evidence that more complex trading strategies based on lagged values of price differences may not yield significant profit improvements relative to a strategy that just conditions on the elements of μ.

5.1   Why not condition on past values of X_d

Because all of the values of the (24 x 1) vector real-time prices for day $d-1$ are not known before offers are submitted to the day-ahead market for day d, there can be first-order autocorrelation between realizations of X_d that cannot be exploited through a feasible trading strategy. Specifically, any trading strategies involving portfolios of the (24 x 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 those are the only realizations of X_{d-k} that are known when a market participant submits bids or offers into the day-ahead market for day d. This logic implies that X_d following a vector MA(1) process is consistent with the lack of a profitable trading strategy that conditions on past values of X_d. Investigate this hypothesis, we would like to estimate a vector MA(1) process for X_d and then test 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.

One measure of the amount of exploitable autocorrelation in the X_d sequence can be derived from computing the sample autocorrelation function for the portfolio of the (24 x 1) vector of price differences for each of the three hypothesis tests for the existence of a profitable trading strategy. Figure 4 presents the sample autocorrelation function and the pointwise 95 percent confidence intervals for these samnple autocorrelations for the three LAP price differences for PG&E, SCE, and SDG&E for the before explicit virtual bidding sample period. Figure 5 computes these same magnitudes for the after explicit virtual bidding sample period. For all of the LAPs, there are very few sample autocorrelations beyond the first-order that appear to be statistically different from zero for either the before or after explicit virtual bidding sample periods. These results are consistent with the view that trading strategies for day d that condition or values of X_{d-k} for k ¿ 1 are unlikely to yield higher expected profits than those that do not condition on lagged values of X_d. For this reason, we only consider trading strategies that depend on, μ, the unconditional mean of X_d.

[Figure 4 about here.]

[Figure 5 about here.]

5.2 Results from Our Trading Costs Hypothesis Tests

We first implement our three hypothesis tests at the LAP level. These results are presented in Table 2. These trading costs are the value at which we would just reject the null hypothesis of the existence of a profitable trading strategy. We would reject the null hypothesis only if the actual trading costs are higher than the ones listed in the table. First note from Table 2 that the implied trading costs from the IU test are much higher than the other two tests, indicating that the IU test is more conservative than the other two, as expected. More importantly, we see that the implied trading costs decrease after the introduction of explicit virtual bidding for all LAPs and all tests. This is consistent with logic outlined in Section 2 that the costs of trading day-ahead versus real-time price differences decreases after the introduction of explicit virtual bidding.

[Table 2 about here.]

To obtain a more formal comparison implied trading costs before versus after explicit virtual bidding, we examine the bootstrap distribution of implied trading costs for each LAP before and after the implementation of explicit virtual bidding. Figure 8 provides the bootstrap distribution of trading costs implied by rejection of the null hypothesis of the existence of a profitable trading strategy. The bootstrap distributions are computed by re-sampling contiguous blocks of length $B = (N^J)^{1/3}$ for $J = pre, post$ of the daily price difference vectors during the pre-explicit virtual bidding period to obtain a sample of size of N^J for $J = pre, post$ and from that bootstrap sample compute the implied trading cost for each test procedure. Re-sampling blocks of contiguous observations ensures that low-order temporal dependence in the original data is preserved in the re-sampled distribution. Repeating this process $L = 10,000$ times for each LAP price and convergence bidding regime yields the histogram of re-sampled implied trading costs for before and after the implementation of explicit convergence bidding. The most obvious result from each of these graphs is that the distribution of implied trading costs is markedly shifted to the left after the introduction of explicit virtual bidding. Recall that for an actual trading cost c, we reject our null hypothesis if c is larger than our implied trading cost measure, which implies that no profitable trading strategy using μ exists. These histograms therefore imply that the introduction of explicit virtual bidding expanded the set of actual trading costs for which we can say that no profitable trading strategy exists. Another claim most readily seen from the Intersection-Union results is that the distribution of implied trading costs has a smaller standard deviation after the introduction of explicit virtual bidding. This claim is consistent with the view that it is more costly for market participants to implicitly virtual bid than the to explicitly virtual bid. These previously high trading cost strategies become much lower cost once the financial product was introduced to trade of elements of X.

We can also compute the implied trading costs for each testing procedure for each node in the California ISO control area. Figure 9 plots the implied trading costs for each node before and after the introduction of explicit virtual bidding.[6] We plot these node-level

---

[6]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

implied trading cost distributions separately for nodes associated with generation units and nodes not associated with generation units. Note that the distribution of implied trading costs across nodes markedly decreases when we calculate it using data after the introduction of explicit virtual bidding (for all three tests and both generation and non-generation nodes). Recall that we reject the null of profitable trading strategies for actual trading costs larger than the plotted implied trading costs. Therefore, we would reject our null hypothesis of the existence of profitable trading strategy for a larger set of potential trading costs at more nodes after the introduction of explicit virtual bidding.

We expect the following two relationships to hold between the means of implied trading costs across generation versus non-generation nodes before versus after explicit virtually bidding. First, because suppliers can implicitly virtual bid at the nodal level before the implementation of the explicit virtual bidding through how they operate their generation units and load-serving entities can only bid in at the LAP level before explicit virtual bidding, we expect the mean of implied trading costs to be higher at non-generation nodes before the implementation of explicit virtual bidding. Second, because the introduction of explicit virtual bidding allows, for the first time, virtual bidding at non-generation nodes, we expect the the mean reduction in implied trading costs for non-generation nodes to be larger than for generation nodes. To test these two hypotheses, we regressed the implied trading cost 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-explicit virtual bidding period, and an indicator variable for whether the observation was from a generation node during the post-explicit virtual bidding period. Table 3 reports the results of estimating these regressions for the implied trading costs from the: (1) Intersection-Union (IU) Test, (2) Square Test, and (3) Ellipse Test. For all three tests, we find that mean trading costs fell significantly after the introduction of explicit virtual bidding. For the IU and Square tests we find strong evidence consistent with both our hypotheses. The mean of implied trading costs before the introduction of explicit virtual bidding is significantly lower for generation nodes and this difference is essentially eliminated after the introduction of virtual bidding. Specifically, for all three measures of implied trading costs regressions, we find that the null hypothesis that the sum of the coefficient on "Generation Node Indicator" and the coefficient on "Interaction Between Geenration and Post EVB Indicator" is zero cannot be rejected. Combining this result with the very large mean reduction in implied trading costs for all nodes after the introduction of explicit virtual bidding, implies that the mean difference in implied trading cost before versus after explicit virtual bidding fell more for non-generation nodes than for generation nodes. For our ellipse test implied trading cost estimates, there is evidence of any difference between generation and non-generation nodes before or after explicit virtual bidding.

