WELFARE IMPACT OF VIRTUAL TRADING ON WHOLESALE ELECTRICITY MARKETS

by

Juan S. Giraldo

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Agricultural Economics

West Lafayette, Indiana

May 2017

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF DISSERTATION APPROVAL

Dr. Paul V. Preckel, Chair
Department of Agricultural Economics
Dr. Otto Doering
Department of Agricultural Economics
Dr. Andrew L. Liu
School of Industrial Engineering
Dr. Douglas Gotham
State Utility Forecasting Group
Dr. Devendra Canchi
Monitoring Analytics

Approved by:
Dr. Gerald Shively
Head of the Departmental Graduate Program

ACKNOWLEDGMENTS

I would like to thank my doctoral advisor Dr. Paul Preckel. His mentorship and support not only made this dissertation possible but also has a had a large influence in my formation as a professional.

I would also like to thank the rest of my committee members for their support and advice. Dr. Douglas Gotham for his dedication to graduate students; Dr. Devendra Canchi for his patience and constant willingness to help; Dr. Andrew Liu and Dr. Otto Doering.

I would also like to thank the State Utility Forecasting Group for providing financial support during my time at Purdue and introducing me to the electric power sector.

Lastly I would like to thank my Mother whose sacrifice made all of this possible and my Grandfather who was always there to encourage me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Author: Giraldo, Juan, S. Ph.D.
Institution: Purdue University
Degree Received: May 2017
Title: Welfare Impact of Virtual Trading on Wholesale Electricity Markets
Major Professor: Paul V. Preckel

Virtual bidding has become a standard feature of multi-settlement wholesale electricity markets in the United States. Virtual bids are financial instruments that allow market participants to take financial positions in the Day-Ahead (DA) market that are automatically reversed/closed in the Real-Time (RT) market. Most U.S. wholesale electricity markets only have two types of virtual bids: a decrement bid (DEC), which is virtual load, and an increment offer (INC), which is virtual generation. In theory, financial participants create benefits by seeking out profitable bidding opportunities through arbitrage or speculation. Benefits have been argued to take the form of increased competition, price convergence, increased market liquidity, and a more efficient dispatch of generation resources. Studies have found that price convergence between the DA and RT markets improved following the introduction of virtual bidding into wholesale electricity markets. The improvement in price convergence was taken as evidence that market efficiency had increased and many of the theoretical benefits realized. Persistent price differences between the DA and RT markets have led to calls to further expand virtual bidding as a means to address remaining market inefficiencies.

However, the argument that price convergence is beneficial is extrapolated from the study of commodity and financial markets and the role of futures for increasing market efficiency in that context. This viewpoint largely ignores details that differentiate

wholesale electricity markets from other commodity markets. This dissertation advances the understanding of virtual bidding by evaluating the impact of virtual bidding based on the standard definition of economic efficiency which is social welfare. In addition, an examination of the impacts of another type of virtual bid, up-to-congestion (UTC) transactions is presented. This virtual product significantly increased virtual bidding activity in the PJM interconnection market since it became available to be used by financial traders in September 2010.

Stylized models are used to determine the optimal bidding strategy for the different virtual bids under different scenarios. The welfare analysis shows that the main impact of virtual bidding is surplus reallocation and that the impact on market efficiency is small by comparison. The market structure is such that it is more likely to see surplus transfers from consumers to producers. The results also show that outcomes with greater price convergence as a result of virtual bidding activity were not necessarily more efficient, nor do they always correct surplus distribution distortions that result from bias in the DA expectation of RT load.

Compared to INCs and DECs, the UTC analysis showed that UTCs do not have the same self-corrective incentives towards price convergence and are less likely to lead to nodal price convergence or correct for surplus distribution distortions caused by uncertainty and bias in the DA expectation of RT load. Additionally, the analysis showed that UTCs allow financial traders to engage in low risk high volume trading strategies that, while profitable, may have little to no impact on price convergence or market efficiency.

CHAPTER 1. INTRODUCTION

1.1 Introduction

Virtual bidding has become a standard feature of multi-settlement wholesale electricity markets in the United States. Virtual bids are financial instruments introduced to increase competition and market efficiency. Virtual bids allow market participants to take a financial position in the Day-Ahead (DA) market that is automatically reversed/closed in the Real-Time (RT) market. Virtual bids are used by physical participants – generators (wholesale suppliers) and load serving entities (wholesale purchasers) – and financial participants – those without physical assets in the market (i.e. banks and hedge funds) that only take financial positions in the market. Physical participants use virtual bids for hedging and arbitraging purposes while financial participants use virtual bids exclusively for arbitraging purposes. Most U.S wholesale electricity markets only have two types of virtual bids, a decrement bid (DEC) which is virtual load, and an increment offer (INC) which is virtual generation. The PJM Interconnection market (PJM) has a third virtual product known as an up-to-congestion transaction (UTC) which is a spread product used for hedging/arbitraging prices differences between two points in the transmission network.

In theory, financial participants create benefits by seeking out profitable bidding opportunities through arbitrage or speculation (henceforth referred to only as arbitrage). Benefits have been argued to take the form of increased competition, price convergence, increased market liquidity, and a more efficient dispatch of generation resources (Isemonger, 2006). Studies by Hadsell and Shawky (2007), Jha and Wolak (2014) and Li et al. (2015) found that price convergence between the DA and RT markets increased following the introduction of virtual bidding into California and New York wholesale electricity markets (CAISO and NYISO respectively). This was taken as evidence that virtual bidding had increased market efficiency. However, the studies also noted that despite the introduction of virtual bidding price differences remained between the DA and RT markets which signified persistent inefficiency in the markets due to an incomplete integration between wholesale electricity markets and financial markets. Their findings were followed by policy recommendations to reduce trading costs to further expand virtual bidding.

Currently there is controversy about the further expansion of virtual bidding in wholesale electricity markets. It has long been recognized that virtual bidding can have negative impacts on the market such as when is used to manipulate the value of positions in other markets, such as the financial transmission rights market (FTRs) (Celebi et al. 2010, and Ledgerwood & Pfeifenberger 2013, Birge et al. 2014), and by the exploitation of bidding strategies that while profitable do not improve system performance (Parsons et al., 2015). However, what really brought a spotlight to the issue and raised question about the value that virtual bidding provides to wholesale electricity markets has been the volatile use of UTCs in the PJM market.

Since made available to be used for arbitraging purposes UTCs have overwhelmingly become the most traded virtual product in the PJM market as seen in Figure 1 with most of the trade volume attributed to financial participants (94 percent in 2014 and 80 percent in 2015) (PJM, 2016a). The volume disparity among virtual products is partly the result of unequal treatment as INCs and DECs are allocated uplift charges while UTCs are not.

The independent market monitor (IMM) for PJM has repeatedly called for the UTC product to be allocated uplift charges in a manner consistent with INCs and DECs (PJM, 2010). Among financial participants there is broad support for UTCs and opposition to the imposition of uplift charges to any of the virtual transactions (FERC, 2015). There has been a push to introduce UTCs into other wholesale electricity markets. A UTC like product has been under consideration for introduction at the Midwest Independent System Operator (MISO) market with support from the IMM for MISO (MISO, 2016). However, the IMM for PJM has called for UTCs to be altogether eliminated from the PJM market. The IMM points to the dearth of evidence that arbitrage UTC trades provide benefits to the market as well as the existence of many documented cases where UTCs have negatively impacted the operation of the market (Monitoring Analytics, 2013).

In September 8, 2014 the Federal Energy Regulatory Commission (FERC) issued an order mandating a study to consider allocating uplift charges to UTCs in manner consistent with the way that uplift is allocated to INCs and DECs- that is, UTC MW would be considered deviations eligible to be allocated uplift charges. While uplift charges were never implemented for UTCs, the possibility that they might be, and in a potentially retroactive manner, caused a precipitous drop in UTC trade volume as seen in Figure 1. UTC trade volume has been slowly recovering since, experiencing a noticeable increase after the expiration of the FERC order on December 2015 (for a detailed explanation of the issue see PJM, 2016¬a).

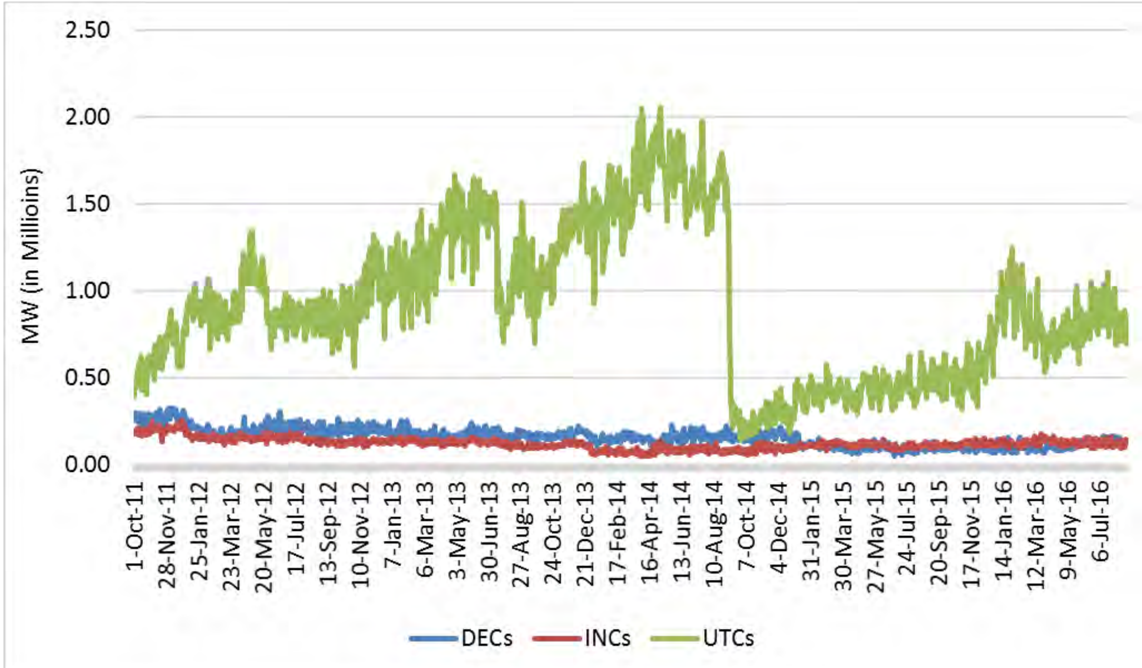


Figure 1-1 Daily cleared virtual transactions in PJM. Source: PJM, 2016[b]

The controversy surrounding virtual bidding is the result of a discrepancy between the theoretical framework of virtual bidding and its assumptions versus the implementation in operating markets. The argument that price convergence is beneficial comes from the study of commodity and financial markets and the role of futures for increasing market efficiency in that context. This viewpoint largely ignores the complexity that differentiates wholesale electricity markets from other commodity markets and imposes expectations about market outcomes and market features that are not feasible or may not complement wholesale electricity markets. For example, from a commodity markets perspective there is an expectation that DA and RT prices should converge much like futures and spot prices for agricultural commodities converge at the time of delivery. This is not truly feasible in wholesale electricity markets as RT prices are not the result of repeated trader interactions but rather an engineering optimization model responding to unpredictable constraints in the system[1]. Price convergence is therefore evaluated in the academic literature in terms of average price convergence over a period of an entire year. As discussed by Parsons et al. (2015) the problem with this metric is that the average may obscure large offsetting price differences.

Another expectation from the financial/commodity markets perspective is the need for market liquidity. To that end the UTC product has been praised for increasing trade volume in the DA market and increasing participation by financial participants. However as pointed out by the IMM for PJM there are several issues with UTCs trades in the PJM market.[2] The UTC product and its impact on the market is largely unknown. In the academic literature the UTC product has not been studied and there is no evidence that UTCs provide the same market benefits that have been associated with INCs and DECs. Before further expansion of the UTC product into other wholesale electricity markets a better understanding of the product and its impact on the market is needed.

The goal of this dissertation is to advance the understating of the impact of virtual bidding on wholesale electricity markets by bridging the gap between market outcomes and their impact on market participants. Instead of taking the financial viewpoint that considers market efficiency as price convergence between the futures and the spot market, the impact of virtual bidding will be evaluated based on the standard definition of economic efficiency which is social welfare. Additionally, the UTC product will be

[1] These include generators tripping, transmission lines becoming unavailable, and ramp limits which are limits on generators ability to increase their output.

[2] The IMM points that there is no evidence that speculative UTC trades provide for price convergence while it has documented several instances where UTCs have negatively impacted the market such as by contributing to FTR underfunding, substantially increasing the time needed to clear the DA market, and forcing PJM to make manual adjustments to unit commitments and transmission line limits in order to accommodate the high volumes of UTCs (Monitoring Analytics, 2013).

thoroughly explored and its market impacts will be also evaluated from a social welfare perspective. This dissertation will test two hypotheses about virtual bidding on wholesale electricity markets:

1. Average price convergence between the day-ahead and real-time markets does not always lead to more efficient market outcomes; and

2. UTCs incentives are not self-corrective towards price convergence and have different market impacts compared to INCs and DECs.

In order to test these hypotheses this dissertation has two main objectives. The first one is to study the different virtual products to try to determine if they have a differential impact on the market and whether UTCs have any inherent advantages that would explain why they are traded at such greater numbers. Specifically, the three different virtual products market (INCs, DECs and UTCs) will be modeled to understand how price convergence incentives motivate bidding strategies for each of the products. The second objective is to provide an alternative method to evaluate the impact of virtual bidding by using social welfare analysis, the standard measure of economic efficiency, to evaluate the impact of bidding strategies on market participants. We believe that this approach provides a more transparent assessment of the efficiency impact of arbitrage/speculative trading.

This dissertation is organized as follows. Chapter 2 presents a brief review of some electricity fundamentals and describes the structure of wholesale electricity markets. Chapter 3 reviews the literature on the impact of virtual bidding on wholesale electricity markets. Chapter 4 presents the optimal bidding model for INCs and DECs and presents a welfare analysis on the optimal INC and DEC bidding strategies. Chapter 5 presents the

UTC optimal bidding model and presents a welfare analysis on the optimal UTC bidding strategy. Chapter 6 summarizes and discusses the findings, compares and contrasts the impacts of INCs and DECs versus UTCs, discusses limitations and future work and provides concluding remarks.

CHAPTER 2. ELECTRICITY AND WHOLESALE ELECTRICITY MARKETS

2.1 Electricity Fundamental Concepts

Electricity is different from any other commodity. Its unique characteristics are driven by two defining attributes, non-storability and loop-flow. Non-storability occurs because there is currently no economically viable way to store electricity at the wholesale level, thus electricity must be consumed as it is produced. This is complicated by the fact that the electricity system must at all times be in balance with supply equal to demand because failure to do so can lead to a system blackout. Additionally, the majority of consumers do not pay the real time price of electricity, creating an almost complete lack of demand side price response in the short term. As such electricity generation must always precisely and instantaneously follow electricity demand, known as load. Given these conditions electricity spot market clearing prices are characterized by a high degree volatility as inventories cannot be used to smooth supply or demand shocks (Bessembinder and Lemmon, 2002).

Loop flow refers to the way electricity flows through a network, which is governed by a set of principles known as Kirchhoff's laws. The main implication of Kirchhoff's laws is that electricity takes the path of least resistance. In an electricity network this means that electricity will travel across many parallel lines and not a single line. The inability to control the path that electricity will travel creates widespread externalities that grow in complexity with the size of the network (Chao and Peck, 1996). In order to manage the complexity of power flows through a network system operators rely on power flow analysis.

2.2 Power Flow Analysis

The electric system is composed of the electric grid which is a large network of transmission and distribution lines, generators that inject power into the grid, and consumers that withdraw power from the grid. The stable operation of the grid requires a precise understanding of the impact of every dispatch decision, change in load, and contingency, on every element within the system. This is accomplished through a modeling exercise known as power flow analysis. Power flow analysis is a model of the flow of electric power in an alternating current (AC) electrical system. It is defined by parameters such as voltages, power flows, currents, phase angles, real power and reactive power. The framework of the model is based on the mathematical relationships among the systems' parameters derived from the physical laws of electricity (Canchi, 2010).

An extension of power flow analysis is the AC optimal power flow (AC-OPF) problem which is constituted by adding an objective function to the power flow model. Objectives include the minimization of generation cost or transmission losses. The AC-OPF is at the heart of the physical operation of wholesale electricity markets. It is used to obtain a set of dispatch instructions for generators that ensure that the system remains in balance and that the normal and emergency limits of the system are always respected.

However, the AC-OPF model is characterized by the non-linear relationships of the power flow parameters of an AC network. Solving the AC-OPF for even a small network presents many computational complexities. Given this fact, a simpler version of the AC-

OPF problem known as the direct current optimal power flow (DC-OPF) problem is used instead. The DC-OPF is a linear approximation of a full alternating current network. The DC-OPF is recognized as a practical alternative that does not widely deviate from the original AC-OPF and yet it is robust and easy to solve because it is a linear programming problem. For a comprehensive review of the topic see Canchi, 2010.

In addition to being critical for the physical operation of the electricity system, power flow analysis is also central to the economical operation of the electricity system. Electricity markets are operated based on the principle of economic dispatch or merit based dispatch. Economic dispatch is defined as the operation of generating facilities to produce energy at the lowest cost to reliably serve consumers, recognizing any operational limits of generation and transmission facilities (FERC, 2005). The basic idea involves using the lowest cost generators first and gradually using the more expensive generators as load increases. Independent system operators (ISOs) use the DC-OPF model with the objective of minimizing generation cost to obtain a set of dispatch instructions consistent with economic dispatch. An extension of the economic dispatch model is the security constrained economic dispatch (SCED) model. The SCED model includes additional constraints to ensure that the emergency limits of the transmission system are not violated in case any one transmission line fails.

2.3 Transmission Line Limits

Transmission lines have a maximum transfer capacity (also known as the thermal limit) which is denoted in MW and refers to the net power flow through the line. The ISO must ensure that the power flows associated with the dispatch of generators remain within the safe operational bounds of every transmission line. Power flows also have a directional component and opposing power flows on the same transmission line will offset each other. For example, consider two nodes, A and B, connected by a transmission line with a 1,000 MW capacity. If there is 1,000 MW flowing from A to B (denoted as the prevalent flow) then the net flow on the line is 1,000 MW and no more power can be transferred in the A to B direction. However, if a 250 MW flow is added in the B to A direction (denoted as the counter flow) then the opposing flows offset each other, and the net flow on the line is 750 MW. The 250 MW counter flow makes it possible for additional MW to be transferred in the prevalent direction (up to 250 MW) without exceeding the line capacity of 1,000 MW. For a more comprehensive review of power flow analysis see (Canchi, 2010).

2.4 Electricity Sector Restructuring

Delivered electricity is a bundle of many services. There is the actual generation of electricity, transmission and distribution for the delivery of power. However, there are other services that are required for operating the system such as voltage control and black start for system reliability and frequency control for power quality. These services in a particular geographic area were all provided, and still are in some regions, by a single regulated firm known as the vertically integrated utility. That started to change following the enactment of the Public Utilities Regulatory Policy Act (PURPA) in 1978 which allowed for the establishment of non-utility generators, opening competition on the generating side of the market.

Wholesale electricity markets were established in the early 1990s following the enactment of FERC Order No. 888 which promoted wholesale competition through open access non-discriminatory use of the transmission system. That is, all public utilities that owned, controlled, or operated transmission infrastructure were required to provide access to their facilities at non-discriminatory rates to all owners of generating equipment. This created a need for a new institution dedicated to the operation of wholesale electricity markets called the independent system operator (ISO). ISOs were established as market administrators for centralized markets based on economic dispatch. Additionally, FERC stablished stricter requirements that must be met in order to obtain the status of a Regional Transmission Organization (RTO). Both ISOs and RTOs are charged with the operation of wholesale electricity markets.

2.5 Market Architecture

Following the directive to create wholesale electricity markets there was a need for a market design. However, electricity's unique characteristics, non-storability and loop-flow, imposed some important constraints on market design. The transmission system is a shared facility and when it becomes congested (i.e. when the flow on a transmission line approaches the capacity which cannot be exceeded without a substantial risk of damaging the line due to high heat – the thermal limit), it can prevent the full use of the lowest cost generators, creating significant opportunity costs. Because of loop flow it is not possible to establish a delivery path, known as a contract path, for electricity transactions. Thus the flow of any electricity transaction will affect other transactions, even in distant areas. This leads to misalignment between private cost and social cost in electricity transactions and causes a potential costly dislocation of resources in the power market (Chao and

Peck, 1996). Thus, there was a need for a system that disciplines the use of the transmission system and internalizes the effects of loop-flow by attributing the transmission opportunity costs to those transactions that induce congestion.

While competing market models were developed, the one that eventually prevailed is known as the nodal pricing approach. This model is based on the theory of locational spot pricing developed by Scheweppe et al. (1988).

2.5.1 Locational Marginal Pricing

Scheweppe et al. (1988) developed the theory of locational spot pricing for electricity, also known as nodal pricing or locational marginal pricing (LMP). The premise for this model is that electricity should be treated as a commodity which can be bought, sold, and traded, taking into account its time- and space-varying values and costs. The goal of the mechanism is to replicate a competitive market outcome in which the market price is set by the marginal unit that clears the market, much like the prices emerging from competitive spot markets for other commodities. In regular commodity markets repeated transactions and the dispersion of information will result in market clearing prices converging towards the cost of producing the marginal unit of output. Because of the complexity of electric power flows in wholesale electricity markets the marginal cost is computed by the system operator.

The operation of markets under the nodal pricing approach can be best described as an auction. The system operator receives bids from generators and load, bids are selected by solving the SCED model, and accepted bids are settled at the corresponding LMPs. LMPs are calculated using the dual variables obtained from solving the SCED linear programming model. LMPs can be decomposed into energy, congestion, and loss components. The energy component represents the system wide marginal cost of supplying one more unit of electricity. The congestion component represents the opportunity cost of limitations imposed by the capacity of the transmission system. This cost occurs because when any one line in the transmission grid gets congested (reaches its thermal limit) this prevents the full use of the cheapest generators. The losses component represents the energy losses that occur due to the use of the transmission system. Thus LMPs reflect information about all the interactions in the network.

2.5.2 Contract Networks

Hogan (1992), developed the theory of contract networks which is a mechanism for defining transmission capacity rights. Contract networks builds on the locational spot pricing theory to provide efficient transmission pricing. Hogan points out that economic dispatch provides for the most efficient allocation of transmission capacity because by definition it maximizes the benefits less costs subject to the availability of plants and constraints in the transmission system. He argued that the short run equilibrium in a competitive market would reproduce both the prices and associated power flows dictated by locational spot pricing theory. Therefore, efficient transmission of power from one node to another would not be priced at anything higher than the difference in the spot prices at the respective nodes. Thus the difference in spot prices is the short run equilibrium price for transmission.

In order to preserve the short-run efficiency of the system Hogan proposed a mechanism that merges short-term pricing and transmission capacity rights traded in a secondary market. The rights are financial in nature given that loop flow makes any physical right system largely impractical. Financial transmission rights (FTRs) are defined by an injection node and a withdrawal node and a specific capacity level (defined in MW). Holders of an FTR are entitled to a rent payment equivalent to the price difference between the withdrawal and injection node times the MW quantity of the FTR.

The way FTRs work to preserve short run efficiency of the system is as follows. If the holder of the right actually uses the transmission system to move power from one node to the other, then the congestion payment that it would incur is just balanced by the rental revenue it receives. However, if the right holder is precluded from using the full capacity of the right, the compensation from the rent payment is just enough to make the right holder indifferent between actually delivering the power or receiving the compensation. If this is the case the right holder can honor any long-term delivery commitments by using the rental payment to purchase power at the point of destination. Compared to a physical rights system FTRs remove the incentives for rights holders to withhold transmission capacity or use transmission in an inefficient way. In this manner the ISO can always seek to schedule the lowest cost generators without disadvantaging any FTR holder.

2.5.3 Two-Settlement Markets

Most U.S. wholesale electricity markets operate as two-settlement markets composed of a forward market, known as the Day Ahead (DA) market, and a spot market known as the Real Time (RT) market. The DA market is a quasi-financial market operated as a centralized auction that is run on the day prior to the operating day. The ISO receives bids from generators, load resources, and financial participants and cleared bids are settled at hourly DA LMPs prices. While there is no physical production or consumption in the DA market, cleared bids for physical participants set the consumption and generation schedules for the operating day, giving the ISO an hourly forecast of demand and the generating resources that will be available to satisfy that demand. This is an essential function of the DA market given the complex task of managing the system in real-time and the fact that generation units are complex machines that may require several hours' notice in order to be brought online. The DA market provides the system operator with some lead time for planning and scheduling resources in order to be able to operate the grid in a physically secure manner, and ideally, at the lowest possible cost. The DA market also allows market participants to secure the less volatile DA prices, giving participants the opportunity to hedge against real time prices which are characterized by a high level of volatility.

