PSERC

Electricity Market Structures to Reduce Seams and Enhance Investment

Final Project Report

Power Systems Engineering Research Center

Empowering Minds to Engineer
the Future Electric Energy System
Since 1996

Electricity Market Structures to Reduce Seams and Enhance Investment

Final Project Report

Project Team

Richard Schuler, Project Leader
Timothy Mount, William Schulze, Ray Zimmerman
Cornell University

Shmuel Oren
University of California at Berkeley

PSERC Document 10-01

February 2010

Information about this report

For information about this report contact:

Richard E. Schuler
Department of Economics and Civil & Environmental Engineering
422 Hollister Hall
Cornell University
Ithaca, NY 14853-3501
Phone: 607-255-7579
Email: res1@cornell.edu

Power Systems Engineering Research Center

The Power Systems Engineering Research Center (PSERC) is a multi-university Center conducting research on challenges facing the electric power industry and educating the next generation of power engineers. More information about PSERC can be found at the Center's website: http://www.pserc.org.

For additional information, contact:

Power Systems Engineering Research Center
Arizona State University
577 Engineering Research Center
Tempe, Arizona 85287-5706
Phone: 480-965-1643
Fax: 480-965-0745

Notice Concerning Copyright Material

Acknowledgements

This is the final report for the Power Systems Engineering Research Center (PSERC) research project entitled "Electricity Market Structures to Reduce Seams and Enhance Investment" (PSERC project M-9). This work was sponsored by the Power Systems Engineering Research Center (PSERC) with additional support from the Consortium for Electricity Reliability Technology Solutions (CERTS). The support provided by PSERC's industrial members and by the United States Department of Energy is greatly appreciated.

Industry Team Members: Jonathan Mayo (NYISO), Art Altman (EPRI), Mark Sanford (GE), Eugene Litvinov (ISO New England) & Jean-Marie Coulondre (RTE) provided valuable input during the conceptual phases of this project, and the suggestions for the final report provided by Art Altman and Henry Chao (NYISO) are appreciated greatly. Throughout, numerous individuals at PSERC member institutions answered endless detailed questions about the operation of their markets and the scheduling of transfers across control areas, including Bob Thompson, Rana Mukerji, Rick Gonzales, Nicole Bouchez and the late Chuck King from the NYISO, Andy Ott from PJM, Bob Ethier from ISO-NE and Terry Boston while he was with TVA.

Colleagues: I appreciate particularly the close, collaborative atmosphere within our research team at Cornell. That requires hard work, plus a special disposition, since we span four colleges and multiple departments. In addition to Tim Mount and William Schulze (both in the Dept. of Applied Economics and Management in the College of Agriculture and Life Sciences) and Ray Zimmerman (School of Electrical and Computer Engineering in the College of Engineering), my own appointments are in the Dept. of Economics (College of Arts and Sciences) and in the School of Civil and Environmental Engineering (College of Engineering). Bob Thomas (also Electrical and Computer Engineering at Cornell) was not a formal member of this research team; nevertheless, he provided comments and intellectual and facilities support throughout. That collaborative spirit also exists with colleagues at other PSERC Universities, and in this instance we were delighted to have Shmuel Oren (Dept. of Industrial Engineering and Operations Research at the University of California, Berkeley) join us in this effort. And of course Dennis Ray, University of Wisconsin-Madison, and PSERC's executive director provided his invaluable editing insights, as well as his customary cheery encouragement.

Graduate Students: None of this would have been possible without the creativity and hard work of our Ph.D. students. Never have I had a threesome (Nodir Adilov, Tom Light and Doug Mitarotonda) working together with me on projects (Ben Hobbs came much earlier) who demonstrate the advantages of interdisciplinary collaboration. All three are economists, but they learned the science and engineering that underlies electricity supply and markets (Tom also mastered the planning and design of highway networks), and they delved into experimental economics to be able to help design, conduct and critique the exercises reported here. And whenever their schedules conflicted, Dan Shawan, one of Bill Schulze's students, pitched in as well. Happily, each of these former students is

gainfully employed, Profs. Adilov and Shawan at Indiana-Purdue Univ., Ft.Wayne and Rennselaer Polytechnic Institute, respectively, Dr. Light at RAND Corp. in Santa Monica and Dr. Mitarotonda at the Brattle Group in San Francisco.

Executive Summary

Today's U.S. high-voltage electricity system is operated and dispatched over a hierarchy of control areas that are defined by individual utilities, their power pools, and independent system operators (ISO's) or regional transmission organizations (RTO's). While electricity is routinely transferred across these boundaries, maintaining the reliability of the grid within each control area has first priority. So external transfers are usually arranged outside of, and before, the least-cost optimization routines are performed to dispatch generation reliably within a particular control area. If subsequent changes in weather, demand or equipment outages occur, "seams" may arise where electricity does not always flow from low- to higher-priced areas even though there is adequate transmission capacity, and so potential economic benefits may be lost.

A second concern, particularly where markets are used to dispatch and price generation, is to ensure that adequate investment in new facilities is forthcoming to maintain system reliability. If either the intensity of competition, or regulated price caps keep energy prices low in these markets, revenue streams may seem inadequate to cover the capital cost of the new generation that is projected to be needed to meet future demand. Most ISO/RTOs have added installed capacity markets to their energy (spot) markets in order to support the needed capacity. But because of the four year or longer planning, permitting and construction lead-time that is required to complete new generation or transmission facilities, those capacity auctions are rarely held prior to the time when new construction has to be committed, and so investment becomes a speculative venture.

Our research explored these two contemporary issues:

- enhancing economically efficient trade across neighboring electricity control areas
- structuring forward markets to facilitate adequate efficient investment in electric supply infrastructure.