Figure 6 contains monthly average hourly virtual supply offered and cleared and virtual demand offered and cleared for October 2011 to December 2012 taken from the California ISO Department of Market Monitoring's Q4 Report on Market Issues and Performance of February 13, 2013. This graph shows that slightly less than 1,000 MWh of virtual supply

25% and 75% quartiles. Finally, the remaining points are outliers outside of the aforementioned range.

clears each hour and approximately the same level of virtual demand clears each hour, with roughly half of the virtual supply and virtual demand offers clearing each hour. Because there are over 5,000 nodes in the ISO system and minimum convergence bid offer is 1 MWh, there are many nodes each hour that do not receive node-level convergence bids. Figure 7 shows the average offer and cleared virtual demand and supply virtual bids by hour of the day for October to December of 2012. Particularly, for demand bids, there are significant higher levels of offered and cleared bids during the peak demand hours of the day, whereas for virtual supply bids the pattern of offers and cleared bids is fairly constant throughout the day.

[Figure 6 about here.]

[Figure 7 about here.]

[Figure 8 about here.]

[Figure 9 about here.]

[Table 3 about here.]

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 result of the introduction of convergence bidding. The p-values corresponding to these tests for each LAP are presented below in Table 4:

[Table 4 about here.]

We cannot reject the null hypothesisl 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. Assuming that $c^{pre} > c^{post}$, these hypothesis testing results provide strong evidence in favor of the view that trading profits fell after the introduction of explicit virtual bidding.

6 Measuring Market Efficiency Implications of Convergence Bidding

This section describes the data used and analysis performed to assess the market efficiency consequences of the introduction of explicit convergence bidding. The three market outcome measures we compare before versus after the introduction of convergence bidding are: (1) the total millions of British Thermal Units (MMBTUs) of natural gas 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

19

of generation units started during that hour. We use a sample period that starts one year before convergence bidding was implemented on February 1, 2011 until one year after it was implemented on January 31, 2012. We nonparametrically control for differences across hours of the day and days of our sample period for differences in the level output of thermal generation units in California, the level of output of intermittent renewable resources (wind and solar resources) and daily wholesale prices of natural gas delivered to both northern and to southern California. To control for the hourly differences in these observable factors as flexibly as possible in computing the difference in the mean values of each performance measure before and after the implementation of convergence bidding, we employ the Robinson (1988) partially linear model to estimate the conditional mean function for each market performance measure.

Constructing the total hourly MMBTUs of energy consumed by all natural gas-fired generation units proceeds as follows. First, the hourly metered output of each natural gas-fired generation unit is obtained from the California ISO's settlement system. This information is combined with the generation unit-level heat rate curve that all natural gas-fired generation unit owners are required to submit as part of the California ISO's local market power mitigation mechanism. This curve is a piecewise linear function that can have ten heat rate level and output quantity pairs up to the full capacity of the generation unit. The vertical axis gives the heat rate denominated in millions of British Thermal Units (MMBTUs) of natural gas burned to produce each additional MWh for the level of output from that generation unit on the horizontal axis. 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 consumed by that generation unit during the hour.

Also included in the total amount of MMBTUs consumed in an hour is the total necessary to start up any generation units that began operating during that hour. Natural gas-fired generation unit owners are also required to file information on the total amount MMBTUs required to start each generation unit with the California ISO as part of its local market power mitigation mechanism. 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. Summing the MMBTUs of energy consumed to produce each unit's metered output in that hour and the MMBTU of energy consumed in that hour to start all units that started during that hour yields, $TOTAL\_ENERGY(t)$, the total amount of energy consumed in hour t by the 228 natural gas-fired generation units in the California ISO control area during our sample period.

The total number of generation units started in an hour t, $STARTS(t)$, is the total number of units in hour t that have zero metered output in hour t-1 and positive output in hour t. The final market performance measure, $TOTAL\_VC(t)$, is the total variable cost of all natural gas-fired 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 price of natural gas for that unit plus the variable operating and maintenance cost that the unit's owner submits to the California ISO for its local market power mitigation mechanism. The total volume variable cost for the unit is computed as the product of the above marginal cost for the unit times its metered output for the hour. For units that start up in hour t, the total energy

to start the unit is converted to a cost by multiplying the MMBTUs of energy consumed to start the unit by the daily price of natural gas. Summing these volume variable costs over all generation units operating in hour t and the start-up costs for all units starting in hour t, yields the value of $TOTAL\_VC(t)$.

We specify nonparametric 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 convergence bidding. All of the hour-of-sample conditional mean functions can be written as $y_t = W_t'\alpha + X_t'\beta + \theta(Z_t) + \epsilon_t$, with $E(\epsilon_t|X_t, W_t, Z_t) = 0$, where y_t is one of our three market performance measures. The function $\theta(Z)$ is an unknown function of the vector Z, W is a (24x1) vector of hour-of-day dummy variables, and α and β are unknown parameter vectors. For all three overall conditional mean functions, X_t is a single dummy variable that takes on the value 1 for all hours after midnight January 31, 2011 and zero otherwise, and Z_t is four dimensional vector composed of the total output in MWhs of all natural gas-fired generation units in California during hour t, the total output in MWhs of all wind and solar generation units in California during hour t, the price of natural gas in northern California (the Pacific Gas and Electric Citygate delivery point) during hour t, and the price of natural gas in Southern California (the Southern California Gas Citygate delivery point) during hour t. For the total starts conditional mean function, y_t equals $STARTS(t)$, for the total energy conditional mean function, y_t equals the natural logarithm of $TOTAL\_ENERGY(t)$, and for the total variable cost conditional mean function, y_t equals the natural logarithm of $TOTAL\_VC(t)$. We also estimate models that allow separate mean differences in each market performance measure by hour of the day. In this case X_t is a (24x1) vector with k^{th} element X_{tk}, which equals one during hour-of-the-day k for all days from February 1, 2011 until the end of the sample period.

Controlling for both the hourly output of thermal generation units and the hourly output of wind and solar generation unit is necessary because the share of energy produced by renewable resources has grown in significantly over our sample period as a result of California's renewables portfolio standard (RPS), which requires all California load-serving entities to procure 33 percent of their energy from qualified renewable sources by 2020. Figure 10 plots the average hourly output of in-state thermal generation resources and in-state renewable generation resources during the year before virtual bidding and year after virtual bidding. Each point on each curve Figure 10(a) is the average over all days during the year before or year after virtual bidding was implemented of the output of all thermal generation units during that hour of the day. Each point on each curve of Figure 10(b) is computed in the same manner using solar and wind generation units. Figure 10 demonstrates that average hourly output of thermal generation units falls substantially, and much of that fall is taken up by the increase in wind and solar energy produced in California. Figure 11 plots the standard deviations of the hourly output for each hour of the day across days in the sample before and after the implementation of convergence bidding. The standard deviation of both thermal and wind and solar output for all hours of the day are higher after virtual bidding. This is particularly the case for wind and solar output. The intermittency of the these resources implies that more thermal resources must be held as operating reserves and stand ready to supply additional energy if the wind or solar resources disappear suddenly. Consequently,

failing to control for both the hourly output of wind and thermal generation units before versus after the implementation of explicit virtual bidding would not account for the significant increase in average wind and solar energy and increased volatility in thermal output and renewable energy output after the implementation of explicit virtual bidding.

