CSEM WP 121

Speculative Trading and Market Performance: The Effect of Arbitrageurs on Efficiency and Market Power in the New York Electricity Market

Celeste Saravia

November 2003

This paper is part of the Center for the Study of Energy Markets (CSEM) Working Paper Series. CSEM is a program of the University of California Energy Institute, a multi-campus research unit of the University of California located on the Berkeley campus.

2547 Channing Way, Suite 5
Berkeley, California 94720-5180
www.ucei.org

Speculative Trading and Market Performance: The Effect of Arbitrageurs on Efficiency and Market Power in the New York Electricity Market

Celeste Saravia*

University of California, Berkeley

Department of Economics

November 24, 2003

Abstract

While the effect speculators have on forward premiums (the difference between forward and expected spot prices) has been widely studied, there has been very little focus on the effect speculators have on competition in the product market. I study the effect speculators have had on production decisions and price levels in New York's deregulated electricity market. For the first two years of its operation, the market, which opened in November 1999, restricted trade to producers and retailers of electricity. During this period, the forward price of electricity in western New York was significantly higher than the expected spot price. I show that, after the market opened to purely speculative traders, the forward premium significantly decreased. In addition, the forward price of transmission (the price difference between two geographically distinct points) ceased to differ significantly from the expected spot price of transmission. I present a theoretical model to help understand these price relationships and other possible effects of speculators on market prices and firms' production decisions. Absent speculators, the model predicts that firms with market power will price discriminate between the forward and spot markets for electricity, resulting in the forward price being higher than the expected spot price. This discrimination in the market for electricity will result in the forward price of transmission under-predicting the spot price of transmission. When speculators that prevent firms from price discriminating are added to the model forward price-cost margins decrease. Using detailed data on the marginal costs of generation units in New York, I test these predictions of the model, and find that, after controlling for other market changes, the forward price-cost margins of firms in western New York did, in fact, significantly decrease after speculators were allowed to enter the market.

---

*E-mail: csaravia@econ.berkeley.edu.

I would like to thank my advisor Severin Borenstein, as well as James Bushnell, Meghan Busse, Richard Gilbert, Matt Lewis, Aviv Nevo, Catherine Wolfram and everyone at the University of California Energy Institute for advice and financial support.

1 Introduction

Some regulators and commodity traders have asserted that speculators destabilize prices. In frustration with this view, Milton Friedman asserted "People who argue that speculation is generally destabilizing seldom realize that this is largely equivalent to saying that speculators lose money, since speculation can be destabilizing in general only if speculators on the average sell when the currency is low in price and buy when it is high"[1953, p.175].[1] The Industrial Organization literature has considered the effect that futures markets have on production decisions and competition in product markets and the Finance literature has considered the effect that speculators have on financial market efficency.[2] But, to my knowledge, neither literature has considered the effect that speculators, who increase futures market efficiency, have on competition in the product market.[3]

The New York electricity market, which began operation in November of 1999, offers a unique opportunity to study the effects that speculators have on futures market efficiency and competition in the product market. Every day, the New York Independent System Operator (NYISO) holds both a day ahead (DA) futures market and a real time (RT) spot market for electricity. During its first two years of operation, trade in the market was limited to what the industry describes as physical participants. A physical participant is a firm that either owns generation or is responsible for procuring electricity for retail consumers.[4] Initially, the market rules also restricted these physical market participants from speculating on futures (DA) and spot (RT) price differences. The "virtual bidding" policy, which was implemented on November 7, 2001, opened the New York market to purely financial players and allowed all traders to speculate on DA and RT price differences.

I begin by demonstrating that since the implementation of the virtual bidding policy,

[1]In response to Friedman's statement, Hart and Kreps [1986] and Stein [1987], along with others, offered models showing that rational speculators trading in a futures market can actually destabilize spot prices. In the Industrial Organization literature, economists have considered whether product market power causes spot prices to be more or less stable. See Carlton [1986] and Slade [1991].

[2]See Allaz and Vila [1991] and Anderson and Sundaresan [1984].

[3]Slade and Thille [2003] look across industries at how liquidity in a futures market and market structure are related, but they do not explicitly examine the effect of an increase in speculators on the ability of firms to exercise market power.

[4]Limiting trade to physical market participants is not a common practice in commodity markets. Most commodities are traded in open exchanges such as the Chicago Board of Trade. This allows speculators to bear some of the risk associated with the market.

the absolute value of the forward premium (difference between the forward and expected spot prices) in the New York market has decreased significantly. In addition, the DA price of transmission (the price difference between two geographically distinct points) ceased to be significantly different than the expected RT price of transmission. To help understand these pre and post-virtual bidding price relationships, I present a duopoly model that predicts, absent speculators, that firms with market power will price discriminate between the DA and RT markets for electricity. The model also predicts that price discrimination in the market for electricity will result in the DA price of transmission under-predicting the RT price of transmission. When speculators that prevent firms from price discriminating are added to the model, DA price-cost margins decrease and there is a small increase in RT price-cost margins. Using detailed data on the marginal costs of generation units in New York, I test these predictions of the model. I find that, controlling for market conditions, the DA price-cost margins of firms in western New York did significantly decrease after the policy was implemented. There is less clear evidence on the effect that the policy has had on RT margins.

A key to understanding the effect of speculators in this market is to understand transmission congestion and transmission pricing. In the same way in which railroad, interstate and shipping transportation networks affect competition in many commodity markets, transmission lines in electricity systems link markets and increase competition. However, unlike other transportation systems, the short run marginal transportation cost of electricity is either very close to zero or infinite. If there is excess capacity on a transmission line, then electricity can be almost costlessly transported across the line.[5] If the flow across the transmission line is at capacity, the line is congested and no more electricity can be transported. In this case, the marginal cost of transportation is infinite.

If a transmission line is congested, the markets at the endpoints of the line clear separately. If the line is not congested, both endpoints clear as one single market. The areas on opposite sides of large transmission lines are called zones. In the New York DA and RT markets, the transmission price between two zones is defined as the price in the

[5]This discussion ignores line losses. The true marginal cost of transporting electricity over an uncongested line is the cost of lost electricity, which is usually less than 2% of the cost of electricity.

importing zone minus the price in the exporting zone. If the transmission line is uncongested, then the transmission price is zero. The NYISO defines a transmission line as congested in the DA market, if the trades scheduled DA would result in the transmission line being congested in RT.

Other researchers have considered the effect that transmission congestion can have on the strategic interaction of generation firms. Joskow and Tirole [2000] analyze the effects of different types of transmission rights when there is a monopoly at the importing zone. Gilbert, Neuhoff and Newbery [2002] extend the work of Joskow and Tirole to include the case of oligopolists at the importing node. Borenstein, Bushnell and Stoft [2000] present a model of two nodes, each of which is served by a monopoly. They find that if a transmission line is sufficiently large it will no longer be profitable for either monopoly to withhold enough to congest the line.

The model I present is unique in that it combines the complexities of market power and transmission congestion with those of market power in multiple temporal markets. Previous work on market power in multiple temporal markets includes that of Allaz and Vila [1993], henceforth AV, who present a duopoly model in which quantity setting firms trade in sequential perfectly-arbitraged forward markets before producing. The AV model predicts that the total quantity produced by duopolists will increase as the number of forward markets increases.

The model presented in this paper combines a two-period AV-like model with a two-zone electricity system in which the procurers of electricity are risk averse.[6] I first analyze the model assuming that only producers and consumers are allowed to trade. Then to analyze the effects of the virtual bidding policy on production decisions and average procurement costs, I add speculators to the model.

If the markets are closed to speculators, the duopolists are able to price discriminate between the DA and RT markets. In the DA market, the duopolists sell a quantity that corresponds to the DA price, P_{da}. Then, in the RT market they are able to move further down the demand curve and sell more at a lower RT price, P_{rt}. Once arbitrageurs enter

[6]Before the virtual bidding policy, the utilities in New York that were responsible for buying electricity for end use consumers followed a risk minimizing procurement strategy. They bid almost all of their forecast demand inelastically into the DA market. The bidding strategy of the utilities is discussed more later.

the market, the duopolists are forced to charge the same price in both the DA and RT markets. When the duopolists can no longer price discriminate, the total quantity they supply decreases. This is similar to the case of a monopolists that is able to perfectly price discriminate. If the monopolist is forced to charge a uniform price, then the quantity sold will decrease and deadweight loss will increase. The same is true in the current situation. Speculators will result in a decrease in production by the duopolists and an increase in deadweight loss.

Price discriminating between the DA and RT markets for electricity may result in the DA price of transmission being less than the RT price of transmission. Generation firms with market power that are located at an exporting zone will earn higher margins if the transmission line is uncongested. These firms may withhold enough sales from the DA market so that trades scheduled through the DA market will not congest the line. In this case, the transmission price in the DA market will be zero and the market clearing price for both zones will be the same. In the RT market, the firms in the exporting zone will increase production, which may result in RT congestion. If the transmission line is congested in RT, the RT price of transmission will be strictly greater than zero. Once speculators are added to the market, the transmission prices in the DA and RT markets will be equal.

While the addition of speculators increases deadweight loss, the model also predicts that they will decrease the average procurement cost of electricity. If, absent speculators, the DA price of transmission was less than the RT price of transmission, the decrease in procurement cost could be quite large. In many deregulated electricity markets, utilities are able to pass through the cost of procuring electricity to end use consumers. This implies that utilities may not have an incentive to minimize procurement costs. If speculators have a large effect on the average procurement cost and a small effect on the total quantity produced by the duopolists, *i.e.* efficiency, regulators may wish to open more electricity markets to speculators. When making this decision, regulators will need to weigh the increases in inefficient production which will result from firms with market power withholding more production with the gains to end-use consumers from lower retail rates.

In the empirical section of this paper, I estimate the effect that the virtual bidding policy has had on the relationship between DA and RT prices and the ability of firms to exercise market power in the Western New York electricity market. I first examine the effect that speculators have had on the relationship between DA and RT prices. There are two reasons speculators may affect this relationship. First, speculators add liquidity to the market which may help to eliminate market inefficiencies. Second, by bearing some of the market risk, speculators should decrease any risk premium in the market.

I find that the DA price in Western New York is significantly higher than the RT price in both the pre and post virtual bidding periods, but after the policy change, the DA bias significantly decreases. I also find that before the introduction of the virtual bidding policy the DA price of transmission was a biased forecast of the RT price of transmission. There is no significant bias in the DA price of transmission in the period after the policy was implemented. If the risk associated with speculating is not correlated with the overall risk in the economy and competitive speculators are trading in a market with no transaction costs, then the DA price should be an unbiased forecast of the RT price. Though I find evidence that the forward premium has decreased, I can still reject that the DA price of electricity is an unbiased forecast of the RT price. As speculative bidders become more familiar with the market and more enter, increasing the liquidity and sharing risk, we may observe further improvements in market efficiency.

By using publicly available engineering data for the marginal cost of each generation unit, I am able to directly calculate the price-cost margins of firms in New York. I use these margin estimates to examine the effect that virtual bidding has had on the ability of firms to exercise market power. In particular, I examine whether the evidence is consistent with the duopoly price discrimination model. The model implies that the DA price-cost margins of firms in the exporting zone should decrease and that the policy should not have a large effect on RT margins. After controlling for changes in demand, I find that the DA margins of the two large firms in western New York, NRG and AES, have significantly decreased while AES's RT margins have not significantly changed. In all specifications, NRG's RT margins appear to have decreased and this decrease is significant in some specifications. However, NRG's DA margins decrease significantly more than its RT margins in all specifications.

The model predicts that, absent speculators, firms in the exporting zone may withhold sales so that the DA market predicts the line will be uncongested. In the RT market, they may increase production so that the line will be congested in RT. If the firms in New York were doing this before the virtual bidding policy, then, after the policy change, the DA price-costs margins of the exporting firms, NRG and AES, should decrease more than that of a firm in the importing zone. The relationship between the RT margins of firms in the Central and West zones should not change after the policy. I test these predictions by estimating the change that has occurred in the difference between West and Central DA and RT margins. I find that the DA price-cost margins of the two large firms in the Western exporting zone, AES and NRG, have decreased significantly when compared to the those of the largest Central firm, Dynegy. The relationship between AES's and Dynegy's RT margins did not significantly change after the policy was implemented. In all specifications, the RT margins of NRG do appear to decrease when compared to those of Dynegy and in some specifications the decrease is significant. However, as in the levels results, the difference between NRG's and Dynegy's DA margins decreased significantly more than the decrease in the differenced RT margins. The result that both NRG's and AES's DA margins decreased significantly more than their RT margins is consistent with the price discrimination model. Another change may have occured that affected the level of both firms' DA and RT margins.

The paper proceeds as follows: in the next section, I give an overview of the New York electricity market's rules and structure. In Section 3, I present evidence that the virtual bidding Policy has improved market efficiency. In Section 4, I present a model of the Western New York electricity market. Section 5 presents the empirical findings on market power and Section 6 concludes.

2 The New York Market

In May of 1996, the New York Public Service Commission (PSC) released an order that called for the restructuring of New York's electricity sector. The restructuring plan required investor-owned utilities to divest the majority of their generation assets, and it called for the development of a competitive wholesale electricity market. The PSC set up the New York Independent System Operator (NYISO), a non-profit public-service entity charged with operating a wholesale electricity market in New York. In this section, I first describe the rules that govern the wholesale electricity market and then I give an overview of the post-divestiture market structure.

2.1 The Market Rules

In November of 1999, the New York Independent System Operator (NYISO) began holding day ahead (DA) futures and real time (RT) spot markets for electricity. Every day, the NYISO holds the DA market, which is an auction for electricity to be delivered in each hour of the following day. The RT market is a residual market that is held at the time of delivery. In both markets, the NYISO accepts supply and demand bids. Supply bids are offers to produce electricity while demand bids are bids to purchase electricity. The NYISO aggregates the supply and demand bids and then clears both the DA and RT markets through uniform price auctions.

The NYISO is a mandatory power pool. This means that all electricity trades must be scheduled through the ISO and every market participant is required to submit bids into the DA auction. Firms may have bilateral contracts, but each firm must still submit either a supply or a demand bid into the market. For example, if a generation firm and a utility have a bilateral contract, then the generation firm will still submit a bid to supply electricity into the NYISO and the utility will still submit a demand bid. The generation firm will be paid the market clearing price and the utility will pay the market clearing price. The firms will then settle the terms of their contract through side payments.

There are difficulties in electricity markets that are not experienced in other commodity markets. First, to avoid blackouts, supply and demand must be balanced at all times. In extreme cases, a large scale imbalance can lead to the type of large scale blackout that occured in the Northeast on August 14, 2003. Second, transmission congestion can make the problem of balancing supply and demand more difficult.

2.1.1 Transmission Pricing in the New York Market

A transmission line is similar to an import quota. If an import quota is binding, the two markets clear separately. Similarly, in an electricity market, if a transmission line is congested, the markets on the two sides of the line will clear separately. If an import quota is not binding, then the price in the importing country should be equal to the price in the exporting country plus the cost of transportation. Since the cost of transporting electricity is zero, if a transmission line is uncongested, then the electricity prices on each side of the transmission line will be the same. The NYISO acts as an arbitrageur and guarantees that the prices on both sides of an uncongested line are equal.

In the New York market, the price of transmission in any hour is the shadow value of increasing the capacity of a transmission line between two distinct points. If the line connecting the two points has excess capacity, the shadow value of capacity is zero. If the line is congested, then the marginal value of increasing the size of the transmission line is the price difference between the importing and exporting regions. The price of transmission is always defined as the price in the importing zone minus the price in the exporting zone, $P_{\text{Transmission}} = P_{\text{Importing}} - P_{\text{Exporting}}$. If the transmission line connecting the two zones is not congested, then the prices in the two zones will be the same, $P_{\text{Exporting}} = P_{\text{Importing}}$, and the price of transmission will be zero.