There is no standard design for DA markets among U.S. ISOs; however in principle they are all very similar. The DA market can be best described as an auction in which the system operator receives bids from generators and load; bids are selected to maximize total surplus; and the accepted bids are settled at the determined market prices. There are two types of DA markets, the power exchange and the power pool, and they differ in the way in which the unit commitment problem is handled. The unit commitment problem is used to determine when individual generators are started. This is important because generators can incur large startup costs and minimum times for which they must be run, and thus once a unit is started it is said to be committed. The unit commitment problem refers to the issue of finding the most economical times to commit and decommit (shut down) all the individual generators in a control area.

In a power exchange the system operator is not responsible for solving the unit commitment problem. Generators submit simple energy-only bids which must indirectly account for startup cost, no load cost (cost of running the generator while producing no output), and the generators' limitations. In a power pool, generators submit multipart bids which cover all important aspects of a generator's operating cost and physical constraints. The system operator then solves the unit commitment and dispatch problem centrally. An important aspect of power pools is that they provide side payments to any generator that is committed but is not able to cover their startup or no load costs as a result of following dispatch instructions from the ISO, these are known as uplift payments. Uplift payments are an incentive to generators to always follow dispatch instructions so that the ISO can operate the system reliably and at the lowest cost. While there are some markets that operate purely as a power exchange, such as the wholesale electricity market in Texas known as the Electric Reliability Council of Texas (ERCOT), most operate as a combination of a power exchange and a power pool.

On the operating day the ISO runs the RT market which is a physical market. The ISO dispatches the generators according to the DA generation schedules plus any additional units necessary to serve load and maintain system reliability. Dispatch levels may be higher or lower depending on load levels or system conditions (for example a large generator unexpectedly tripping). In the RT market start-up times restrict the number of units available for dispatch and only a subset of additional units (those that can start up quickly) are available to supply power. The RT market produces hourly LMPs that are used to settle RT market transactions. However, RT market transactions do not include all of the power that was consumed and produced in the RT market but rather deviations from the DA schedule and unscheduled production and consumption. In essence the RT market is for balancing any deviations from the DA schedules, and thus it is also referred to as the balancing market.

2.5.4 Bilateral Transactions and Transmission Bids

An LMP based market provides market participants the flexibility to enter into private contracts, known as bilateral transactions, for the purchase or sale of power outside of the auction market operated by the ISO. However, because these transactions take place within the ISO's control area and require the use of the shared transmission infrastructure they have to pay the cost of transmission that the associated power flows incur. This automatically occurs in an LMP based market. Take for example a generator who enters into a private contract to deliver X MW to a load serving entity (LSE). To fulfill the contract, the generator injects X MW of power at its node, getting paid the corresponding LMP, and is responsible to pay for X MW of withdrawals at the LSE's node. The generator pays to the ISO the difference in LMPs between the injection and withdrawal nodes, which is the cost of transmission.

Participants may hedge the RT price of transmission by using a transmission bid in the DA market. Transmission bids specify an injection and withdrawal node, the quantity of MWs being moved and the reservation price per MW. The bid will clear as long as the transmission price is less than or equal to the reservation price. If the bid clears, the participant can use the transmission grid to fulfill the bilateral transaction and only pay the DA price of transmission. If the transmission bid does not clear it might be cheaper to fulfill the contract by purchasing power from the ISO in the RT market.

2.6 Virtual Bids

Virtual bids allow market participants to arbitrage price differences between the DA and the RT market by allowing them to take a position in the DA market without having to physically inject or withdraw power from the transmission network. The three virtual bids that will be considered are DECs, INCs, and UTCs.

2.6.1 Decrement Bids (DECs) and Increment Offers (INCs)

A DEC is a bid to purchase a specified amount of power in the DA market conditional on the DA price being equal or lower than the bid price. Cleared DECs create additional total demand (real plus virtual) in the system, which causes additional generation to be scheduled and consequently may cause an increase in the DA price. An INC is a bid to sell a specified amount of power in the DA market conditional on the market price being equal or higher than the bid price. INCs increase the total available supply (real plus virtual), and are equivalent to a rightward shift in the supply curve at prices equal to or greater than the reservation price. When an INC offer clears it can displace physical generators and may cause the DA price to decrease.

In the RT market INCs and DECs are automatically treated as deviations from the DA scheduled and settled at the RT market price. Thus a DEC is equivalent to taking a long position on the market – buying power in the forward market and selling it on the spot market. Hence, it will be profitable whenever RT prices are greater than DA prices (buy low sell high). On the other hand, an INC is equivalent to taking a short position on the market – selling power in the forward market and buying it on the spot market. That transaction will be profitable whenever RT prices are less than DA prices (sell high buy low). Financial participants who correctly predict price differences between the DA and

RT markets will have profitable trades, with the possible result that the DA and RT prices will be closer. The phenomenon of narrowing the gap between DA and RT prices is called price convergence.

2.6.2 Up-to-Congestion Transactions (UTCs)

In the PJM market transmission bids, or bids on the difference between LMPs at two nodes in the network, are known as up-to-congestion transactions (UTCs). While UTCs were available since the inception of the PJM market the requirement that they be accompanied by the procurement of physical transmission and that one of the nodes had to be external to PJM limited their use to physical participants who were importing, exporting, or moving power through PJM. When that requirement was dropped it made UTCs available to be used for arbitrage and speculative purposes by both physical and financial participants. In the PJM market there are 400 nodes available to be injection and withdrawal points for UTCs.

Whether used for hedging or arbitrage/speculative purposes UTCs are bid in the same manner. A UTC bid specifies a source (injection) node, a sink (withdrawal) node, a MW quantity, and a reservation price per MW. If the price difference between the sink and the source is less than the reservation price, then the UTC will clear at the DA price of transmission. This represents a cost if the UTC is in the prevalent direction (sink LMP > source LMP) or a revenue if the UTC is in the counter flow direction (sink LMP < source LMP). In the RT market UTC bids are settled at the RT price of transmission (i.e. the difference between the RT LMP at the sink and RT LMP at the source). This will represent revenue if the UTC is in the prevalent flow direction and a cost if the UTC in the counter flow direction. Thus on a particular pair of nodes a UTC in the prevalent direction is used to take a long position and will be profitable when the RT price of transmission is greater than the DA price of transmission, and a UTC in the counter flow direction is used to take a short position and will be profitable when the RT price of transmission is lower than the DA price of transmission.

As with INCs and DECs, there is also an expectation that, in seeking profitable bidding opportunities with UTCs, financial participants will cause DA and RT transmission prices to converge. Cleared UTCs in the prevalent direction, placed in expectation of higher RT transmission prices, may increase the DA price of transmission if the line is already congested (reached its transfer limit) or if the UTC causes the line to become congested. Similarly cleared UTCs in the counter flow direction, placed in expectation of lower RT transmission prices, may decrease DA transmission prices by relieving congestion in the prevalent direction. Financial participants who correctly predict price differences will have profitable trades, with the possible result that the DA and RT transmission prices will be closer.

However, profitable UTCs do not necessarily lead to nodal price convergence. This is because the profitability of a UTC depends on the combination of the separate outcomes when the UTC is broken down into its INC (source) and DEC (sink) components. A UTC may incur a loss at one node and a profit at the other node, potentially causing price divergence at the loss making node and price convergence at the profitable node. The UTC can be net profitable if the gains on the profitable node are large enough to offset the losses on the loss making node.

Unlike INCs and DECs, UTCs do not have an impact on unit commitment as they are only incorporated in the optimal dispatch model used to determine optimal power flows after the unit commitment decisions have been made. This is an important distinction because it means that UTCs cannot result in commitment improvements the way that INCs and DECs can. UTCs can only have an impact on the scheduling of units that were already committed, and thus nodal prices.

CHAPTER 3. LITERATURE REVIEW

Several studies in the academic literature have investigated convergence between DA and RT prices in wholesale electricity markets. Convergence is typically reported as the average difference between the DA and RT prices (DA-RT) and is referred to as the DA/RT spread. These studies have covered the major wholesale electricity markets in the United States across different time periods (Longstaff and Wang (2004), Douglas and Popova (2008), Ullrich (2007), Haugom and Ullrich (2012) and Pirrong and Jermakyan (2008) study the DA/RT spread in PJM; Saravia (2003), Hadsell and Shawkey (2006) and Hadsell and Shawkey (2007) study the DA/RT spread in NYISO; Hadsell (2008), Hadsell (2011), and Werner (2014) study the DA/RT spread in NE-ISO; Bowden el al. (2009) and Birge et al (2014) study the DA/RT spread in MISO; and Borenstein et al. (2008), Jha and Wolak (2014) and Li et al. (2015) study the DA/RT spread in CAISO). Overall these studies have consistently found that across different time periods the average DA/RT spreads are statistically different from zero. The spreads vary in magnitude and may be positive or negative depending on variables such as season and peak and off-peak hours.

From a commodity markets perspective there is an expectation that the DA/RT spread should be zero. As described by Borenstein et al. (2008) in efficient commodity markets with risk neutral traders, all contracts - forward and spot - for the same good, to be delivered at the same time and location, should on average trade at the same price. This definition implies that an efficient wholesale electricity market is one where, on average, DA LMPs are equal to RT LMPs. Thus many of the studies also investigated possible reasons as to why the spreads were non-zero.

As discussed by Parsons et al. (2015) risk premia have been observed in other commodity markets and they may very well be present in wholesale electricity markets. Spot prices are more volatile than DA prices and some market participants may be willing to pay a price premium in the DA market to avoid the risk of RT (spot) prices. Bessembinder and Lemmon (2002) developed an equilibrium model that explains price differences between the DA and the RT market as a result of risk and risk premia. The model predicts that the price premium may be positive or negative for different hours. An empirical analysis of the model using PJM data for years 1997 through 2000 and CAISO and California Power Exchange (CALPX) data for the years 1998 through 2000 found that forward prices contain a positive risk premium when either expected demand or demand variance is high, and a low or negative risk premium when expected demand is low and demand risk is moderate. Bessembinder and Lemmon's model was further empirically tested by Longstaff and Wang (2004) and Haugom and Ullrich (2012). Longstaff and Wang analyze PJM price data for the years 2000 through 2012 and found further evidence of the presence of risk premiums in the DA price. Haugom and Ullrich analyze PJM price data for the years 2000 through 2010 and found a lack of evidence to support the claim that the forward price contains a risk premium. Instead they claim that forward prices have converged to unbiased predictors of the subsequent spot prices. They claim that this change is the result of increased market efficiency, or reduced price premia, or both as agents have gained experience.

Saravia (2003) explains the DA/RT spread as a result of the exercise of market power. The NYISO started operating in 1999 and for two years only physical participants were allowed to participate, and market rules restricted these physical participants from speculating on DA and RT price differences. In 2001 virtual bidding was introduced and physical and financial participants were allowed to speculate on DA and RT price differences. Saravia shows that the absolute value of the differences between the DA and RT prices decreased significantly in the New York Western zone following the introduction of virtual bidding, and the DA price of transmission ceased to be significantly different from the RT price of transmission following the introduction of virtual bidding. To help understand the pre- and post-virtual bidding price relationships, the author presents a duopoly model that predicts that, absent speculators, firms with market power will price discriminate between the DA and RT markets, and this price discrimination results in the DA price of transmission under-predicting the RT price of transmission. When speculators are added to the model the DA price cost-margins decrease. The model is empirically tested using detailed engineering data on the marginal cost of generation of units in New York. The results indicate that, controlling for other market conditions, the DA-price cost margins for generators in western New York significantly decreased after virtual bidding was implemented.

Borenstein et al. (2008) study the price differences between the CALPX market (equivalent to the DA market for CAISO) and the CAISO balancing (RT) market in the years 1998 through 2000. The authors find that significant price differences persisted and that a simple trading strategy, based only on prior prices, would have made money. They find that the common explanations for persistent price differences—risk aversion and differential trading costs across markets—are not consistent with the data. The authors assert that if one firm sees a profitable trading opportunity, its trading will tend to reduce the profitability of the strategy, but it will not trade to the point that the marginal trade by itself breaks even. The firm will have market power in the trading strategy and it will take into account its effect on the strategy's profitability when it decides how much trading to do. The authors' claim that the price differences observed between the CALPX and the CAISO market were the result of market power in financial arbitrage given that institutional or legal constraints, or asymmetric information limited the number of agents that recognize trading opportunities and are in a position to exploit it.

Jha and Wolak (2014) argue that in a market with risk neutral traders and zero transactions costs there would be no difference between the DA and RT prices. However, in the presence of transaction costs there will be price differences between the DA and RT prices. Without financial instruments to explicitly arbitrage price differences, physical participants will leverage their assets to implicitly arbitrage price differences (i.e. withhold either load or generation in the DA market). Implicit arbitrage may involve costly actions (high transaction costs) that may have an adverse impact on system reliability. The authors claim that virtual bidding reduces the costs to market participants to exploit price differences between the DA and RT markets as it makes it unnecessary for physical participants to employ costlier distortions in their DA transactions in order to exploit the price differences. To test their hypothesis, the authors use CAISO price data for the years 2009 through 2012. Virtual bidding was introduced in CAISO in 2011 which allows for the comparison between the period before and after the introduction of virtual bidding. The authors find that after the introduction of virtual bidding 1) transaction costs decreased, making it less costly for physical participants to engage in price arbitrage, 2) profitability of trading strategies decreased suggesting that price differences between the DA and the RT market decreased (improved price convergence) and 3) RT price volatility decreased. The authors further find that following the introduction of virtual bidding there was a reduction in the total MMBTU/MWh in CAISO which lead to a reduction in both total variable cost and CO2 emissions which resulted from an improvement in the commitment of generation resources.

Li et al. (2015) define market efficiency in terms of trading profitability where a zero-profit competitive equilibrium implies market efficiency. The zero profit condition is achieved when there is convergence between the DA and RT prices. The authors view price convergence and convergence bidding (arbitrage through the use of virtual bidding) as beneficial to the market. The authors claim that price convergence reduces incentives for market participants to defer their physical resources to the RT market in expectation of favorable RT prices which increases the stability of the DA market. A more stable DA market incurs fewer uplift costs as the ISO does not have to make out market transactions to procure generation units to maintain system reliability. Convergence bidding allows for the revelation of true economic costs, which allows the ISO to allocate resources efficiently and optimally. The authors' test whether CAISO's DA and RT markets are efficient in the sense of eliminating trading profits, and if not, to what extent virtual bidding improves market efficiency. The efficiency test was conducted by developing a bidding strategy model and applying it to CAISO market data from 2010 to 2012 (virtual bidding was introduced 2011). The authors find that following the introduction of virtual bidding the profitability of trading strategies decreased, which they concluded was an indication that the market had become more efficient.

Parsons et al. (2015) provide a different perspective on wholesale electricity markets and virtual bidding. The authors claim that virtual bidding, introduced to increase competition and pricing, may not work as advertised and profits from virtual bidding may be a purely parasitic transfer from consumers and producers. This can occur because the DA and RT market clearing process is thought of as successive versions of the same bidding and auction process and virtual bidding relies on the very strong assumption that the different stages of the market operate identically. Under this assumption any price difference between the DA and RT markets is the result of either a deficiency in supply or demand bid into the DA market and virtual bidding through price convergence incentives are able to correct this by adding to the net demand or net supply in the DA market. However, the high level of complexity required to solve the security constrained economic dispatch (SCED) model requires many approximations, decompositions, and engineering judgements which are applied differently across the DA and RT markets. Because of this, price differences between the DA and the RT market may arise even when there are no demand and supply imbalances. Thus virtual bidding cannot help with price differences that result from modeling differences between the DA and RT markets. Virtual bidding may converge prices but not the underlying issue causing the price differences, in some cases virtual bidding can negatively impact the unit commitment and dispatch, adding costs to the system.

As described by Ledgerwood and Pfeifenberger (2013) and Celebi et al. (2010) it has also been recognized that virtual bidding can and has been used for purposes of market manipulation. The main form of market manipulation using virtual bids is engaging in uneconomic trading. Uneconomic trading consists of placing bids that consistently lose money on a stand-alone basis with the intention to move the DA price and increase the value of a complementary position in another market. Uneconomic trading is profitable when the additional profit in the complementary position more than offsets the loss in the virtual trades. As discussed by Ledgerwood and Pfeifenberger (2013) FERC has prosecuted several cases in which companies were using virtual bids to engage in uneconomic trading to enhance the value of financial transmission rights. Uneconomic trading can lead to price divergence between the DA and RT market and negatively impact the unit commitment leading to higher generation costs and decrease in system reliability.

3.1 Summary

Numerous studies have shown that wholesale electricity markets have persistent price differences between the DA and RT markets. These differences have been ascribed to a wide array of causes including a lack of integration with financial markets, risk premia, market power by generators, market power in financial arbitrage, transaction costs for arbitraging price differences. This led to conclusions that price differences between the DA and RT markets were evidence that wholesale electricity markets were inefficient. Some of these studies found that price convergence increased following the introduction of virtual bidding into wholesale electricity markets. The increase in price convergence was taken as evidence that market efficiency had improved and that the ascribed causes for price divergence had been mitigated.

However, there is no guarantee that price convergence is an indication that the markets are more efficient. The assumption that price convergence is an indication of market efficiency comes from the study of commodity markets and the role of futures and financial futures in that context. As Parsons et al. (2015) point out the viewpoint of wholesale electricity markets as commodity markets relies on the very strong assumption that the DA and RT markets are a successive iteration of the same bidding in an auction process. The purpose and implementation of the DA market is different from the RT market. It is important to keep in mind that LMPs are not set by repeated interactions by traders but rather they are shadow prices of an optimization model. The idea of price convergence requires that the shadow prices of two different large and complex optimization models match each other. In that context price convergence is not an informative metric of market performance. The purpose of this research is to provide an alternative means to evaluate the impact of virtual bidding using welfare analysis which is the standard measure of economic efficiency. Using this metric, a change in market performance is said to be efficiency improving only when it increases total surplus, that is the summation of consumer and producer surplus. We believe that this approach provides a more transparent assessment of the efficiency impact of arbitrage trading and contributes to the ongoing discussion about the role of virtual bids in wholesale electricity markets.

CHAPTER 4. STYLIZED ELECTRICITY MARKET MODEL FOR INCS AND DECS

In this chapter, we introduce a stylized model that includes many of the salient features of the DA and RT markets. We introduce virtual bidding and derive the optimal bidding strategy for INCs and DECs. We then derive the impact on price convergence as well as social welfare.

4.1 Optimal Bidding Strategy Model

The unit commitment and dispatch problem that the ISO solves in order to determine the DA dispatch schedule is a highly complex problem. In order to make the problem analytically tractable we make a series of simplifications. This analysis abstracts from the network and considers a market that occurs at a single node, analogous to a zone within an ISO defined by an electric distribution company (EDC). Demand is assumed to be perfectly inelastic. In addition, we abstract from the lumpiness associated with the commitment of generators due to fixed costs and minimum operating limits. We assume that market power is not being exercised by either generators or load serving entities, and the only reason for price differences between the DA and the RT markets are demand and supply imbalances due to imperfect knowledge of demand in the DA market. While acknowledging that this is an over simplification of the market we believe that it is sufficient to capture the qualitative efficiency impacts of virtual arbitrage trades with INCs and DECs.

Demand and supply imbalances are driven by uncertainty in the model. We assume that a fixed amount of demand is bid into the DA market, D_{DA}, and that this is the expected level of demand in the RT market. Demand in the RT market, D_{RT}, is the sum of D_{DA} plus a random deviation Δ which represents the uncertainty in the RT demand (1). The random variable Δ has a known density $f(\Delta)$ with an expected value of zero, and the support of Δ is such that D_{DA} is always positive and finite. Thus, the DA demand bid is an unbiased forecast of the RT demand (this assumption will later be relaxed in the numerical analysis).

$$D_{RT} = D_{DA} + \Delta \tag{1}$$

For the DA supply curve, S_{DA}, we will abstract from the lumpy nature of the commitment process and focus only on the marginal supply offers that make up the aggregate supply curve which is continuous in nature. The typical generation fleet is composed of a large quantity of baseload generation available at a low marginal cost and smaller amounts of cycling and peaking generation available at rapidly increasing marginal costs – that is, costs are increasing at an increasing rate. This suggests that the supply curves in our stylized model should be convex. While aggregate supply curves faced by ISOs are step functions for convenience we will use a constant elasticity function characterized by elasticity ε_1 and parameter α_1 to represent S_{DA}. In order to obtain convex curvature elasticity parameters of only one or less will be considered which gives the supply curve either no curvature (for ε_1=1) or convex curvature (for ε_1<1). The DA market price P_{DA} is determined by the inverse supply function in (2).

$$P_{DA} = D_{DA}^{\frac{1}{\varepsilon 1}} \alpha_1^{\frac{-1}{\varepsilon 1}} \tag{2}$$

As previously stated only a subset of generating resources – excess capacity in already committed units and those units that can be brought online very quickly – are available to be dispatched in the RT market. In order to characterize these attributes the RT supply curve, $S_{RT,}$ will have three additional features compared to the DA supply curve. First S_{RT} will be represented by a piecewise constant elasticity function with elasticity ε_2 and parameter α_2. This feature was added to represent that the market will only clear along S_{RT} for higher than expected levels of demand (positive deviations in demand) because it is only in this case that the additional resources are needed. If demand is lower than expected (negative deviations in demand) the market clears along S_{DA} to represent the price being set by units that were committed in the DA market (see Figure 4-1). Second, in order to represent a particular commitment level in the DA market S_{RT} will begin at the point where D_{DA} intersects S_{DA}. This is modeled by making the α_2 parameter a function the DA market outcome as shown in equation (3).

$$\alpha_2 = \frac{D_{DA}}{(P_{DA})^{\varepsilon_2}} = \frac{D_{DA}}{\left((D_{DA})^{1/\varepsilon_1}(\alpha_1)^{-1/\varepsilon_1}\right)^{\varepsilon_2}} \tag{3}$$

Third, S_{RT} will be more inelastic than S_{DA} ($\varepsilon_2 \leq \varepsilon_1$) to represent the higher cost of the fast response resources. The RT market price P_{RT} is determined by the inverse supply function in (4) and the realization from the Δ distribution represented by the parameter δ.

$$P_{RT} = \begin{cases} {D_{RT}}^{1/\varepsilon_1} {\alpha_1}^{-1/\varepsilon_1} & if\ D_{RT} < (D_{DA}) \\ {D_{RT}}^{1/\varepsilon_2} \left[\frac{D_{DA}}{\left((D_{DA})^{1/\varepsilon_1}(\alpha_1)^{1/\varepsilon_1}\right)^{\varepsilon_2}}\right]^{-1/\varepsilon_2} & if\ D_{RT} \geq (D_{DA}) \end{cases} \tag{4}$$

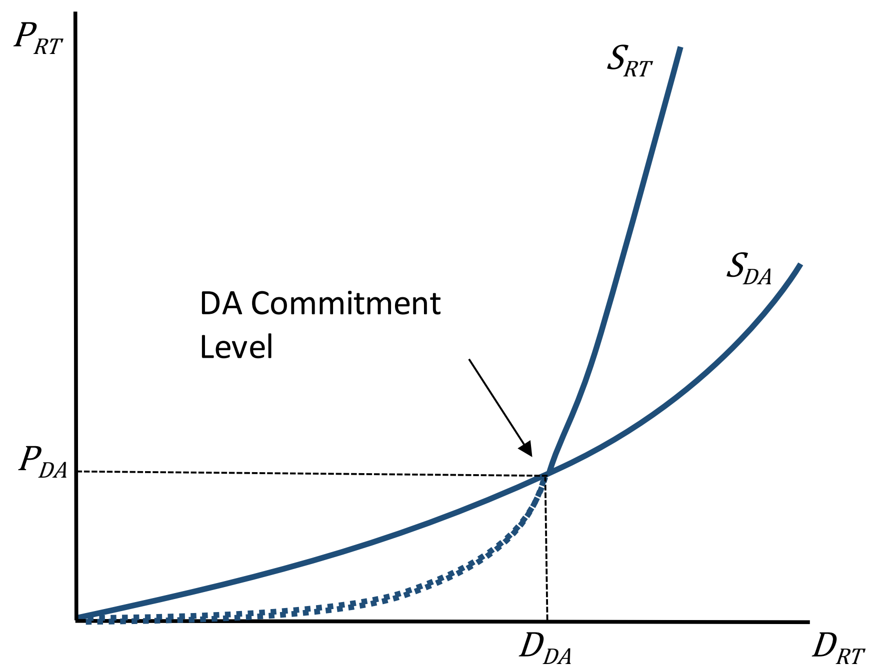


Figure 4-1 Day-Ahead and Real-Time Supply Curves

Financial participants engage in price arbitrage by purchasing and selling virtual MW V in the DA market that are resolved as deviations in the RT market. That is, the virtual bidder must settle a cleared bid through either purchase or sale of energy in the RT market. In this model if V is positive the transaction is equivalent to a DEC bid and if V is negative the transaction is equivalent to an INC offer. While financial participants can specify a reservation price, for simplicity it will be assumed that they are price takers, accepting the DA market clearing price. Thus, we treat selecting the level of V as the sole choice of the virtual bidder.