[Figure 10 about here.]

[Figure 11 about here.]

We employ a two-step estimation procedure that recognizes that $\theta(Z_t) = E(y_t - W_t'\alpha + X_t'\beta|Z_t)$ and estimates it using $\hat{\theta}(Z_t, h) = \frac{\sum_{t=1}^{T}(y_t - W_t'\alpha - X_t'\beta)K((z-Z_t)/h)}{\sum_{t=1}^{T}K((z-Z_t)/h)}$ to estimate both α and β. The first step finds the values of h, α, and β that minimize $\sum_{j=1}^{T}[y_j - W_j'\alpha - X_j'\beta - \hat{\theta}_{-j}(Z_j, h)]^2$, where $\hat{\theta}_{-j}(Z_j, h)$ has the same form as $\hat{\theta}(z, h)$ evaluated at $z = Z_j$ except that $\sum_{t=1}^{T}$ in the numerator and denominator is replaced with $\sum_{t=1, t \neq j}^{T}$. The second step is a least squares regression of $[y_t - \hat{\theta}(Z_t, h^*)]$ on W_t and X_t, where h^* is the optimized value of h from the first step. Robinson (1988) demonstrates that $\sqrt{T}([\begin{array}{cc} \hat{\alpha} & \hat{\beta} \end{array}]' - [\begin{array}{cc} \alpha & \beta \end{array}]')$, where $\hat{\alpha}$ and $\hat{\beta}$ are the second-stage estimates of α and β, has an asymptotic normal distribution. Standard error estimates are constructed using the expression for the estimated asymptotic covariance matrix given in Robinson (1988).

6.1 Empirical Results

Table 5 reports the results of estimating the conditional mean function, $y_t = W_t'\alpha + X_t'\beta + \theta(Z_t) + \epsilon_t$, for each measure of market performance for the case that 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 hourly energy (controlling for the total hourly output from all natural gas-fired units, the total hourly output of wind and solar resources, the prices of natural gas in northern and southern California and the hour of the day) is 2.8 percent lower after January 31, 2011. The conditional mean of total hourly starts (controlling for the same variables) is 0.63 starts higher after January 31, 2011. The conditional mean of total variable costs is 2.6 percent lower after January 31, 2011.

[Table 5 about here.]

Figures 12 plots the estimates of hour-of-the-day change in the conditional mean of the three hourly market performance measures after the implementation of convergence bidding along with the pointwise the 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 are also increase after the implementation of convergence 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 convergence bidding and $2.2 billion the year after convergence 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 convergence 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 milliion 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.

[Figure 12 about here.]

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 location in the transmission network. In addition, these results demonstrate economically significant economic and global environmental benefits associated with the introduction of convergence 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 consumer make in forward markets can provide important information and feedback to market participants that improves the subsequent performance of short-term physical markets. Although convergence 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 prof-

itable 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.

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. "Testing inequality constraints in linear econometric models." *Journal of econometrics*, 41(2): 205–235.

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

List of Figures

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

(a)

CAISO Wholesale Electricity Day-Ahead - Real-Time Price Spread
Time-weighted Means for PGE, By Hour of Day

(b)

CAISO Wholesale Electricity Day-Ahead - Real-Time Price Spread
Time-weighted Means for SCE, By Hour of Day

(c)

CAISO Wholesale Electricity Day-Ahead - Real-Time Price Spread
Time-weighted Means for SDGE, By Hour of Day

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

(a)

CAISO Wholesale Electricity Price Spread: Before Virtual Bidding
Time-weighted Means for PGE, By Hour of Day

(b)

CAISO Wholesale Electricity Price Spread: Before Virtual Bidding
Time-weighted Means for SCE, By Hour of Day

(c)