In the New York market, generators in the exporting region are paid the market clearing price in that region while buyers in the importing region pay the market clearing price in the importing region. This means that if a transmission line is congested, the NYISO will collect congestion revenues equal to the price of transmission times the capacity of the line. Congestion revenues that are collected in the DA market are distributed to transmission rights owners. The NYISO retains any congestion revenues that are collected in the RT market and uses them to offset operating costs.

The DA market is cleared by running the bids of all participants through an algorithm that takes account of the transmission constraints within the NY system. If the trades scheduled through the DA market would result in a transmission line exceeding its physical limit, then the transmission line will be congested in the DA market. This implies that the market on each side of the line will clear separately in the DA market. If a line is congested in the DA market, aggregate supply and demand curves are constructed with bids on each side of the line and each side is cleared separately. In the DA market, owners of transmission rights receive the price difference between what is paid in exporting zones and the price demand pays for electricity.

The RT market is a residual market. This means that the only trades that occur in the market are deviations from DA positions. Retail rates for electricity are usually preset for extended periods of time and only loosely tied to the RT price in any hour.[7] This means retail consumers do not respond to RT prices, which implies that all of the demand in the RT market is inelastic. In particular, utilities must purchase the difference between the quantity they purchased DA and the actually quantity consumed by retail consumers, regardless of the RT price. Generators only submit bids to the RT market if they are willing to deviate from the quantity they sold in the DA market. In order to ensure that supply and demand are always in balance, a separate RT auction is held at least once every five minutes. This means that at least once every five minutes the ISO clears the supply bids with the inelastic quantity demanded in that interval.

The actual flow of electricity determines the congestion status of transmission lines in the RT market. If a line is operating at its capacity, then it will be congested in RT and the markets on each side of the line will clear separately.[8] Financial transmission rights are only tied to congestion revenues collected in the DA market. If there is congestion in the RT market that was not present DA, the ISO will keep the transmission rents. Since the RT market is a residual market that only accounts for about 5% of total trades, the quantity over which the ISO will earn the transmission price will generally be small.

2.1.2 The Virtual Bidding Policy

From November 1999 to November 2001, many rules discouraged market participants from speculating on price differences between the DA and RT markets. Only generators and load serving entities (LSEs), such as investor-owned utilities and municipal power authorities, were allowed to trade in the DA and RT markets.[9] All bids submitted to the NYISO had to be resource-specific and feasible. A resource specific bid is either a supply bid which is tied to a specific generation unit or a demand bid which is tied to a geographic area in which

[7]Retail rates for some customers are tied to the weighted average DA price. But, very few actual observe or have to pay the RT price.

[8]System operators do not actually know the true physical capacity of the line. Rather, they use historical limits to determine the capacity and once the historical limit is met, they will not schedule any more electricity to flow over the line

[9]In electricity markets, the quantity consumed is referred to as load. This terminology is also used in auctions for electricity. Demand bids that are submitted to electricity markets are often referred to as load bids.

there is physical demand for electricity. The feasibility rule prohibited participants from attempting to speculate on DA and RT price differences. Firms that were responsible for procuring electricity were only allowed to bid demand and generation firms were only allowed to bid supply. Feasibility also prohibited generation firms from selling a quantity higher than a unit's capacity. This rule limited a generators ability to sell electricity short. All of these rules severely limited generators ability to arbitrage price differences between the two markets. However, LSEs were allowed to purchase electricity in either the DA or RT market without incurring any penalties or additional trading charges. Thus, a cost minimizing, risk neutral LSE would purchase electricity in the market with the lower expected price.

Even though before the virtual bidding policy, there were no rules prohibiting or even discouraging LSEs from purchasing electricity in the RT market, most LSEs procured all of their electricity in the DA market. This was even true in zones in which the DA price was on average higher than the expected RT price. During this time, the retail rates of some utilities were tied to the DA market. Since these utilities could pass the DA price onto their retail consumers, procuring electricity DA would have been their risk minimizing strategy.

The virtual bidding policy that took affect on November 7, 2001 opened the market to purely financial trades. After the rule change, any entity that passed the credit requirements could submit financial bids into the DA market. Bids submitted to the NYISO that are not tied to a physical resource are referred to as "virtual bids." Each virtual bid submitted to the NYISO corresponds to a specific zone and if accepted, virtual positions must be reversed at the zonal RT price. In the DA market, virtual bids are treated the same as physical bids. If virtual trades scheduled through the DA market would result in a transmission line being congested in RT, then the line will be congested in the DA market.

Virtual bids that are submitted into the DA market are either bids to supply or consume electricity. If a bid is accepted, the bid must reversed in the RT market. The DA market does not distinguish between virtual and physical bids. This means that virtual transactions may induce or relieve congestion in the DA market. Consider the following example. Suppose a virtual supply bid of 50 MWh is submitted into the DA market at a price of $40. This virtual supply bid will be aggregated with all of the other DA supply bids to construct an aggregate market supply curve. This aggregate market supply curve will be intersected with the aggregate market demand curve. If the DA market clearing price is above \$40 then the virtual bid will be accepted. This means in RT the virtual bidder will have to reverse his position by purchasing 50 MWh. Virtual bids are considered "price takers" in the RT market because they do not submit price responsive bids. They must reverse their position at the RT market clearing price.

The previous example ignored congestion constraints. Consider the case of two zones that are connected by a transmission line with a of capacity 300 MW. Suppose that if the virtual bid to supply 50 MWh at a price of \$40 were not submitted, then the DA market clearing price in both zones would be \$45 and that the exporting zone would be sending 275 MWh to the importing zone. This implies that there are 25 MW of excess capacity on the line. Now, suppose the virtual bid to supply 50 MWh in the exporting zone at a price of \$40 were submitted. Now the trades implied by aggregating the bids on both sides of the transmission line would predict RT congestion. This implies that the DA markets will be cleared separately. In this example, the virtual supply bid induced congestion. In the RT market this virtual bidder will still have to reverse his position by purchasing 50 MWh of electricity.

2.2 The Market Structure

New York has an installed generation capacity of a little over 36,000 megawatts (MW) of electricity and a summer peak of slightly less than 31,000 megawatt hours (MWhs). The New York market can be divided into four zones: West, Central, New York City and Long Island.[10] Table 1 presents the capacity ownership of the largest nine firms by zone. The incentives that a firm, which owns generation, will have to exercise market power depends on its load serving obligations and/or long-term contracts. For example, Keyspan owns 17% of all the generation capacity in New York and 80% of the capacity on Long Island. However, Keyspan's incentives to raise the price on Long Island are limited by a long-term contract

[10] The New York market has 11 zones which correspond to LSE service territories and major transmission lines. I aggregate the five most western zones and the 4 central zones into the West and Central zones. Within these aggregated zones, the cost of transmission congestion is low compared to the across-zone transmission costs.

to serve all the Long Island Power Authority (LIPA) customers. When Keyspan purchased the generation assets on Long Island, it signed a buy-back-agreement. This means that Keyspan is required to provide electricity for LIPA retail consumers. Keyspan does not own enough generation assets to fulfill this obligation, which means it is a net buyer. Given that Keyspan is a net buyer, the firm will not have incentive to increase market clearing prices. The second largest firm with 12% of instate capacity is the New York Power Authority (NYPA) a state-owned public power enterprise. Niagara Mohawk and Consolidated Edison are both investor-owned utilities with obligations to supply electricity. These two firms are also net buyers of electricity and thus, do not benefit from high prices.

A merchant generator is a firm that owns generation assets and is not vertically integrated. This means merchant generators do not have obligations to supply electricity to end use consumers. These merchant generators will benefit from high prices as long as they have some un-contracted generation capacity. With 4,453 MW, NRG Energy is the largest merchant generating firm in the NY system. Orion, Dynegy, Sithe and AES are the other large merchant generators.

Table 1: The New York Market Structure

Firm	West	Central	NYC	Long Island	Total
AES	1380	0	0	0	1380
ConEd	0	948	542	0	1490
Dynegy	0	1708	0	0	1708
Keyspan	0	0	2165	3900	6065
Niagara Mohawk	2279	291	0	0	2570
NRG	2990	0	0	1463	4453
NYPA	3339	1032	883	0	5254
Orion	393	0	1899	0	2292
Sithe	1476	0	0	0	1476

In the empirical section of this paper, I will focus on the effect that the virtual bidding policy has had on the behavior of AES and NRG in the Western zone.[11] There

[11] I am forced to leave Sithe out of the analysis because the Sithe plants are co-generation plants. This means that the actual cost of producing electricity is not a simple engineering formula and that the actual production of the plants is not available. However, Sithe has operated these plants since well before the

are two reasons for this focus. First, I am interested in the ability of firms in an exporting region to exercise market power. The Western zone has an installed capacity of 15,000 MW an average load of 6,450 MWhs and a peak load of 9,100 MWh. The transmission line between the Central and Western zone which has a capacity of 6,400 MW is congested 25% of the time, virtually always from the West zone to the Central zone. Thus, the Western zone offers a unique opportunity to examine the ability of firms in an exporting region to exercise market power.

The second reason I choose not to include firms in NYC and Long Island in the analysis is that these zones have additional market power mitigation rules. When the transmission lines into Long Island and NYC are congested, generators inside the market face a much less elastic residual demand curve and hence have more market power. To mitigate this additional market power, generators in NYC and Long Island face additional market power mitigation rules. In particular, if the DA market is congested and the price in NYC is more than 5% greater than the price in the Central zone, the in-city mitigation rules come into effect. Under these rules, the NYISO may change the bid of a generator that has submitted an unusually high bid or the ISO may pay a generator a price higher than the market clearing price. This implies that the market clearing price in Long Island and NYC may not accurately reflect the prices being earned by the generators in those zones.

3 Virtual Bidding and DA Market Efficiency

There are at least two ways in which the addition of speculators may affect the relationship between forward and spot prices. In a market in which trade is restricted to a few participants, these participants will have "market power" in the speculation market. Like firms with product market power, firms with speculative market power, will produce (trade) until the marginal cost of production (marginal transaction cost) is equal to marginal revenue. If firms are able to affect the market clearing price (the expected price difference between the two markets) then marginal revenue will not equal price. By opening the market to more speculators, the virtual bidding policy should increase competition in the speculation mar-

introduction of the New York market. This suggests that the generation may be under a contract.

ket. This should result in a decrease in the marginal expected speculation revenue, which is the difference between the DA and expected RT prices.[12] In the New York electricity market, the marginal transaction cost is very close to zero.[13] This means that, if speculation is competitive and traders are risk neutral, then the DA price should equal the expected RT price.

A second way in which speculators may affect the spot-futures relationship is by bearing some of the risk in the market. If the risk associated with speculating is uncorrelated with the overall risk in the economy, then competitive speculators should compete away any risk premium. If the speculation is partially diversifiable, meaning that the speculation risk is not perfectly correlated with the overall risk in the economy, then speculators will compete away the part of the risk premium that is diversifiable. In general, as long as the speculation risk is not perfectly correlated with the overall risk in the economy, the addition of competitive speculators to a market should decrease the risk premium.

The risk in electricity markets includes the risk of a generation unit having an unexpected outage, meaning that it needs repairs, or demand being higher than expected. The risk between the DA and RT market for electricity will not include risk associated with the long run growth of electricity demand or trends in the prices of fuel oils. This means the risk should not be correlated with the overall risk in the economy, which implies speculators should compete away any risk premium.[14]

In this section, I examine the effect that the virtual bidding policy has had on the relationship between the DA and RT prices for electricity and transmission. In an efficient futures market with competitive traders, completely diversifiable risk and no transaction costs, the futures price for delivery at time t should be an unbiased forecast of the spot

[12] For a more detailed discussion of market power in electricity speculation markets see Borenstein, Bushnell, Knittel and Wolfram [2003].

[13] However, there may be significant fixed costs involved in trading. For example, new traders would need to learn the complex rules that govern the market.

[14] A long term electricity contract may be correlated with the overall risk in the economy because the growth in the economy does affect the demand for electricity. The price of natural gas and oil, which affect the overall economy, are major inputs into electricity generation and thus affect the cost of electricity. But, neither of these variables, overall economic growth or fuel prices, change significantly between the DA and RT markets and thus they are not important for this type of short run trading.

price at time t. I compare the forward premium in the New York market to the forward premium in this type of market, which should be zero. I test the hypothesis that the DA price of electricity is an unbiased forecast of the RT price of electricity and if there has been a change in the bias since the virtual bidding policy was implemented. Similarly, I test if the DA price of transmission is an unbiased forecast of the RT price of transmission. I first explain the predictions of an efficient market in more detail. Then, I test to see if the virtual bidding policy has had a significant effect on the predictive power of the DA prices of electricity and transmission.

Other researchers have also studied the relationship between future and spot prices in electricity market. Bessimbinder and Lemmon [2002] present a model of a perfectly competitive electricity market in which speculators are not allowed to trade. Longstaff and Wang [2002] examine the relationship between DA and RT prices in the Mid Atlantic Electricity Market. They find that mean DA prices were lower than mean RT prices. This relationship is attributed to risk aversion. Borenstein, Bushnell, Knittle and Wolfram [2003] examine the relationship between DA and RT prices in the California electricity market. They find evidence that firms speculating in the California market were exercising market power in the speculation activity.

3.1 Speculators and Futures Market Efficiency

If risk neutral traders are permitted to trade in a futures market with no transaction costs, then the futures price for delivery at time t should be an unbiased forecast of the spot price at time t. If this is not the case, for example if the DA price of electricity is higher than the expected RT price, a risk neutral trader could realize positive expected profits by selling short DA and buying back in RT. If enough, speculators are making this type of trade, then the DA price will decrease and the expected RT price will increase until the two converge.[15] The DA price should incorporate all information which is available to traders at the time of the DA market. Notationally this implies that $P_{\mathsf{rt}} = P_{\mathsf{da}} + \epsilon$ where ϵ is a mean zero error

[15]Even if traders are risk averse, the DA price should still be an unbiased forecast of the RT price if the β (correlation with the market portfolio) of the DA RT price difference is zero.

and $P_{\mathsf{da}} = E[P_{\mathsf{rt}}|\Omega_{\mathsf{da}}]$; where Ω_{da} is the information set at the time of the DA market.[16]

A transmission line of capacity k connecting zones i and e will be denoted as congested from zone e to zone i if zone e is exporting k MWs to zone i.[17] Let $\alpha_{\mathsf{da}}^{\mathsf{e\rightarrow i}}$ be the probability that a transmission line will be congested in RT from zone e to i conditional on all the information available DA. Similarly, let $\alpha_{\mathsf{da}}^{\mathsf{e\leftarrow i}}$ be the probability of the transmission line being congested from i to e in RT. If in RT the transmission line between the two zones is uncongested then the price in the two zones will be the same. Define $P_{\mathsf{rt}}^{\mathsf{uc}} = E[P_{\mathsf{rt}}|\Omega_{\mathsf{da}}, UC]$ to be the expected RT price in both zones conditional on the transmission line being uncongested. Let $P_{\mathsf{zrt}}^{\mathsf{e\rightarrow i}} = E[P_{\mathsf{zrt}}|\Omega_{\mathsf{da}}, e \rightarrow i]$, and $P_{\mathsf{zrt}}^{\mathsf{e\leftarrow i}} = E[P_{\mathsf{zrt}}|\Omega_{\mathsf{da}}, e \leftarrow i]$ be the expected RT price in zone $z = e, i$ conditional on the line being congested from e to i and i to e respectively. If the market is efficient, then the DA price in zone z, P_{zda}, will be equal to the expected RT price conditional on all the information that is available at the time of the DA market:

$$P_{\mathsf{zda}} = \alpha_{\mathsf{da}}^{\mathsf{e\rightarrow i}} P_{\mathsf{zrt}}^{\mathsf{e\rightarrow i}} + \alpha_{\mathsf{da}}^{\mathsf{e\leftarrow i}} P_{\mathsf{zrt}}^{\mathsf{e\leftarrow i}} + (1 - \alpha_{\mathsf{da}}^{\mathsf{e\rightarrow i}} - \alpha_{\mathsf{da}}^{\mathsf{e\leftarrow i}}) P_{\mathsf{rt}}^{\mathsf{uc}}$$

Suppose $\alpha_{\mathsf{da}}^{\mathsf{e\leftarrow i}} = 0$ and $\alpha_{\mathsf{da}}^{\mathsf{e\rightarrow i}} > 0$. In this case, the DA price in each zone will simply be $P_{\mathsf{zda}} = \alpha_{\mathsf{da}}^{\mathsf{e\rightarrow i}} P_{\mathsf{zrt}}^{\mathsf{e\rightarrow i}} + (1 - \alpha_{\mathsf{da}}^{\mathsf{e\rightarrow i}}) P_{\mathsf{rt}}^{\mathsf{uc}}$. If the line is congested, then the price in the exporting region will be lower than importing region: $P_{\mathsf{ert}}^{\mathsf{e\rightarrow i}} < P_{\mathsf{irt}}^{\mathsf{e\rightarrow i}}$. If the line is uncongested, then electricity can be costlessly transported between the two zones and the two zones will have the same price, $P_{\mathsf{rt}}^{\mathsf{uc}} = P_{\mathsf{ert}}^{\mathsf{uc}} = P_{\mathsf{irt}}^{\mathsf{uc}}$. If there is a positive probability that a line will be congested in RT, then the DA price in the exporting region must be lower than the DA price in the importing zone, $P_{\mathsf{eda}} < P_{\mathsf{ida}}$. These prices will only differ if the transmission

[16]This argument for price convergence differs from cost-of-carry models which link the delivery price in period t to the forward price at t for delivery in period $t + j$. In these models the relationship between current spot and forward prices depends on the cost of storage. If arbitrageurs are able to buy in the period t spot market, sell in the forward market for delivery at period $t + j$, and store the good until the delivery date, then in an efficient market the futures price at time t for delivery at time $t + j$ must equal the spot price at time t plus the cost of storage for j periods. Currently there is no way to buy and store electricity on any large scale. Pump storage and batteries can be used to a very small extent to arbitrage inter temporal price differences, but the cost of these types of storage are extremely high. It can take twice as much electricity to pump water, as the water will release.