The objective for financial participants is to maximize profit π by choosing to buy or sell V (MW) in the DA market given that the V that they choose will have an impact on the price in both the DA and the RT markets. By increasing or decreasing the clearing level of demand in the DA market to D_{DA} + V, virtual bids set the DA market price P_{DA}^{V} as shown in (5).

$$P_{DA}^{V} = (D_{DA} + V)^{1/\varepsilon_1} {\alpha_1}^{-1/\varepsilon_1} \tag{5}$$

The bid or offer V impacts the RT market price through its impact on the commitment level in the DA market. The impact on commitment is a shift of S_{RT} to the intersection between S_{DA} and $D_{DA} + V$ (see Figure 4-2 for the case where $V > 0$, or a DEC bid clears). The DA and RT supply curves will intersect because the outcome of the DA market determines the α_2 parameter as represented by equation (6).

$$\alpha_2 = \frac{D_{DA}+V}{\left(P_{DA}^{V}\right)^{\varepsilon_2}} = \frac{D_{DA}+V}{\left((D_{DA}+V)^{1/\varepsilon_1}(\alpha_1)^{-1/\varepsilon_1}\right)^{\varepsilon_2}} \tag{6}$$

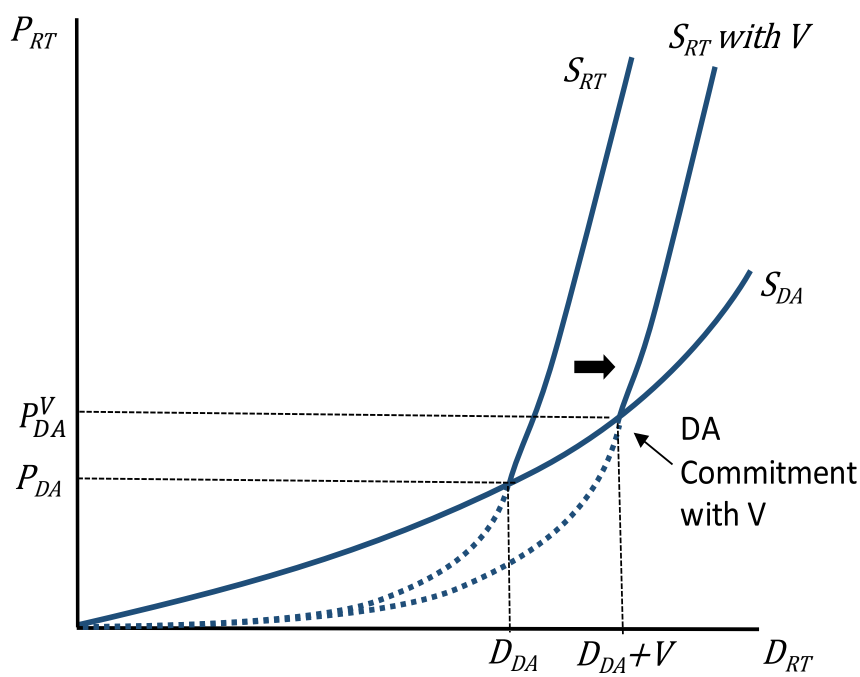


Figure 4-2 Real-Time Supply Curve Shift

The RT price with virtuals P_{RT}^{V}, represented by equation (7), is a piecewise constant elasticity function.

$$P_{RT}^{V} = \begin{cases} {D_{RT}}^{1/\varepsilon_1} {\alpha_1}^{-1/\varepsilon_1} & if\ D_{RT} < (D_{DA} + V) \\ {D_{RT}}^{1/\varepsilon_2} \left[\frac{D_{DA}+V}{\left((D_{DA}+V)^{1/\varepsilon_1}(\alpha_1)^{1/\varepsilon_1}\right)^{\varepsilon_2}}\right]^{-1/\varepsilon_2} & if\ D_{RT} \geq (D_{DA} + V) \end{cases} \tag{7}$$

If RT demand is less than the DA demand plus virtual bids, the price is set along the portion of the RT supply curve that is identical to the DA supply curve, representing price being set by already committed units. If RT demand is greater than or equal to DA demand plus virtual bids, price is set along the portion of the RT supply curve that is less elastic, representing price being set by newly committed units.

A financial participant's objective function for either DECs or INCs can be represented by equation (8).

$$\max_{V} \pi : -{\alpha_1}^{1/\varepsilon_1}(D_{DA} + V)^{-1/\varepsilon_1} V$$

$$+ \int_{dlow}^{V} (D_{DA} + \Delta)^{1/\varepsilon_1} (\alpha_1)^{-1/\varepsilon_1}\, V df(\Delta)$$

$$+ \int_{V}^{dhigh} (D_{DA} + \Delta)^{1/\varepsilon_2} \left[\frac{D_{DA}+V}{\left((D_{DA}+V)^{1/\varepsilon_1}(\alpha_1)^{-1/\varepsilon_1}\right)^{\varepsilon_2}}\right]^{-1/\varepsilon_2} V df(\Delta) \tag{8}$$

The first term represents the cost or revenue from the DA transaction. The second term represents the settlement outcome in the RT market for the case when demand in the RT market is lower than the quantity that cleared the DA market, a negative deviation. When this occurs P_{RT} is set by the DA supply curve S_{DA}. This represents the case of having to ramp down resources that were committed in the DA market. The third term represents the settlement outcome in the RT market for the case when demand in the RT market is higher than the quantity that cleared in the DA market, a positive deviation. When this occurs P_{RT} is set along the less elastic portion of the RT supply curve, representing the case of having to commit more expensive fast-start resources to satisfy the higher demand in the RT market. The optimal bidding strategy V^* is found by solving for the first order conditions and solving for V, where *dlow* denotes the lower limit on the support for delta and *dhigh* denotes the upper limit on the support for delta. Note that this objective can be viewed as $-P_{DA}(D_{DA}+V)V + E[P_{RT}(D_{DA}+\Delta)]V$. In the case where $E[\Delta]=0$ and P_{RT} is convex, a DEC will always be profitable as is shown in the following theorem.

<u>Theorem:</u>

If P_{RT} is strictly convex and $E[\Delta]$ is zero, then there exists $V > 0$ that is optimal.

<u>Proof:</u>

By strict convexity of P_{RT} and Jensen's Inequality,

(1) $E[P_{RT}(D_{DA}+\Delta)] > P_{RT}(D_{DA}+E[\Delta]) = P_{DA}(D_{DA})$

By continuity of P_{RT}, there exists a V that is small enough so that

(2) $E[P_{RT}(D_{DA}+\Delta)] - P_{DA}(D_{DA}+V) > 0,$

and for $V > 0$

(3) $E[P_{RT}(D_{DA}+\Delta)]V - P_{DA}(D_{DA}+V)V > 0.$

Hence a strictly positive DEC is profitable when the day ahead demand forecast is unbiased. Q.E.D.

4.2 INC and DEC Bidding Model Results

Even in the case where the distribution of the demand deviation between the DA and RT markets is assumed to be uniform on [*dlow*, *dhigh*], the first order condition of the virtual bid optimization model does not appear to have a closed form solution. Thus to analyze the behavior of the model we make this distributional assumption and solve the model numerically. For this analysis the DA supply curve was benchmarked at a price of $100 for a 100 MW load. The model is solved for a bidding strategy that maximizes the expected revenue from the virtual bid minus its cost. The model was solved for different combinations of DA and RT supply elasticities in order to explore how changing these affects the optimal bidding strategy. The initial calculations are for the case where the expected demand deviation between DA and RT markets is zero – i.e., *dlow* = –*dhigh*. We take as our base case the situation where load is bid at 100 MW in the DA market by the physical participants. Demand uncertainty is ± 10 percent or 10 MW represented by a uniform distribution function, making the expected demand deviation in the real time market zero – thus, DA demand is an unbiased forecast of RT demand. Figure 4.3 displays the numerical solution for the optimal bidding strategy V^*. The two horizontal axes are for the elasticity parameters ε_1 and ε_2, and the vertical axis is for the optimal level of virtual demand, which represents a DEC if positive, or an INC if negative. Only cases in which the RT supply curve was at least as inelastic as the DA supply curve ($\varepsilon_2 \leq \varepsilon_1$) were considered.

Consistent with the theorem in the previous section, the results show that under the circumstances of our base case it is always optimal to bid DEC MW. The results show that the more inelastic the RT supply curve is relative to the DA supply curve (lower ε_2 values relative to ε_1 values), the greater the optimal quantity of DEC MW. The existence of this strategy is the result of the convex nature of the RT supply curve. Because of this if the RT deviations are symmetric around zero, prices changes will be greater in magnitude for positive deviations compared to price changes for negative deviations.

Thus while bidding a DEC incurs a loss half the time, the gains from the other half are greater in magnitude, leading to a positive expected profit. Thus as the RT supply curve becomes more inelastic, expected profits increase and the size of the optimal DEC bid *V** also increases. Similarly, a more elastic DA supply curve, which is the same as the RT supply curve below the DA demand, also leads to greater price changes for positive relative to negative deviations. Hence as the DA supply becomes more elastic (ε_1 increases), the expected profits and size of the optimal DEC bid both increase. This relationship is illustrated in Figure 4.3.

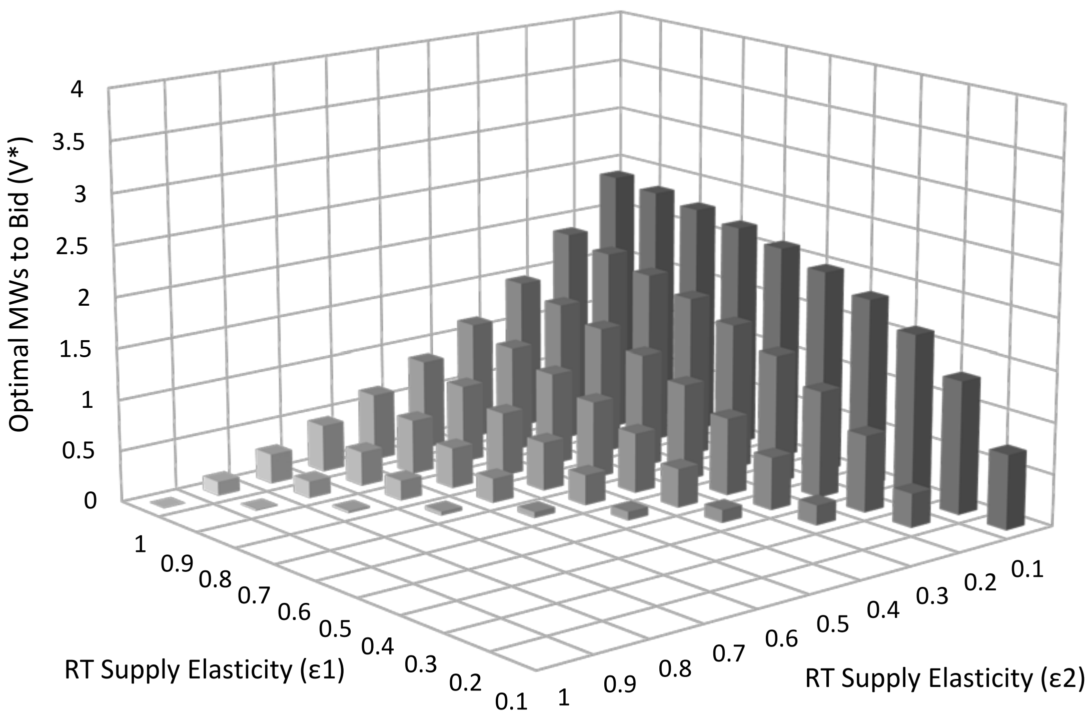


Figure 4-3 Unbiased Day-Ahead Demand: Real-Time Deviation Δ~U(-10,10)

Figure 4.4 shows the results for a smaller demand uncertainty (plus or minus five percent of DA demand), and Figure 4.5 shows that results for greater demand uncertainty (plus or minus fifteen percent). The results indicate that increasing or decreasing the uncertainty has no effect on whether DECs are profitable. However, uncertainty does affect the magnitude of the optimal position with higher uncertainty leading to a larger DEC bid and lower uncertainty leading to a lower DEC bid.

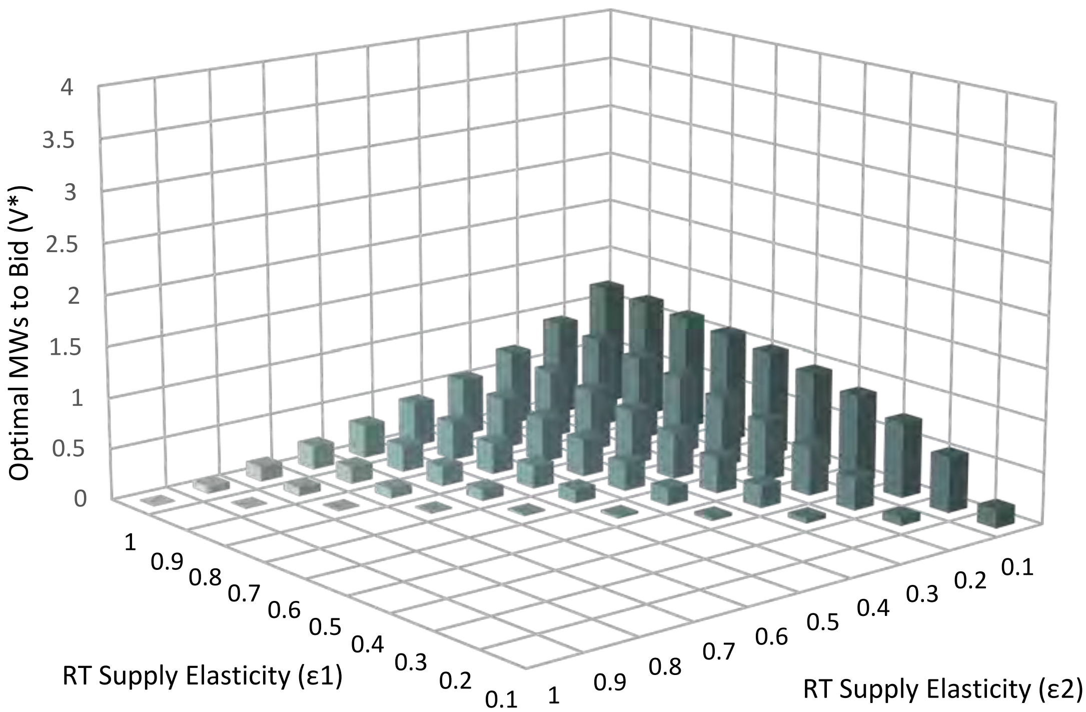


Figure 4-4 Unbiased Day-Ahead Demand: Real-Time Deviation Δ~U(-5,5)

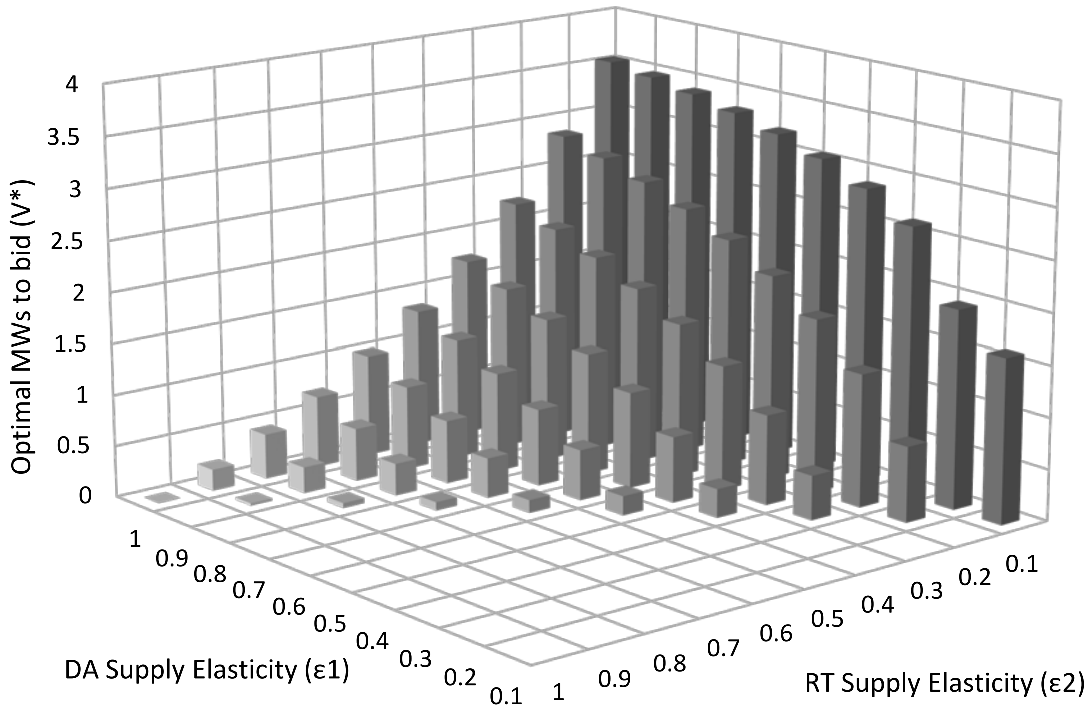


Figure 4-5 Unbiased Day-Ahead Demand: Real-Time Deviation Δ~U(-15,15)

Next we assess the impact of bidding strategies on price convergence. This analysis was conducted with the DA elasticity parameter ε_1 equal to 0.6, and the RT elasticity parameter ε_2 equal to 0.3. Figure 4.6 shows the expected DA and RT prices for cases with an unbiased DA expectation of RT load. Given that real time supply is more inelastic than day ahead supply, it was optimal to bid DEC MW. The results show that the optimal bidding strategy leads to greater expected price convergence compared to cases without virtual bidding. These results are consistent with the theoretical and empirical findings of Woo et al. (2015).

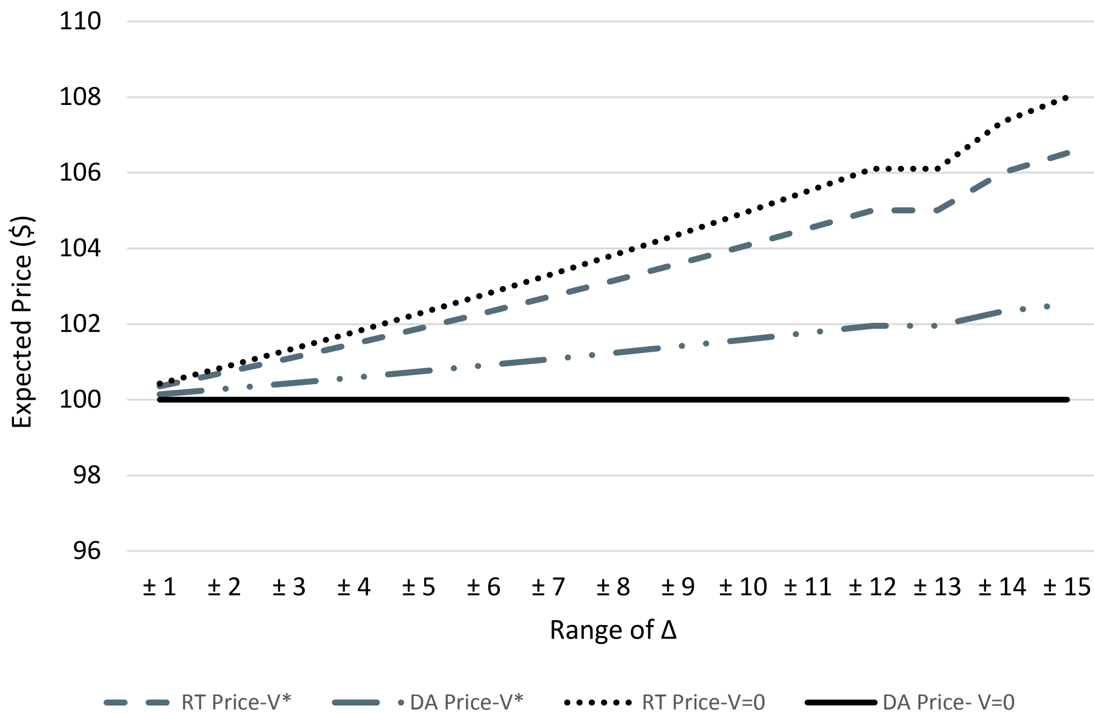


Figure 4-6 Expected Market Prices with and Without Virtual Bidding (V=0 and V* respectively) for Cases with an Unbiased Day-Demand (E(Δ)=0).

All results so far have focused on the case where the expected demand deviation has mean zero. We now consider cases where there is bias in the DA expectation of demand. Figure 4.7 displays the results for the case when there is a negative bias in the DA expectation of demand (leading to a positive expected deviation in the RT market) equivalent to three percent of the load (i.e. Δ is distributed uniformly on [-7,13]). The results show that for almost all combinations of ε_1 and ε_2 it is optimal to bid a DEC. Figure 4.8 displays the optimal bid for the case when there is a positive bias in the DA expectation of demand (leading to a negative expected deviation in the RT market) equivalent to three percent of the load (i.e. Δ is distributed uniformly on [-13,7]). In this case it is optimal to bid an INC ($V < 0$) for almost all combinations of ε_1 and ε_2.

Even though the magnitude of the biases are equivalent we find that it is not optimal to bid as many INC MW when there is a positive bias compared to DEC MW when there is a negative bias. This is the result of our assumptions about the convexity of the supply curves, resulting in price spikes that are greater for positive deviations in demand compared to negative deviations. For cases when there is a bias in the DA demand, the results showed that for the vast majority of elasticity combinations it is not optimal to bid a quantity of virtual MW equivalent to the expected magnitude of the bias. This is because profitability for financial participants depends on the quantity of virtual MW and the price difference between the DA market and the RT market. Bidding in a quantity of virtual MW equivalent to the expected deviation would reduce the price difference and consequently profits as well. Thus financial participants have an incentive to only partially eliminate differences between the DA and the RT load. This is important because it means that if the market is not competitive and/or there are transactions costs it cannot be expected for financial participants to completely eliminate price differences.

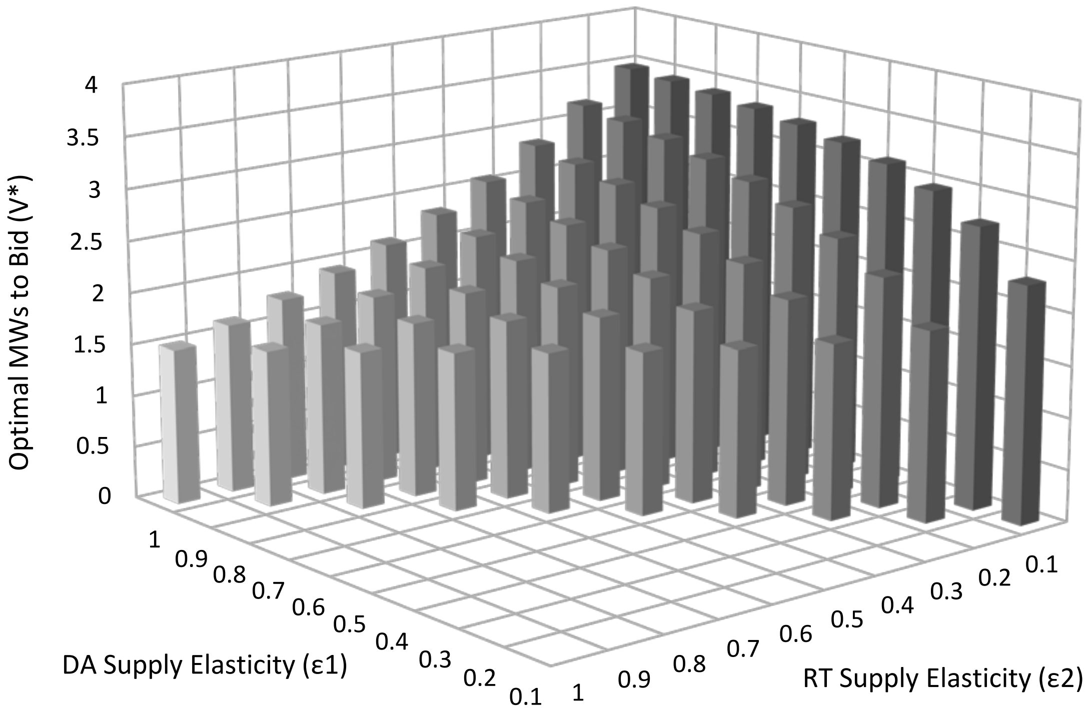


Figure 4-7 Biased Day-Ahead Demand E(Δ)=3: Real-Time Deviation Δ~U(-7,13)

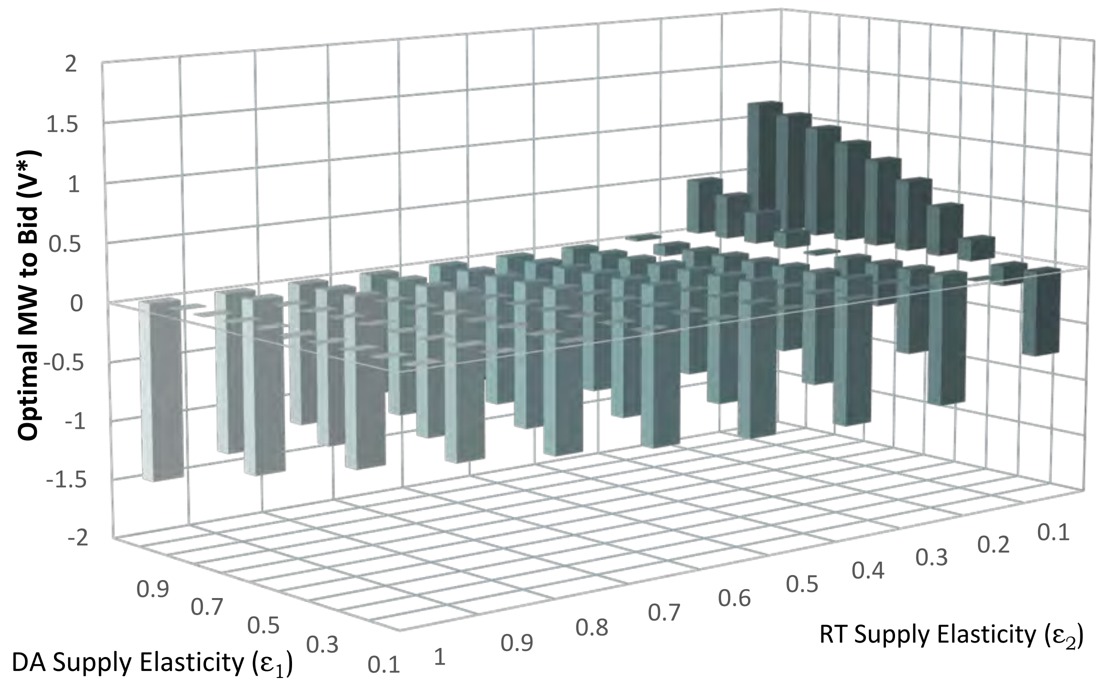


Figure 4-8 Biased Day-Ahead Demand E(Δ)=-3: Real-Time Deviation Δ~U(-13,7)

Figure 4.9 shows the expected market prices for cases when there is a bias in the DA expectation of RT load. In these cases, the optimal bidding strategy also leads to greater expected price convergence compared to a case without virtual bidding. For positive deviations (negative bias) prices converge at a higher level because the optimal bidding strategy requires DEC MW. For cases with negative deviations (positive bias) prices converge at a lower level because the optimal bidding strategy requires INC MW.