CAISO Wholesale Electricity Price Spread: Before Virtual Bidding
Time-weighted Means for SDGE, By Hour of Day
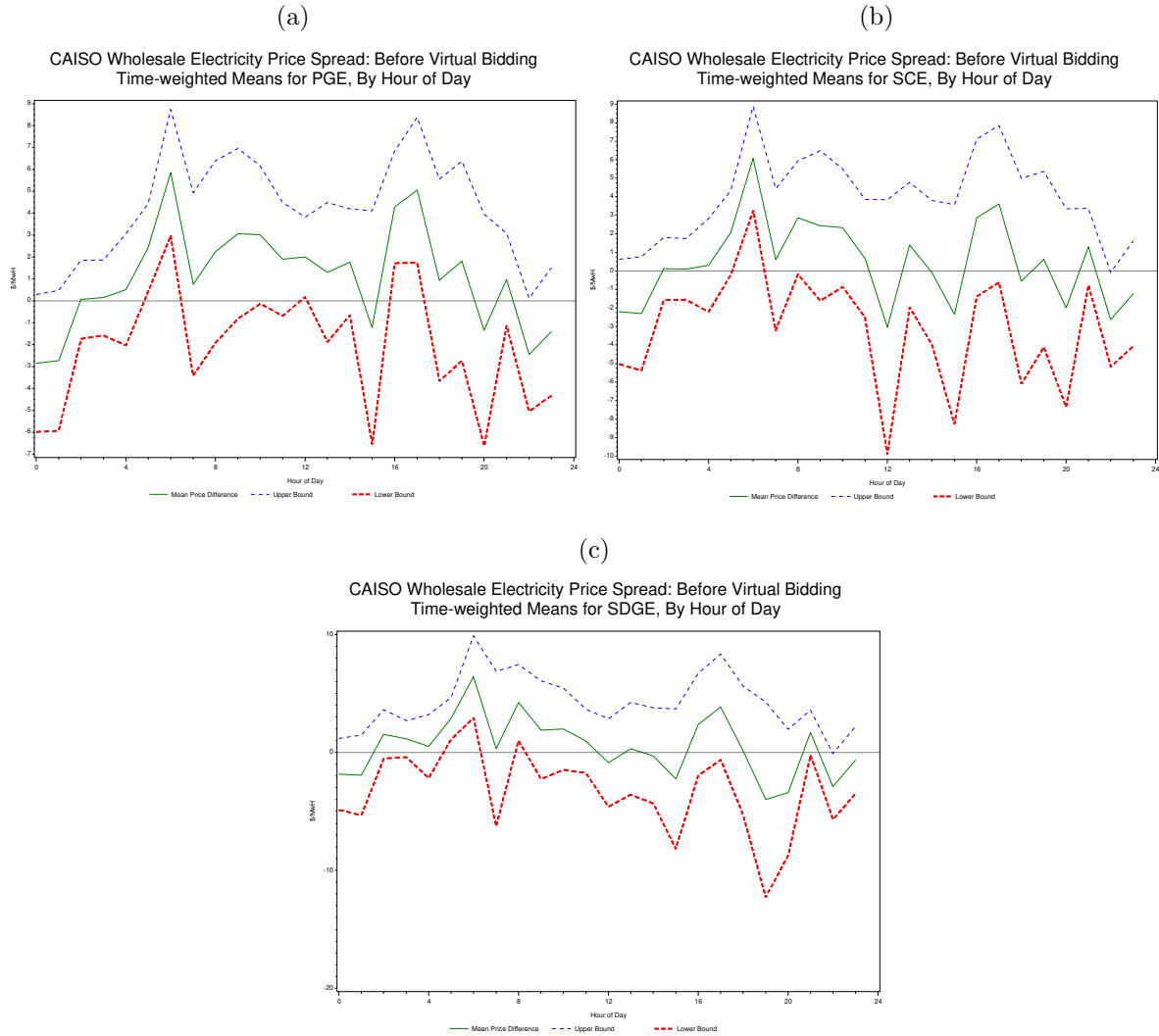
Figure 3: Graphical Intuition of the Rejection Regions For Hypothesis Tests
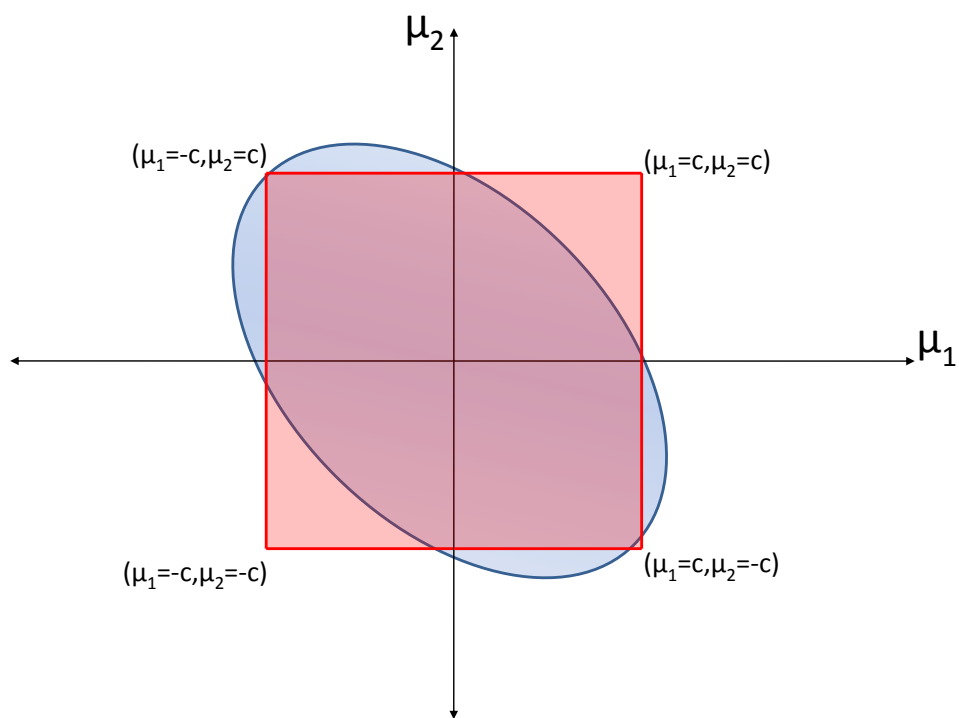
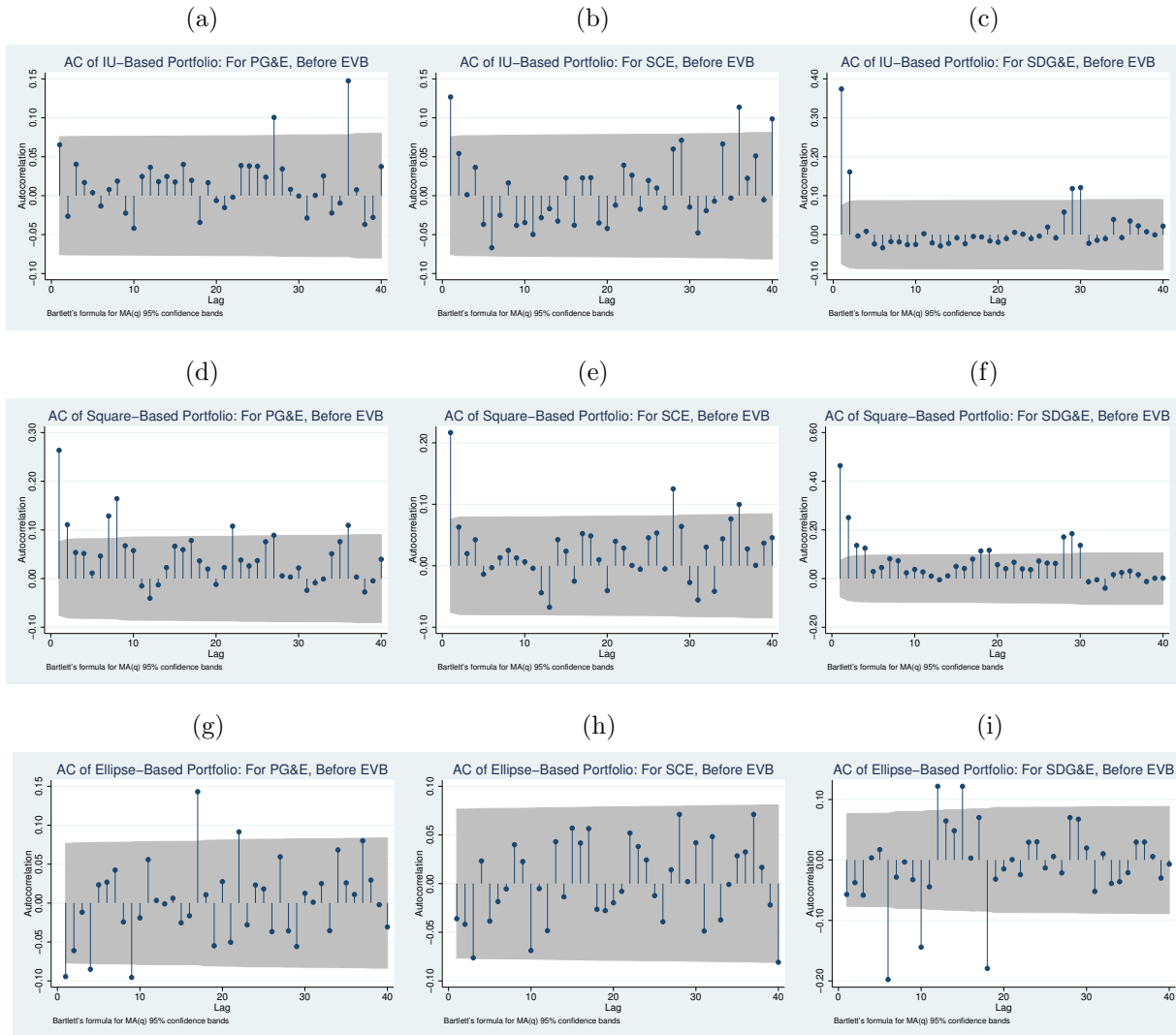Figure 4: LAP-level Daily Autocorrelations for Portfolios: Before EVB