[17]The line will be congested from i to e if i is the exporting zone and the transmission line is congested.

line connecting the two zones is congested in the DA market.

This argument can be extended to the case in which $\alpha_{\mathsf{da}}^{\mathsf{e}\leftarrow\mathsf{i}} > 0$ and $\alpha_{\mathsf{da}}^{\mathsf{e}\rightarrow\mathsf{i}} > 0$. If there is positive probability of RT congestion in both directions and the expected RT prices in the two zones are not equal, then, in an efficient market, the line connecting the two zones will be congested in the DA market. Furthermore, in the DA market, the line will be congested from the zone with the lower expected RT price to the zone with the higher expected RT price.

The price of transmission is the shadow value of adding additional capacity to a line, which is the price difference on the two sides of the line. Notationally, this implies that the price of transmission is the difference in the prices in the two zones: $P_{\mathsf{e}\rightarrow\mathsf{i}}^{\mathsf{Transmission}} = P_{\mathsf{i}} - P_{\mathsf{e}}$. If a market is efficient then the DA price of transmission from zone e to zone i, $P_{\mathsf{ida}} - P_{\mathsf{eda}}$, should be an unbiased forecast of the RT price of transmission, $E[P_{\mathsf{irt}} - P_{\mathsf{ert}}] = E[P_{\mathsf{irt}}] - E[P_{\mathsf{ert}}]$. This implies that in an efficient market $P_{\mathsf{ida}} - P_{\mathsf{eda}} - E[P_{\mathsf{irt}} - P_{\mathsf{ert}}] = \epsilon$, where ϵ is a mean zero error.

3.2 Tests of Market Efficiency

The previous subsection outlined testable implications of an efficient market. In this section, I use data from the NYISO to test the effect the virtual bidding policy has had on the predictive power of the DA price of electricity and the DA price of transmission. Although the ISO divides the New York region into 11 zones, I aggregate the five most western zones and the four central zones. This aggregation divides the New York market into the West, Central, New York City (NYC) and Long Island (LI) zones.[18] As was discussed in the previous section, this paper only considers the effect that the virtual bidding policy has had on the West and Central zones.

Table 2 presents separate summary statistics for the periods before and after the virtual bidding (VB) policy was implemented. As would be expected, the RT zonal price

[18]The ISO reports most data at the level of each of the 11 zones. The West and Central DA and RT prices are created by taking the average price of all of the member zones. The forecast and load variables are created by aggregating the zonal load and forecasts provided by the ISO. Congestion in the aggregated zones will be equal to 1 if the lines out of the aggregated zones are congested.

Table 2: Pre and Post Virtual Bidding Summary Statistics

Pre Virtual Bidding 16,221 Observations					Post Virtual Bidding 15,543 Observations				
West Zone Pre VB					West Zone Post VB				
Variable	Mean	Std. Dev.	Min	Max	Variable	Mean	Std. Dev..	Min	Max
DAZP	36.49	20.54	0.01	915.02	DAZP	38.28	18.17	8.27	165.33
RTZP	32.21	30.80	-844.38	951.18	RTZP	36.49	26.69	-120.79	992.58
Central Pre VB					Central Post VB				
DAZP	44.65	33.26	-0.03	1075.50	DAZP	44.33	20.13	9.04	201.30
RTZP	43.00	45.14	-879.87	1089.32	RTZP	43.09	31.37	-135.21	1404.56
Congestion West $\rightarrow$ Central Pre VB					Congestion West $\rightarrow$ Central Post VB				
DAC	0.28	0.45	0	1	DAC	0.26	0.44	0	1
RTC	0.35	0.48	0	1	RTC	0.16	0.37	0	1

(RTZP) in each zone is far more volatile than the DA zonal price (DAZP). The average Central zonal prices are higher than the average West zonal prices. This reflects the fact that the West zone exports electricity to the Central zone and the transmission lines connecting the two zones are frequently congested. DA and RT congestion (DAC and RTC) are zero-one variables that indicate the transmission lines between two zones are congested.[19] A transmission line is defined as congested in the DA market, if the trades scheduled through the DA market would result in the transmission line being congested in RT.

I do not separately report DAC and RTC from the Central zone to the West because it is extremely rare, occuring in less than 2% of hours. When such "reverse" congestion does occur, the cost of transmission is negligible. It averages less than 50 cents which is less than one sixtieth of the average cost of electricity. Before the virtual bidding policy, the transmission line between the Central and West zones was congested 28% of the time in the DA market and 35% of the time in the RT market. This is evidence that the market was not functioning efficiently. Since the congestion between the West and Central zone rarely reverses direction, if the RT market has a positive probability of congestion then the DA market should be congested.

I first test the efficient market prediction that the DA price should be an unbiased

[19] Several transmission lines connect the two zones. It is almost always the case that either all the lines are congested or uncongested.

forecast of the RT price. I define the prediction error in the DA price at time t in zone z as $DAError_{zt} = P_{zda} - P_{zrt}$. In an efficient market, conditional on all the information available DA, the expected error should be zero: $E[DAError_{zt}|\Omega_{da}] = 0$. Even in an efficient market, $DAError_{zt}$ may be serially and contemporaneously correlated across zones. In particular, the following error structure is consistent with the market being efficient:

$$DAError_{zt} = \epsilon_{zt} = \mu_t + \eta_{z,t}$$

where μ_t and $\eta_{z,t}$ are mean zero shocks. μ_t is common across all of the zones and η_{zt} is a serially correlated zone specific shock. This error structure reflects the fact that in an efficient market a shock might occur after the close of the DA market that affects all of the zones and that shocks might have an effect on RT prices in multiple hours.[20]

To test the efficiency of the DA price in the pre and post virtual bidding period, I estimate the following equation:

$$DAError_{zt} = \sum_{z=1}^{4} \alpha_z Pre\_Zone_{zt} + \sum_{z=1}^{4} \gamma_z Post\_Zone_{zt} + \epsilon_{zt}$$

where Pre_Zone_{zt} ($Post\_Zone_{zt}$) is a dummy variable for the zone z before (after) the introduction of the virtual bidding policy; $z \in \{West, Central\}$. This model allows me to test if the DA price was a biased predictor of the RT price and if the bias changed after the introduction of virtual bidding. If the market is efficient, then the average $DAError$ should not be significantly different from zero. A test of the market being efficient before the policy change is a test of α_z being equal to zero. Similarly, a test of the market being efficient after the policy change is a test of γ_z being equal to zero.

The error, ϵ_{zt}, is modeled as contemporaneously correlated across zones and serially correlated within a zone. I estimate the model using the Prais-Winsten estimator which corrects for the error structure described above. The results are presented in column 1 of Table 3. The DA price in the West averaged \$3.97 more than the RT price. This implies

[20] In the estimate, I treat $\eta_{z,t}$ as an AR(1) process. I do this for ease of estimation and in order to get convergence. Under an efficient market, the actually structure of $\eta_{z,t}$ is rather complicated. The error will range from an moving average with 23 lags to a moving average with 48 lags. Borenstein, Bushnell, Knittle and Wolfram [2001] attempt to estimate the more complicated error structure when testing for convergence in California DA and RT market and are not able to get convergence of the maximum likelihood estimator.

Table 3: Tests of Price Convergence

Dependent variable for (1) & (2) is DAZP-RTZP by zone and hour

Variable	(1)	(2)
Pre VB West	3.97	3.86
	(0.39)**	(0.79)**
Pre VB West trend		0.41
		(2.62)
Post VB West	1.79	-0.25
	(0.41)**	(2.30)
Post VB West trend		2.69
		(2.98)
Pre VB Central	1.40	1.21
	(0.61)**	(1.23)
Pre VB Central trend		0.73
		(4.06)
Post VB Central	1.24	3.06
	(0.64)*	(3.56)
Post VB Central trend		-2.39
		(4.61)
West AR(1) Coefficient	0.55	0.55
Central AR(1) Coefficient	0.63	0.63
Observations	66266	66266

Praise-Winston coefficients are reported. Standard errors are given in parenthesis.
* Denotes significance at the 10% level, ** 5% level

that firms that purchased electricity in the West DA market were paying a 13% premium over those that waited to purchase in the spot market. The post virtual bidding DA bias in the West was \$1.79 or 5%. A t-test conclusively rejects that the pre and post virtual bidding coefficients in the West are the same.[21] The DA bias in the Central zone, which was \$1.40 before the policy did not significantly change in the post virtual bidding period.

When the market opened in November of 1999, most market participants were new to trading in deregulated electricity markets. Even those who had previously traded in other markets needed to learn the rules of the New York market. The DA price may have gradually become a better predictor of the RT price as the market participants learned more about how the market functioned. The previous model would have missed any long term trend toward price convergence. I build on the previous model by including a trend term to

[21]The p-value of the test is 0.0001.

capture the effect of learning on price convergence. I also interact the trend term with the virtual bidding dummy in order to pick up any changes in the learning process that may have resulted from the virtual bidding policy. I estimate the following model:

$$DAError_{zt} = \sum_{z=1}^{4} \alpha_z Pre\_Zone_{zt} + \sum_{z=1}^{4} \gamma_z Post\_Zone_{zt} + ...$$

$$+ \sum_{z=1}^{4} \beta_z trend_t * Pre\_Zone_{zt} + \sum_{z=1}^{4} \xi_z trend_t * Post\_Zone_{zt} + \epsilon_{zt}$$

where $trend_t$ is the number of days the market has been open. These results are presented in column 2 of Table 3. None of the time trends are significant. The previous two models provide evidence that before the virtual bidding policy the DA price was a biased forecast of the RT price in both zones and that after the policy, the bias significantly decreased in the West zone. Furthermore, these models do not show evidence of a long term trend toward convergence.

Table 2 reports that before the virtual bidding policy, there was real-time transmission congestion between the West and Central zone in 35% of the hours. However, the DA market only predicted congestion in 28% of the hours. This under prediction of congestion may have resulted in inefficient pricing of transmission in the DA market. The DA price of transmission is the DA zonal price in the importing zone minus the DA zonal price in the exporting zone $P_{ida} - P_{eda}$. If the DA market predicts that there will not be RT congestion between the two zones, then the DA price of transmission will be zero. If there is congestion in RT, then the RT price in the importing zone will be greater than the real time price in the exporting which means that the RT price of congestion, $P_{irt} - P_{ert}$ will be strictly greater than zero. To test the hypothesis that the under prediction of congestion in the RT market contributed to the DA price bias, I estimate the two previous models using $Transmission\_Error_{c,w,t} = P_{ct}^{da} - P_{wt}^{da} - [P_{ct}^{rt} - P_{wt}^{rt}]$ as the dependent variable where zone c is the Central zone and w is the West zone. The results for the two models are presented in Table 4 columns 1 and 2.

In the model with no trends, the pre virtual bidding coefficient is negative and significant. This implies that before the virtual bidding policy the DA price of transmission under predicted the RT price of transmission. The Central-West DA price of transmission was on average, \$2.57 less than the RT price before the policy change. After the virtual

Table 4: Tests of Congestion Efficiency

Dependent variable for (1) & (2) is $P_{\mathsf{cda}} - P_{\mathsf{wda}} - [P_{\mathsf{crt}} - P_{\mathsf{wrt}}]$ by zone and hour

w refers to the West zone and c the Central zone.

Variable	(1)	(2)
Pre VB West-Central	-2.57	-2.66
	(0.49)**	(0.99)**
Pre VB West-Central trend		0.32
		(3.27)
Post VB West-Central	-0.56	3.36
	(0.51)	(2.86)
Post VB West-Central trend		-5.13
		(3.70)
AR(1) Coefficient	0.70	0.70
Observations	33133	33133

Praise-Winston coefficients are reported. Standard errors are given in parenthesis.

* Denotes significance at the 10% level, ** 5% level

bidding policy was implemented this bias decreased to -$.56 and it is no longer significant. A t-test rejects, at the 1% level, that the pre and post virtual bidding coefficients are equal. Neither of the trend terms are significant.

The previous results have indicated that before the virtual bidding policy, the price of congestion in the DA market was less than that in the RT market. The owners of transmission rights only receive the transmission revenues which are collected in the DA market. This means that before the policy change, owners of transmission rights were earning less rents than they would have if the DA market were efficient.

4 A Model of the New York Electricity Market

In the previous section, it was shown that before the virtual bidding policy the price of transmission in the DA market was a biased predictor of the RT price of transmission between the West and Central zones and that the average DA price in each zone was significantly greater than the average RT price. In this section, I present a duopoly, two zone (Central and West), two market (DA and RT) model which suggests that the congestion inefficiencies documented in the previous section may be the result of firms in the exporting zone strategically keeping the transmission line uncongested in the DA market. I first analyze the model under the assumption that speculators are not allowed to trade in the market and then I add speculators to the model to analyze the affects of the virtual bidding policy on market power.

I present this model to gain insight into the inefficiencies in the New York market. To achieve this goal, I make a number of simplifying assumptions and many assumptions that are specific to the New York market. I first describe these production assumptions and solve the model for the one market case in which there is only a spot market. Then in Section 4.2, I outline the possible equilibria in the two market game in which the firms trade in both a DA and RT market. In Section 4.3, I describe how the equilibria change once speculators are added to the model. I give the intuition for the model here and leave many of the details and proofs for the Technical Appendix.[22]

4.1 Production Assumptions and The One Market Model

4.1.1 Production Assumptions

In electricity markets, retail rates are usually set for months at a time. In any given hour, an end use consumer's demand will depend on this pre-set retail rate and not on the price in the spot market. I incorporate this lack of consumer response into the model by assuming that each zone has inelastic demand Q_{d}.