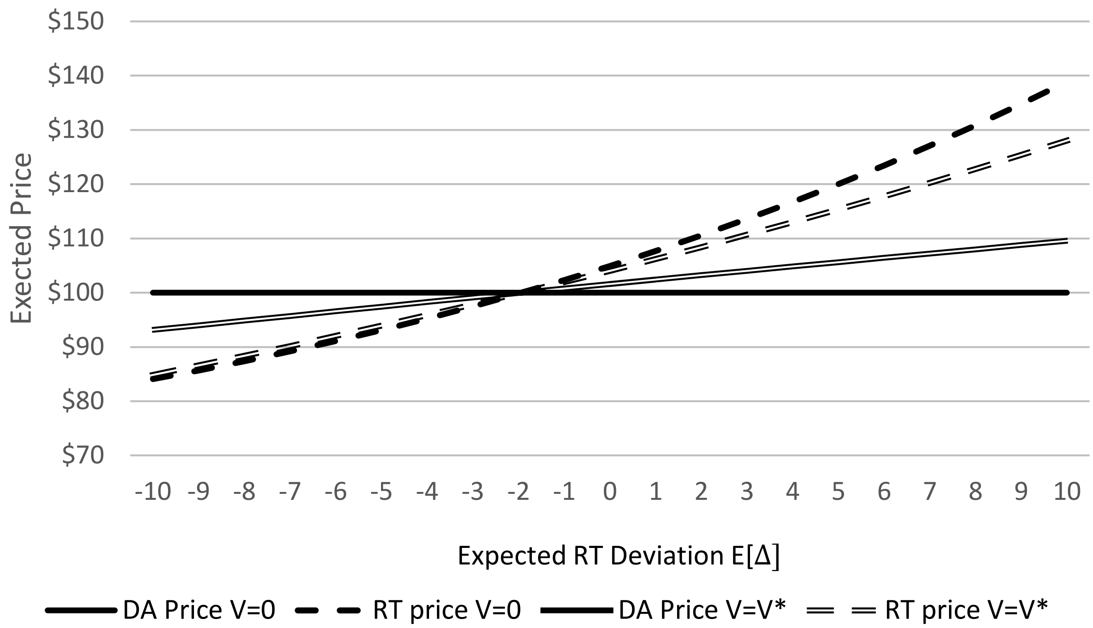


Figure 4-9 Expected market prices with and without virtual bidding (V* and V=0 respectively) for cases with a biased DA demand.

Our stylized model shows the possible existence of bidding strategies for virtual transactions whose profitability depends not on providing additional information to the market but rather on taking advantage of asymmetric price changes. While these strategies can provide for price convergence on average, the expected prices in both the DA and RT markets may be higher than in the absence of the virtual transactions. This is why it is important to look beyond average price convergence to assess the impact of virtual trading. The welfare analysis, addressed in the next section, will take into account the impact of virtual bidding both when it causes price convergence and price divergence.

4.3 Welfare Calculations for INC and DEC Model

This section presents the formulas for calculating expected social welfare given the uncertainty of final demand in the RT market and the optimal bidding strategy by financial participants. Expected consumer surplus is calculated as the expected value to consumers minus expected DA and RT market payments made by consumers as represented by equation (9).

$$E[CS] = E[Consumer\ Value] - E\big[[DA\ Payments_{consumers}] + E[RT\ payments_{consumers}]\big] \tag{9}$$

Expected producer surplus is calculated as expected DA and RT market payments to producers by both consumers and financial participants minus expected generation cost as represented by equation (10).

$$E[PS] = E\big[[DA\ Payments_{consumers}] + E[RT\ payments_{consumers}]\big] + E\left[\big[DA\ Payments_{financial}\big] + E\big[RT\ payments_{fianancial}\big]\right] - E[Generation\ Cost] \tag{10}$$

Since we have assumed that electricity demand is fixed, the value of electricity to consumers will be represented by the reservation price *RP* which will be assumed to have a large fixed value so that it is always greater than the electricity price. Equation (11) presents the formula for calculating the expected value of electricity to consumers given the uncertainty parameter Δ and the quantity of physical MW bid in the DA market *d*.

$$E[Consumer\ Value] = \frac{\int_{dlow}^{dhigh}(d+\Delta)RPd\Delta}{dhigh - dlow} \tag{11}$$

DA payments by consumers are equivalent to the DA market price (which is set by the physical and virtual demand) times the quantity of physical demand. Equation (12) represents the DA market payments given the quantity of physical MW *d* and the virtual MW *V* that are bid into the DA market.

$$DA\ Payments_{consumers} = (d+V)^{1/\varepsilon_1}\alpha_1^{-1/e_1}d \tag{12}$$

DA payments made by financial participants are equivalent to the DA market price times the quantity of virtual MW as represented by equation (13).

$$DA\ Payments_{financial} = (d+V)^{1/\varepsilon_1}\alpha_1^{-1/e_1}V \tag{13}$$

In the RT market payments by consumers are known as balancing payments. If there is a negative deviation in demand these will be negative since the ISO pays consumers back for any electricity purchased in the DA market that was not consumed in the RT market. If there is a positive deviation in demand, balancing payments will be positive since this represents additional purchases of electricity. All deviations are quoted at the RT market price. The expected RT payments is represented by equation (14).

$$E[RT\ Payments_{consumers}] = \frac{\int_{dlow}^{V}(d+\Delta)^{1/\varepsilon_1}\alpha_1^{-1/\varepsilon_1}\Delta d\Delta + \int_{V}^{dhigh}(d+\Delta)^{1/\varepsilon_2}\alpha_2^{-1/\varepsilon_2}\Delta d\Delta}{dhigh - dlow} \tag{14}$$

For cases without virtuals V would be equal to zero. For these cases the first integral represents payments for negative deviations and for which the RT price is set by the DA supply curve. The second integral represents payments for positive deviations for which the RT price is set by the RT supply curve. Non-zero values of V will shift the intersection of the DA supply curve and the RT supply curve and thus the boundaries at which the RT price is set.

Positive values of V (DEC MW) cause a rightward shift of the RT supply curve. The price for positive deviations up to V is now set by the DA supply curve. This lowers the payments for positive deviations while having no effect on the price for negative deviations. Negative values of V (INC MW) cause a leftward shift of the RT supply curve. This will increase the price for all positive deviations and negative deviations up to

V. The price for negative deviations below V are not affected and their price is set by the DA supply curve.

In the RT market the DA positions of financial participants are automatically settled by the ISO at the RT market price. Thus payments by financial participants are equivalent to the opposite position taken in the DA market ($-V$) times the RT market price as presented in equation (15).

$$E\left[RT\ Payments_{financial}\right] = \frac{\int_{dlow}^{V}(d+\Delta)^{\frac{1}{\varepsilon_1}}\alpha_1^{-\frac{1}{\varepsilon_1}}(-V)d\Delta + \int_{V}^{dhigh}(d+\Delta)^{\frac{1}{\varepsilon_2}}\alpha_2^{-\frac{1}{\varepsilon_2}}(-V)d\Delta}{dhigh-dlow} \qquad (15)$$

The generation cost is calculated using the RT supply curve which has a kink but it is continuous. The kink occurs at the DA clearing demand d. For realized demand values up to d the cost is calculated as the area under the inverse supply function with parameters α_1 and ε_1 which corresponds to the DA supply curve as seen by the grey area in Figure 4.10. For realized demand values above d the cost is calculated as the area under the DA inverse supply up to d, plus the area under the inverse supply function with parameters α_2 and ε_2 for values from d to the realized demand value as seen in by the grey area in Figure 4.11.

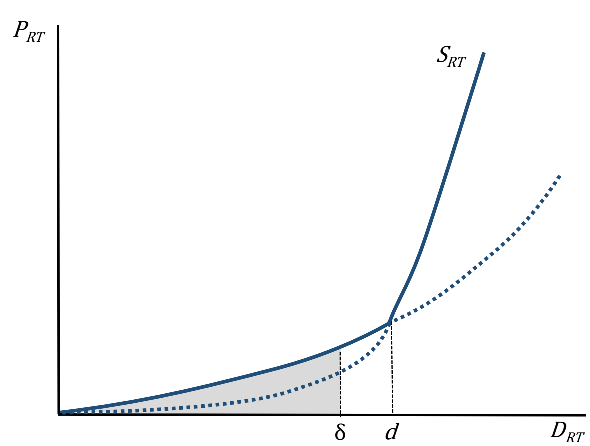


Figure 4-10 Cost for realized demand value δ≤d

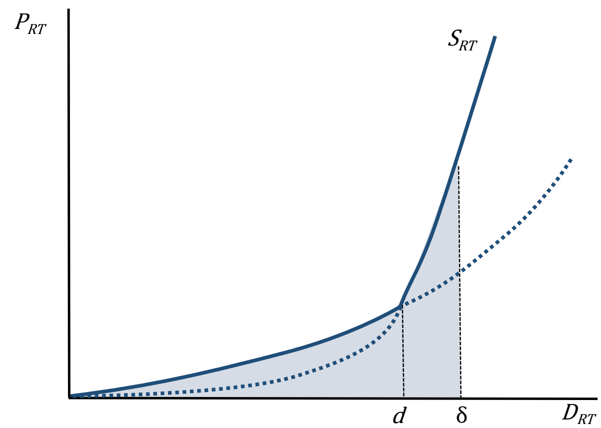


Figure 4-11 Cost for realized demand value δ>d

The expected generation cost is the expectation taken over Δ of the area under the inverse RT supply function as represented by the grey area in Figure 4.12 for a case without virtuals. The expected generation cost is represented by equation (12).

$$E[Generation\ \ Cost] = \frac{1}{(dhigh - dlow)}\left[\int_{dlow}^{V} \frac{\varepsilon_1 (d+\Delta)^{\frac{1+\varepsilon_1}{\varepsilon_1}} \alpha_1^{\frac{-1}{\varepsilon_1}}}{1+\varepsilon_1} d\Delta + \int_{V}^{dhigh}\left(\frac{\varepsilon_2 (d+\Delta)^{\frac{1+\varepsilon_2}{\varepsilon_2}} \alpha_2^{\frac{-1}{\varepsilon_2}}}{1+\varepsilon_2} - \frac{\varepsilon_2 (d+V)^{\frac{1+\varepsilon_2}{\varepsilon_2}} \alpha_2^{\frac{-1}{\varepsilon_2}}}{1+\varepsilon_2} + \frac{\varepsilon_1 (d+V)^{\frac{1+\varepsilon_1}{\varepsilon_1}} \alpha_1^{\frac{-1}{\varepsilon_1}}}{1+\varepsilon_1}\right) d\Delta\right] \quad (12)$$

For cases without virtuals V is equal to zero and the first integral captures the cost for cases when there are negative deviations in demand while the second integral captures the cost for cases when there are positive deviations in demand. Non-zero values of V impacts both the place at which the kink occurs and how much area is captured under the different parts of the inverse RT supply function. The former impact is captured through the α_2 parameter as described in equation (6) and the latter impact is captured through the inclusion of V in the limits of the integrals.

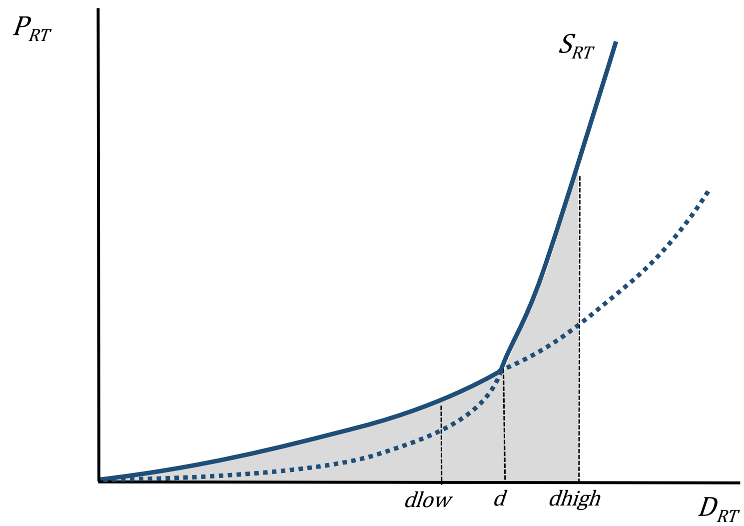


Figure 4-12 Expected generation cost with no virtuals.

4.4 Welfare Results for INC and DEC Model

This section explores the welfare impact of virtual bidding for different scenarios. This section will use as an input the optimal bidding strategy V^* from the model in the previous section using a uniform distribution to model the RT demand uncertainty parameter Δ. Given the many possible combinations of elasticities for the DA and RT supply curves, a base case with the elasticity for DA supply curve equal to 0.6 and the elasticity for the RT supply curve at 0.3 was chosen to estimate V^* and the resulting market prices. These were chosen to capture the more inelastic nature of the RT supply curve relative to the DA supply curve. The results show that virtual bidding has a marked effect on welfare. Higher or lower amounts of virtual bids resulting from different combinations of DA and RT supply curve elasticities would amplify or reduce the effects respectively. The results are estimated for a representative hour in the DA and RT markets where 100 MW of physical load are bid in the DA market and there is some uncertainty about the final realization of demand. The DA supply curve is benchmarked at a price of $100 for the 100 MW load.

The first analysis involves comparing the welfare outcomes for cases when the DA physical load is an unbiased expectation of the RT load ($E[\Delta]=0$) but there is uncertainty about the final realization of demand. Figure 4.13 displays the changes in consumer and producer surplus (measured in dollars) for cases with and without virtuals for different ranges of demand uncertainty compared to an outcome without uncertainty (thus there are no deviations in the RT market). The results indicate that for outcomes without virtuals ($V=0$) demand uncertainty causes surplus transfers from consumers to producers and this effect increases with the level of uncertainty. This occurs because the price increase for positive deviations is larger than the price decrease for negative deviations.

Given the base case assumption where the RT supply curve is more inelastic than the DA supply curve it is optimal to bid DEC MW for all cases when the RT deviations are unbiased. The results show that the optimal bidding strategy ($V=V^*$) exacerbates the welfare transfers from consumers to producers caused by the demand uncertainty. This occurs because the DEC MW increases the DA market price which is the price at which the majority of electricity is traded. The results show that this effect increases with the magnitude of the uncertainty as it is optimal to bid higher quantities of DEC MW.

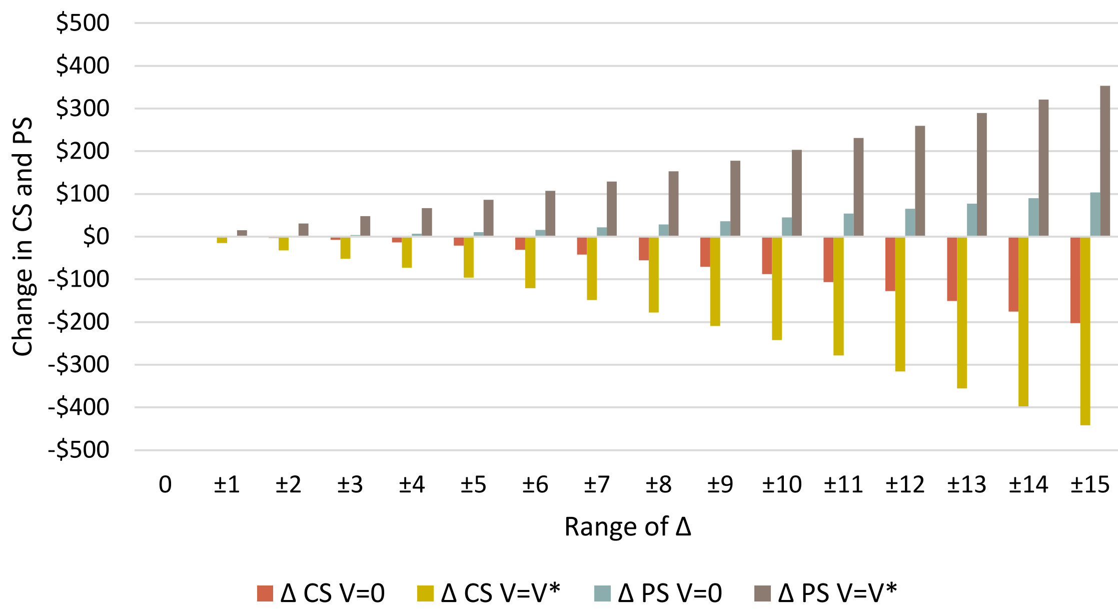


Figure 4-13 Changes in consumer and producer surplus for outcomes with (V=V*) and without (V=0) virtual bidding for different ranges of unbiased demand uncertainty compared to outcomes without demand uncertainty.

Figure 4.14 displays the changes in total surplus (measured in dollars) for cases with and without virtuals for different ranges of demand uncertainty compared to an outcome without uncertainty as well as the profit for financial participants when they bid V*. The results show that in general demand uncertainty causes a decrease in market efficiency as measured by total surplus (results for V=0), and this effect increases with the level of uncertainty. For cases with the optimal bidding strategy (V=V*) there is a small increase in market efficiency compared to outcomes without virtuals. Figure 4.14 also shows that profit for financial participants is relatively low, however profits increase with increasing uncertainty.

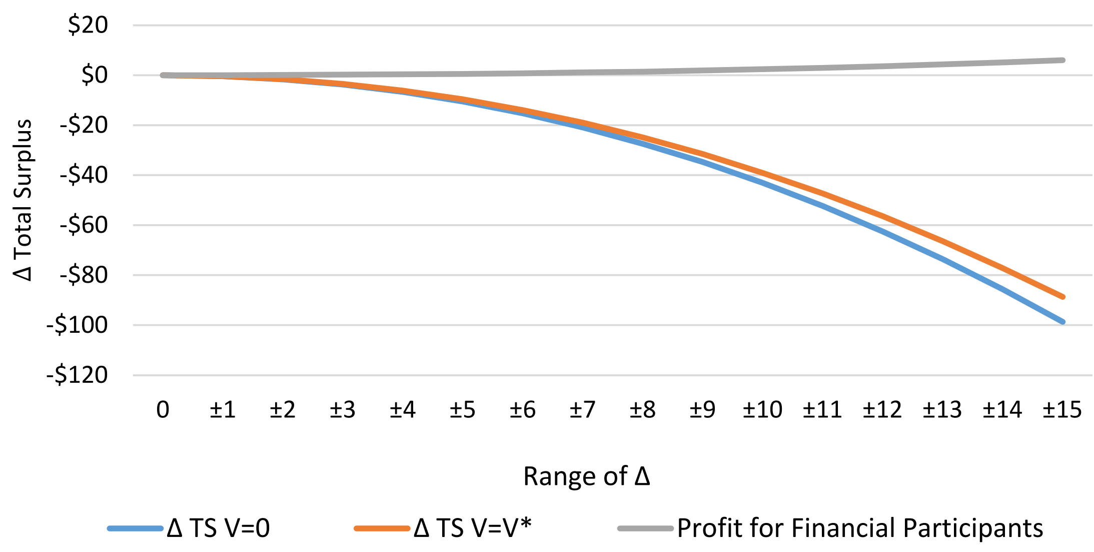


Figure 4-14 Changes in total surplus for outcomes with (V=V*) and without (V=0) virtual bidding for different ranges of unbiased demand uncertainty compared to outcomes without demand uncertainty and profit for financial participants when they bid.

Next we analyze the impact of virtual bidding for cases when there is a bias in the RT load deviations which occurs when the DA load is consistently under or over bid. This analysis will compute changes in surplus for cases with and without virtuals and compare it to cases with no bias and no uncertainty but the same quantity of MW consumed. For example, for all of the analyses we have assumed that 100 MW of physical load are always bid into the DA market. For a case with an expected deviation of 2 MW ($E[\Delta]=2$) the expected consumption is 102 MW. Thus changes in welfare with and without virtuals are compared to a case where the DA physical load is 102 MW with no deviations in the RT market. In this manner we are comparing outcomes with and without virtuals to a perfect foresight “ideal” outcome.

First we analyze the impact of virtual bidding for cases when there is physical load underbidding in the DA market. Load underbidding leads to positive deviations in the RT market (E[Δ]>0). Figure 4.15 displays changes in consumer surplus and producer surplus measured in dollars for different levels of expected positive demand deviations and a range of 20 MW for Δ representing the level of uncertainty. The results for cases without virtual bidding (V=0) show that load underbidding in general causes welfare transfers from producers to consumers. These transfers occur because the lower DA clearing MW lowers the DA market price at which most of the electricity is purchased, thus benefiting consumers over producers. The larger the proportion of load underbidding the larger the surplus transfers.

Given our assumptions about the supply elasticities, load underbidding causes RT prices to be higher than DA prices as more inelastic generation resources have to be used in the RT market to meet the unscheduled demand. In this situation the optimal bidding strategy is to bid DEC MW. The higher the magnitude of load underbidding, the higher the optimal quantity of DEC MW to bid. The results with the optimal bidding strategy (V=V*) show a large reduction in the surplus transfers caused by the load underbidding. This occurs because the DEC MW increase the demand level and thus the DA price. Thus the results indicate that the optimal virtual bidding strategy provides for a correction in the distribution of surplus that brings the market outcome closer to an outcome without bias and demand uncertainty.

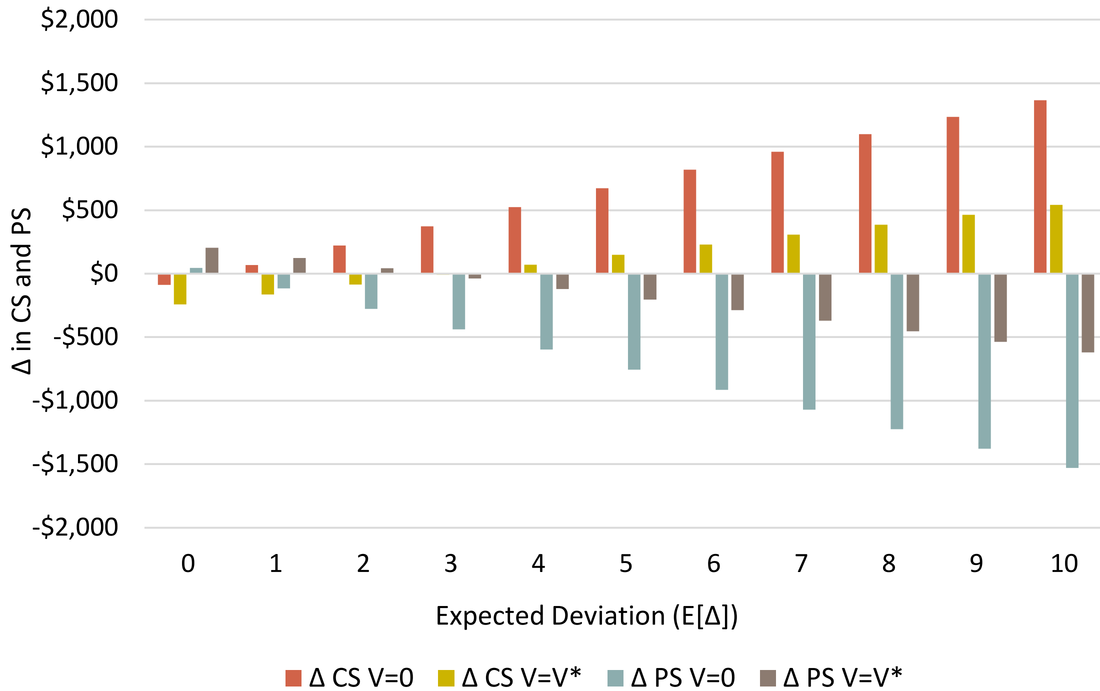


Figure 4-15 Changes in consumer and producer surplus for outcomes with (V=V*) and without (V=0) virtual bidding and load underbidding compared to the "ideal" outcomes without deviations.