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

(i)

Figure 5: LAP-level Daily Autocorrelations for Portfolios: After EVB

(a)

AC of IU−Based Portfolio: For PG&E, After EVB

(b)

AC of IU−Based Portfolio: For SCE, After EVB

(c)

AC of IU−Based Portfolio: For SDG&E, After EVB

(d)

AC of Square−Based Portfolio: For PG&E, After EVB

(e)

AC of Square−Based Portfolio: For SCE, After EVB

(f)

AC of Square−Based Portfolio: For SDG&E, After EVB

(g)

AC of Ellipse−Based Portfolio: For PG&E, After EVB

(h)

AC of Ellipse−Based Portfolio: For SCE, After EVB

(i)

AC of Ellipse−Based Portfolio: For SDG&E, After EVB

31

Figure 6: Average Hourly MW Virtual Supply and Demand Offered and Cleared: Monthly
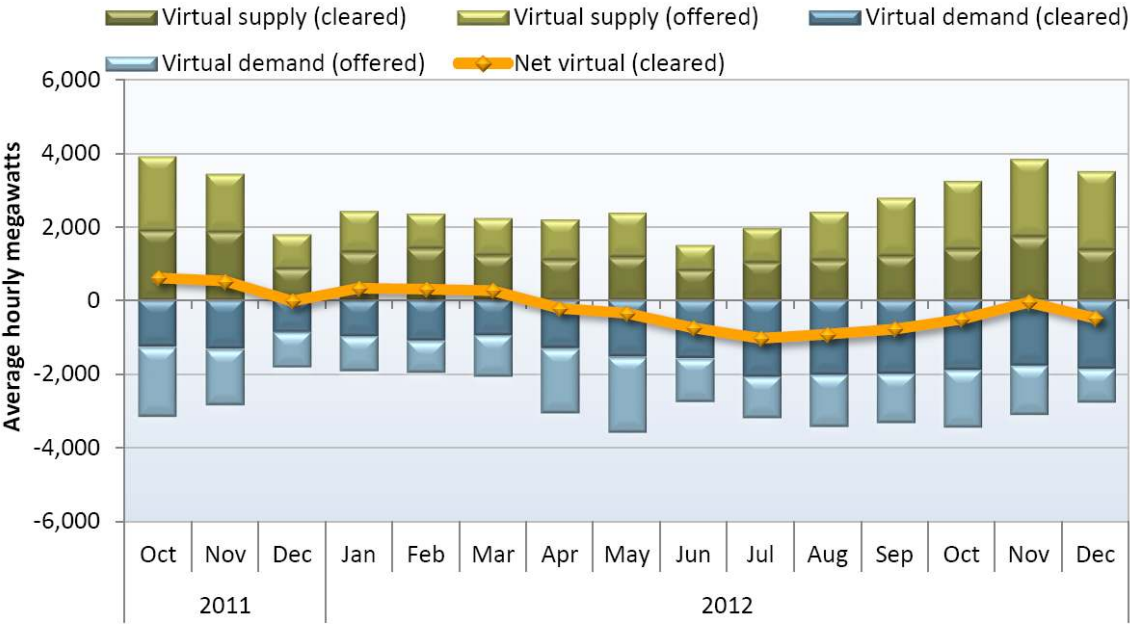
Figure 7: Average Hourly MW Virtual Supply and Demand Offered and Cleared: Hourly
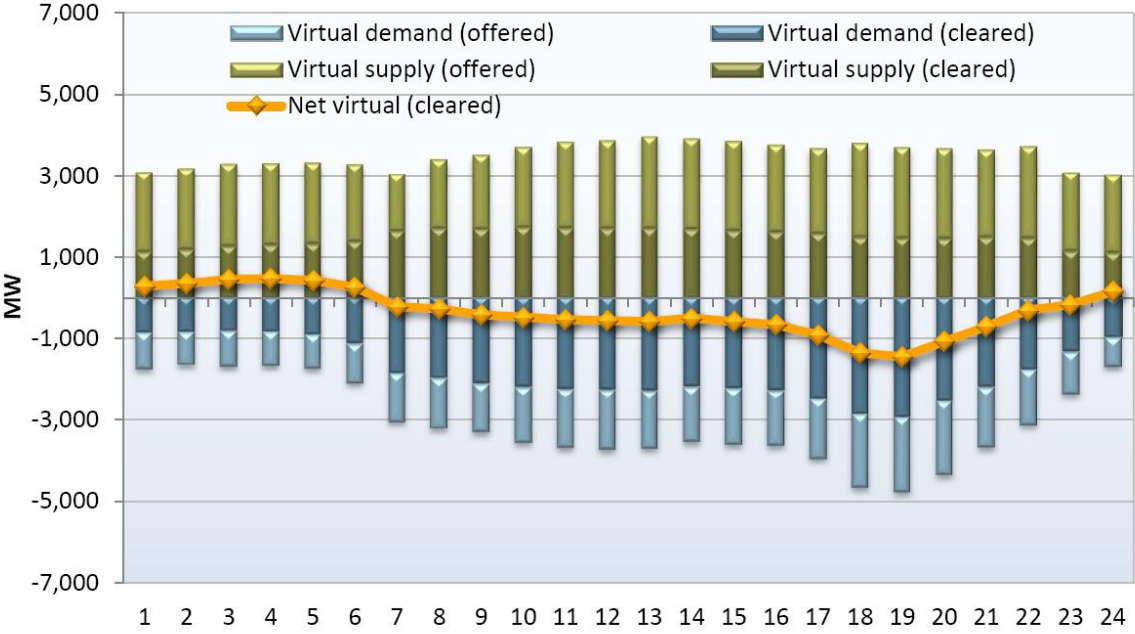
Figure 8: LAP-level Bootstrap Distribution of Implied Trading Costs: By Hypothesis Test
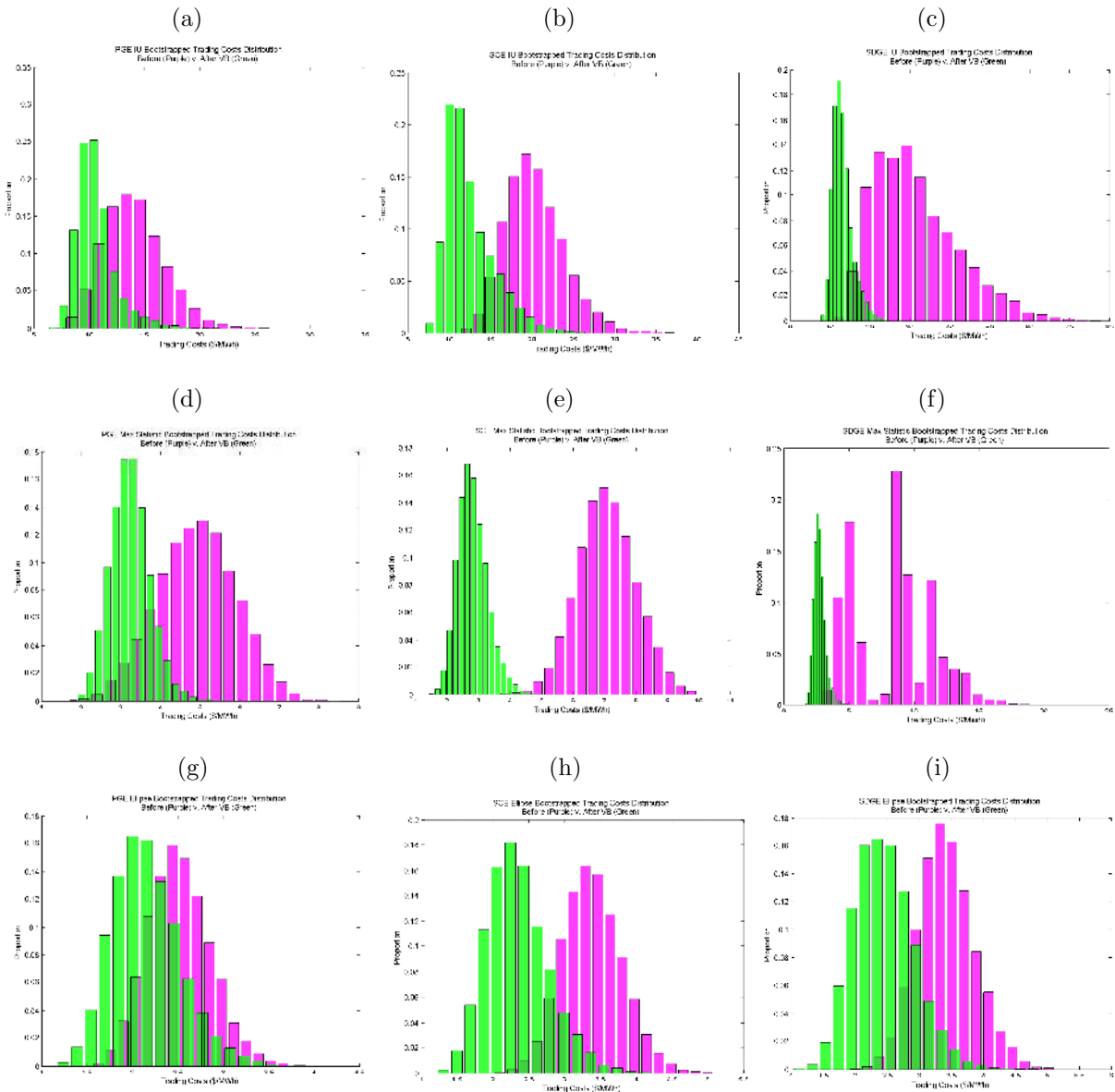
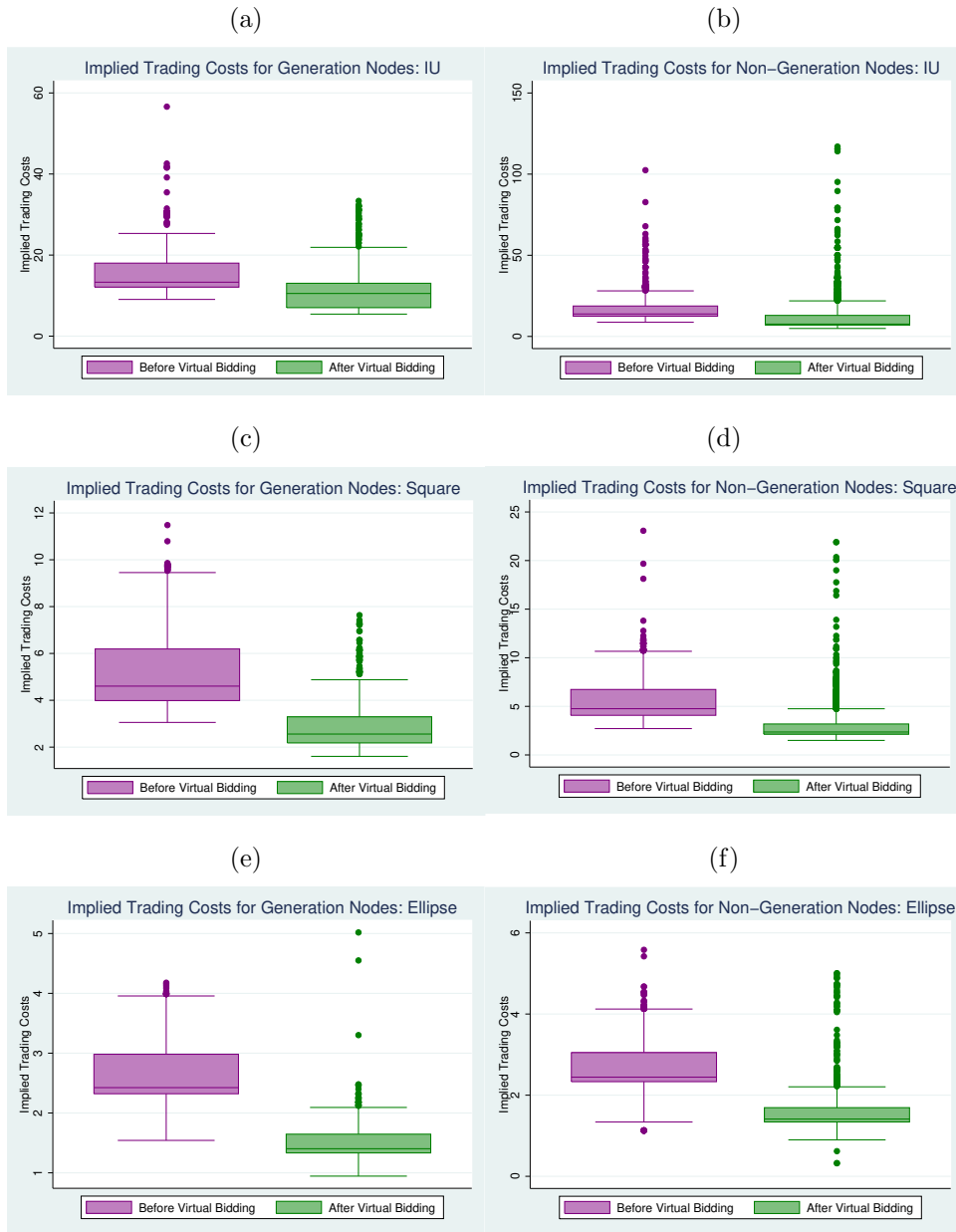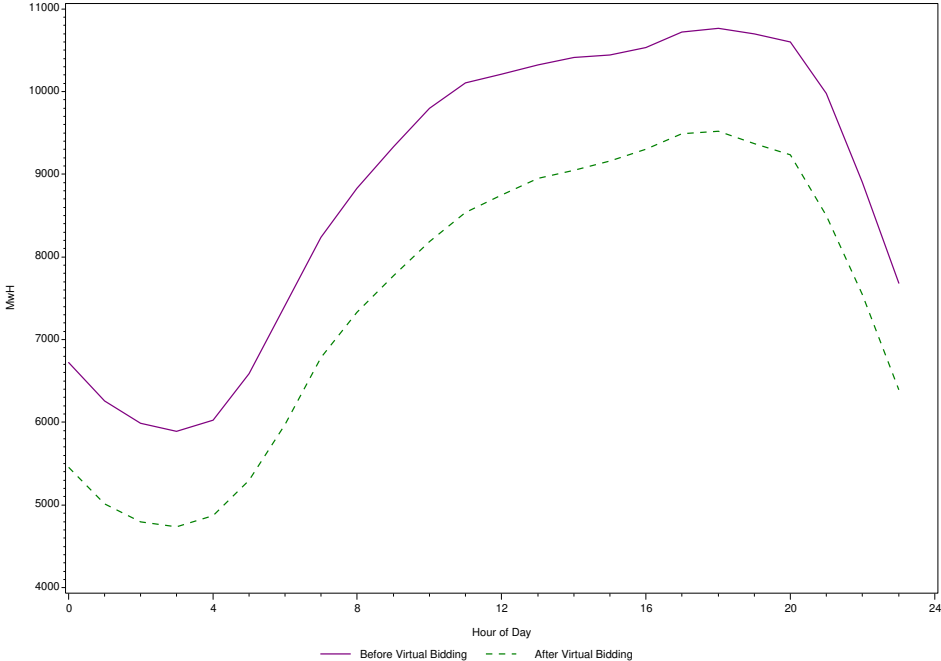Figure 9: Boxplot of Nodal-level Implied Trading Costs: By Type of Node