I assume the West zone is supplied by symmetric duopolists with constant marginal cost that are normalized to be zero. A set of competitive fringe generators are also located in the Western zone. The Western fringe are assumed to have linear marginal cost $mc^{\mathsf{w}}(q_{\mathsf{f}}^{\mathsf{w}}) = q_{\mathsf{f}}^{\mathsf{w}}$ and a finite capacity F_{k}. The Central zone is assumed to be supplied by a set of high cost competitive generators with marginal cost $mc^{\mathsf{c}}(q_{\mathsf{f}}^{\mathsf{c}}) = bq_{\mathsf{f}}^{\mathsf{c}}$ where $b > 1$. These cost assumptions reflect the cost structure that is present in the New York electricity market. The West zone in New York is supplied primarily by low cost coal units. The marginal cost of coal units ranges from about \$15 to \$25 MWh. Most of the generation plants located in the Central zone use either natural gas or oil for fuel. The marginal cost of oil and gas units

[22] The Technical Appendix is available at www.http://www.ucei.berkeley.edu/Staff/saravia.html.

ranges from \$35 to \$80 MWh. The cost assumptions that I make reflect that the average marginal cost of gas and oil units is higher than the average marginal cost of coal units and that there is more variation in marginal cost across oil and gas units.

The two zones are assumed to be connected by a transmission line of capacity T_k. In New York, the transmission lines connecting the West and Central zones are not large enough for the less costly coal units in the West to supply all of the demand in the eastern part of the state. I incorporate this into the model by assuming that the demand in the Central zone is greater than the capacity of the line; $Q_d > T_k$. If the total production in the West is $Q_d + T_k$, then the transmission line will be congested and the two markets will clear separately.

If the duopolists produce at least $Q_d - F_k + T_k$, then the total production in the West will be $Q_d + T_k$ and the transmission line will be congested. Once the transmission line is congested, any increase in production by the duopolists will not affect the price in the Central zone. The increased production will displace production by low cost western fringe units and drive the price down in the West zone. In general, if the line is congested from the West to the Central zone, changes in production in the West will not effect the market clearing price in the Central zone unless the change results in the line becoming uncongested. This is similar to the situation in which an importing country has a binding quota on some good. Changes in demand or supply in the exporting country will not affect the price in the importing country unless the changes result in the quota no longer binding.

If the duopolists produce nothing, then the line will not be congested. This assumption has two purposes. First, it guarantees that there will always be adequate supply in the West, even if the duopolists produce nothing. Since demand is perfectly inelastic, if production from either duopolist were necessary to satisfy demand, there would be no equilibrium price.[23] Second, this assumption guarantees that there will never be congestion from the Central zone to the West zone. In the New York market, this type of congestion is extremely rare, occuring in less than 2% of hours. When this type of reverse congestion does occur the price differences in the two zones is not significant. In over 75% of hours with reverse congestion, the price difference in the two zones is less than 25 cents.

[23]Since the duopolists are assume to compete in quantities, they could withhold until the price is infinite

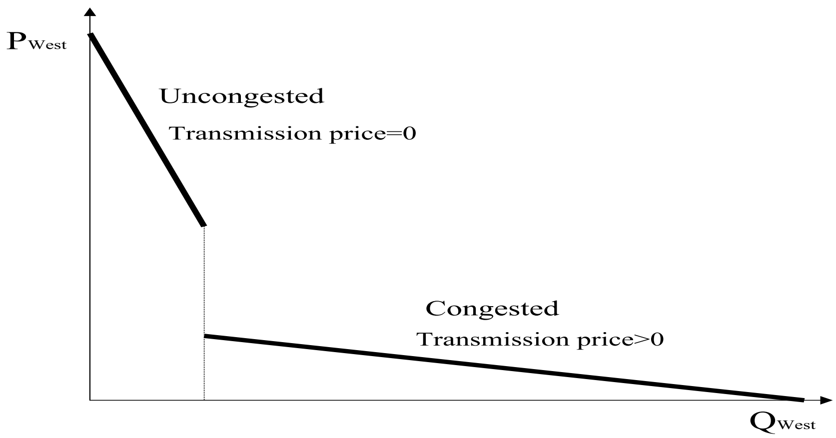

Figure 1: Residual Demand Curve

The residual demand curve facing the duopolists in the West is illustrated in Figure 3. The residual demand curve is steep over the range of quantities that would not congest the transmission line, $Q_{\mathsf{West}} < Q_{\mathsf{d}} + T_{\mathsf{k}} - F_{\mathsf{k}}$. The uncongested section has a steep slope because each unit produced by the duopolists will displace a unit of production from the high cost Central fringe and thus decrease the price by b. The jump in the residual demand curve represents the difference in the marginal cost of gas and coal units.[24] The congested section of the residual demand curve is less steep because once the line is congested any increase in production by the duopolists will displace a low cost coal unit and there is less variation in marginal costs across coal units. The residual demand curve in the West is as follows:

$$P^{\mathsf{west}}(Q) = \begin{cases} P^{\mathsf{c}}(Q) = Q_{\mathsf{d}} + T_{\mathsf{k}} - Q & if \quad Q > Q_{\mathsf{d}} + T_{\mathsf{k}} - F_{\mathsf{k}} \\ P^{\mathsf{uc}}(Q) = b(2Q_{\mathsf{d}} - F_{\mathsf{k}} - Q) & if \quad Q \leq Q_{\mathsf{d}} + T_{\mathsf{k}} - F_{\mathsf{k}} \end{cases}$$

The price of transmission is $P^{\mathsf{Transmission}} = P_{\mathsf{Central}} - P_{\mathsf{West}}$. If the transmission line is uncongested then $P_{\mathsf{Central}} = P_{\mathsf{West}}$ and the price of transmission is zero. If the line is congested, then the market clearing price in the Central zone will be the marginal cost of supplying $Q_{\mathsf{d}} - T_{\mathsf{k}}$ in the central zone, $mc^{\mathsf{c}}(Q_{\mathsf{d}} - T_{\mathsf{k}}) = b(Q_{\mathsf{d}} - T_{\mathsf{k}}) = P_{\mathsf{Central}}$. In this case, the price of transmission will be $P^{\mathsf{Transmission}} = b(Q_{\mathsf{d}} - T_{\mathsf{k}}) - P_{\mathsf{West}}$.

[24] The assumption that $mc^{\mathsf{w}}(F_{\mathsf{k}}) < mc^{\mathsf{c}}(Q_{\mathsf{d}} - T_{\mathsf{k}})$ guarantees that there will be a jump in the residual demand curve at the point at which the fringe run out of capacity.

4.1.2 The One Market Model

To solve the one market model, I will derive firm 1's best response function $q^{1*}(q^2)$. Since the two firms are symmetric, firm 2's reaction function will be the same as firm 1's. Also, since the firms are symmetric, any equilibrium in this game will be symmetric. Firm 1's profit function is as follows:

$$\Pi^1(q^1, q^2) = \begin{cases} \pi^{\mathsf{uc}}(q^1, q^2) = q^1 b(2Q_{\mathsf{d}} - F_{\mathsf{k}} - q^1 - q^2) & if \quad q^1 \leq Q_{\mathsf{d}} + T_{\mathsf{k}} - F_{\mathsf{k}} - q^2 \\ \pi^{\mathsf{c}}(q^1, q^2) = q^1(Q_{\mathsf{d}} + T_{\mathsf{k}} - q^1 - q^2) & if \quad q^1 > Q_{\mathsf{d}} + T_{\mathsf{k}} - F_{\mathsf{k}} - q^2 \end{cases}$$

$\pi^{\mathsf{uc}}(.,.)$ denotes firm 1's profit function conditional on the transmission line being uncongested. Similarly, $\pi^{\mathsf{c}}(.,.)$ refers to firm 1's profit function conditional on the line being congested.

Define $q^{1\mathsf{uc}}(.)$ and $q^{1\mathsf{c}}(.)$ to be firm 1's Cournot best response function conditional on the line being uncongested and congested, respectively. The status of the transmission line is determined by the quantity sold by the duopolists. For any given quantity produced by firm 2, q^2, firm 1 knows for each possible output level q^1 whether or not the line will be congested. Conditional on the line status $q^{1\mathsf{uc}}(.)$ and $q^{1\mathsf{c}}(.)$ emit the highest profits by definition. This implies that firm 1's best response function must correspond to either $q^{1\mathsf{uc}}(.)$ or $q^{1\mathsf{c}}(.)$ for all q^2.[25] This argument allows me to restrict my attention to the two conditional Cournot reaction functions.

Figure 4 illustrates firm 1's reaction function. The shaded area is the set of values for which $q^1 + q^2$ is not large enough to congest the transmission line. The two lines are the conditional Cournot best response functions, $q^{1\mathsf{c}}(.)$ and $q^{1\mathsf{uc}}(.)$. These functions are parallel with $q^{1\mathsf{c}}(q^2)$ lying above $q^{1\mathsf{uc}}(q^2)$.[26] For small values of q^2, $q^2 + q^{1\mathsf{c}}(q^2)$ may not be large

[25] Suppose this is not the case. Then, there exists some q^{2*} such that q^{1*} maximizes $\Pi^1(., q^{2*})$ and $q^{1*} \neq q^{1\mathsf{k}}(q^{2*})$ for $k = uc, c$. It must be the case that either $q^{1*} + q^{2*}$ is large enough to congest the line, or it is not. Suppose $q^{1*} + q^{2*} \geq Q_{\mathsf{d}} + T_{\mathsf{k}} - F_{\mathsf{k}}$, so that the line is congested. Then playing q^{1*} in response to q^{2*} must emit higher profits than the Cournot best response along the congested portion of the demand curve: $\pi^{\mathsf{c}}(q^{1\mathsf{c}}(q^{2*}), q^{2*}) < \pi^{\mathsf{c}}(q^{1*}, q^{2*})$. But this contradicts the definition of the Cournot best response function. $q^{1\mathsf{c}}(.)$ is defined such that $\pi^{\mathsf{c}}(q^{1\mathsf{c}}(q^2), q^2) \geq \pi^{\mathsf{c}}(q^1, q^2)$ for all q^1. The same argument holds if $q^{1*} + q^{2*}$ does not congest the line. Thus, $q^{1*}(.)$ must correspond to either $q^{1\mathsf{c}}(.)$ or $q^{1\mathsf{uc}}(.)$ for all q^2.

[26] The two reaction functions are parallel because I have assumed the duopolists have constant marginal cost.

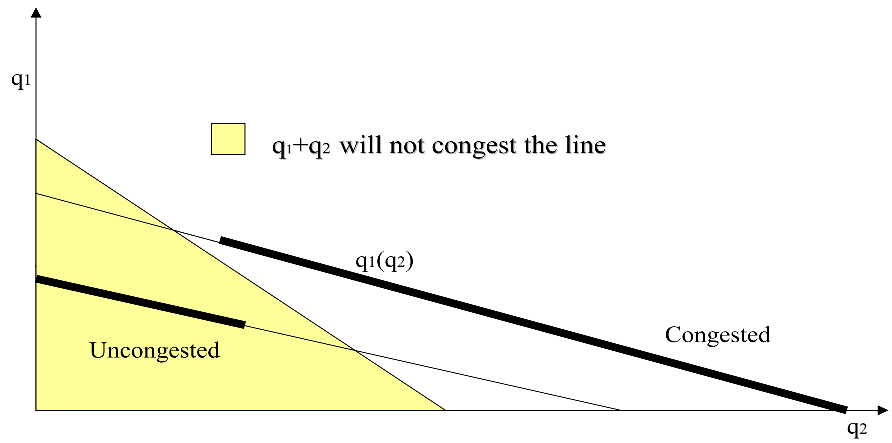

Figure 2: Firm 1's Congested and Uncongested Reaction functions

enough to congest the line. If this is the case, then it will not be feasible for firm 1 to play his congested best response function, $q^{1c}(q^2)$. In figure 4, this is illustrated by the portion of $q^c(.)$ which is in the shaded area. In this case, $q^{1*}(.)$ must correspond to firm 1's Cournot uncongested best response function $q^{1uc}(.)$. For very large values of q^2, $q^{1uc}(.)$ will not be feasible because $q^{1uc}(q^2) + q^2$ will congest the line. In figure 4, this is illustrated by the portion of the $q^{1uc}(.)$ that is not in the shaded area.

There are some values of q^2 for which both $q^{1uc}(.)$ and $q^{1c}(.)$ are feasible. These are the values of q^2 such that $q^{1uc}(q^2)$ is in the shaded area of Figure 4 while $q^{1c}(q^2)$ is in the unshaded area. For values of q^2 in this range, $q^{1*}(.)$ will correspond to the condition best response function which emits the highest profits. In the technical appendix, I show that firms 1's reaction function is at most discontinuous at one point and at the one point of discontinuity, the reaction function jumps from the uncongested Cournot reaction function to the congested reaction function. Since the congested Cournot reaction function lies above the uncongested Cournot reaction function, this will be a jump up to a higher production quantity for firm 1.

Any pure strategy equilibrium in this game must be symmetric.[27] The pure strategy

[27] In any asymmetric equilibrium, the firms would be on different conditional Cournot best response functions. This cannot be possible because either the total quantity produced congests the line or it does not. This implies that in any asymmetric equilibrium one of the firms would be playing a reaction function which was not feasible.

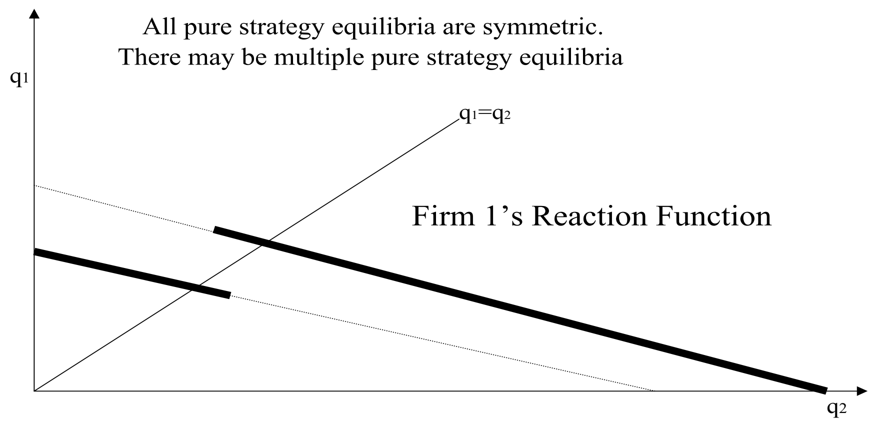

Figure 3: Pure Strategy Equilibrium in the One Market Game

equilibrium quantities can be found by intersecting firm 1's reaction function with the 45% line. Since the only jump in the reaction function is a jump up, the 45% line will intersect the reaction function which means an equilibrium exists. As is illustrated in Figure 5, there may be multiple pure strategy equilibria.

If the transmission line is large relative to the size of fringe, the duopolists will withhold enough so that in the unique equilibrium the transmission line will be uncongested and the market clearing price will be set by the high cost Central fringe. If the size of the line is small relative to the size of the Western fringe, the duopolists will not find it profitable to withhold enough to uncongest the line. In this case, there will be a unique equilibria in which the transmission line is congested. For intermediate size transmission lines relative to the size of the fringe, there will be three equilibria in the game; two pure strategy equilibria, one in which the line is uncongested and one in which it is congested, and a mixed strategy equilibrium.

The pure strategy equilibria in this game are defined by the same prices and quantities as the Cournot game conditional on the status of the transmission line. This is because the pure strategy equilibria are defined by the intersection of the conditional Cournot best response functions. If the transmission line is uncongested, then the transmission price will be zero. If the transmission line is congested, then the transmission price will be the congested price in the Central zone, which is always $b(Q_{\mathsf{d}} - T_{\mathsf{k}})$ minus the equilibrium price in the West.

4.2 The Two Market Game with No Speculators

In order to analyze the model in which there is both a DA and RT market for electricity, I first describe the bidding strategies of each of the market participants. In what follows, I will define the transmission line to be congested in the DA market if the market clearing quantities from the DA market would result in the line being congested in RT. This is the definition of DA congestion used by the NYISO.