Cases with physical load overbidding lead to negative expected deviations in the RT market (E[Δ]<0). Figure 4.16 displays changes in consumer surplus and producer surplus measured in dollars for different levels of expected negative demand deviations and a range of 20 MW for Δ representing the level of uncertainty. The results for cases without virtual bidding (*V*=0) show that load overbidding in general causes welfare transfers from consumers to producers. These transfers occur because the higher DA clearing MW increases the DA market price above the marginal cost of generation, benefiting producers over consumers. It is also the case here that the larger the proportion of load overbidding the larger the surplus transfers.

With load overbidding DA prices are higher than RT prices for which the optimal bidding strategy is to bid INC MW. The greater the magnitude of load overbidding, the greater the optimal quantity of INC MW to bid. The results with the optimal bidding strategy (*V*=*V**) also show a reduction in the welfare transfers caused by the load overbidding. These occur because the INC MW reduce DA demand which lowers the DA prices and thus DA payments made by consumers. For these cases the optimal bidding strategy also provides for a correction in the distribution of surplus that brings the market outcome closer to an outcome without bias and demand uncertainty.

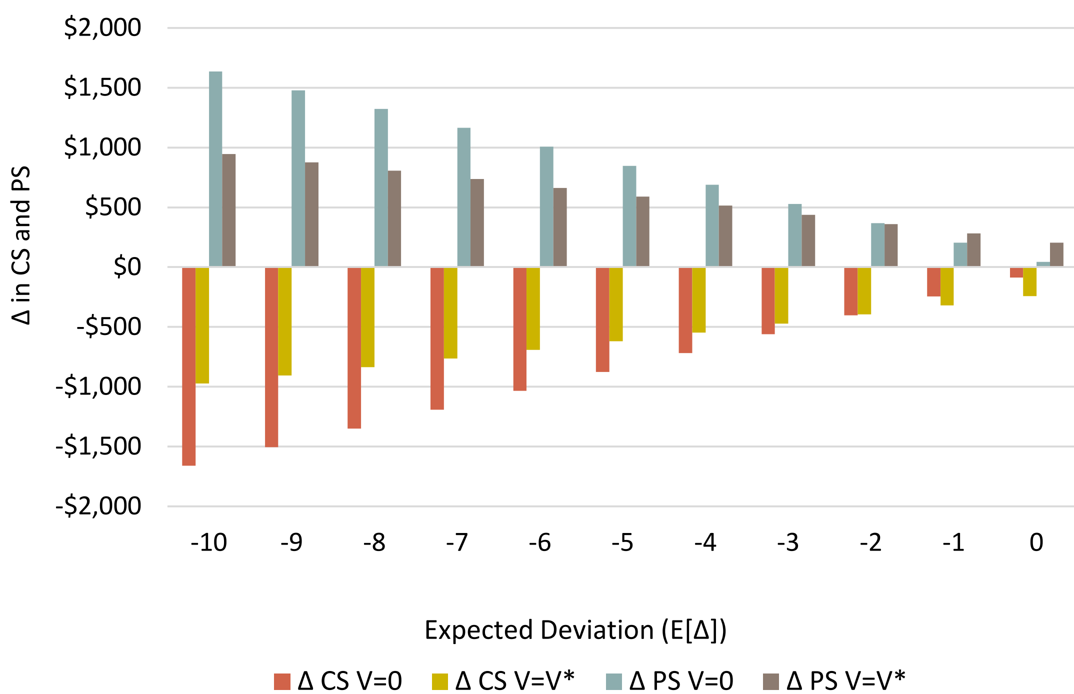


Figure 4-16 Changes in consumer and producer surplus for outcomes with (V=V*) and without (V=0) virtual bidding and load overbidding compared to the “ideal” outcomes without deviations.

Figure 4.17 displays the changes in total surplus for cases without virtual bidding and cases with the optimal bidding strategy compared to the “ideal” outcomes for different levels of expected RT deviations as well as profit for financial participants when they bid

V*. The results show that without virtual bids (*V*=0) the over and under bidding of demand along with demand uncertainty result in lower levels of total surplus compared to outcomes with no uncertainty. This effect is more pronounced for cases with load underbidding (E[Δ] >0) compared to cases with load overbidding (E[Δ] <0). This occurs because load underbidding leads to a higher reliance in the more expensive fast start-up generators to cover unscheduled demand, leading to higher generation costs and higher prices for consumers.

With the optimal bidding strategy (*V*=*V**) there are no significant differences in changes in total surplus for cases with demand overbidding. Thus the only impact for these cases is a redistribution of surplus and not any increase in market efficiency. For cases with demand underbidding the results show that the optimal bidding strategy increases total surplus compared to outcomes without virtuals. This occurs because by increasing demand in the DA market virtuals contribute to an improvement in the scheduling of resources which leads to a decrease in generation costs. For cases with and without virtual bidding the impact on total surplus of under- and overbidding is much smaller in magnitude compared to the welfare transfers that these cause.

Profits for financial participants is greater the larger the magnitude of the bias. For cases when there is load overbidding (E[Δ]) < 0) profit for financial participants is positive despite there being no increase in total surplus as a result of arbitrage activity. For cases when there is load underbidding (E[Δ] > 0) profit for financial participants is greater compared to cases of load overbidding. The results indicate that while arbitrage activity can increase efficiency by improving commitment for cases where load underbidding is occurring, financial participants keep a bounty of approximately half the increase in welfare in the form of profits.

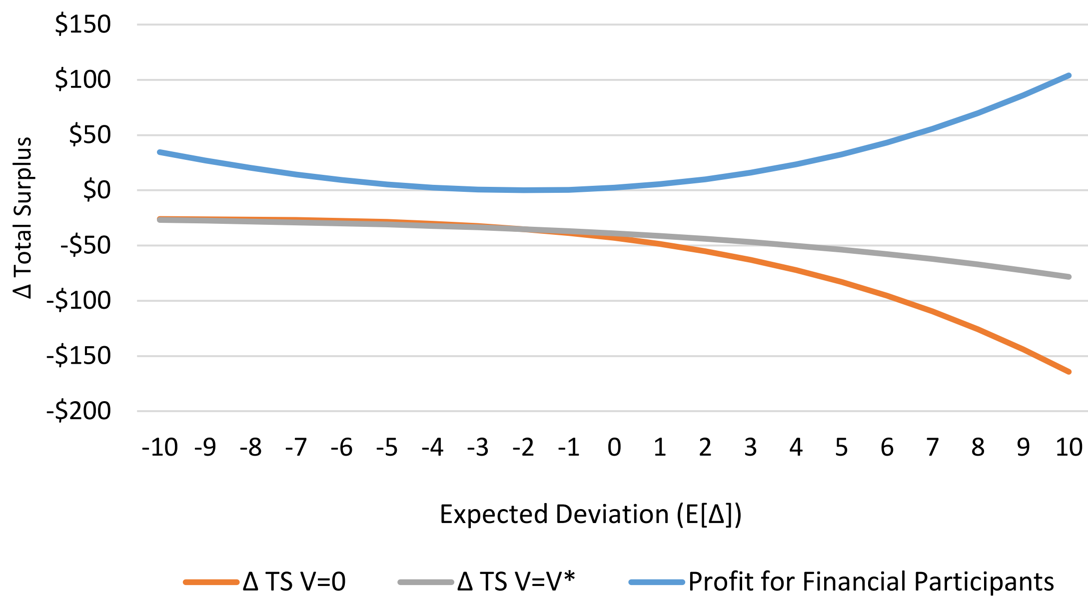


Figure 4-17 Changes in consumer, producer, and total surplus for outcomes with (V=V*) and without (V=0) virtual bidding and load under and over bidding compared to the "ideal" outcomes without deviations and profit for financial participants when they bid V.

4.5 Discussion of INC and DEC Model Results

The welfare analysis in our research provides a very different perspective on the impact of virtual bidding on wholesale electricity markets compared to the existing literature. In general, the results show that the main impact of virtual bidding is surplus reallocation and less so an impact on total surplus or efficiency. There are several reasons why this is the case. When load serving entities consistently bid the vast majority of their demand into the DA market (95 percent on average in the case of PJM, (PJM 2015)), there is very little room for any increases in efficiency that result from improvements in the scheduling and commitment of generation units in the DA market. When virtual bidding helps to improve the commitment of resources, some of the benefits from the gains in efficiency may accrue to consumers or generators, but some of the efficiency gains are also captured by financial participants in the form of profits from arbitrage activity.

The impact of virtual bidding on surplus allocation is strong in our modeling results. DECs raise prices and benefit generators at the expense of consumers while INCs have the opposite effect. The results show that for cases when there is a bias (either positive or negative) in the expectation of RT load in the DA market, virtual bidding does help to reduce the distortions in surplus allocation that the bias may cause. However, for cases when the DA demand is an unbiased expectation of RT load (and even for cases with small positive biases) the market incentives, which benefit DECs, skew the market outcomes in favor of generators at the expense of consumers. For either case the actual impact on market efficiency (total surplus) is very limited. The results indicate that virtual bidding has an unequal impact that is potentially detrimental to consumers especially considering the limited participation of consumers in wholesale electricity markets.

The bidding model results show that price differences that create profitable arbitrage opportunities are driven by uncertainty, the shape of the supply curves, and consistent differences between demand in the DA and RT markets. The optimal bidding strategy always leads to greater average price convergence, which is claimed as evidence of more efficient market outcomes. However, the welfare analysis shows that outcomes with greater price convergence as a result of virtual bidding activity are not necessarily more efficient, nor do they always correct surplus distribution distortions that result from bias in the DA expectation of RT load.

While the results do show that the most profitable bidding opportunities occur when arbitrageurs correct for consistent load differences in the DA and RT market, the results also show that there are bidding strategies that do not have to consistently predict the correct price difference in order to be both profitable and provide for price convergence in expectation. These strategies do not rely on improving the expectation of RT load but rather on taking advantage of the expected asymmetric price changes for positive and negative deviations. The possible existence of these strategies that nonetheless result in greater expected price convergence suggests that in the context of wholesale electricity markets price convergence is an unreliable measurement of market performance. As such price convergence should neither be a policy objective in it of itself, nor be used as the principle metric for evaluating market performance because there is no guarantee that it will lead to more efficient outcomes or provide evidence that market efficiency has increased.

CHAPTER 5. A STYLIZED TWO NODE ELECTRICITY MARKET MODEL FOR UTCS

In this section, we introduce a stylized model that includes many of the salient features of the DA and RT markets. We introduce virtual bidding with UTCs and derive the optimal bidding strategy. We then derive the impact on price convergence as well as social welfare.

5.1 Optimal UTC Bidding Strategy Model

The unit commitment and dispatch problem that the ISO solves in order to determine the DA dispatch schedule is a highly complex problem. In order to make the problem analytically tractable we make a series of simplifications. The analysis occurs in a two node network connected by a single transmission line. Demand is assumed to be perfectly inelastic. In addition, we abstract from the lumpiness associated with the commitment of generators due to fixed costs and minimum operating limits. We assume that market power is not being exercised by either generators or load serving entities, and the only reason for price differences between the DA and the RT markets are demand and supply imbalances due to imperfect knowledge of demand in the DA market. While acknowledging that this is an over simplification of the market we believe that it is sufficient to capture the qualitative efficiency impacts of virtual speculative trades using UTCs

The two nodes in our stylized UTC model are denoted A and B, and they are connected by a transmission line with capacity T (MW). All demand occurs at node B, but generation can occur at both nodes A and B. For the representative hour, we assume that a fixed amount of demand, D_{DA}, is bid into the DA market. Demand in the RT market, D_{RT}, is the sum of D_{DA} plus a random deviation Δ which represents the uncertainty in the RT demand ($D_{RT} = D_{DA}+\Delta$). The random variable Δ is represented by a uniform distribution with a lower value *delta_low* and upper value *delta_high* ($\Delta \sim U$[*delta_low*, *delta_high*]). Given D_{DA} and Δ demand in the RT market D_{RT} is characterized by the uniform distribution δ with lower value *dlow* and upper value *dhigh* ($\delta \sim U$[*dlow*, *dhigh*]). The supply curves for both generators are represented by linear functions, and assuming a perfectly competitive market, the inverse supply curve (price as a function of supply) is such that price is equal to marginal cost of generation. Thus, the marginal cost for Gen_i (i = A, B) as a function of the MW_i supplied is represented by a slope parameter α_i and an intercept β_i, as in equation (1):

$$MC_i = \alpha_i MW_i + \beta_i. \tag{1}$$

Gen_A is the low cost generator, producing output MW_A, is located at node A, and the high cost generator, Gen_B producing output MW_B, is located at node B. The low cost generator Gen_A is representative of base load generation which is characterized by large fixed startup costs and low marginal costs. The high cost generator Gen_B is representative of cycling and peak generation which is characterized by low fixed startup costs and high marginal costs. In order to represent the lower marginal costs and greater fixed cost of Gen_A compared to Gen_B the slope parameter α_A is smaller than α_B and the intercept β_A is larger than β_B. The scheduling of generators for both the DA and the RT markets is done according to the principles of economic dispatch in which the load must be served by the lowest cost combination of output from Gen_A and Gen_B given available transmission capacity *T*.

Next we will present the formulas for solving the economic dispatch problem, which are the same for both the DA and the RT market. Demand will be denoted generally by *d* as it may represent D_{DA} or D_{RT} depending on whether the DA or RT market is being modeled. The first task is to solve for the optimal output from each generator by solving the system cost minimization problem. The cost function for each generator *i* is represented by equation (2), and it is defined by the integral of the marginal cost function with respect to output:

$$Cost_i = 0.5\alpha_i MW_i^2 + \beta_i MW_i. \tag{2}$$

Given the cost functions for each generator and demand at node B, denoted by *d*, and assuming no line losses or transmission constraints, the cost minimization problem is represented by equation (3).

$$\min_{0 \le MW_A, 0 \le MW_B} 0.5\alpha_A MW_A^2 + \beta_A MW_A + 0.5\alpha_B MW_B^2 + \beta_B MW_B \tag{3}$$

$$s.t. \qquad MW_A + MW_B = d$$

In this setup, all generation at node A flows through the transmission system to node B. The Karush-Kuhn-Tucker conditions can be used to determine the optimal solution for output from the low cost generator Gen_A represented by equation (4), and the optimal output from high cost generator Gen_B represented by equation (5):

$$MW_A = \frac{\alpha_B d - \beta_A + \beta_B}{\alpha_A + \alpha_B} = \frac{\alpha_B d}{\alpha_A + \alpha_B} - \frac{\beta_A - \beta_B}{\alpha_A + \alpha_B} \tag{4}$$

$$MW_B = d - MW_A = d - \frac{\alpha_B d - \beta_A + \beta_B}{\alpha_A + \alpha_B}. \tag{5}$$

Note that these solutions assume that $0 < MW_A < d$.[3]

These optimal output quantities are for a system with no transmission constraints or line losses and the marginal cost, and hence price, at the two nodes will always be the same. A limit on the available amount of transmission capacity, $MW_A \leq T$, is introduced in order to admit the possibility of congestion and price differences across the network. Thus, the transmission constrained problem is represented by equation (6).

$$\min_{0 \leq MW_A, 0 \leq MW_B} 0.5\alpha_A MW_A^2 + \beta_A MW_A + 0.5\alpha_B MW_B^2 + \beta_B MW_B \tag{6}$$

$$s.t. \qquad MW_A + MW_B = d$$

$$MW_A \leq T$$

Karush-Kuhn-Tucker conditions can again be used to show that (7) represents the output from the low cost generator Gen_A given transmission capacity T for the transmission constrained case:

$$MW_A = Min\left(T, \frac{\alpha_B d - \beta_A + \beta_B}{\alpha_A + \alpha_B}\right). \tag{7}$$

Thus, Gen_A will be scheduled to generate at the lowest system cost optimal quantity up until the point where the optimal quantity meets the transmission capacity, after that point the output of Gen_A is held at the transmission capacity limit. Equation (8) represents the output for the high cost generator Gen_B given transmission capacity T.

$$MW_B = d - Min\left(T, \frac{\alpha_B d - \beta_A + \beta_B}{\alpha_A + \alpha_B}\right) \tag{8}$$

---

[3] We are only considering cases where $0 < MW_A < d$ as it only in these cases where congestion can occur. UTCs profit from congestion differences and thus if there is no congestion there is no incentives to place UTCs. In addition, satisfaction of these restrictions is implied by the conditions $\alpha_i > 0$ and $\beta_A > \beta_B$.

Gen_B will also be scheduled to generate at the lowest system cost optimal quantity up until the point where the transmission capacity is reached. After that point Gen_B is used to serve the rest of the load.

To calculate the DA market solutions *d* is replaced by D_{DA}. The DA price at node A is determined by the marginal cost function for GenA and the optimal output that GenA is scheduled for given D_{DA} is given by equation (9):

$$P_A^{DA} = \alpha_A \left(Min\left(T, \frac{\alpha_B D_{DA} - \beta_A + \beta_B}{\alpha_A + \alpha_B} \right) \right) + \beta_A \tag{9}$$

Similarly, the DA price at node B is determined by the marginal cost function for Gen_B and the optimal output that Gen_B is scheduled for given D_{DA} is given by equation (10).

$$P_B^{DA} = \alpha_B \left(D_{DA} - Min\left(T, \frac{\alpha_B D_{DA} - \beta_A + \beta_B}{\alpha_A + \alpha_B} \right) \right) + \beta_B \tag{10}$$

In order to calculate RT prices given the range of possible demand realizations (*dlow*, *dhigh*) in the RT market, we must know at which load level within the uncertainty range, the transmission line will become congested. This value will be referred to as the critical value (*CritVal*) and is represented by equation (11):

$$CritVal = max\left(dlow, min\left(dhigh, \left(\frac{T(\alpha_A + \alpha_B) + \beta_A - \beta_B}{\alpha_B} \right) \right) \right) \tag{11}$$

The expression $[T(\alpha_A + \alpha_B) + \beta_A - \beta_B)]/\alpha_B$ calculates the load value at which the transmission line becomes congested and will be referred to as the threshold load. The expression is obtained by solving for the load value in the minimum function in equations (7) and (8) that causes a switch for the second term to the first term. If the threshold load is less than *dlow* then the critical value is *dlow* and the transmission line will always be congested. If the threshold load is greater than *dhigh* then the critical value is *dhigh* and the transmission line will never be congested. If the threshold load is in between *dlow* and *dhigh* then the transmission will not be congested from *dlow* to the threshold load and congested from the threshold load to *dhigh*. The expected RT price at node A is given by equation (12)

$$E[P_A^{RT}] = \frac{\int_{dlow}^{CritVal} \alpha_A \left(\frac{\beta(D_{RT} - \beta_A + \beta_B)}{\alpha_A + \alpha_B}\right) + \beta_A \, \mathrm{d}\Delta + \int_{CritVal}^{dhigh} \alpha_A(T) + \beta_A \, \mathrm{d}\Delta}{dhigh - dlow} \tag{12}$$

The first integral calculates the expectation of price over the range where there is no congestion and the second integral calculates the expectation of price over the range where there is congestion. The expected RT price at node B is calculated in the same manner as represented by equation (13).

$$E[P_B^{RT}] = \frac{\int_{dlow}^{CritVal} \alpha_B \left(D_{RT} - \frac{\beta(D_{RT} - \beta_A + \beta_B)}{\alpha_A + \alpha_B}\right) + \beta_B \, \mathrm{d}\Delta + \int_{CritVal}^{dhigh} \alpha_\beta(D_{RT} - T) + \beta_B \, \mathrm{d}\Delta}{dhigh - dlow} \tag{13}$$

Financial participants can place UTC MW bids, denoted by *V*, between nodes A and B in order to speculate on the difference in congestion value between the DA and the RT markets. In order to place UTC bids, financial participants have to specify a MW quantity, the reservation price for how much congestion/MW they are willing to pay for in the DA market, and the injection and withdrawal nodes. In order to simplify our model, we will assume that financial participants are price takers and only specify a MW quantity. (This effectively ignores the case where the bid does not clear.) The expected profit for a UTC transaction is represented by equation (14).

$$E[Profit(V)] = -\left[\alpha_B\left(D_{DA} - Min\left(T - V, \frac{\alpha_B D_{DA} - \beta_A + \beta_B}{\alpha_A + \alpha_B}\right)\right) + \beta_B - \alpha_A\left(Min\left(T - V, \frac{\alpha_B D_{DA} - \beta_A + \beta_B}{\alpha_A + \alpha_B}\right)\right) + \beta_A\right]V + (E[P_B^{RT}] - E[P_A^{RT}])V \quad (14)$$

UTCs are modeled as either exacerbating or relieving congestion which depends on the direction of the UTC and the direction of the prevalent flow. Given the way that the network is configured with the less costly generator node A and both the load and the expensive generator at node B, the prevalent direction of power flows will be from A to B. Thus a UTC in the prevalent direction (injection at node A and withdrawal at node B) would reduce transfer capability in the prevalent direction and is therefore modeled as a positive *V* value in (14). A UTC in the counter flow direction (injection at node B and withdrawal at node A) would increase transfer capability in the prevalent direction and is therefore modeled as a negative *V* value.

The optimal bidding strategy *V** is determined by solving the first-order optimality conditions of the expected profit function and solving for *V*. However, because the profit function contains a minimum function, it is not differentiable at the point where $V = T - \frac{\alpha_B D_{DA} - \beta_A + \beta_B}{\alpha_A + \alpha_B}$. This point is where *V* is equal to the "spare" capacity on the transmission line in the DA market which will be referred to as *SpareT*. At this point there is a regime change in the optimal solution formula and thus the optimal bidding strategy is defined by a piecewise differentiable function. The solution obtained by solving for V in the first-order optimality conditions will be referred to as $V`$. The optimal bidding strategy is defined by equation (15).

$$V^* = \begin{cases} V` & for\ V` > SpareT \\ SpareT & for\ V` \leq SpareT \end{cases} \tag{15}$$

Because UTCs have an impact on the transfer capability between nodes A and B, they may impact the scheduling of generators in the DA market and consequently the DA market prices. Taking UTCs into account, the DA dispatch of Gen A is given by (16) and the dispatch of GenB is given by (17):

$$MW_A = Min\left(T - V, \frac{\alpha_B d - \beta_A + \beta_B}{\alpha_A + \alpha_B}\right) \tag{16}$$

$$MW_B = d - Min\left(T - V, \frac{\alpha_B d - \beta_A + \beta_B}{\alpha_A + \alpha_B}\right) \tag{17}$$

Taking UTCs into account, the price at node A is given by (18) and the price at node B is given by (19):

$$P_A^{DA} = \alpha_A\left(Min\left(T - V, \frac{\alpha_B D_{DA} - \beta_A + \beta_B}{\alpha_A + \alpha_B}\right)\right) + \beta_A \tag{18}$$

$$P_B^{DA} = \alpha_B\left(D_{DA} - Min\left(T - V, \frac{\alpha_B D_{DA} - \beta_A + \beta_B}{\alpha_A + \alpha_B}\right)\right) + \beta_B \tag{19}$$

5.2 UTC Bidding Model Results

In a system as large and complex as PJM's there are many possible different system conditions that can affect congestion differences between the DA and the RT markets and hence UTC bidding strategies. In this research were are primarily interested in learning about the qualitative nature of UTC bidding strategies and their impact on market participants. In order to provide evidence in that regard, we will solve the model in the previous section numerically, choosing model parameters to capture scenarios representing a variety of possible general system conditions.

In our model, congestion differences are dependent on differences in load levels between the DA and the RT market and available transmission capacity relative to the optimal dispatch of Gen_A (this is important since all of the output from Gen_A is transferred to the load center at node B). Thus in order to capture a range of possible system conditions we will explore three different load scenarios which are: (1) expected RT load equals DA load (no bias), (2) expected RT load is greater than DA load (DA load underbidding), and (3) expected RT load is less than DA load (DA load overbidding); and three different levels of transmission capacity which are: (1) low (the line is always congested), (2) intermediate (the line is congested about half the time), and (3) high (the line is only congested for the highest levels of RT demand).