(a)

(b)

(c)

(d)

(e)

(f)

Figure 10: Average Total Output By Type of Resource: By Hour of Day

(a)

CAISO Wholesale Electricity Total Natural Gas Fueled Output
Average: By Hour of Day

Before Virtual Bidding   - - -   After Virtual Bidding

(b)

CAISO Wholesale Electricity Total Wind and Solar Output
Average: By Hour of Day

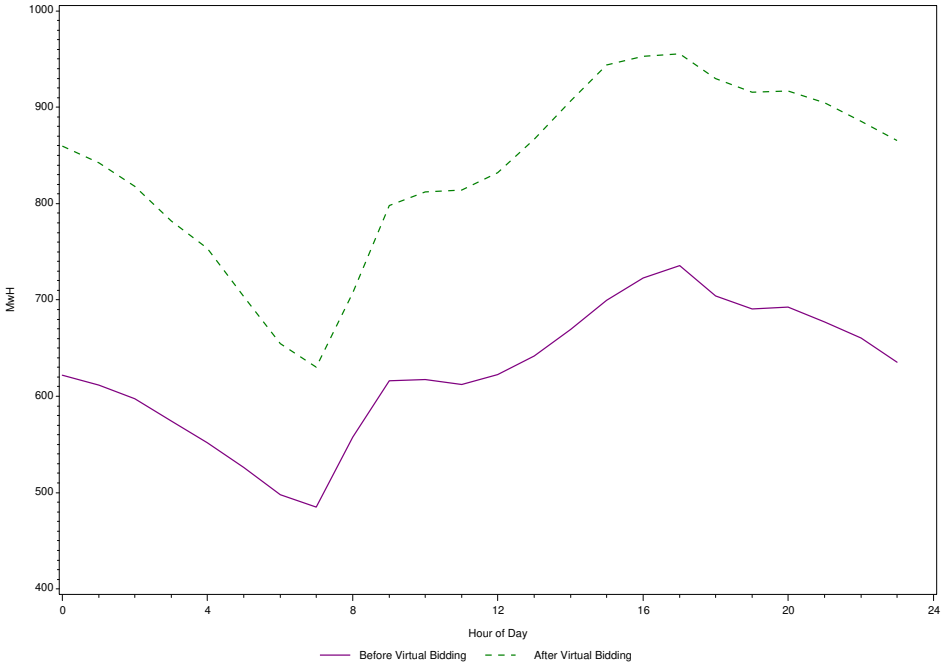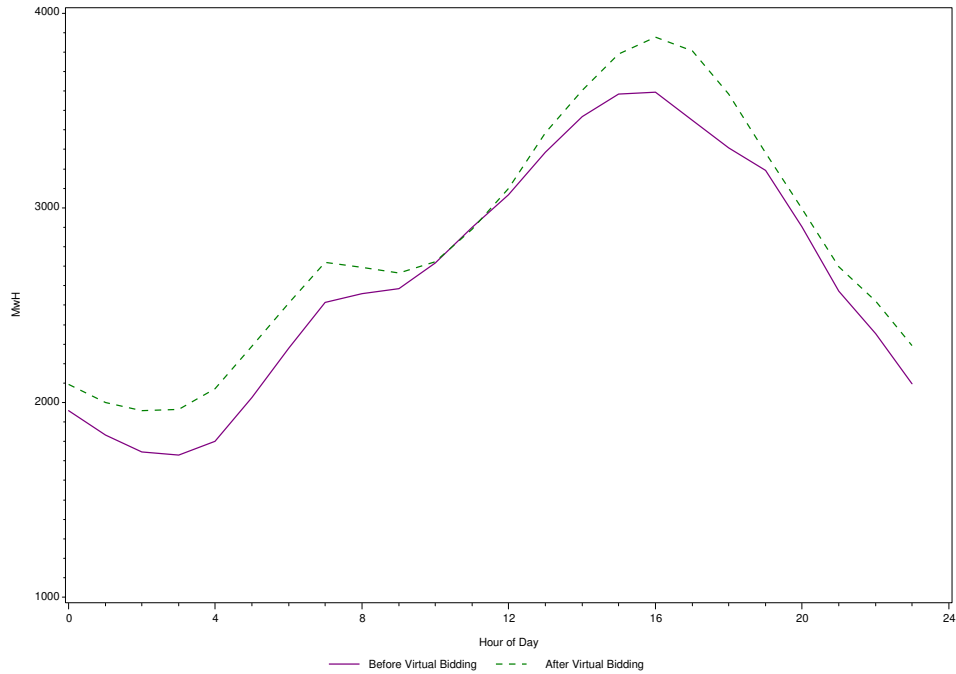Before Virtual Bidding   - - -   After Virtual Bidding

Figure 11: Standard Deviation of Total Output By Type of Resource: By Hour of Day

(a)

CAISO Wholesale Electricity Total Natural Gas Fueled Output
Standard Deviation: By Hour of Day

(b)