4.2.1 Demand and Fringe Bidding Strategies

Before the virtual bidding policy, the utilities in New York were buying almost all of their power in the higher priced DA market. Some of the utilities have contracts that are tied to the DA market which makes purchasing DA their risk minimizing strategy. I incorporate the utilities' risk aversion into the model by assuming that the total quantity demanded by retail consumers, $2Q_{\mathsf{d}}$, is purchased in the DA market.

The fringe generators at both nodes are assumed to bid into the DA market at the maximum of marginal cost and the expected RT price. If the RT price is less than the DA price, $P_{\mathsf{rt}} < P_{\mathsf{da}}$, any fringe generator that sold DA and has a marginal cost, mc, greater than the RT price, $P_{\mathsf{rt}} < mc \leq P_{\mathsf{da}}$, will buy back electricity in the RT market. This bidding strategy guarantees that fringe generators will receive the highest possible price for any electricity sold and a fringe generator will never produce if the RT price is less than its marginal cost.

Under the assumed bidding strategy, each fringe generator offers to sell electricity in the DA market at its marginal cost. If the RT price is greater than the DA price, $P_{\mathsf{da}} < P_{\mathsf{rt}}$, the generator's marginal cost is the opportunity cost of selling in the RT market, P_{rt}. A fringe generator with $P_{\mathsf{rt}} < mc \leq P_{\mathsf{da}}$ that sells one unit in the DA market will earn $P_{\mathsf{da}} - P_{\mathsf{rt}}$ as an arbitrage profit. Since the fringe generators earn these arbitrage profits by bidding their marginal cost into the DA market, the profits will be denoted as passive arbitrage profits. The fringe generators do not actively arbitrage DA and RT price differences because it was against the market rules before the virtual bidding policy.

Since all retail demand is purchased in the DA market, the only players in the RT

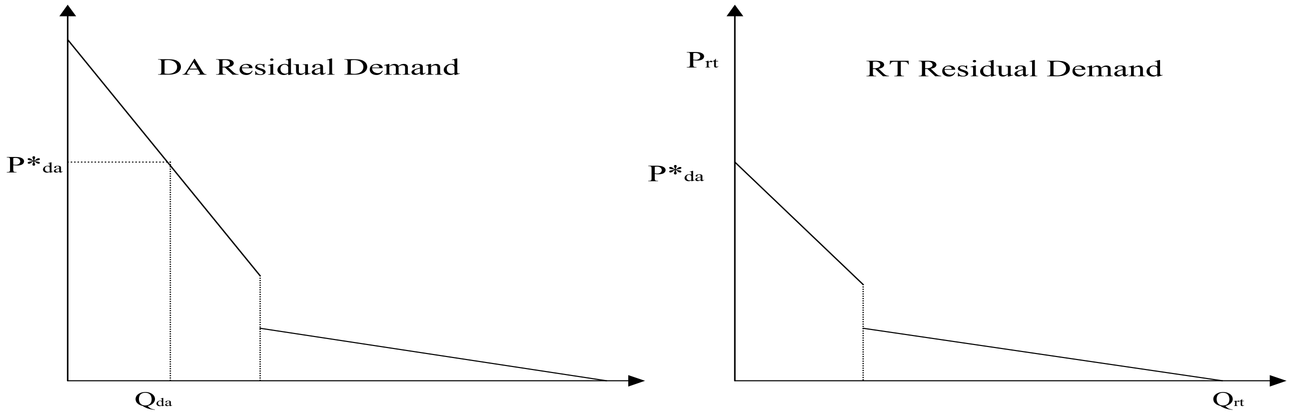

Figure 4: DA and RT Demand Curves

market are the fringe generators and the duopolists. If $P_{\mathsf{rt}} > P_{\mathsf{da}}$, the fringe will sell electricity in the RT market. Any electricity sold by the fringe in RT must be purchased by the duopolists. This cannot occur in equilibrium, because the duopolists would earn negative RT profits. This implies that in any subgame perfect equilibrium the DA price must be at least as high as the RT price, $P_{\mathsf{da}} \geq P_{\mathsf{rt}}$, and the fringe buy in RT.

Denote total DA and RT duopoly sales by $Q_{\mathsf{da}} = q^1_{\mathsf{da}} + q^2_{\mathsf{da}}$ and $Q_{\mathsf{rt}} = q^1_{\mathsf{rt}} + q^2_{\mathsf{rt}}$, respectively. In the case of no speculators, the inverse residual demand curves facing the duopolists at the West zone are as follows:

$$P^{\mathsf{west}}_{\mathsf{da}}(Q_{\mathsf{da}}) = \begin{cases} P^{\mathsf{c}}_{\mathsf{da}}(Q_{\mathsf{da}}) = Q_{\mathsf{d}} + T_{\mathsf{k}} - Q_{\mathsf{da}} & if \quad Q_{\mathsf{da}} > Q_{\mathsf{d}} + T_{\mathsf{k}} - F_{\mathsf{k}} \\ P^{\mathsf{uc}}_{\mathsf{da}}(Q_{\mathsf{da}}) = b(2Q_{\mathsf{d}} - F_{\mathsf{k}} - Q_{\mathsf{da}}) & if \quad Q_{\mathsf{da}} \leq Q_{\mathsf{d}} + T_{\mathsf{k}} - F_{\mathsf{k}} \end{cases}$$

$$P^{\mathsf{west}}_{\mathsf{rt}}(Q_{\mathsf{rt}}, Q_{\mathsf{da}}) = \begin{cases} P^{\mathsf{c}}_{\mathsf{rt}}(Q_{\mathsf{rt}}, Q_{\mathsf{da}}) = Q_{\mathsf{d}} + T_{\mathsf{k}} - Q_{\mathsf{da}} - Q_{\mathsf{rt}} & if \quad Q_{\mathsf{rt}} > Q_{\mathsf{d}} + T_{\mathsf{k}} - F_{\mathsf{k}} - Q_{\mathsf{da}} \\ P^{\mathsf{uc}}_{\mathsf{rt}}(Q_{\mathsf{rt}}, Q_{\mathsf{da}}) = b(2Q_{\mathsf{d}} - F_{\mathsf{k}} - Q_{\mathsf{da}} - Q_{\mathsf{rt}}) & if \quad Q_{\mathsf{rt}} \leq Q_{\mathsf{d}} + T_{\mathsf{k}} - F_{\mathsf{k}} - Q_{\mathsf{da}} \end{cases}$$

Figure 6 illustrates the DA and RT residual demand curves. The DA residual demand curve is the same as the residual demand curve in the one market game. The RT demand curve is the DA demand curve shifted to the left by the quantity sold in the DA market, Q_{da}. Because there are no speculators in this model, the duopolists are able to sell Q_{da} in the DA market and then sell more in the spot market at a lower price.

4.2.2 The Spot Market Subgame

The model is solved backwards by first solving for the spot market equilibria as a function of DA sales and then solving the entire model. Like in the one market model, in order to solve for the spot market equilibrium, I first characterize each firm's best response function. Firm 1's spot market profit function is as follows:[28]

$$\Pi^1_{\mathsf{rt}}(q^1_{\mathsf{rt}}, q^2_{\mathsf{rt}}, Q_{\mathsf{da}}) = \begin{cases} \pi^{\mathsf{uc}}_{\mathsf{rt}}(q^1_{\mathsf{rt}}, q^2_{\mathsf{rt}}) = q^1_{\mathsf{rt}} b(2Q_{\mathsf{d}} - F_{\mathsf{k}} - Q_{\mathsf{da}} - q^1_{\mathsf{rt}} - q^2_{\mathsf{rt}}) & if \quad q^1_{\mathsf{rt}} \leq Q_{\mathsf{d}} + T_{\mathsf{k}} - F_{\mathsf{k}} - Q_{\mathsf{da}} - q^2_{\mathsf{rt}} \\ \pi^{\mathsf{c}}_{\mathsf{rt}}(q^1_{\mathsf{rt}}, q^2_{\mathsf{rt}}) = q^1_{\mathsf{rt}}(Q_{\mathsf{d}} + T_{\mathsf{k}} - Q_{\mathsf{da}} - q^1_{\mathsf{rt}} - q^2_{\mathsf{rt}}) & if \quad q^1_{\mathsf{rt}} > Q_{\mathsf{d}} + T_{\mathsf{k}} - F_{\mathsf{k}} - Q_{\mathsf{da}} - q^2_{\mathsf{rt}} \end{cases}$$

If the transmission line is congested in the DA market, $\pi^{\mathsf{c}}_{\mathsf{rt}}$, the profit function conditional on congestion, will be the only relevant portion of the profit function.[29] This implies that firm 1's reaction function will correspond to the Cournot reaction function derived from $\pi^{\mathsf{c}}_{\mathsf{rt}}$. Denote this reaction function as $q^{1\mathsf{c}}_{\mathsf{rt}}(.)$, where $q^{1\mathsf{c}}_{\mathsf{rt}}(q^2_{\mathsf{rt}}) = argmax_{q^1_{\mathsf{rt}}} \pi^{\mathsf{c}}_{\mathsf{rt}}(q^1_{\mathsf{rt}}, q^2_{\mathsf{rt}})$. Similarly define $q^{1\mathsf{uc}}_{\mathsf{rt}}(q^2_{\mathsf{rt}})$, to be the Cournot best response function of firm 1 along $\pi^{\mathsf{uc}}_{\mathsf{rt}}(., q^2_{\mathsf{rt}})$. Like in the one market case, these two functions are parallel to one another with $q^{1\mathsf{c}}_{\mathsf{rt}}(.)$ lying above $q^{1\mathsf{uc}}_{\mathsf{rt}}(.)$.

If the transmission line is congested in the DA market, then the spot market reaction function will correspond to $q^{1\mathsf{c}}_{\mathsf{rt}}(.)$ for all values of q^2_{rt}. If the DA market is uncongested, then firm 1's spot market reaction function will correspond to $q^{1\mathsf{uc}}_{\mathsf{rt}}(.)$ for small values of q^2_{rt} and $q^{1\mathsf{c}}_{\mathsf{rt}}(.)$ for larger values. As was the case in the one market game, if there is a point of discontinuity in firm 1's spot market reaction function it will be a jump from the uncongested conditional reaction function to the congested conditional reaction function.

A pure strategy equilibrium is guaranteed to exists, but it may not be unique. In particular, the three possible cases outlined in the one market game can occur in the spot market subgame. There may be a unique equilibrium in which the transmission line connecting the two zones is either congested or uncongested. Alternatively, there may be three equilibria, two pure strategy, one in which the line is congested and one in which it is not, and a mixed strategy equilibrium.

In the one market game, the equilibrium was determined by the capacity of the western fringe relative to the capacity of the transmission line. If the transmission line was

[28] Firm 1's spot market profits are symmetric in the quantity sold DA by firm 1, q^1_{da}, and firm 2, q^1_{da}. This is because firm 1's RT sales, q^1_{rt}, will not affect the revenue firm 1 receives on the DA sales q^1_{da}. This implies that the quantity sold DA is not infra-marginal in the RT market. Firm 1's profit function does depend on the total quantity sold in the DA market, Q_{da}, because that quantity determines the shift in the spot market residual demand curve.

[29] $\pi^{\mathsf{uc}}_{\mathsf{rt}}$ would only be relevant if $q^2_{\mathsf{rt}} < 0$. This cannot occur because firm 2 would be earning negative spot market profits.

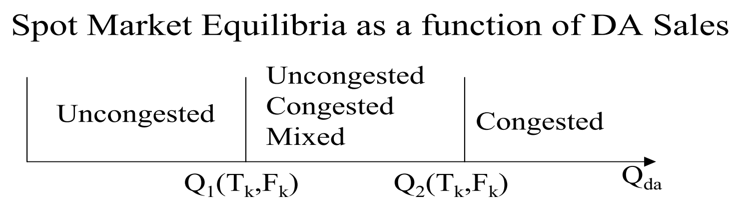

Figure 5: DA and RT Demand Curves

large relative to the capacity of the fringe, then the transmission line would be uncongested in the one market model. In the spot market subgame, the quantity sold in the DA market also affects which equilibria will occur. For given values of the fringe capacity, F_{k}, and the size of the line, T_{k}, if the total quantity sold in the DA market, Q_{da}, is very small then there may be a unique equilibrium in the spot market in which the line will be uncongested. For a larger Q_{da}, all three equilibria may be possible. If Q_{da} is large enough to congest the line, then in the unique spot market equilibria the line will be congested. Figure 6 illustrates the possible equilibria as a function of the total quantity sold DA. The cut off points for the possible equilibria, $Q_1(T_k, F_k)$ and $Q_2(T_k, F_k)$, are determined by the parameter values of the model. $Q_1(T_k, F_k)$ and $Q_2(T_k, F_k)$ are decreasing in the size of the fringe, F_k and increasing in the size of the transmission line, T_k. It is shown in the appendix that $Q_1(T_k, F_k) < Q_2(T_k, F_k)$.

4.2.3 The DA Market Absent Speculators

A pure strategy equilibrium was guaranteed to exist in the spot market subgame because at the only point of discontinuity in the reaction function there was a jump up to a higher output level. The reaction function for the collapsed game is not as nicely behaved. In the technical appendix, I present the possible pure strategy equilibria in the game and give necessary conditions for each to be an equilibria. I also show in the appendix that each possible pure strategy equilibrium is the unique equilibrium for some set of parameter values. In what follows, I describe the possible equilibria and give intuition for the circumstances under which each may occur.

Let $\pi^1_{c/c}$ denote firm 1's total profits in the game conditional on the transmission line being congested in both the DA and RT markets. Similarly let $\pi^1_{uc/c}$ and $\pi^1_{uc/uc}$ denote the

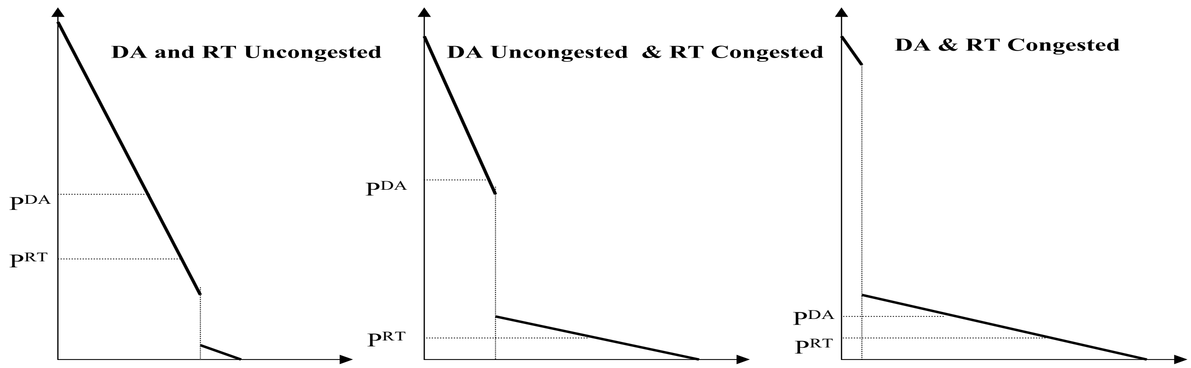

Figure 6: Equilibria in Game with No Speculators

profits of firm 1 if the transmission line is uncongested DA/congested in RT and uncongested in both markets, respectively.[30] As in the spot market, the only possible pure strategy equilibria are symmetric and each corresponds to a line status $(c/c, uc/c, uc/uc)$.[31] These possible equilibria are obtained by solving the model conditional on the line status. For very large demand levels relative to the capacity of the fringe, there will be a unique equilibrium in which the transmission line is uncongested in both the DA and RT market. This means that the duopolists sell some quantity Q_{da} in the DA market. In the RT market, they sell more, but the total quantity sold by the duopolists, $Q_{\mathsf{da}} + Q_{\mathsf{rt}}$, will be less than $Q_{\mathsf{d}} + T_{\mathsf{k}} - F_{\mathsf{k}}$, the quantity that is necessary to congest the line. For small demand levels, it will not be profitable for the duopolists to withhold enough to leave the line uncongested in the DA market. In this case, the line will be congested in both the DA and RT market. For some intermediate demand levels, there will a unique equilibrium in which the DA trades will leave the transmission line uncongested and the increased RT trades will congest the line. Figure 8 illustrates the type of demand curve which corresponds to each pure strategy equilibrium.