We assume that 100 MW of load are bid into the DA market for all scenarios, the difference between load scenarios is captured by different distributions of outcomes in the RT market. For the unbiased load deviation scenarios, the UTC model is solved for increasing ranges of demand uncertainty, all with an expected value of 100 MW. For the load underbidding scenarios the RT load has a constant uncertainty range of 20 MW, the UTC model is solved for RT load realizations with an expected value from 100 MW to 110 MW – that is, for underbidding ranging up to 10 percent of load. For the load overbidding scenarios the RT load has also has a constant range uncertainty range of 20

MW, the UTC model is solved for RT load realizations with an expected value from 90 MW to 100 MW – up to 10 percent overbidding of load. The different transmission capacity levels, *T*, are correspondingly meant to capture general system conditions in the DA market where the system is (1) always congested, (2) congested for about half the possible load outcomes, and (3) experience no congestion except for few cases when the highest loads occur.

The parameters for the low cost generation where chosen to be \$0.2/MW for the slope parameter α_A and \$20 for the intercept parameter β_A. The parameters for the high cost generator where chosen to be \$0.8/MW for the slope parameter α_B and \$5 for the intercept parameter $\beta_{B.}$ The parameters for the low cost generator located at node B were chosen to be representative of a base load generator with high startup costs and low marginal cost, and the parameters for the high cost generator located at node A were chosen to be representative of a cycling or peaking unit with low startup cost and high marginal cost. The three levels of available transmission capacity are 60 MW for the low case, 65 MW for the intermediate case, and 70 MW for the high case.

The quantitative results for the optimal bidding strategy, its profits, and congestion differences with and without the optimal bidding strategy for each of the explored scenarios can be found in Appendix I. A qualitative summary of the results is presented in Table 5-1. UTCs are a product for speculating on transmission price differences between the DA and RT markets, and thus it would be expected that these differences would drive UTC volumes and profits with greater differences leading to higher UTC volumes and profits, and vice versa. Consequently, such bidding behavior would lead to a reduction in the transmission price difference, i.e. congestion convergence. The results show that this is the case for UTCs only for situations in which the system is congested or close to being congested (the low and intermediate transmission cases). For these cases the results show that the optimal UTC MW, profits, and convergence impact are proportional to the congestion value differences without UTCs.

Table 5-1. Bidding model results summary for the different load and transmission scenarios.

		Low Transmission	Intermediate Transmission	High Transmission
No Bias	Optimal UTC MW	Very small qty. that increases with increasing load uncertainty	Small qyt. that increases with increasing load uncertainty	Large qty. equal to spare transmission capacity
	Expected Profit	Minimal; Increases with load uncertainty	Moderate; Increases with load uncertainty	Large; Increases with load uncertainty
	Transmission Price Difference (V=0)*	Minimal	Small	Minimal
	Transmission Price Difference (V=V*)**	Half compared to case where V=0	Half compared to case where V=0	Same as V=0
Load Underbidding	Optimal UTC MW	Moderate qty. that increases with the bias	Moderate qty. that increases with the bias	Large qty. equal to spare transmission capacity
	Expected Profit	Moderate; Increases with DA load bias	Moderate; Increases with DA load bias	Large; Increases with DA load bias
	Transmission Price Difference (V=0)*	Large; Increases with DA load bias	Moderate; Increases with DA load bias	Small
	Transmission Price Difference (V=V*)**	Half compared to case where V=0	Half compared to case where V=0	Same as V=0
Load Overbidding	Optimal UTC MW	Small qty of counter flow UTC MW that increases with the bias	Small qyt. that decreases to zero with the bias	Large qty. equal to spare transmission capacity
	Expected Profit	Moderate; Increases with DA load bias	Small; Decreases with DA load bias	Small; Decreases with DA load bias
	Transmission Price Difference (V=0)*	Negative congestion that increases with the bias	Minimal	Minimal
	Transmission Price Difference (V=V*)**	Half compared to case where V=0	Half compared to case where V=0	Same as V=0

* Expected Transmission price difference between the DA and RT market for the case with no UTCs

** Expected Transmission price difference between the DA and RT with the optimal UTC bidding strategy V* compared to the case with no UTCs

However, for scenarios where there is ample transmission capacity, leading to no congestion in the DA market and only a small possibility of congestion in the RT market (the high transmission capacity case), the results show that there is a different regime motivating bidding strategies. For these cases the optimal bidding strategy involves clearing a quantity of UTC MW equivalent to the spare transmission capacity, and therefore the results show that it is optimal to bid and clear a large quantity of UTC MW despite little congestion value differences between the DA and the RT market.

This strategy can be accomplished by bidding a large number of UTC MW with either a zero or very low reservation price. Only the quantity equivalent to the spare transmission capacity will clear because anything greater would cause congestion to increase beyond the reservation price. These bidding strategies can be very profitable as the cost to take these positions is very small (DA congestion times UTC MW) and any congestion in the RT market represents a profit to them.

There is evidence that this is in fact the most popular trading strategy based on a report by PJM that states that 51.1 percent of all cleared UTCs had bid offers between ±$1/MW (PJM, 2015). The results also show that these strategies would have little or no impact on congestion convergence because the profitability of the strategies rely on creating little or no congestion in the DA market.

5.3 Nodal Price Convergence

We also examine the impact of the UTC trading strategies on nodal price convergence. Keeping in mind that congestion between two nodes is simply the difference in LMPs between the two nodes, congestion convergence (which is supposed to be one of the outcomes of UTC introduction) should lead to nodal price convergence. Ultimately congestion convergence is desirable because it leads to nodal price convergence. Given that market transactions are settled at nodal prices it is the impact on these that drives the efficiency gains.

Appendix II contains the figures that show the impact on nodal price convergence of the optimal UTC biding strategy for all of the load and transmission scenarios. The results show that UTCs do not consistently provide for nodal price convergence. For all high transmission capacity cases UTCs have no impact on price convergence at either the source or sink node. For the low and intermediate transmission capacity cases UTCs provide for price convergence at both the source and the sink only in the no bias load scenario. In the load underbidding and overbidding cases UTCs caused convergence at one node and either had no impact or caused price divergence at the other node. Ultimately the results suggest that the ability of UTCs to provide for nodal price convergence does not depend on profitable speculation incentives but rather what was causing the price differences between the DA and the RT market. If price differences are the result of unbiased demand uncertainty, then UTCs can possibly provide for nodal price convergence. If price differences are the result of under or over bidding, then UTCs can only provide for price convergence at either the source or sink node only. These results stand in contrast with the results of Giraldo et al. (2016) which found that independent optimal INCs and DECs consistently provide for nodal price convergence in cases with no bias load uncertainty, load underbidding and load overbidding.

5.4 Welfare Calculations for UTC Model

This section presents the formulas for calculating expected social welfare given the uncertainty of final demand in the RT market and the optimal bidding strategy by financial participants. While power is only produced and consumed on the operating day, given the structure of the market there are two settlements over which power is bought and sold, one for the DA market and one for the RT market. Expected consumer surplus is calculated as the expected value to consumers minus total payments made by consumers which are the sum of DA payments and expected RT payments (known as balancing payments) as represented by equation (20).

$$E[CS] = E[Consumer\ Value] - DA\ Pay - E[Bal\ Pay] \tag{20}$$

Since we have assumed that electricity demand is fixed, the value of electricity to consumers will be represented by the reservation price *RP* which will be assumed to have a large fixed value so that it is always greater than the electricity price up to the level of demand, and zero beyond that level. Consumers only consume power on the operating day, thus consumer value is calculated over the integral of the possible RT realizations of demand δ as shown in equation (21).

$$E[Consumer\ Value] = \frac{\int_{dlow}^{dhigh}(\delta)RPd\delta}{dhigh - dlow} \tag{21}$$

DA payments by consumers are equivalent to the DA demand times the DA price at node B (which is where load is located) as shown in equation (22).

$$DA\ Pay = \left[\alpha_B\left(D_{DA} - Min\left(T - V, \frac{\alpha_B D_{DA} - \beta_A + \beta_B}{\alpha_A + \alpha_B}\right)\right) + \beta_B\right] D_{DA} \tag{22}$$

Balancing payments by consumers are equivalent to the deviation from the DA demand times the RT price at node B as represented by equation (23).

$$E[Balancing\ Pay] = \int_{dlow}^{CritVal} \left(\left(\alpha_B \left(\delta - \frac{\alpha_B \delta - \beta_A + \beta_B}{\alpha_A + \alpha_B} \right) + \beta_B \right) (\delta - D_{DA}) \right) \frac{d\delta}{dhigh - dlow}$$

$$+ \int_{CritVal}^{dhigh} \left((\alpha_B(\delta - T) + \beta_B)(\delta - D_{DA}) \right) \frac{d\delta}{dhigh - dlow} \quad (23)$$

If RT load is greater than DA load ($\delta - D_{DA} > 0$) balancing payments are positive and if RT load is less than the DA load ($\delta - D_{DA} < 0$) then the balancing payments are negative. The first integral represents balancing payments for RT load realizations where there is no congestion and the price paid by consumers is the unconstrained price. The second integral represents balancing payments for RT load realizations where there is congestion and the price paid by consumers is the constrained price set by high cost generator GenB. Expected producer surplus for generator *i* is calculated as expected DA revenues for generator *i* plus balancing revenues for generator *i* minus generating cost for generator *i* as seen in equation (24).

$$E[PS]_i = DA\ Rev_i + E[Bal\ Rev_i] - E[Gen\ Cost_i] \quad (24)$$

DA revenues for generators is for power sold in the DA market and is calculated as the DA MW quantity cleared by each generator times the price at its node. Revenue for Gen_A is represented by equation (25) and revenue for Gen_B is represented by equation (26).

$$DA\ Rev_A = Min\left(T - V, \frac{\alpha_B d - \beta_A + \beta_B}{\alpha_A + \alpha_B} \right) \left(\alpha_A \left(Min\left(T - V, \frac{\alpha_B D_{DA} - \beta_A + \beta_B}{\alpha_A + \alpha_B} \right) \right) + \beta_A \right) \quad (25)$$

$$DA\ Rev_B = d - Min\left(T - V, \frac{\alpha_B d - \beta_A + \beta_B}{\alpha_A + \alpha_B} \right) \left(\alpha_B \left(D_{DA} - Min\left(T - V, \frac{\alpha_B D_{DA} - \beta_A + \beta_B}{\alpha_A + \alpha_B} \right) \right) + \beta_B \right) \quad (26)$$

The expected balancing revenue for GenA is described by equation (27) and expected balancing payment for GenB is described by equation (28).

$$E[Bal\ Rev_A] = \int_{dlow}^{dhigh} \left[\left(min\ \left(T, \frac{\alpha_B \delta - \beta_A + \beta_B}{\alpha_A + \alpha_B} \right) - min\ \left(T - V, \frac{\alpha_B D_{DA} - \beta_A + \beta_B}{\alpha_A + \alpha_B} \right) \right) \left(\alpha_A \left(Min \left(T, \frac{\alpha_B \delta - \beta_A + \beta_B}{\alpha_A + \alpha_B} \right) \right) + \beta_A \right) \right] \frac{d\delta}{dhigh - dlow} \tag{27}$$

$$E[Bal\ Rev_B] = \int_{dlow}^{dhigh} \left[\left(\left(\delta - min\ \left(T, \frac{\alpha_B \delta - \beta_A + \beta_B}{\alpha_A + \alpha_B} \right) \right) - \left(D_{DA} - min\ \left(T - V, \frac{\alpha_B D_{DA} - \beta_A + \beta_B}{\alpha_A + \alpha_B} \right) \right) \right) \left(\alpha_B \left(\delta - Min \left(T - V, \frac{\alpha_B \delta - \beta_A + \beta_B}{\alpha_A + \alpha_B} \right) \right) + \beta_B \right) \right] \frac{d\delta}{dhigh - dlow} \tag{28}$$

For both generators the expected balancing revenue payment is calculated as the difference between what was generated in the RT market and what was scheduled in the DA market (represented as the first two terms in the equation) times the RT market price (the third term in the equations). Balancing revenue payments will be negative if there is a negative deviation and positive if there is a positive deviation. The expected generation cost for GenA is described in equation (29) and the expected cost for GenB is described by equation (30).

$$E[GenCost_A] = \int_{dlow}^{dhigh} \left(0.5\alpha_A min \left(T, \frac{\alpha_B \delta - \beta_A + \beta_B}{\alpha_A + \alpha_B} \right)^2 + \beta_A min \left(T, \frac{\alpha_B \delta - \beta_A + \beta_B}{\alpha_A + \alpha_B} \right) \right) \frac{d\delta}{dhigh - dlow} \tag{29}$$

$$E[GenCost_B] = \int_{dlow}^{dhigh} \left(0.5\alpha_B \left(\delta - min \left(T, \frac{\alpha_B \delta - \beta_A + \beta_B}{\alpha_A + \alpha_B} \right) \right)^2 + \beta_B \left(\delta - min \left(T, \frac{\alpha_B \delta - \beta_A + \beta_B}{\alpha_A + \alpha_B} \right) \right) \right) \frac{d\delta}{dhigh - dlow} \tag{30}$$

Given that each agent settles their transactions (purchase or sale) at their node's price when the system is congested more revenue will be collected from consumers than what is paid out to generators. This extra revenue is known as congestion rent as it is allocated to owners of financial transmission rights (FTRs).[4] Total congestion rent is calculated as the difference between the revenue collected by the ISO from consumers and the sum of what is paid to generators *i* as described by equation (31).

$$Congestion\ Rent = DA\ Pay + E[Bal\ Pay] - \sum_i (DA\ Rev_i + E[Bal\ Rev_i]) \qquad (31)$$

5.5 Welfare Results for UTC Model

For the welfare analysis we want to determine if arbitraging congestion differences with UTCs increases market efficiency in terms of correcting the distribution of surplus towards the optimal distribution of surplus as identified by Chao and Peck (1996). This socially optimal distribution is predicated on the price being equal to the marginal cost and marginal benefit at each node. However, uncertainty about RT demand and possible biases in DA load quantities perturb the distribution of welfare away from the social optimum because the DA market clearing prices usually do not represent the marginal costs of providing electricity in the RT market. Therefore, we will calculate a socially optimal distribution of welfare in which all of the consumption and generation is settled at the marginal price at each node. This would only occur if there were no deviations between the DA and the RT markets, i.e. if the DA load was a perfect, deterministic prediction of RT load (this situation would also occur in a single settlement market). For

[4] FTRs are a financial instrument that entitles the holder to receive a share of the congestion rent payments collected on a specific source and sink path. The share is based on the MW quantity of the FTR contract.

each scenario we will calculate the social optimum consumer surplus, producer surplus (Gen_A and Gen_B separately) and congestion rent to be used as a benchmark.

The surplus distribution or each uncertainty and bias scenario without UTCs and with the optimal UTC bidding strategy will also be calculated. These will be compared to the social optimum benchmark to determine the impact of uncertainty and biases on the distribution of welfare and whether UTCs make any corrections or whether they cause further distortions. For the no bias scenario, we assume that 100 MW is bid into the DA market and in the RT market there is uncertainty about the final realization of demand, represented by the uncertainty ranges from ±1 MW to ±10 MW, but the expected load is 100 MW. The social optimum benchmark is calculated for 100 MW in the DA market with no deviations in the RT market. For the load underbidding and load underbidding scenarios we assume that 100 MW is bid into the DA market but the expected RT load ends up being a different quantity from the 100 MW bid in the DA market. For load underbidding/overbidding cases load is greater/smaller by up 10 MW. These cases also have an uncertainty range of 20 MW. The social optimum benchmark for these cases is calculated by assuming the load bid into the DA market is the same as the expected RT load; thus, the surplus comparison is made relative to the same quantity of MW consumed (for example if expected RT load is 105 MW the social optimum is calculated assuming that 105 MW was bid in the DA market)

Table 2 presents a qualitative summary of the welfare analysis results for the no bias load bidding scenario and all transmission scenarios. Two types of welfare results are presented for each of the agents in the model. The first set of results are displayed under the “No UTCs” column headers and indicate the impact on surplus allocation of introducing demand uncertainty (and bias for the load under and overbidding scenarios) compared to the deterministic (social optimum which we represent by the single settlement market) allocation of surplus. Demand uncertainty can either increase, decrease, or have no impact on the surplus allocation of the agents. The second set of results are displayed under the "UTC Impact" column headers and indicate the impact of the optimal UTC bidding strategy on the surplus allocation, indicating whether UTCs corrected (moved the welfare measure for the agent closer to the socially optimal welfare distribution), distorted, or had no impact on the surplus allocation. UTCs can distort the allocation of surplus in three different ways; they can either exacerbate the distortion caused by the uncertainty, overcorrect for the distortion caused by the uncertainty, or create a distortion where none had been caused by the uncertainty. When the distortion results in the agent being allocated a greater amount of surplus compared to the social optimum or the case with no UTCs then the result is represented by the ↑Distortion symbol. When the distortion results in the agent being allocated a smaller amount of surplus compared to the social optimum or the cases with no UTCs then the result is represented by the ↓Distortion symbol. Figures with the numerical results can be found in Appendix III.

The results in table 2 show that for the no bias load scenario there is only once instance in which UTCs provide for a correction toward the social optimum. This occurs with congestion rents in the high transmission scenario. The results show that compared to social optimum congestion rents are higher when uncertainty is introduced. The results with UTCs show that UTCs bidders are capturing, as profits, the additional congestion rent caused by the uncertainty. This congestion rent would have been allocated to agents who hold FTRs on the A to B path. In the low and intermediate transmission cases UTCs systematically make consumers and GenA worse off and GenB better off.

Table 5-2 presents the qualitative summary of the welfare analysis results for the load underbidding scenario and all transmission scenarios. The results display the impact on surplus allocation of introducing demand uncertainty and load underbidding compared to the deterministic (social optimum) allocation of surplus and whether the optimal UTC bidding strategy corrected, further distorted, or had no impact on the surplus allocation. Figures with the numerical results can be found in Appendix III. The results show that for the low and intermediate transmission cases UTCs correct the surplus allocation to the social optimum for consumers and GenB which is a positive impact on the market; however, for GenA UTCs further distort the surplus allocation. For the high transmission case UTCs have no impact on correcting surplus allocation to consumers or generators, and again they correct the congestion rents towards the social optimum.

Table 5-2. Qualitative summary of the surplus allocation impact of demand uncertainty with no bias and UTCs' surplus reallocation impact.

Transmission Level	GenA		GenB		Consumers		Congestion Rent	
	No UTCs*	UTC Impact**	No UTCs*	UTC Impact**	No UTCs*	UTC Impact**	No UTCs*	UTC Impact**
Low	—	↓Distortion	Increase	↑Distortion	Decrease	↓Distortion	—	↑Distortion
Intermediate	Increase	↓Distortion	Increase	↑Distortion	Decrease	↓Distortion	—	↑Distortion
High	Increase	—	Increase	—	Decrease	—	Increases	**Corrects**

*Impact of introducing demand uncertainty on surplus distribution compared to the social optimum (deterministic case) with no UTCs.
1) Increase = Increase in surplus
2) Decrease = Decrease in surplus
3) — = No distortion

**Impact of the optimal bidding strategy V*
Impacts are:
1) ↑Distortion = Increase in surplus
2) ↓Distortion = Decrease in surplus
3) Corrects = Corrects to social optimum
4) — = No correction or distortion

Table 5-3. Qualitative summary of the surplus allocation impact of demand uncertainty and load underbidding and UTCs surplus reallocation impact.

Transmission Level	GenA		GenB		Consumers		Congestion Rent	
	No UTCs*	UTC Impact**	No UTCs*	UTC Impact**	No UTCs*	UTC Impact**	No UTCs*	UTC Impact**
Low	—	↓Distortion	Decrease	**Corrects**	Increase	**Corrects**	Decrease	**Corrects**
Intermediate	—	↓Distortion	Decrease	**Corrects**	Increase	**Corrects**	Decrease	—
High	Decrease	—	Decrease	—	Increase	—	Increases	**Corrects**

* Results for the case with no UTCs are presented in the same maner as Table 5-2.
** Result with the optimal bidding starategy V* are presented in the same manner as Table 5-2.

Table 5-4 presents the qualitative summary of the welfare analysis results for the load overbidding scenarios and all transmission scenarios. The results display the impact on surplus allocation of introducing demand uncertainty and load underbidding compared to the deterministic (social optimum) allocation of surplus and whether the optimal UTC bidding strategy corrected, further distorted, or had no impact on the surplus allocation. Figures with the numerical results can be found in Appendix III. The results show that for the low and intermediate transmission cases UTCs correct surplus distribution to the social optimum in about half the cases. There is no consistency in the impact of optimal UTC bidding on the different market participants in the low and intermediate cases. However, for the high transmission case UTCs again have no impact on the surplus allocation to consumers or generators, but shift the congestion rents in the direction of the social optimum.

Table 5-4. Qualitative summary of the surplus allocation impact of demand uncertainty and load overbidding and UTCs surplus reallocation impact.

Transmission Level	GenA		GenB		Consumers		Congestion Rent	
	No UTCs*	UTC Impact**	No UTCs*	UTC Impact**	No UTCs*	UTC Impact**	No UTCs*	UTC Impact**
Low	Increase	↑Distortion	Decrease	**Corrects**	Decrease	**Corrects**	Increases	**Corrects**
Intermediate	Increase	**Corrects**	Increase	↑Distortion	Decrease	↓Distortion	—	—
High	Increase	—	Decrease	—	Decrease	—	Increases	**Corrects**

* Results for the case with no UTCs are presented in the same maner as Table 5-2.
** Result with the optimal bidding starategy V* are presented in the same manner as Table 5-2.

In general, the results showed that for less than half of the cases UTCs correct the surplus distribution towards the social optimum. For the majority of scenarios UTCs either have no impact or exacerbate the surplus distribution distortions caused by uncertainty and biases. In particular, for all load scenarios in the high transmission case UTC bidding had no impact on surplus distribution. UTCs seem to be helping more in cases when the distribution of surplus is affected by biases in the DA load compared to cases when is affected by uncertainty. UTCs also do not seem to have a consistent impact across market participants. For a given scenario UTCs may have a beneficial or corrective impact on one participant and a detrimental impact on another participant. The impact of UTCs on congestion rents show that in the high transmission cases they are able to consistently extract congestion rents as profits. In the low and intermediate transmission cases the impact of UTCs on congestion rents is more inconclusive. On some cases they increased congestion rents, on others they corrected towards the optimum, and on others they had no impact.

5.6 Discussion of UTC Model Results

This research showed that one of the particularly concerning aspects of the particularly concerning aspects of UTCs is the ability that it gives traders to engage in low risk high volume trading strategies that exist when there is ample transmission capacity and a small possibility of RT market congestion, analogous to the high transmission capacity scenarios explored with our model. The UTC product allows for the maximum exploitation of such circumstances by taking the guesswork/risk out of trying to identify the maximum quantity of UTC MW to bid that creates little or no DA congestion. Traders can simply submit a bid for a very large quantity of UTC MW with a very small reservation price, this guarantees the lowest risk for traders and that there will be little to no impact on price convergence. Our analysis indicated that this strategy is one of the most profitable which is counter intuitive given that for a product that arbitrages transmission prices it is most profitable when congestion is least likely to occur. Our results shed light and support the claims by the IMM for PJM that there is little evidence that UTCs cause nodal price convergence.

The results also showed that UTCs have an inconsistent impact on both transmission and nodal price convergence. UTCs caused transmission price convergence only for the low and intermediate transmission cases, there was no impact for the high transmission cases. Where transmission price convergence occurred for most cases it was driven by nodal price convergence at only one of the nodes, at the other node there was either no impact or price divergence. For the high transmission cases there was no impact on nodal price convergence.

By design UTCs do not have an impact on unit commitment. Thus UTCs cannot provide for improvements in unit commitments that would lead to increases total surplus. Through their impact on prices UTCs can only redistribute surplus. As a direct impact of UTCs inconsistent impact of price convergence, the results showed that UTCs do not consistently redistribute surplus to the social optimum. On most of the cases UTCs either made one market participant better off at the expense of another market participant or had no impact on surplus distribution.

CHAPTER 6. DISCUSSION AND CONCLUSION

This chapter briefly summarizes the results for the examination of incentives for bidding and welfare implications for virtual transactions in electricity markets. In particular, the INC/DEC model and the UTC model are compared and contrasted focusing on the differences in bidding and market impacts among these contracts. The hypotheses presented in the introduction are revisited and conclusions are drawn given the model results. Finally, a discussion and concluding remarks regarding the overall role and impact of virtual bidding in wholesale electricity markets is presented along with the limitations and future work.

6.1 INC and DEC Model Results Summary

The INC and DEC model showed that load uncertainty, biases, and the shape of the supply curve will typically create bidding opportunities that are profitable in expectation. The optimal bidding strategy always lead to greater price convergence; however, the impact on market efficiency (total surplus) was very limited. INCs and DECs did have a strong impact on surplus reallocation between consumers and producers. When price differences were the result of biases, INCs and DECs reallocated surplus in a manner that corrected the surplus allocation distortions caused by the biases. However, when price differences were the result of demand uncertainty alone, market incentives benefited DECs which skewed the market outcomes in favor of generators at the expense of consumers.