CAISO Wholesale Electricity Total Solar and Wind Output
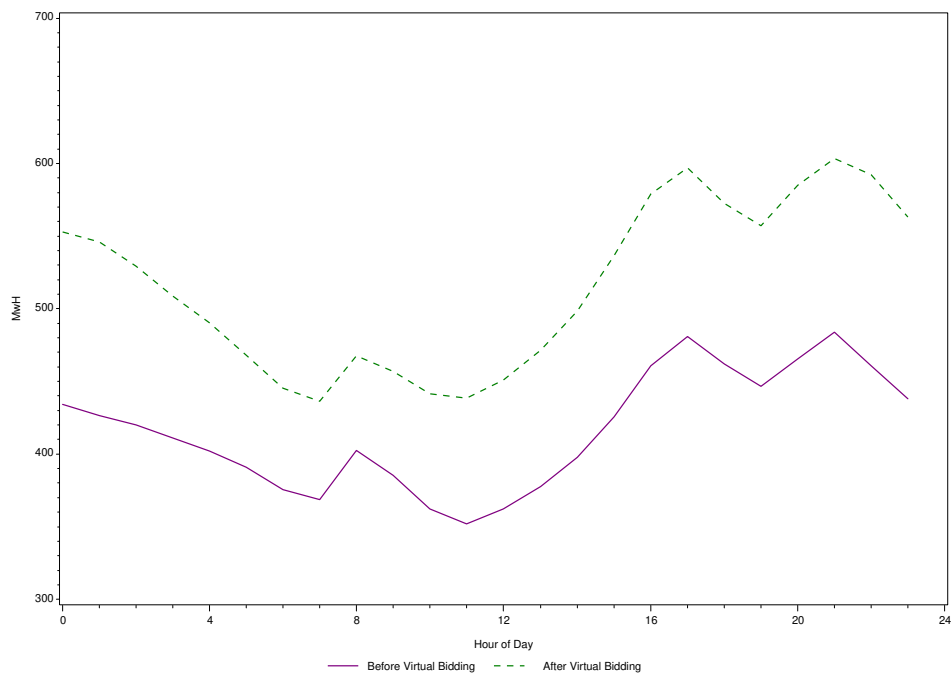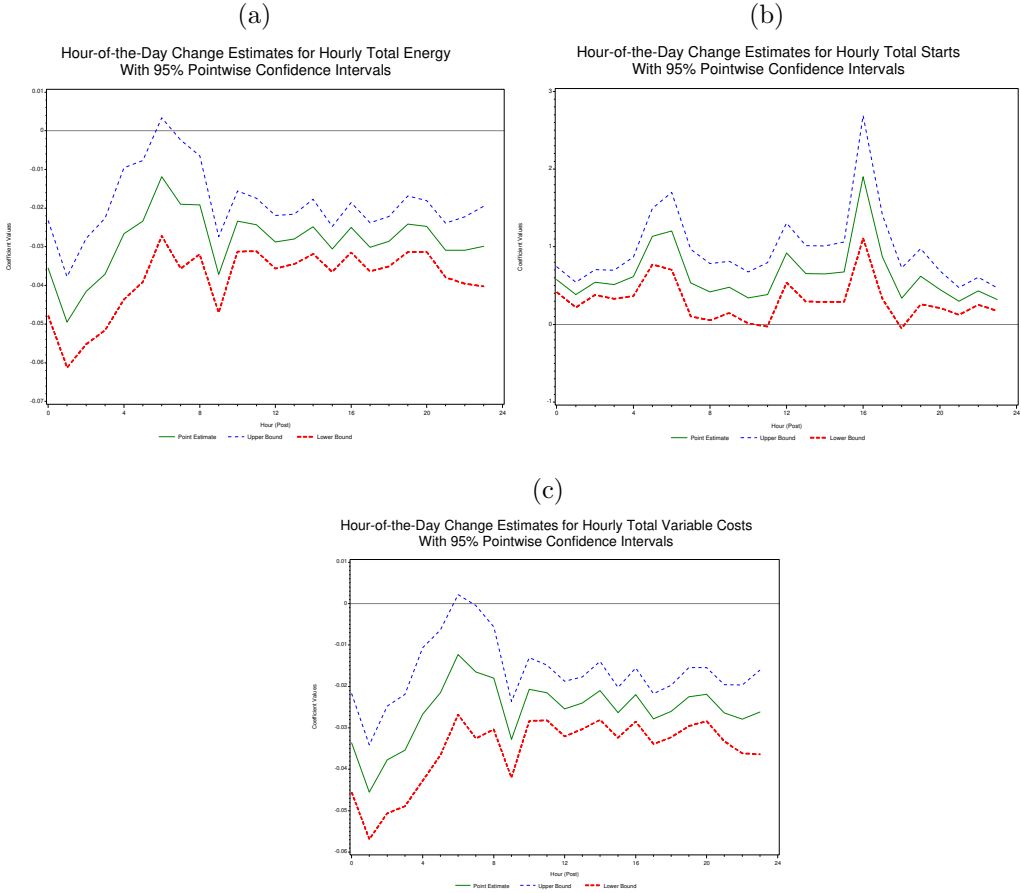Standard Deviation: By Hour of Day

37

Figure 12: Hour-of-the-Day Percent Change Estimates from Semi-Parametric Regressions

(a)

Hour-of-the-Day Change Estimates for Hourly Total Energy
With 95% Pointwise Confidence Intervals

(b)

Hour-of-the-Day Change Estimates for Hourly Total Starts
With 95% Pointwise Confidence Intervals

(c)

Hour-of-the-Day Change Estimates for Hourly Total Variable Costs
With 95% Pointwise Confidence Intervals

38

List of Tables

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

Table 2: LAP level Implied Trading Costs

		Before EVB	After EVB
	PG&E	12.882	9.692
IU Test	SCE	18.176	9.730
	SDG&E	30.307	11.067
	PG&E	4.265	3.028
Square-Based Test	SCE	6.737	2.656
	SDG&E	9.742	2.984
	PG&E	2.436	2.139
Ellipsoidal Test	SCE	3.716	1.848
	SDG&E	3.093	1.844

Table 3: Regression Results Associated with Implied Trading Costs

Dependent Variable: Implied Trading Costs	(1) IU	(2) Square	(3) Ellipse
Generation Node Indicator	-1.402	-0.271	-0.000936
	(0.241)	(0.0745)	(0.0230)
Post Explicit Virtual Bidding (EVB) Indicator	-5.741	-2.642	-1.104
	(0.164)	(0.0388)	(0.0102)
Interaction Between Generation and Post EVB Indicators	1.532	0.251	-0.00982
	(0.345)	(0.0888)	(0.0276)
Constant	17.13	5.607	2.650
	(0.110)	(0.0313)	(0.00832)
Observations	9,780	9,780	9,773
R^2	0.126	0.364	0.587

Heteroscedasticity-consistent standard errors in parentheses

Table 4: P-values associated with the Absolute Difference Tests

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

Table 5: Semiparametric Coefficient Results

Dependent variable	$ln(TOTAL\_ENERGY(t))$	$STARTS(t)$	$ln(TOTAL\_VC(t))$
β	-0.0284	0.6328	-0.0257
Standard error	0.0015	0.0496	0.0015