I will use the status of the line as shorthand for each of the equilibria illustrated in Figure 8. I will refer to the equilibrium in which the line is congested in both the DA and RT markets as C/C. Similarly, I will refer to the equilibria in which the transmission line

[30] Since $\pi^{1}_{\mathsf{c/uc}}$ cannot occur in equilibrium, I leave it out of the following analysis

[31] The argument for symmetric pure strategy equilibrium is the same as that in the spot market sub game. Any asymmetric equilibrium would require the two firms to be on different sections of the reaction function. But, this would imply that one of the firms is violating the congestion constraints of the reaction function.

is uncongested DA and congested in RT and uncongested in both markets as UC/C and UC/UC respectively.

The UC/C equilibrium is the equilibrium of most interest for a couple of reasons. First, the predictions on the price of electricity and transmission are consistent with the results reported in the previous section for some hours. In this equilibrium, the difference between the DA and RT prices may be quite large. The DA market clearing price is set by the marginal cost of the high cost Central fringe. In the RT market, the low cost Western fringe will set the price. Since the transmission line is uncongested DA, the DA price of transmission will be zero while the RT price of transmission will be greater than zero. From a policy point of view, the equilibrium is interesting because the firms with market power are able to extract the rents from the transmission owners. Transmission rents are only tied to the DA market. By keeping the transmission line uncongested in the DA market, the firms with market power earn the shadow value of additional capacity on the line.

4.3 The Addition of Speculators

To analyze the effect of the virtual bidding policy, I add speculators to the model. Competitive speculators are assumed to compete away any differences between the DA and RT prices. This implies that speculators will impose a no-arbitrage constraint that the DA price must equal the RT price: $P_{\mathsf{da}} = P_{\mathsf{rt}} = P^*$. Speculators that require the DA price in each zone to equal the RT price in that zone will also force the DA price of transmission to be equal to the RT price of transmission which implies that the transmission line will have the same congestion status in both markets.[32]

The spot market subgame in this model will be the same as in the model without speculators. In the model with no speculators, the duopolists were able to price discriminate between the DA and RT markets. By arbitrating away any DA and RT price differences, speculators prevent the duopolists from price discriminating. The equilibrium price in the

[32] If there were uncertainty in the model, then the congestion status would not necessarily be the same. Rather, the DA market would be congested whenever there was a positive probability of RT congestion, the DA price would equal the expected RT price, and the price of transmission would be equal to the RT price of transmission.

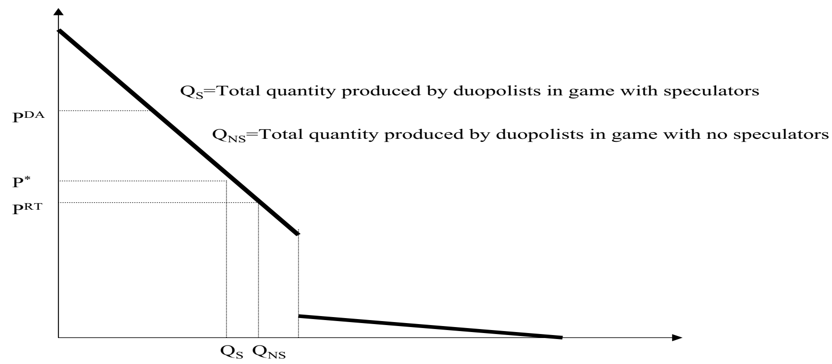

Figure 7: Equilibrium in Game with and without speculators

game with speculators, P^*, will be less than the DA price absent speculators and greater than the RT price. This implies that the total production by the duopolists will decrease. This is similar to the case of a perfect price discriminating monopolist. If the monopolists is prevented from price discriminating, the uniform price will be between the highest and lowest prices charged when price discriminating. The quantity produced by monopolist will be less under uniform pricing than under perfect price discrimination and dead weight loss will be higher. This is true regardless of which equilibrium existed in the model without speculators.

If the transmission line was either congested or uncongested in both the DA and RT markets before the introduction of speculators, then it will have the same status once speculators enter the market. However, the total quantity sold by the duopolists will be less than in the game absent speculators. Figure 9 illustrates the equilibrium prices and total duopoly production with and without speculators for a case in which the equilibrium is UC/UC absent speculators.

In the appendix, it is shown that the average procurement costs of electricity will also decrease regardless of the pre speculators equilibrium. Procurement costs decrease because absent speculators the majority of trades occur in the higher priced DA market. Although more is sold at the lower RT price, the quantity sold is small relative to the DA quantity.

Speculators prevent an equilibrium in which the line is uncongested DA and congested in RT. For parameter values in which, absent speculators, the equilibrium was UC/C, once speculators enter the market the transmission line will either be congested or uncongested in both markets. In either case, the total quantity produced by the duopolists will decrease and so will the average procurement costs of electricity. If the market is congested once speculators are added and there is a large cost difference between the West and Central fringe, the decrease in average procurement cost will be large.

Consider an example in which $T_{\mathsf{k}} = \frac{2}{3}Q_{\mathsf{d}}$, $F_{\mathsf{K}} = \frac{1}{2}Q_{\mathsf{d}}$ and $b = 2$. In this example if there were no futures market, the equilibrium production by the duopolists would be Q_{d} and the transmission line will not be congested. In the two market model without speculators, there is a unique equilibrium in which the transmission line is uncongested in the DA market and congested in RT. The total production by the duopolists will be $1.43Q_{\mathsf{d}}$. The DA and RT prices will be $1.2Q_{\mathsf{d}}$ and $0.24Q_{\mathsf{d}}$, respectively. Once speculators are added, in the unique equilibrium, the transmission line will be congested in both markets, the total production by the duopolists will decrease to $1.33Q_{\mathsf{d}}$ and the market price will be $0.33Q_{\mathsf{d}}$. The addition of speculators decreases the average procurement cost of electricity from $0.77Q_{\mathsf{d}}$ to $0.33Q_{\mathsf{d}}$.

This example illustrate the general results of the model. The existence of a futures market increases the production of Cournot duopolists. Allowing speculators to trade in the market will decrease the quantity produced by the firms with market power, but this decrease will be small relative to the increase which results from the futures market. Once speculators are added, the DA margins of firms with market power decrease. If, absent speculators, the line was uncongested in the DA market and congested in the RT market, then the decrease in margins may be large. The average procurement cost of electricity should decrease in all cases, and the decreases may be quite large if before speculators enter the transmission line was uncongested in the DA and congested in RT.

5 The Virtual Bidding Policy and Market Power

In this section, I estimate the change in DA and RT margins that has occured since the implementation of the virtual bidding policy. The model presented in Section 4 of this paper, offers two testable implications. First, the model predicts that the DA margins of firms in the exporting zone should decrease while there should not be a large effect on RT margins. Second, the DA margins of firms in the exporting zone should decrease more than those of firms in the importing zone. To test this second prediction, I compare the effect of the policy on the price cost margins of Dynegy, the largest firm in the Central (importing) zone, to those of NRG and AES the two large firms in the Western (exporting) zone.

I use publicly available, detailed data on each thermal unit's cost function and hourly production to estimate hourly firm level price-cost margins. After controlling for observable changes in market conditions, I estimate the change in each firm's margin that has occured since the virtual bidding policy was implemented. I will consider the policy to have had an effect on market power, if for a given residual demand curve, the equilibrium price-cost margins are different under the policy. The residual demand curve facing the duopolists in any given hour is not known. Therefore,I will attempt to control for changes in the residual demand when estimating the effect of the virtual bidding policy. The results of this event study will be sensitive to other un-observable market changes that occured around the time that the virtual bidding policy was implemented. When testing the price discrimination model, I control for observable changes directly and for unobservable changes in market conditions by using Dynegy as a control for the two Western firms, NRG and AES.

The incentive a firm has to withhold production in order to increase the spot price is a function of the firm's contract position. If a firm has a contract to supply Q_{c} at a fixed price P_{c}, then Q_{c} are not part of the firm's infra-marginal quantity in the spot market. Since the firm has less infra-marginal quantity in the spot market, it has more of an incentive to increase production. This implies that, for a given production capacity, a firm's total production should be increasing in the quantity it has contracted forward. Thus, a firm's price cost margin is decreasing in the quantity it has contracted forward. This means that if the firms in New York signed more contracts around the time that the virtual bidding policy was implemented, then their DA margins would also decrease.

Around the period that the virtual bidding policy took place many changes were occuring in both the economy as a whole and the energy sector. The overall downturn in the economy that occured around this time may have lead firms to sign more long term contracts. Alternatively, and possibly more plausibly, the Enron collapse which began to enter the news in October of 2001 may have changed the liquidity and risk attitudes of firms in the energy industry. There is anecdotal evidence that energy firms in general felt a bit of a fallout from the Enron collapse. In particular, Dynegy took out a full page ad in the *Wall Street Journal* claiming that Dynegy was different than Enron because it had 'real assets.' This ad provides some evidence that the firms in New York felt they needed to protect themselves from any Enron fallout. One way in which they could have sought protection would be to sign long term contracts for power produced in the New York market. This would have provided them with a more certain revenue stream.

I have looked at many financial filings and at the industry press and have not found evidence that the firms in New York changed their contract positions around the time of the virtual bidding policy, but I am not able to rule the possibility out. However, the effect that a change in contract position should have on margins is different than the effect that the price discrimination model predicts the virtual bidding policy will have on margins. If firms sign more long term contracts, then their total production should increase and margins in both the DA and RT markets should decrease. This effect is different than the predictions of the price discrimination model that DA margins should decrease while the RT margins should not.[33]

I estimate the marginal cost of each thermal unit by using engineering data. The marginal cost of unit i is assumed to be constant up to the unit's capacity, cap_i. The following formula is used to estimate the marginal cost of firm f's generation unit i at time t.

$$mc_{it}^{f} = heatrate_i * (P_{it}^{fuel} + P_t^{NOx} RateNOx_i + P_t^{SO_2} RateSO_{2i}) + VOM_i,$$

where $heatrate_i$ is a unit-specific measure of the efficiency with which the unit transforms fuel into electricity. P_{it}^{fuel} is the price at time t of the type of fuel used by unit i. Generation units in New York participate in two pollution permit programs. The Acid Rain Program requires units with a capacity greater than 25 MWs to acquire a permit for each ton of SO_2 the unit releases. Similarly, the Ozone Transport Commission requires a permit for each ton of NO_x released during the summer months. P_t^{NOx} and $P_t^{SO_2}$ are the average monthly prices for NOx and SO_2 pollution permits. $RateNOx_i$ and $RateSO_{2i}$ are the average quantity of pollution unit i releases per unit of fuel input. VOM_i is the variable operating and

[33]The model actual predicts that RT margins should increase, but this increase is small relative to the decrease in DA margins and thus, would probably not be observable in the data.

maintenance cost of the unit. As a result of years of regulatory oversight, all of the data used in this calculation are publicly available.[34]

Hourly production data are available at the generation unit level from the EPA's Continuous Emissions Monitoring Systems (CEMS). I combine these hourly production data with the marginal cost estimates above to calculate each firm's hourly margin. Following Puller [2003], I define firm f's DA (RT) margin at time t to be the DA (RT) zonal price less the marginal cost of the firm's highest cost unit which is currently operating with excess capacity; $Margin^{da}_{ft} = P^{da}_{zt} - max_i(mc^{f}_{it}(q_{it})|0 < q_{it} < cap_i)$ and $Margin^{rt}_{ft} = P^{rt}_{zt} - max_i(mc^{f}_{it}(q_{it})|0 < q_{it} < cap_i)$. If a firm is operating its highest cost unit that is currently running at full capacity, then I define the firm's margin in that hour to be zero. This definition of margin recognizes that there are significant startup costs to operating a generation unit.[35]

The largest generation firm in Western New York, NRG, filed for bankruptcy protection on June 30, 2002. In other industries, researches have found evidence that the financial condition of firms may affect how they compete in the product market.[36] I would like to separate any changes in behavior that have occurred as a result of financial distress from changes that have resulted from the virtual bidding policy. I am currently using each firm's average stock price from the previous month as a control for the firm's financial condition. I use the lagged stock price because it reflects changes in expectations about the future profitability of the firm. These changes in the expected future profits should be correlated with a firm's access to new debt and risk of defaulting on old debt. I am currently investigating other possible controls for the firm's financial condition that have at least monthly variation.

[34]For a detailed discussion of the data used to estimate marginal, see the Data Appendix in Mansur [2003].

[35]The cost of turning on a generation unit is many times the marginal cost described above. Given these cost non-convexities, when making production decisions firms solve a dynamic optimization problem. If firms behave competitively, then they will turn on a unit if its expected revenues will cover the startup cost. Ignoring these non-convexities will only affect my results if firms solve the dynamic problem differently before and after the introduction of virtual bidding.

[36]See Busse [2000], Chevalier[1995] and Chevalier and Scharfstein [1996].

The stock price of a firm will depend on its expected future profits. If financial markets had perfect foresight, then a firm's price-cost margin in any hour would be incorporated into the firm's lagged stock price. Although financial markets do not have perfect foresight, it still may be true that a firms price cost margins is correlated with its lagged stock price. To correct for this potential endogeneity, I use the previous month's average S&P 500 index as an instrument.

The terrorist attacks of September 11, 2001 occurred less than two months before the virtual bidding policy was implemented. Even if the attacks did not change the business practices of firms operating in the New York market, they did change the relative demand of electricity throughout the state. I control for this by including a dummy and a time trend for the period after the attacks.

I estimate the following model separately for each of the three firms:

$$Margin^{\mathsf{k}}_{\mathsf{ft}} = \alpha VB_{\mathsf{t}} + \gamma_1 Trend_{\mathsf{t}} + \gamma_2 Sept11_{\mathsf{t}} + \gamma_3 Sept11Trend_{\mathsf{t}} + \gamma_4 StockPrice_{\mathsf{ft}} + \Omega Market_{\mathsf{t}} + \epsilon_{\mathsf{ft}},$$

where $Margin^{\mathsf{k}}_{\mathsf{ft}}$ is defined as the zonal price at time t less firm f's marginal cost at time t in market k, for k =DA, RT. VB_{t} is a dummy variable for the period after the policy change. $Trend_{\mathsf{t}}$ is a time trend which is normalized to be between zero and one. $Sept11_{\mathsf{t}}$ is a dummy for the period after the terrorist attacks. $Sept11Trend_{\mathsf{t}}$ is a time trend for the period after the attacks which ranges from zero to one. $StockPrice_{\mathsf{ft}}$ is firm $f's$ average stock price in the previous month.

$Market_{\mathsf{t}}$ is a matrix of controls for market conditions. To test if the virtual bidding policy has had an effect on price-cost margins, I need to control for changes in the residual demand facing firms in New York. Residual demand is a function of the quantity demanded and the marginal cost of the competitive fringe. For the DA model, $Market_{\mathsf{t}}$ includes zonal forecasted and log forecasted demand to control for the level of demand. To control for the marginal cost of fringe units, $Market_{\mathsf{t}}$ also includes the price of natural gas.[37] To control for seasonality, $Market_{\mathsf{t}}$ includes month of year and hour of day dummies. In the RT model, $Market_{\mathsf{t}}$ also includes the zonal and log zonal realized demand.[38] I include both

[37]All of the units owned by Dynegy and AES use either fuel oil or coal as the primary fuel. NRG has one plant which uses natural gas as its primary fuel, but it was not operational during this time period.