6.2 UTC Model Results Summary

The UTC model showed that load uncertainty and biases create profitable bidding opportunities, and that the optimal bidding strategy and its impact on the market was highly dependent on the quantity of available transmission capacity. For cases where the transmission system was at or near capacity the optimal bidding strategy was proportional to the transmission price differences between the DA and RT markets. For cases when there was only a small likelihood of transmission congestion the optimal bidding strategy was to bid a quantity of UTC MW so as to take a position equivalent to the excess transmission capacity. The results showed that UTCs have an inconsistent impact on transmission and nodal price convergence. The most concerning finding was that the most profitable bidding strategies (those in cases where there is excess transmission capacity) have no impact on nodal price convergence and for most other cases UTCs cause price convergence at one node and price divergence on the other node.

The impact of UTCs on market efficiency was inconclusive. UTCs cannot increase market efficiency (total surplus) given the way that they are modeled, and since they are profitable it means that they are withdrawing surplus from the market and thus decreasing market efficiency. While there is some value in correcting for surplus distortions caused by uncertainty and biases, the results showed that UTCs did not redistribute surplus in a manner that corrects surplus distribution to the ideal social optimum. For most cases, UTCs had either no impact on surplus distribution or exacerbated surplus distribution distortions caused by uncertainty and biases.

6.3 Comparing the Market Impacts of INCs and DECs versus UTCs

The three aspects of INCs, DECs, and UTCs that were investigated were their ability to 1) increase market efficiency as measured by total surplus, 2) correct the surplus distribution towards the ideal social optimum when load and uncertainty and biases caused distortions in the distribution of welfare, and 3) provide for price convergence. The results showed that UTCs have a different market impact compared to INCs and DECs. While INCs and DECs were able to increase market efficiency in some scenarios, UTCs did not increase market efficiency in any of the scenarios considered. The reason for this is that INCs and DECs have an impact on unit commitment and they can increase market efficiency by improving the commitment of units leading to lower generation costs. UTCs are modeled only once the unit commitment has already been done (as its done in PJM), and thus they only have an impact on the dispatch level of units that are already committed.

The LMP based markets were implemented with the expectation that efficiency in the electricity markets would be maximized when the price of electricity paid by consumers and received by generators equaled the marginal cost of generation. However, given the unique attributes of electricity generation it was necessary to implement a two-settlement market structure. Given that load clearing quantities can vary between the DA and RT markets the price paid by consumers and received by generators may differ from the marginal cost of generation. This causes distortions in the ideal socially optimal distribution of surplus. Through their impact on prices, virtual bidding can correct these distortions. The results showed that while the optimal bidding strategy for INCs and DECs did not fully bring the distribution of surplus to the ideal social optimum, it always led to a correction of the distribution of surplus for both consumers and generators towards the ideal social optimum. In the case of UTCs the results showed that the optimal UTC strategy had an inconsistent impact of surplus redistribution towards the ideal social optimum. For the high transmission cases there was no impact on surplus redistribution. For the low and intermediate transmission cases for most scenarios there was a correction for some market participants and an increase in the distortion for other market participants.

The benefit most commonly attributed to virtual bidding is price convergence. One of the appealing aspects of INCs and DECs is that the incentives are self-correcting towards nodal price convergence. Speculators who correctly predict price differences will make a profit and cause prices to converge. Conversely speculators who incorrectly predict the price differences will lose money and cause prices to diverge. The results showed that the optimal INC and DEC bidding strategy always causes price convergence in expectation. With INCs and DECs it would not be possible to have a profitable bidding strategy that causes price divergence in expectation. The results for the optimal UTC bidding strategy showed that UTCs have a very different impact on price convergence compared to INCs and DECs. For UTCs the most profitable bidding strategies (high transmission cases) did not have an impact on price convergence. For the other cases the optimal bidding strategy caused price convergence at one node and price divergence or no impact on the other node. There were very few instances where UTCs caused price divergence at both the source and sink nodes.

UTCs are product for speculating on congestion price differences. The results showed that UTCs provided for transmission price convergence in the low and intermediate transmission cases but not in the high transmission cases. Since the transmission price for a specific path is simply the LMP difference between the two nodes, nodal price convergence leads to transmission price convergence. Since the optimal INCs and DECs bidding strategy consistently provided for nodal price convergence it can hence provide for consistent transmission price convergence. Thus INCs and DECs may be more effective at providing transmission price convergence than UTCs.

Our results indicate that the market benefits attributed to INCs and DECs by previous studies should not be extended to UTCs. Gauging by their impact on price convergence and market efficiency the results indicate that UTCs are an inferior product compared to INCs and DECs. PJM's uplift allocation policy that favors UTCs over INCs and DECs is probably having a detrimental impact on the market. This policy has caused financial traders to overwhelmingly use UTCs instead of INCs and DECs. The results of this research show that UTCs may not have the ability to provide value to the market in return for the profits being made from their trading. This stands in contrast with INCs and DECs which do have the possibility to increase market efficiency or at least consistently provide for price convergence and surplus redistribution towards the ideal social optimum.

6.4 Conclusions Regarding Hypotheses

The following two hypotheses were presented at the beginning of this dissertation:

1) Average price convergence between the day-ahead and real-time markets does not always lead to more efficient market outcomes; and

2) UTCs incentives are not self-corrective towards price convergence and have different market impacts compared to INCs and DECs.

Results from the optimal bidding models as well as the social welfare analysis allow for conclusions to be made about the hypotheses. For the first hypothesis the social welfare analysis for the market impacts of INCs, DECs, and UTCs showed that price convergence does not indeed always lead to more efficient market outcomes, nor is it evidence that market efficiency improved. There were some situations in which price convergence was accompanied by an increase in market efficiency, however there were also instances where price convergence was accompanied by a decrease in market efficiency.

For the second hypothesis the results showed that UTCs' incentives are not self-correcting towards nodal price convergence and a comparison with the market impacts of INCs and DECs showed that UTCs have different market impacts. The optimal INC and DEC bidding strategies always to lead to price convergence in expectation while the optimal UTC strategies may have no impact on prices or lead to price divergence at one of the nodes. While the main impact of INCs and DECs was surplus reallocation this always occurs towards the ideal social optimum. The only impact of UTCs is surplus reallocation, and this impact does not consistently occur towards the ideal social premium.

6.5 Discussion

The social welfare analysis in our research provides a very different perspective on the impact of virtual bidding on wholesale electricity markets compared to the existing literature. In general, the results show that the main impact of virtual bidding is surplus reallocation and less so an impact on total surplus or efficiency. While the results did show that virtual bidding led to an increase in average price convergence, the welfare analysis indicated that market outcomes with greater average price convergence were not necessarily more efficient.

The impact of virtual bidding on surplus allocation is strong in our modeling results. Whether the surplus allocation impacts are distorting or beneficial depends largely on what was causing the price differences that were incentivizing the bidding behavior. If the price differences were the result of a consistent positive or negative bias in the DA demand, then virtual bidding helped correct the surplus distribution distortions caused by the bias. However, if the price differences were the result of unbiased uncertainty, then virtual bidding skewed market outcomes in a manner that benefits one market participant at the expense of another. In the case of INCs and DECs it benefits generators at the expense of consumers and in the case of UTCs it benefits one of the generators at the expense of consumers and other generators.

A load bidding analysis by the IMM for PJM showed that in the PJM market there were small differences between the DA and the RT loads. These differences, which were analyzed by year, showed either a very small underbidding bias for some years (less than half a percent of RT load) and not statistically different from zero for other years (need reference). This suggests that the no bias scenarios in this research are more likely to characterize real world market outcomes at least in the PJM market. For virtual bidding these were the cases where they were least likely to increase market efficiency or provide for a beneficial redistribution of surplus.

Over all, the results call into question the reliability of average price convergence as a metric of market efficiency in wholesale electricity markets. The results showed that bidding strategies do not have to consistently predict the correct price difference, and hence the load deviation, in order to be both profitable and provide for price convergence in expectation. These strategies do not rely on improving the expectation of RT load but rather on taking advantage of predictable price differences given the structure of the market. For cases with unbiased DA load, as seems to be the case in PJM, these strategies will deteriorate unit commitment and scheduling half of the time. If load behavior is similar at other ISOs this calls into this calls into question claims by Jha and Wolak (2014) that virtual bidding led to a reduction in generation costs and emissions as a result in improvements in unit commitment as virtual bidding may not really be helping improve unit commitment. The possible existence of these strategies that nonetheless result in greater expected price convergence suggests that in the context of wholesale electricity markets price convergence is an unreliable measurement of market performance. As such, price convergence should neither be a policy objective in it of itself, nor be used as the principle metric for evaluating market performance because there is no guarantee that it will lead to more efficient outcomes or provide evidence that market efficiency has increased.

Regarding UTCs, the analysis showed a particularly concerning aspect which is the ability it gives traders to engage in low risk high volume strategies that exist when there is ample transmission capacity and a small possibility of RT market congestion, analogous to the high transmission capacity scenarios explored with our model. The UTC product allows for the maximum exploitation of such circumstances by taking the guesswork/risk out of trying to identify the maximum quantity of UTC MW to bid that creates little or no DA congestion. Traders can simply submit a bid for a very large quantity of UTC MW with a very small reservation price, guaranteeing the low risk for traders and little to no impact on price convergence. Our analysis indicated that this strategy is one of the most profitable which is counterintuitive given that for a product that speculates on transmission prices is most profitable when congestion is least likely to occur. An analysis by PJM found that these are indeed the most popular strategies as 51.1 percent of all cleared UTCs had bid offers between ±$1/MW (PJM, 2015). Our results showed that these strategies will have the least impact on prices which support claims by the IMM for PJM that there is little evidence that UTCs cause nodal price convergence.

The perspective of wholesale electricity markets as financial markets with the ISO as the clearing house, as presented in the existing literature, can lead to an extrapolation of what the markets can reasonably accomplish. UTCs are a prime example of this. It was a more complex financial product introduced in response to calls to further integrate wholesale electricity markets with financial markets. However, the result was a product that while highly profitable for traders does not provide for nodal price convergence and degrades the operation of the market. High volumes of UTCs which require substantially increase market clearing time because they require PJM to make manual adjustments to unit commitment and transmission line limits clearly decreases market transparency and predictability. The current market structure has the ISO as the clearing house with load and generation as the unwitting counterparties to the speculative trades by financial participants. If virtual bidding cannot produce obvious market benefits the merits of this market structure should be revisited. All of the combined evidence suggest that products such as UTCs might be better suited to be traded in a separate commodity exchange where they do not disrupt the physical operation of the market and where participation and exposure to speculative behavior is more consensual.

6.6 Conclusion

The analysis presented in this research is based on parsimonious, stylized models of two-settlement electricity markets that include the different virtual bidding products. The analysis shows that welfare analysis provides a more comprehensive assessment of the impact of virtual bidding by financial participants in wholesale electricity markets. The analysis also shows that price convergence as the commonly accepted measurement of market performance in the academic literature while convenient is not very informative, not all circumstances that result in price convergence are necessarily accompanied by greater market efficiency. This means that the profits made by virtual bidding did not buy the physical market participants any benefit in return. Net payoffs for virtual transaction amounted to $259 million for the 2014-2015 planning year in the PJM Interconnection market (PJM, 2016a). With such large sums of money at stake, it is important to ensure that there are benefits to the market. Social welfare analysis should play a central role in the consideration of any further expansion of virtual bidding for electricity markets.

This research was the first to study the UTC product for use as a financial instrument to speculate on differences between transmission prices in the DA and RT markets. The analysis showed that, compared to INCs and DECs, UTCs have an inconsistent impact on price convergence. When they have an impact on prices they can cause price convergence on one node and price divergence on the other node. There is evidence that this is exactly what is happening as the IMM for PJM reported that if separated into their INC and DEC components 95 percent of cleared UTCs have one end that makes money and one end that loses money (need citation). The most important finding regarding UTCs is that the nature of the product makes it so that the most profitable bidding opportunities do not rely on seeking arbitrage opportunities that lead to price convergence but rather on taking low risk high volume positions that have little impact on price convergence. Overall UTCs as a financial arbitrage instrument do not seem to complement existing market products, features, and goals.

The models employed in this research ignored several important features of wholesale electricity markets. The transmission system, which is one of the features that distinguishes wholesale electricity markets from other markets is not considered in the INC and DEC model and only a simple two node system is considered in the UTC model. Expanding the analysis to include a more complex transmission system to capture the externalities caused by loop-flow is a priority for future work. In this research it was assumed that the DA and RT models are implemented identically, however as pointed out by Parsons et al. (2015) this is not the case as the sheer complexity involved in solving the market models necessitates approximations that are applied differently between the DA and RT market models. Evaluating the impact of virtual bidding in the context of a different implementation of the DA and RT market models that create predictable modelling differences would also be an important extension this work. This research briefly covered the impact of UTCs on congestion rents however the impact on FTR holders was not considered. There is a large interaction between UTCs and FTRs as they both gain their revenue from congestion rents which as pointed out by the IMM for PJM has contributed to FTR underfunding. Incorporating FTRs into the analysis would provide a more complete assessment of the impact of UTCs.

REFERENCES

Bessembinder, H., Lemmon, M.L., 2002. Equilibrium pricing and optimal hedging in electricity forward markets. Journal of Finance. 3, 1347-1382.

Birge, J., Ali H., Ignacia M. & Michael P. (2014). Limits to arbitrage in electricity markets, Working Paper, October 2, 2014.

Bowden, N., Su H. & James P. (2009). Day-ahead premiums on the Midwest ISO, Electricity Journal 22(2), 64-73, March.

Borenstein, S., James B., Christopher K. & Catherine W. (2008). Inefficiencies and market power in financial arbitrage: a study of California's electricity markets, Journal of Industrial Economics 56(2), 347-378.

Canchi, D., (2010). Financial transmission rights: valuation and auction structure. Dissertation. Purdue University, 2010

Celebi, M., Hajos, A., Hanser, P.Q., (2010). Virtual bidding: the good, the bad and the ugly. The Electricity Journal 23, 16–25.

Chao, H., Peck, S. (1996). A market mechanism for electric power transmission. Journal of Regulatory Economics. 10, 10-59.

Douglas, S. & Popova, J. (2008). Storage and the electricity forward premium. Energy Economics, 30, 1712-1727.

FERC, 2005. Economic dispatch: concepts, practices and issues. Palm Springs California, 11.

FERC, (2015). Technical Conference on PJM Interconnection. Federal Energy Regulatory Commission. January 7, 2015. Available at: http://www.ferc.gov/EventCalendar/EventDetails.aspx?ID=7625&CalType=%20&CalendarID=116&Date=01/07/2015&View=Listview

Hadsell, L., & Shawky, H. A. (2006). Electricity price volatility and the marginal price of congestion: and empirical study of peak hours on the NYISO market. Energy Journal, 27(2), 157-179.

Hadsell, L., & Shawky, H. A. (2007). One-day forward premiums and the impact of virtual bidding on the New York wholesale electricity market using hourly data. J. Fut. Mark. Journal of Futures Markets, 27(11), 1107-1125.

Hadsell, L. (2008). Day-ahead forward premiums on the New England ISO. Electricity Journal 21(4), 51-57.

Hadsell, L. (2011). Inefficiency in deregulated wholesale electricity markets: the case of the New England ISO, Applied Economics 43, 515-525.

Hogan, W., 1992. Contract networks for electric power transmission. Journal of Regulatory Economics. 4.3, 211-242.

Haugom, E., Ullrich, C., 2012. Market efficiency and risk premia in short-term forward prices. Energy Economics 34, 1931-1941.

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

Jha, A., Wolak, F. (2014). Testing for market efficiency with transaction costs: An application to convergence bidding in wholesale electricity markets. Technical report, Working Paper.

Ledgerwood, S. D., & Pfeifenberger, J. P. (2013). Using virtual bids to manipulate the value of financial transmission rights. The Electricity Journal, 26(9), 9-25.

Longstaff, F. & Wang, A. (2004). Electricity forward prices: a high frequency empirical analysis. The Journal of Finance, 59(4), 1877-1900.

Li, R., Svoboda, A. J., & Oren, S. S., (2015). Efficiency impact of convergence bidding in the California electricity market. Journal of Regulatory Economics, 48(3), 245-284.

MISO, (2016). 2015 State of the market report for the MISO electricity markets, Potomac Economics, Independent Market Monitor for MISO, June 2016.

Monitoring Analytics. (2013). MISO workshop on virtual spread proposal November 18, 2013. available at: https://www.misoenergy.org/_layouts/MISO/ECM /Redirect.aspx?ID=163938.

Parsons, J.E., Colbert, K., Larrieu J., Martin, T., Mastrangelo, E. (2015). Financial arbitrage and efficient dispatch in wholesale electricity markets. MIT Center Energy and Environmental Policy Research.

Pirrong, C., & Jermakyan, M. (2008). The price of power: the valuation of power and weather derivatives, Journal of Banking and Finance 32:2520-2529.

PJM. (2010). State of the market report for PJM, 2009. Volume 2: detailed analysis. Monitoring Analytics, Independent Market Monitor for PJM, March 2010.

PJM, (2015). Virtual Transactions in the PJM Energy Market. PJM Interconnection. October 12, 2015.

PJM. (2016[a]). State of the market report for PJM, 2015. Volume 2: detailed analysis. Monitoring Analytics, Independent Market Monitor for PJM, March 2016.

PJM. (2016[b]). Data directory. Daily cleared INC, DECs, and UTCs. Available at: http://pjm.com/markets-and-operations/data-dictionary.aspx. Accessed on August 24, 2016.

Saravia, C. (2003). Speculative Trading and Market Performance: The Effect of Arbitrageurs on Efficiency and Market Power in the New York Electricity Market.

Schweppe, F.C. (1988). Spot pricing of electricity. Boston: Kluwer Academic. Print.

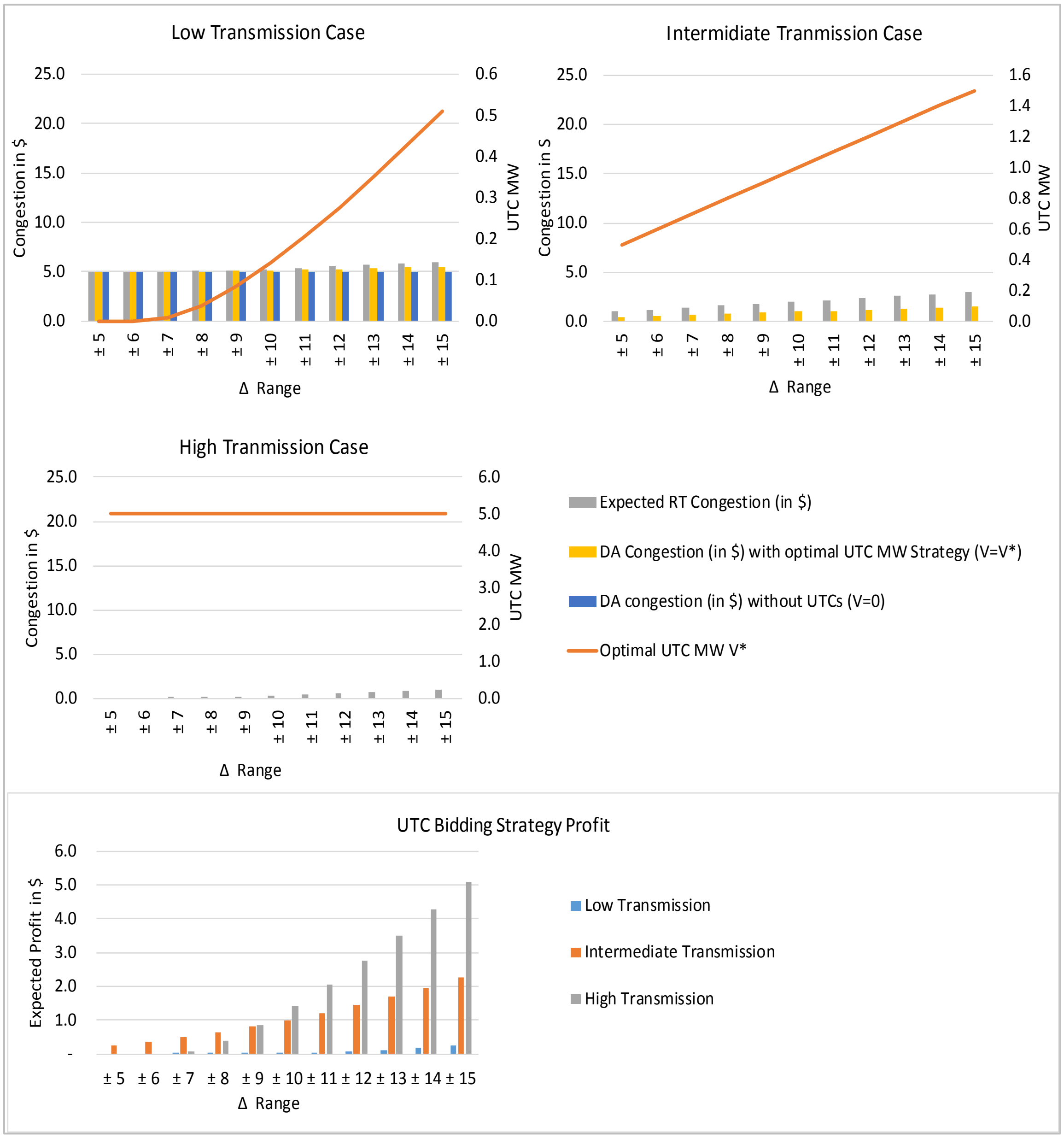
Ullrich, C. (2007). Constrained capacity and equilibrium forward premia in electricity markets, Working paper.

Werner, D. (2014). Electricity market price volatility: the importance of ramping costs, Working Paper.

Woo, C., Zarnikau, J., Cutter, E., Ho, S., & Leung, H. (2015). Virtual bidding, wind generation and California's day-ahead electricity forward premium. The Electricity Journal, 28(1), 29-48.