[38]Since end use consumers do not respond to the wholesale price of electricity, the actual demand in the

forecast and actual demand in the RT model because deviations from forecasts may affect the RT prices.[39] I instrument $Stockprice_{ft}$ using the previous month's average level of the S&P index. I report three sets of results for each firm in each market (DA and RT). Since a firm's stock price may not a good proxy for financial distress, I first present the OLS estimates which exclude the variable to test the robustness of the results. The second set of results are the OLS with stock price and the third are the IV results. The results for the DA margins are reported in Tables 6, 7 and 8.

For the case when the DA margin is the dependent variable, in the two OLS specifications, the VB coefficient for AES is around -3.50. This means that after controlling for factors that affect residual demand, AES's average DA margin decreased by $3.50 in the period after the policy was implemented. This decrease is almost 10% of the DA zonal price during the post-virtual bidding period. In the IV specification, the predicted decrease in AES's average DA margin after the policy change is $6.28. NRG's average margin decreases by almost $5.00 in the specification without stock price and $7 to $9 in the specifications which include stock price. In all three specifications for Dynegy, the large Central firm, the coefficient on VB is positive but not significant. In the IV specification, the point estimate implies that Dynegy's DA margins increased by $2.25, but this estimate is extremely noisy with a standard error of 2.97.

The results for the RT market are presented in Tables 9, 10 and 11. In all three specifications for AES, the VB coefficient is negative but the estimates are very noisy. All of the point estimates for NRG are also negative. In the two specifications which include the stock price, the coefficients on VB are significant for NRG. In the OLS specification which includes lagged stock price, the coefficient on VB suggests that NRG's average RT margin decreased by $3.73 while the IV specification implies a decrease of more than five dollars. In the RT results, the point estimates for Dynegy are all positive but none is significant.

The price discrimination model predicted that the virtual bidding policy should result in a decrease in DA margins and not have a large effect on RT margins. I find that the DA and RT margins decreased for both of the western firms. However, it appears as though

system will not be endogenous.

[39]Many lower costs generation units require a longer time to start. If demand is significantly higher than forecasted, higher cost units with faster start times may have to produce.

the DA margins decreased more than the RT margins. I explicitely test this prediction of the model by stacking the DA and RT margins and then re-estimate the previous models. I allow for separate DA and RT coefficients for each of the regressors. I estimate the following model separately for each of the three firms.

$$Margin_{\mathsf{kft}} = \alpha_{\mathsf{da}} DA\_VB_{\mathsf{kt}} + \Omega_{\mathsf{da}} DA\_Controls_{\mathsf{kt}} + \alpha_{\mathsf{rt}} RT\_VB_{\mathsf{kt}} + \Omega_{\mathsf{rt}} RT\_Controls_{\mathsf{kt}} + \epsilon_{\mathsf{ft}},$$

DA_VB_{kt} is a dummy variable which is equal to one if the margin is for the DA market and it is the post-virtual bidding period. $DA\_Controls_{\mathsf{kt}}$ is a matrix of all of the controls from the previous model interacted with a dummy for the DA market. RT_VB_{kt} and $RT\_Controls_{\mathsf{kt}}$ are defined in the same manner except each variable is interacted with a dummy for the RT market. After estimating this model, I test the hypothesis that the change in the DA margins was equal to the change in the RT margins. This test is a t-test that $\alpha_{\mathsf{da}} = \alpha_{\mathsf{rt}}$. The p-values from this test are presented in the first three rows of Table 5. The three columns of results correspond to the three models: OLS excluding the regressor *stockprice* (OLS (1)), OLS including *stockprice* (OLS(2)), and IV.

The p-values for the test that the post-virtual bidding change in AES's average DA margin is equal to the change in its average RT margins range from 0.08 to 0.16. These tests suggest that although AES's DA and RT margins both decreased, the firm's DA margins decreased significantly more. For NRG, the p-values range from 0.06 to 0.23, again providing evidence that the DA margins decreased more than the RT margins. For the large Central firm, Dynegy, the p-values of the test range from 0.50 to 0.91. I cannot reject the hypothesis that there has been no change in the relationship between Dynegy's DA and RT margins.

The price discrimination model presented in this paper predicts that the DA margins of the West firms should decrease when compared to those of the Central firms. The previous results suggest that this may be the case. I test this hypothesis explicitly by estimating the following model separately for AES and NRG.

$$MarginDiff^{\mathsf{k}}_{\mathsf{ft}} = \alpha VB_{\mathsf{t}} + \gamma_1 Trend_{\mathsf{t}} + \gamma_2 Sept11_{\mathsf{t}} + \gamma_3 Sept11Trend_{\mathsf{t}} + \gamma_4 StockPrice_{\mathsf{ft}} + \Omega Market_{\mathsf{t}} + \epsilon_{\mathsf{ft}}$$

$MarginDiff^{\mathsf{k}}_{\mathsf{ft}} = Margin^{\mathsf{k}}_{\mathsf{ft}} - Margin^{\mathsf{k}}_{\mathsf{Dynegyt}}$ where f refers to firms AES and NRG and $Margin^{\mathsf{k}}_{\mathsf{Dynegyt}}$ is the margin of the Central firm Dynegy. The results from this model are presented in Tables 12 through 15. The DA margins of both Western firms, AES and

Table 5: p-values of Test:

Post-VB ΔDA Margins=Post-VB ΔRT Margins

Firm	OLS (1)	OLS (2)	IV
AES levels	0.16	0.14	0.08
Dynegy levels	0.91	0.50	0.87
NRG levels	0.23	0.06	0.09
AES-Dynegy	0.09	0.19	0.10
NRG-Dynegy	0.17	0.03	0.09

NRG, decreased significantly more than those of Dynegy. The results imply that AES's DA margins decrease by $4.00 more than Dynegy's. The NRG point estimates are even larger, suggesting that NRG's DA margins decreased at least $5.00 more than Dynegy's. In the RT margin results, the point estimates on VB are negative in all specifications for both NRG and AES. None of the AES results are significant. In the IV specification, the coefficient on VB is negative and significant in the results for NRG. This specification suggests that NRG's average RT margins decreased by almost $6.00 more than Dynegy's after the virtual bidding policy was implemented.

For this differenced model, I also test if the relationship between West-Central DA margins have changed significantly more than the relationship between West-Central RT margins. I stack the DA and RT differenced margins and re-estimate each model allowing for different DA and RT coefficients for each regressor. Then, I test if the DA coefficient on virtual bidding is significantly different than the RT coefficient. The p-values for these tests are presented in the bottom two rows of Table 5. The p-values for the AES test range from .09 to .19 and for NRG they range from .03 to .17.

The difference between the coefficient on the change in DA margins and the change in RT margins averages about $3.00 for both NRG and AES. This implies that DA margins decreased about $3.00 more than RT margins. The price discrimination model predicts that DA margins should always decrease relative to RT margins and that the decrease will be large if, absent speculators, the DA price of transmission was zero and the RT price was greater than zero. Before the virtual bidding policy this situation occurred in 16% of all hours (the DA market predicted no congestion and in RT there was congestion). After the policy change, this occured in 5% of the hours. In both the pre and post-virtual bidding hours in which the DA market under-predicted congestion, the DA price of electricity averaged \$10.50 more than the RT price of electricity. This suggests that pre virtual bidding \$1.6 or 41% of the \$3.97 forward premium can be explained by the 16% of hours in which the DA market under-predicted RT congestion. Post-virtual bidding \$.50 or 27% of the \$1.80 forward premium can be attributed to the 5% of hours in which the DA market under-predicted RT congestion. The decrease in the percentage of hours in which the DA market under-predicts congestion is consistent with the model. As more speculators enter the market, the percentage of hours in which the DA market under-predicts congestion may further decrease.

The model also predicts that DA margins should decrease in hours in which the DA market did not under-predict RT congestion. Before the virtual bidding policy the forward premium averaged \$2.70 in the 84% of hours in which the DA market did not under-predict congestion. After the policy change, the average premium was \$1.35 in the 96% of the hours in which the DA market did not under-predict congestion.

A decrease in the forward premium does not mean that the DA price decreased. Rather, a decrease in the forward premium could also result from an increase in the RT price. The result that DA margins decrease relative to RT margins suggests that the price convergence was a result of a decrease in the DA price and not an increase in the RT price. Again, this is consistent with the price discrimination model.

All of the previous results imply that the DA margins of the two Western firms, AES and NRG, have decreased more than the RT margins. In particular, the DA margins of AES and NRG appear to have decreased by about \$3.00 more than the RT margins. This magnitude appears to be consistent with the observed changes in the forward premium. Almost half of the decrease in DA margins (relative to RT margins) could be explained by a decrease in the under-prediction of congestion in the DA market. The other half of the decrease in margins could be explained by a decrease in the DA prices in hours in which the DA market does not under-predict congestion.

The decrease in DA margins relative to the decrease in RT margins, is consistent with the price discrimination model. The fact that RT margins have also decreased in the post virtual bidding period suggest that some other change has also occurred. As was previously discussed, one very plausible explanation for the decrease in RT margins is that firms may have increased their contract positions in the period after the policy was implemented. Although, I have not found evidence of an increase in contract position in the regulatory filings or the industry press, these results are consistent with increased contracts.

6 Conclusion

In this paper, I have provided evidence that before the virtual bidding policy the DA market in the New York electricity market was not functioning efficiently. In particular, the DA zonal price of electricity was a biased forecast of the RT zonal price of electricity. The DA price of transmission was also a biased predictor of the RT price of transmission. These inefficiencies decreased once the virtual bidding policy was implemented.

I have presented a model of a two zone, two market electricity system that suggests the types of inefficiencies observed in the New York market could result from generation firms located at an exporting zone exercising market power. In particular, these firms may find it profitable to withhold sales so that the DA market predicts the line will be uncongested. In the RT market, these firms will increase production which may result in the transmission line being congested. This strategy would allow the generation firms to extract the transmission rents.

The model predicts that once speculators are added to the market the margins of the firms in the exporting zone should decreases. If the firms were successfully expropriating the transmission rents before speculators were allowed to trade, then this decrease in margins could be quite large. By analyzing the change in DA margins of the firms in the Western New York electricity market, I find results that are consistent with the predictions of the model. In particular, I find that the DA margins of firms in the Western (exporting) zone of New York have significantly decreased and that this decrease is greater than that of the large firm in the Central (importing) zone.

Since the evidence is consistent with the model it is important to consider the policy implications of the model. First, the model predicts that the addition of speculators to an electricity market in which buyers are risk averse and sellers have market power will increase productive inefficiencies. This occurs because speculators prevent the firms with market power from price discriminating. If the firms with market power are forced to charge the same price in the DA and RT markets, they will produce less. This is similar to the case a monopolist that is able to perfectly price discriminate; forcing the monopolist to charge a uniform price will decrease sales and increase deadweight loss. The same results holds in the current situation.

Second, speculators will decrease the average procurement cost of electricity. This occurs because, absent speculators, the majority of electricity is traded at the higher DA price. Speculators result in a decrease of the DA price which leads to lower average procurement costs. This should be of importance to regulators because the decrease in procurement costs may be reflected by a decrease in retail rates. When deciding whether or not to allow speculators to trade in an electricity market, regulators should weigh these two effects.

References

[1] Allaz, Blaise and Jean-Luc Vila (1991) "Cournot Competition, Forward Markets and Efficiency." *Journal of Economic Theory* 59: 1-16

[2] Anderson, Ronald and Mahadevan Sundaresan (1984) "Futures Markets and Monopoly." In R. Anderson ed., *The Industrial Organization of Futures Markets.* Lexington, Mass., and Toronto: Heath, Lexington Books 1984; 75-105

[3] Bessembinder, Hendrik and Michael L. Lemmon (2002) "Equilibrium Pricing and Optimal Hedging in Electricity Forward Markets." *The Journal of Finance* 57: 1347-1382

[4] Borenstein, S., Bushnell, J., Knittel, C., and C. Wolfram (2001) "Trading Inefficiencies in California's Electricity Markets." University of California Energy Institute POWER Working Paper no. PWP-086.

[5] Borenstein, Severin, Bushnell, James, and Steven Stoft (2000) "The Competitive Effects of Transmission Capacity in a Deregulated Electricity Industry." *RAND Journal of Economics* 31 (Summer):294-325

[6] Busse, Meghan (2000) "Firm Financial Condition and Airline Price Wars." *RAND Journal of Economics* 33 (Summer): 298-318

[7] Carlton, Dennis W. (1986) "The Rigidity of Prices." *The American Economic Review* 76: 637-658.

[8] Carlton, Dennis W. (1983) "Futures Trading, Market Interrelationships and Industry Structure."

[9] Chevalier, Judith A. (1995) "Do LBO Supermarkets Charge More? An Empirical Analysis of the Effects of LBOs on Supermarket Pricing." *The Journal of Finance* 50: 1095-1112.

[10] Chevalier, Judith A., and David S. Scharfstein (1996) "Capital Market Imperfections and Countercyclical Markups: Theory and Evidence." *The American Economic Review* 84 : 703-725.

[11] Friedman, Milton (1953) *Essays in Positive Economics.* Chicago: Univ. Chicago Press 1953.

[12] Gilbert, Richard, Neuhoff, Karsten and David Newbery (2002) "Allocating Tranmission to Mitigate Market Power in Electricity Networks." *Univeristy of California Berkeley mimeo.*

[13] Hart, Oliver.D., and David M. Kreps (1986) "Price Destabilizing Speculation." *The Journal of Political Economy* 94: 927-952.

[14] Joskow, Paul and Jean Tirole (2000) "Transmission rights and market power on electricity power networks." *RAND Journal of Economics* 31 (Autumn):450-487

[15] Longstaff, Francis and Ashley Wang (2002) "Electricity Forward Prices: A High-Frequency Empirical Analysis."

[16] Klemperer, Paul D. and Margaret A. Meyer (1989)"Supply Function Equilibria in Oligopoly under Uncertainty." Econometrica 57: 1243-1277

[17] Mansur, Erin T. (2003)"Vertical Integration in Restructured Electricity Markets: Measuring Market Efficiency and Firm Conduct." University of California Energy Institute CSEM Working Paper no. CSEM WP-117.

[18] Oren, S.S. (1997) "Economic Inefficiency of Passive Transmission Rights in Congested Electricity Systems with Competitive Generation." The Energy Journal 18: 63-83.

[19] Powell, Andrew (1993) "Trading Forward in an Imperfect Market: The Case of Electricity in Britain." The Economic Journal 103 (March): 444-453

[20] Puller, Steven (2002) "Pricing and Firm Conduct in California's Deregulated Electricity Market." University of California Energy Institute POWER Working Paper no. PWP-080.

[21] Slade, Margaret E. (1991) "Market Structure, Marketing Method and Price Instability." The Quarterly Journal of Economics 106 : 1309-1340.

[22] Slade, Margaret E. and Henry Thille (2003) "Commodity Spot Prices: An Exploratory Assessment of Levels and Volatilities." University of Warwick mimeo.

[23] Stein, Jeremy. C. (1987) "Informational Externamlities and Welfare-Reducing Speculation." The Journal of Political Economy 95 : 1123-1145.

[24] Wolfram, Catherine D. (1999) "Measuring Duopoly Power in the Brittish Electricity Spot Market." The American Economic Review 89: 805-826.

Table 6: AES DA Margins

Dependent Variable Margin: $DAZP_{\mathsf{t}} - MC^{\mathsf{f}}_{\mathsf{t}}$

Model	OLS	OLS	IV
VB	-3.50	-3.60	-6.28
	(1.13)**	(1.16)**	(1.50)**
Stock Price	.	0.15	4.06
	.	(0.15)	(1.18)**
Sept11	-6.74	-6.18	8.46
	(1.92)**	(2.05)**	(4.31)*
Sept11Trend	-21.03	-32.80	-344.35
	(24.49)	(24.67)	(110.15)**
Trend	36.88	64.66	800.51
	(17.82)*	(30.43)**	(232.44)**
Natural Gas Price	2.28	2.27	2.18
	(0.46)**	(0.46)**	(0.49)**
Log(West Forecast)	-167.69	-167.76	-169.50
	(74.03)*	(74.03)**	(77.45)* *
Log(Cent. Forecast)	-29.69	-30.68	-57.03
	(49.48)	(49.34)	(56.88)
Log(NYC Forecast)	-110.03	-111.52	-151.14
	(16.39)*	(16.26)**	(23.79)**
Log(LI Forecast)	38.50	41.04	108.23
	(5.32)**	(6.03)**	(20.24)**
West Forecast	3.03	3.02	2.99
	(1.33)**	(1.33)**	(1.37)**
Cent. Forecast	2.11	2.12	2.26
	(1.39)	(1.39)	(1.49)
NYC Forecast	1.71	1.71	1.60
	(0.28)**	(0.28)**	(0.32)**
LI Forecast	-1.00	-1.00	-0.92
	(0.14)**	(0.14)**	(0.16)**
Observations	17515	17515	17515
R-squared	0.50	0.50	0.43

Robust standard errors are in parenthesis

* Denotes significance at the 10% level, ** 5% level

Table 7: NRG DA Margins

Dependent Variable Margin: $DAZP_t - MC_t^f$

Model	OLS	OLS	IV
VB	-4.91	-7.11	-8.53
	(1.15)**	(1.18)**	(1.24)**
Stock Price	.	-0.51	-0.84
	.	(0.07)**	(0.13)**
Sept11	-10.19	-11.04	-11.59
	(2.24)**	(2.33)**	(2.44)**
Sept11Trend	11.18	45.07	66.93
	(25.02)	(23.71)*	(22.56)**
Trend	15.74	-7.69	-22.79
	(19.40)	(18.66)	(18.39)
Natural Gas Price	-0.60	-1.00	-1.26
	(0.46)	(0.48)**	(0.50)**
Log(West Forecast)	-158.69	-186.22	-203.98
	(75.00)**	(75.96)**	(78.05)**
Log(Cent. Forecast)	-55.91	-29.11	-11.82
	(50.96)	(49.00)	(46.95)
Log(NYC Forecast)	-103.29	-111.61	-116.98
	(18.34)**	(18.18)**	(18.52)**
Log(LI Forecast)	34.76	40.28	43.83
	(5.81)**	(5.69)**	(5.65)**
West Forecast	3.25	3.54	3.72
	(1.35)**	(1.36)**	(1.38)**
Cent. Forecast	2.34	2.09	1.93
	(1.45)	(1.42)	(1.40)
NYC Forecast	1.33	1.26	1.22
	(0.32)**	(0.31)**	(0.31)**
LI Forecast	-0.86	-0.84	-0.83
	(0.16)**	(0.15)**	(0.15)**
Observations	17515	17515	17515
R-squared	0.41	0.42	0.42

Robust standard errors are in parenthesis

* Denotes significance at the 10% level, ** 5% level

Table 8: Dynegy DA Margins

Dependent Variable Margin: $DAZP_t - MC_t^f$

Model	OLS	OLS	IV
VB	0.65	0.36	2.25
	(2.19)	(2.22)	(2.97)
Stock Price	.	-0.08	0.42
	.	(0.14)	(0.45)
Sept11	0.63	0.24	2.79
	(2.72)	(2.71)	(2.80)
Sept11Trend	-31.44	-44.35	38.47
	(27.36)	(34.58)	(62.92)
Trend	25.26	27.51	13.08
	(22.05)	(21.85)	(19.05)
Natural Gas Price	-0.05	-0.00	-0.30
	(0.32)	(0.33)	(0.42)
Log(West Forecast)	-118.34	-118.09	-119.70
	(90.60)	(90.56)	(91.31)
Log(Cent. Forecast)	-83.49	-83.46	-83.66
	(60.57)	(60.58)	(60.68)
Log(NYC Forecast)	-150.68	-150.93	-149.31
	(24.38)**	(24.37)**	(24.29)**
Log(LI Forecast)	45.04	44.78	46.47
	(8.13)**	(8.16)**	(8.08)**
West Forecast	2.06	2.06	2.07
	(1.60)	(1.60)	(1.61)
Cent. Forecast	3.62	3.61	3.68
	(1.72)**	(1.73)**	(1.75)**
NYC Forecast	2.00	2.02	1.91
	(0.41)**	(0.41)**	(0.43)**
LI Forecast	-0.90	-0.91	-0.85
	(0.20)**	(0.21)**	(0.22)**
Observations	17473	17473	17473
R-squared	0.35	0.35	0.35

Robust standard errors are in parenthesis

* Denotes significance at the 10% level, ** 5% level

Table 9: AES RT Margins

Dependent Variable Margin: $RTZP_t - MC_t^f$

Model	OLS	OLS	IV
VB	-0.98	-0.96	-2.47
	(1.72)	(1.78)	(2.15)
Stock Price	.	-0.03	1.76
	.	(0.27)	(1.25)
Sept11	-9.82	-9.91	-3.65
	(2.81)**	(3.16)**	(5.46)
Sept11Trend	-63.63	-61.60	-206.29
	(19.45)**	(23.06)**	(101.40)**
Trend	59.83	55.01	397.40
	(19.93)**	(49.42)	(237.25)*
Natural Gas Price	1.90	1.90	1.90
	(0.33)**	(0.33)**	(0.33)**
Log(West Load)	57.23	57.30	52.15
	(28.05)**	(28.18)**	(29.45)*
Log(Cent. Load)	-164.41	-164.34	-168.98
	(84.78)*	(84.79)*	(86.01)**
Log(NYC and LI Load)	63.99	64.47	30.25
	(63.27)	(64.94)	(68.57)
West Load	-0.25	-0.26	-0.22
	(0.36)	(0.36)	(0.36)
Central Load	6.64	6.64	6.68
	(2.68)**	(2.68)**	(2.70)**
NYC and LI Load	-0.40	-0.40	-0.06
	(0.76)	(0.77)	(0.81)
Observations	17505	17505	17505
R-squared	0.29	0.29	0.28

Robust standard errors are in parenthesis

* Denotes significance at the 10% level, ** 5% level

Table 10: NRG RT Margins

Dependent Variable Margin: $RTZP_t - MC_t^f$

Model	OLS	OLS	IV
VB	-2.83	-3.73	-5.22
	(1.95)	(1.70)*	(2.11)**
Stock Price	.	-0.21	-0.54
	.	(0.09)**	(0.15)**
Sept11	-13.81	-14.12	-14.63
	(3.11)**	(3.13)**	(3.15)**
Sept11Trend	-36.85	-22.77	0.47
	(28.15)	(28.66)	(30.04)
Trend	44.68	35.80	21.15
	(25.22)*	(25.51)	(26.50)
Natural Gas Price	-0.99	-1.13	-1.37
	(0.73)	(0.75)	(0.78)*
Log(West Load)	82.00	78.68	73.18
	(32.55)**	(31.68)**	(30.72)**
Log(Cent. Load)	-148.00	-146.34	-143.61
	(90.72)	(90.60)	(90.08)
Log(NYC and LI Load)	31.27	27.80	22.07
	(66.15)	(65.66)	(66.03)
West Load	-0.38	-0.37	-0.33
	(0.43)	(0.41)	(0.39)
Central Load	6.14	5.99	5.75
	(2.87)**	(2.87)**	(2.84)**
NYC and LI Load	-0.01	0.03	0.09
	(0.80)	(0.80)	(0.80)
Observations	17505	17505	17505
R-squared	0.24	0.24	0.23

Robust standard errors are in parenthesis

* Denotes significance at the 10% level, ** 5% level

Table 11: Dynegy RT Margins

Dependent Variable Margin: $RTZP_t - MC_t^f$

Model	OLS	OLS	IV
VB	0.71	1.90	1.05
	(2.79)	(2.98)	(3.46)
Stock Price	.	0.33	0.09
	.	(0.30)	(0.56)
Sept11	-5.23	-3.67	-4.79
	(3.78)	(4.16)	(4.34)
Sept11Trend	-53.82	0.70	-38.33
	(37.38)	(73.32)	(89.30)
Trend	24.86	16.94	22.61
	(35.93)	(39.81)	(35.35)
Natural Gas Price	-0.79	-0.98	-0.85
	(0.75)	(0.85)	(0.81)
Log(West Load)	-65.87	-67.36	-66.29
	(63.00)	(63.55)	(62.79)
Log(Cent. Load)	-322.19	-314.62	-320.04
	(136.07)**	(136.62)**	(134.70)**
Log(NYC and LI Load)	135.73	123.27	132.19
	(95.07)	(94.53)	(98.40)
West Load	0.45	0.47	0.46
	(0.55)	(0.55)	(0.55)
Central Load	13.08	12.86	13.02
	(4.35)**	(4.34)**	(4.31)**
NYC and LI Load	-0.70	-0.57	-0.66
	(1.12)	(1.12)	(1.15)
Observations	17463	17463	17463
R-squared	0.20	0.20	0.20

Robust standard errors are in parenthesis

* Denotes significance at the 10% level, ** 5% level

Table 12: AES Differenced DA Margins

Dependant Variable: $Margin^{\mathsf{DA}}_{\mathsf{AES}} - Margin^{\mathsf{DA}}_{\mathsf{Dynegy}}$

Model	OLS	OLS	IV
VB	-4.17	-4.29	-6.18
	(2.16)*	(2.16)**	(2.40)**
Stock Price	.	0.17	2.86
	.	(0.21)	(1.09)**
Sept11	-7.40	-6.76	3.27
	(2.21)**	(2.24)**	(4.33)
Sept11Trend	10.53	-3.09	-216.82
	(19.63)	(25.61)	(93.47)**
Trend	11.77	43.98	549.43
	(17.76)	(43.19)	(210.93)**
Natural Gas Price	2.32	2.32	2.24
	(0.48)**	(0.48)**	(0.49)**
Log(West Forecast)	-49.62	-49.71	-51.12
	(38.52)	(38.44)	(39.88)
Log(Cent. Forecast)	54.07	52.93	35.07
	(25.07)**	(25.21)**	(28.05)
Log(NYC Forecast)	40.56	38.82	11.42
	(20.25)**	(20.47)*	(23.46)
Log(LI Forecast)	-6.40	-3.45	42.85
	(6.67)	(7.58)	(19.39)**
West Forecast	0.97	0.96	0.95
	(0.61)	(0.61)	(0.63)
Cent. Forecast	-1.51	-1.50	-1.41
	(0.74)**	(0.74)**	(0.78)*
NYC Forecast	-0.29	-0.30	-0.37
	(0.35)	(0.35)	(0.38)
LI Forecast	-0.11	-0.10	-0.05
	(0.16)	(0.16)	(0.18)
Observations	17473	17473	17473
R-squared	0.36	0.36	0.29

* Denotes significance at the 10% level, ** 5% level

Robust standard errors are in parenthesis

Table 13: NRG Differenced DA Margins

Dependant Variable: $Margin^{\mathsf{DA}}_{\mathsf{NRG}} - Margin^{\mathsf{DA}}_{\mathsf{Dynegy}}$

Model	OLS	OLS	IV
VB	-5.60	-7.54	-8.56
	(2.22)**	(2.32)**	(2.36)**
Stock Price	.	-0.45	-0.69
	.	(0.09)**	(0.14)**
Sept11	-10.85	-11.61	-12.01
	(2.73)**	(2.80)**	(2.86)**
Sept11Trend	42.80	72.74	88.58
	(22.35)*	(23.12)**	(23.10)**
Trend	-9.38	-30.08	-41.02
	(20.84)	(21.15)	(21.32)*
Natural Gas Price	-0.55	-0.91	-1.10
	(0.52)	(0.54)*	(0.54)**
Log(West Forecast)	-40.57	-64.94	-77.83
	(42.26)	(42.30)	(42.41)*
Log(Cent. Forecast)	27.80	51.53	64.08
	(27.57)	(27.76)*	(27.78)**
Log(NYC Forecast)	47.25	39.87	35.97
	(21.07)**	(20.77)*	(20.84)*
Log(LI Forecast)	-10.12	-5.23	-2.64
	(6.86)	(6.82)	(6.86)
West Forecast	1.19	1.44	1.58
	(0.68)*	(0.67)**	(0.67)**
Cent. Forecast	-1.29	-1.51	-1.62
	(0.83)	(0.82)*	(0.82)**
NYC Forecast	-0.67	-0.73	-0.76
	(0.36)*	(0.36)**	(0.36)**
LI Forecast	0.03	0.05	0.06
	(0.17)	(0.17)	(0.17)
Observations	17473	17473	17473
R-squared	0.36	0.36	0.29

* Denotes significance at the 10% level, ** 5% level

Robust standard errors are in parenthesis

Table 14: AES Differenced RT Margins

Dependant Variable: $Margin^{\mathsf{RT}}_{\mathsf{AES}} - Margin^{\mathsf{RT}}_{\mathsf{Dynegy}}$

Model	OLS	OLS	IV
VB	-1.62	-2.17	-2.88
	(2.58)	(2.64)	(3.02)
Stock Price	.	0.63	1.46
	.	(0.36)*	(1.75)
Sept11	-4.61	-2.40	0.50
	(2.39)*	(2.36)	(6.13)
Sept11Trend	-9.72	-61.00	-127.93
	(31.38)	(50.27)	(154.53)
Trend	34.75	156.27	314.88
	(30.80)	(88.48)*	(346.00)
Natural Gas Price	2.69	2.69	2.68
	(0.69)**	(0.69)**	(0.70)**
Log(West Load)	122.40	120.51	118.05
	(69.63)*	(68.99)*	(68.85)*
Log(Cent. Load)	155.92	154.23	152.02
	(77.89)**	(78.55)**	(79.34)*
Log(NYC and LI Load)	-69.70	-81.74	-97.45
	(65.66)	(68.07)	(76.91)
West Load	-0.70	-0.69	-0.67
	(0.71)	(0.70)	(0.70)
Central Load	-6.38	-6.37	-6.35
	(2.43)**	(2.44)**	(2.47)**
NYC and LI Load	0.28	0.40	0.56
	(0.70)	(0.72)	(0.81)
Observations	17463	17463	17463
R-square	0.15	0.15	0.15

* Denotes significance at the 10% level, ** 5% level

Robust standard errors are in parenthesis

Table 15: NRG Differenced RT Margins

Dependant Variable: $Margin^{\mathsf{RT}}_{\mathsf{NRG}} - Margin^{\mathsf{RT}}_{\mathsf{Dynegy}}$

Model	OLS	OLS	IV
VB	-3.49	-4.24	-5.72
	(2.61)	(2.67)	(2.82)**
Stock Price	.	-0.17	-0.51
	.	(0.14)	(0.20)**
Sept11	-8.60	-8.86	-9.37
	(2.86)**	(2.87)**	(2.97)**
Sept11Trend	17.09	28.83	52.13
	(30.13)	(34.08)	(29.01)*
Trend	19.60	12.21	-2.47
	(31.26)	(33.87)	(30.16)
Natural Gas Price	-0.20	-0.32	-0.56
	(0.56)	(0.60)	(0.56)
Log(West Load)	147.21	144.43	138.91
	(71.88)**	(72.64)**	(70.91)*
Log(Cent. Load)	172.39	173.79	176.57
	(76.09)**	(76.28)**	(76.36)**
Log(NYC and LI Load)	-102.50	-105.40	-111.15
	(66.64)	(65.71)	(66.85)*
West Load	-0.83	-0.82	-0.78
	(0.76)	(0.76)	(0.74)
Central Load	-6.89	-7.01	-7.25
	(2.38)**	(2.38)**	(2.39)**
NYC and LI Load	0.67	0.70	0.76
	(0.72)	(0.71)	(0.72)
Observations	17463	17463	17463
R-square	0.16	0.16	0.16

* Denotes significance at the 10% level, ** 5% level

Robust standard errors are in parenthesis