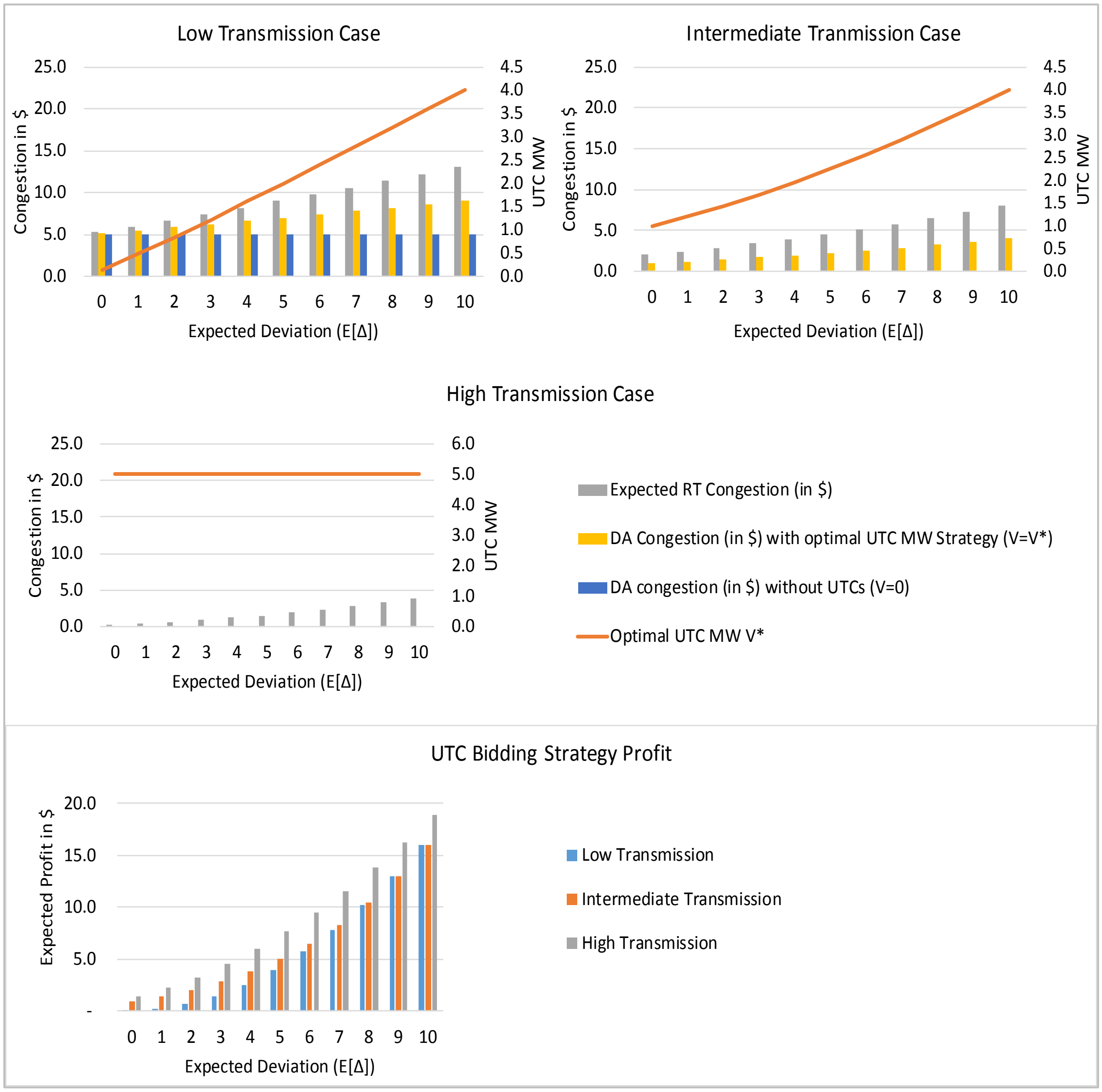
APPENDIX A. NUMERICAL RESULTS FOR THE OPTIMAL BIDDING STRATEGY MODEL

A) No bias scenario



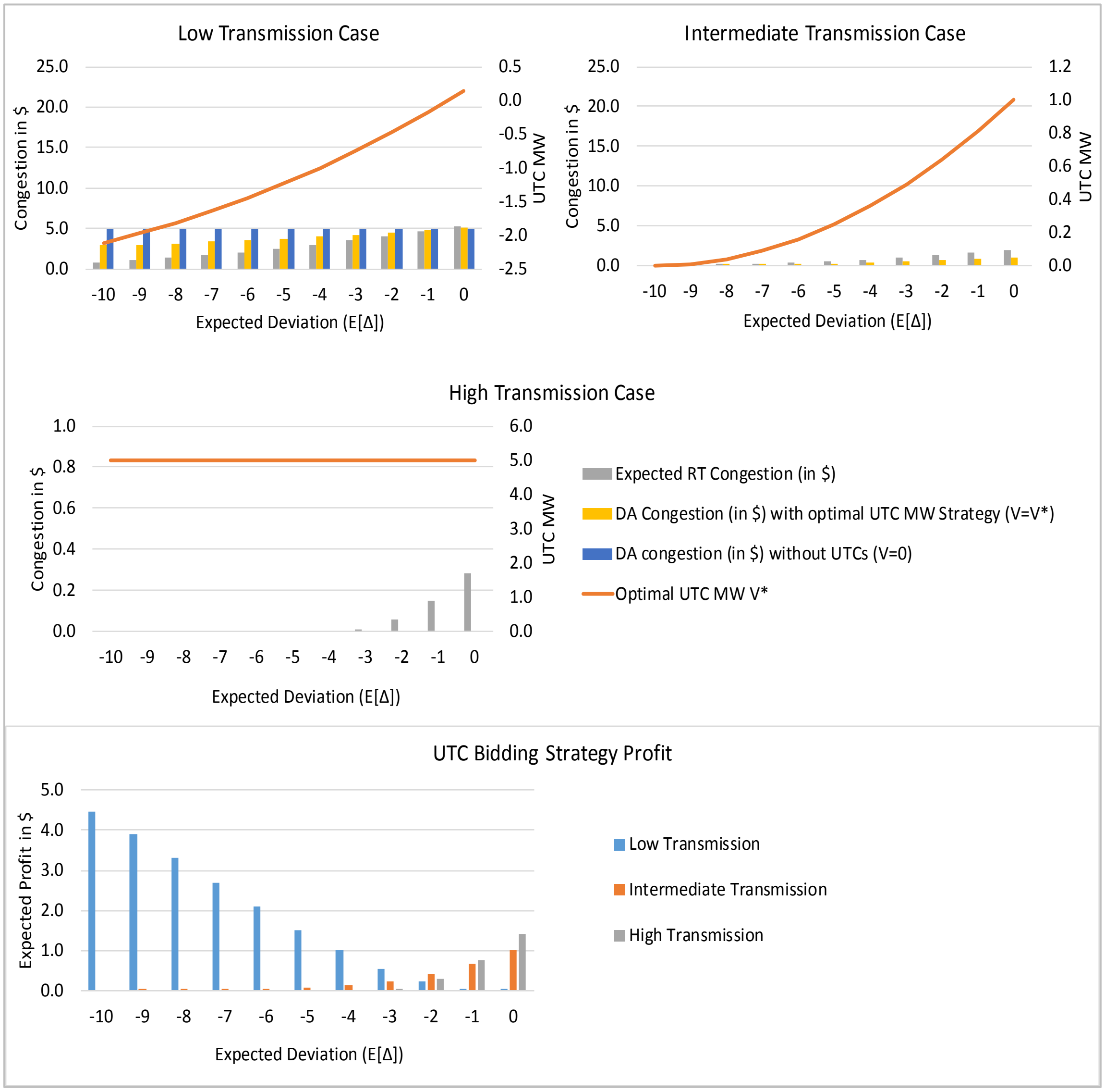
Optimal UTC bidding strategy, congestion, and profits for different levels of available transmission capacity and different ranges of unbiased demand uncertainty.

B) Load underbidding scenario



Optimal UTC bidding strategy, congestion, and profits for different levels of available transmission and different demand underbidding cases.

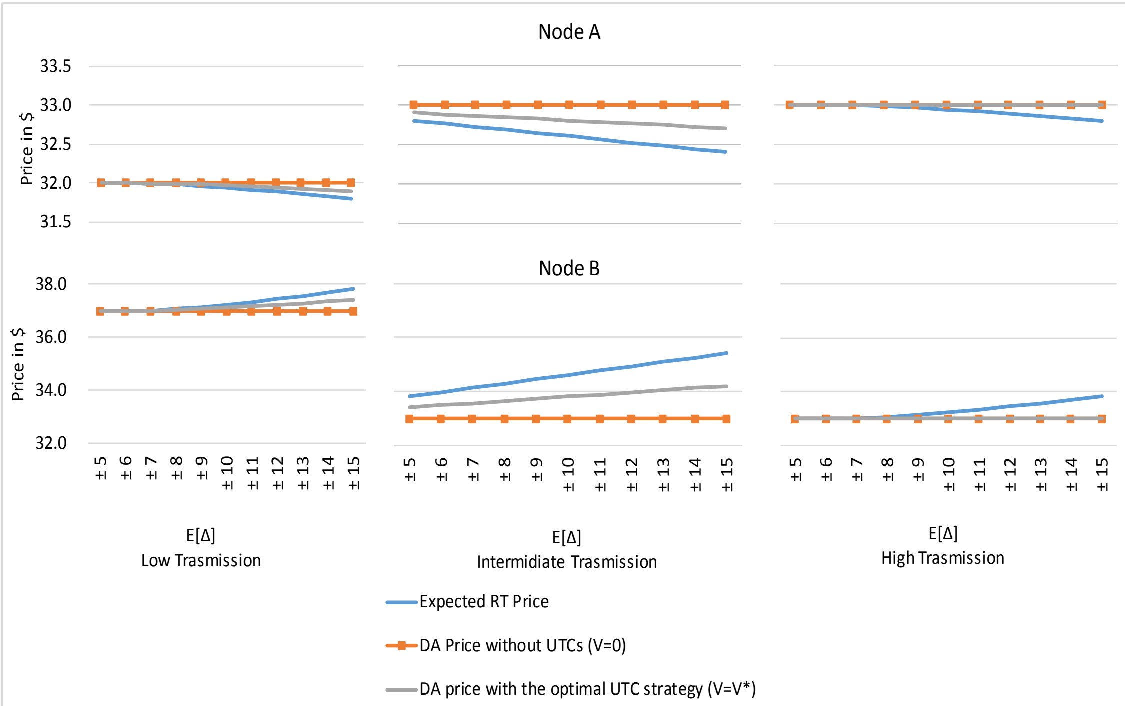
C) Load overbidding scenario



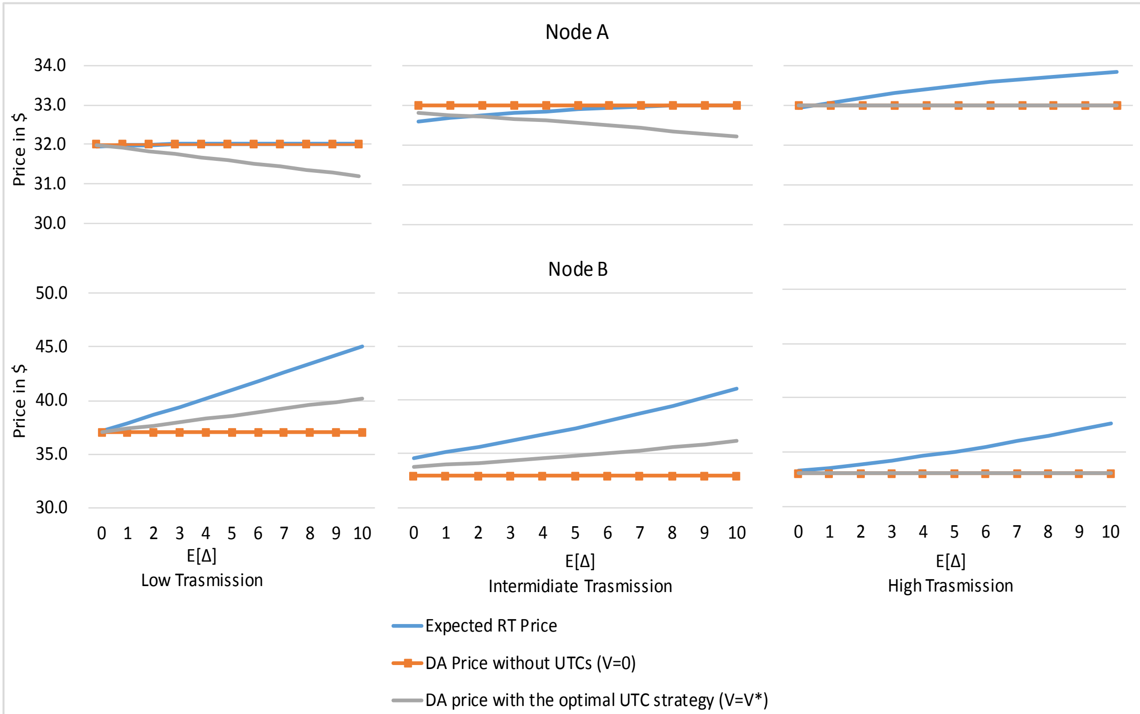
Optimal UTC bidding strategy, congestion, and profits for different levels of available transmission and different demand overbidding cases.

APPENDIX B. NODAL PRICES FROM OPTIMAL UTC BIDDING STRATEGY

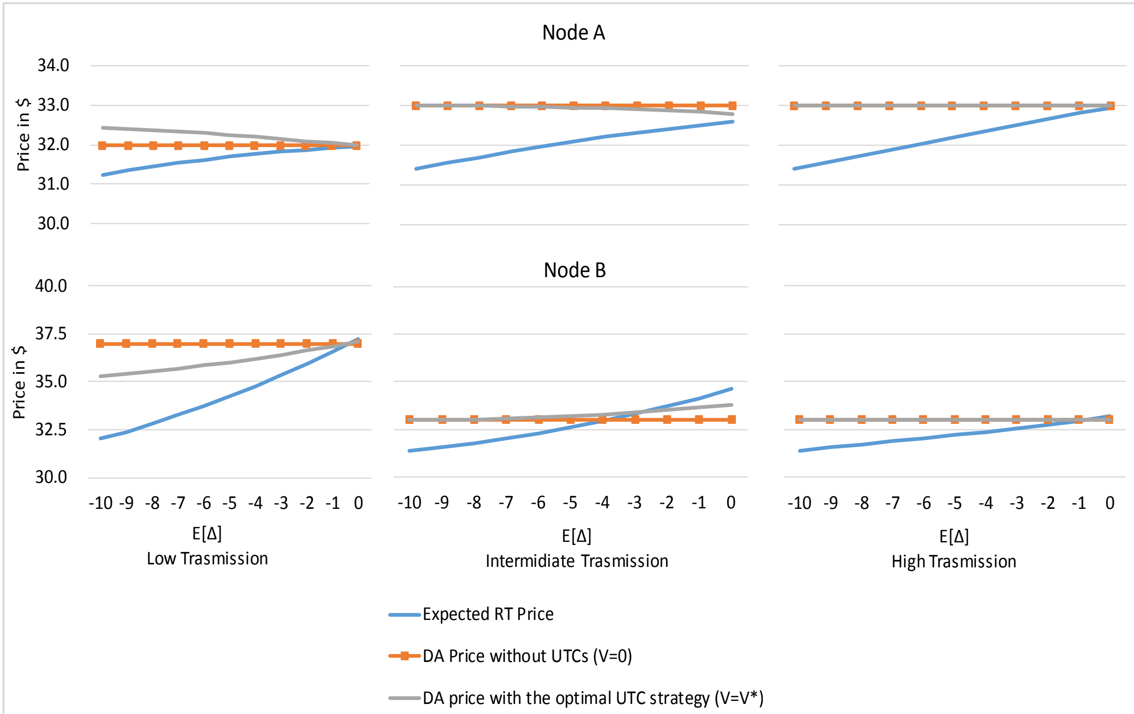
A) No bias scenarios



B) Load underbidding scenarios

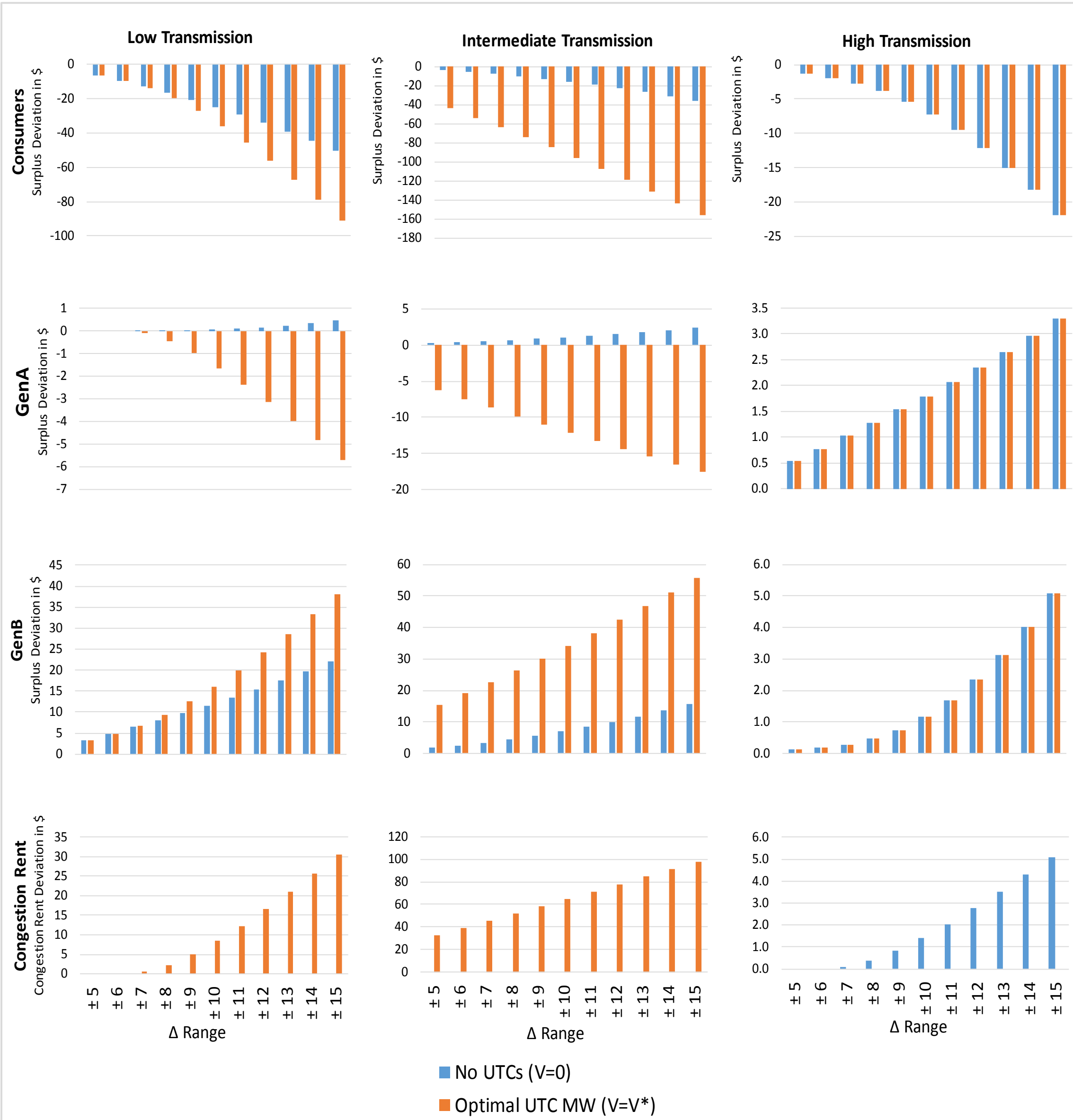


C) Load overbidding scenarios

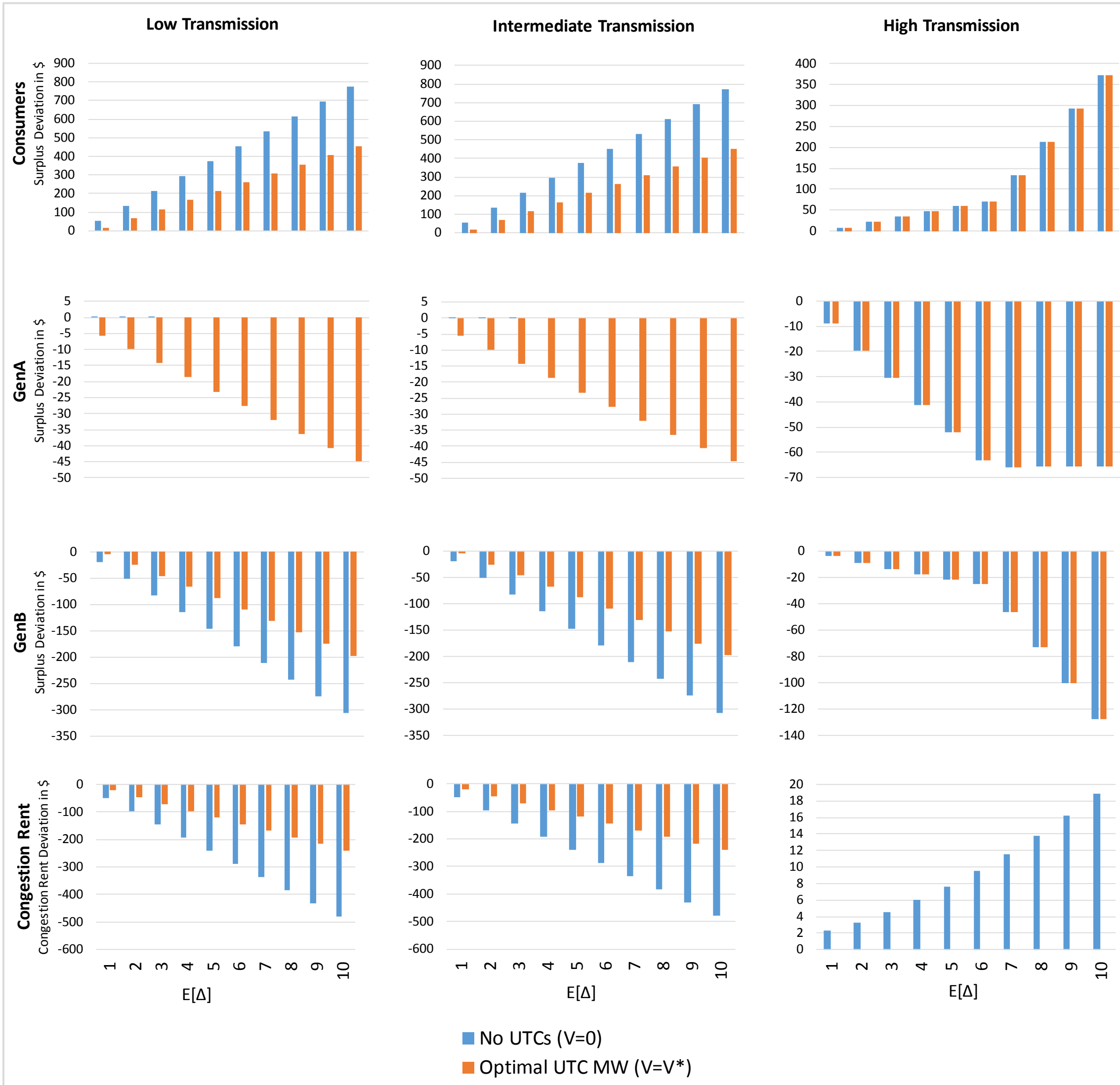


APPENDIX C. WELFARE RESULTS FROM OPTIMAL UTC BIDDING STRATEGY

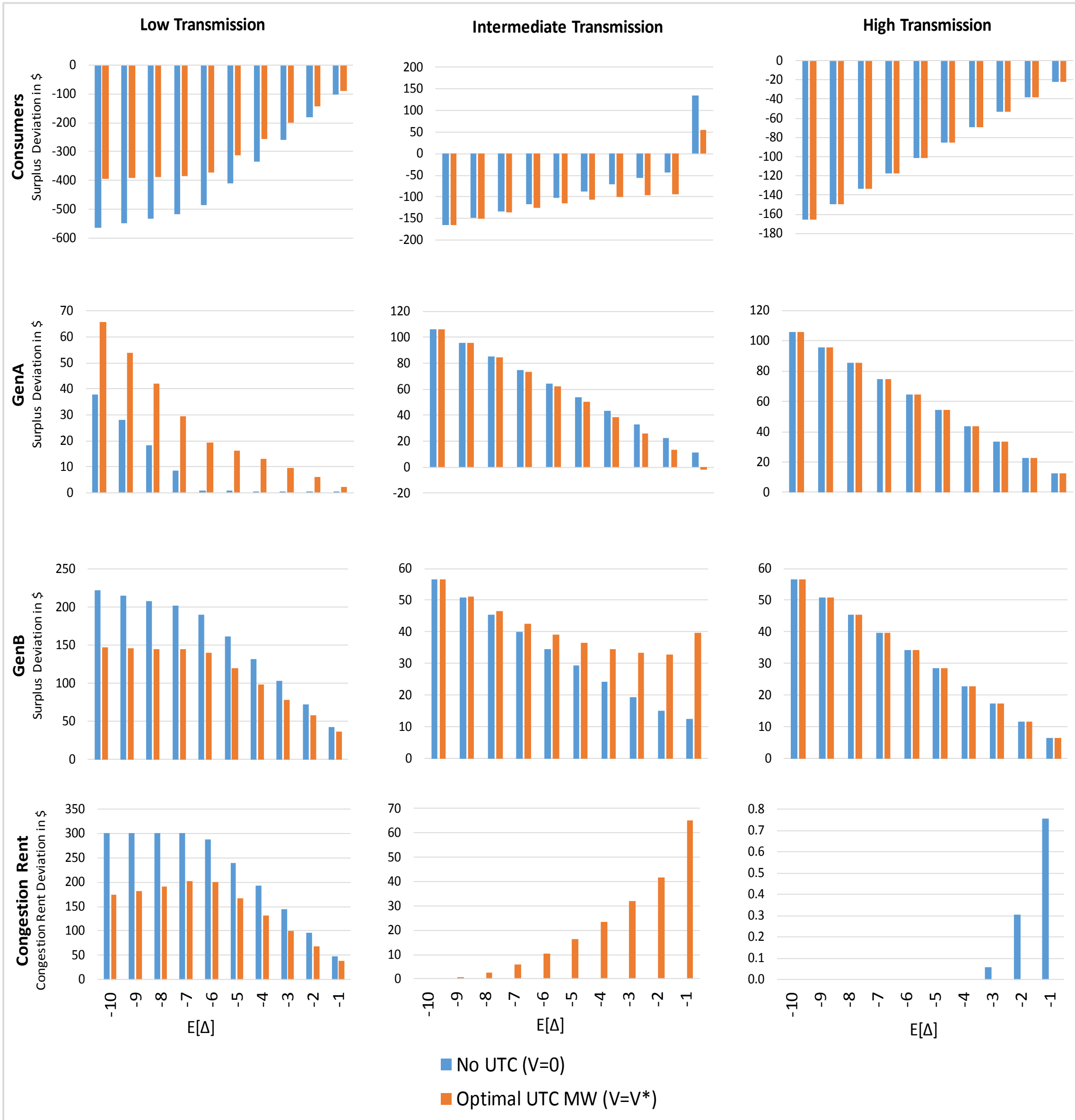
A) No bias scenarios



B) Load underbidding scenarios



C) Load overbidding scenarios



VITA

JUAN S GIRALDO

Education

Bachelor of Science in Business Administration, University of Florida 2009

Master of Science in Agricultural Economics, Purdue University 2013

Employment

03/16-	Economist, Monitoring Analytics, Eagleville, PA, USA.
08/11-02/16	Research Assistant, State Utility Forecasting Group, Purdue University, West Lafayette IN, USA.