Both topics raise fundamental questions about the proper design of markets when substantial costs are incurred for developing new production capacity and/or for transporting the product. Economic theory is sparse on these subjects. Nearly all other markets have evolved over centuries of trial and error experiments of the whole, but electricity markets have been widely available for little more than a decade. Given two unique attributes of electricity, that it can't be stored (optimal supply over time is simpler with storage) and that its reliability has public-good aspects that must be regulated, the need and opportunities for further improvement in electricity market designs are not surprising.

Theory and Experiments

Much of the economic theory about efficient markets assumes perfect competition where no individual can affect the market's outcome by their own behavior. Because electricity markets are conducted thousands of times a year over many periods in which similar supply and demand patterns occur, it is reasonable to presume that both learning by participants and subsequent strategic behavior may be possible. Also the typical assumption in finance theory that arbitrage is perfect (forward and spot prices are always equal) may not necessarily be true in these repeated markets. Since analytical methods do not exist that reflect all of these behavioral possibilities, particularly when considering the complexity of electricity flow over a complex network, experiments were designed and conducted with human participants (Cornell graduate students who were paid in proportion to the profits they earned) in order to test the validity of hypotheses.

Conclusions Based on Theory and Experiments

Our theoretical analysis of spatial competition suggests that introducing arbitrage across boundaries can improve the competitiveness of adjacent markets, but that perverse flows from high- to low-priced areas may not be totally eliminated by that spatial competition when transport costs matter. These hypotheses were confirmed by our subsequent experimental trials. Theoretical analyses of the effect of forward markets on investment suggests that if they are conducted before the lead time needed to plan and construct new physical facilities, they can enhance the competitiveness and lower prices in the subsequent spot markets. Again our experimental trials confirmed these results where the forward markets are voluntary and accommodate financial arbitrage.

Methodology and Results

Experiments on Arbitrage Markets to Reduce Seams (Over Space).

Two IEEE simulated 30 bus AC PowerWeb networks, each with six generator busses, with and without a connecting tie line, were used as the test bed for clearing simulated wholesale markets. Suppliers were represented by three and then six individuals in each control area who operated the generators, made price and quantity offers into the wholesale power market and could bid for directional transfer rights over the tie-line. Each system was dispatched by its own system operator according to a least-cost AC optimal power flow (OPF), subject to all line and voltage constraints. A separate uniform clearing price was set by the last accepted supply offer in each region, and each supplier was paid in proportion to the profits they made, adjusted for transmission costs from each generator node. Suppliers also paid or received the revenue from transfers over the tie-line (the price difference across the terminals times the directional quantity flow-rights they acquired), when it was connected. A combined OPF using cost-based offers was also computed as a benchmark for socially-optimal conditions. In the experiments, the demand for energy varied between periods but it was assumed not to be price sensitive, thus the trials focused on supplier behavior. The initial network constraints and generator costs were calibrated so that under a prearranged bi-lateral flow across sixty percent of the tie-line's capacity, perverse flows would arise during some demand conditions.

The exercises yielded the following statistically significant results:

- Increasing the number of competitive suppliers from three to six by connecting two neighboring regions (each with three suppliers) through a tie line with arbitrage across it reduced both the prices that buyers paid and the suppliers' profits in both regions.

- An alternative way providing six competitors within each region – by separating the ownership of existing generation – had an even greater effect on reducing prices and profits; although even further gains were obtained by connecting these two regions, now each with six generators, using the tie-line and arbitrage market (providing twelve competitors, total).
- In the case with three generators in each region but connected by the tie-line under a pre-arranged bilateral contract, perverse flows arose, and adding an arbitrage market was only partially successful in further reducing prices and profits.

While adding the tie-line and spatial arbitrage market improved the competitiveness of these power markets, perverse flows remained in most instances because of continued speculative behavior by some suppliers. Thus the results never reached the socially optimal condition of least-cost generation because of continued speculation; although longer trials may have led to greater improvement.

Experiments on Forward Markets and Generation Investment

Three Cornell graduate students each played the role of a physical generator who made sequential capacity investments and then price-quantity offer decisions into a spot market where the quantities could not exceed previously built capacity. The initial baseline treatment had no forward markets in which all participants were given experience. The two treatments with forward markets (one before, and one after the investment decision, but always prior to the spot market) added three individuals to play the role of intermediaries. These arbitragers were not constrained to sell their voluntary purchases from the forward market into the spot market, since they were allowed to dump any residual into an external market at a very low price. Thus efficient arbitrage was not imposed! The demand schedule in the spot market had a slight price-responsiveness and varied randomly by up to plus or minus ten percent. All generators were assigned identical production and capital costs, roughly proportional to that in the electric industry, and their production capacity depreciated by ten percent per market sequence but they were free to invest in as much new capacity as they wanted. The physical owners of generation were always free to speculate with any supply not committed in the forward market by withholding it from the spot market.

Statistically significant results from experimental exercises showed that:

- Forward markets were efficient (forward and spot prices were not significantly different); although the intermediaries always lost money.
- Forward markets conducted before the investment commitment led to increased investment and reduced spot prices. (These markets differ from the forward procurement markets established by ISO-NE and PJM, since the intermediaries in these experiments were not physical suppliers or demand response providers and their participation was voluntary.)

- Forward markets placed after the investment commitments, like the NYISO's ICAP market, had little significant effect on overall investment or spot prices.
- The shapes of the aggregate spot-market offer curves were different under these three treatments. With a far-forward market, spot offers shift out (increased quantity without a significant change in slope) compared to the case without any forward markets. But the spot-market offer curve following the shorter-run forward market results in a statistically significant steeper (less competitive) aggregate offer structure.

Suggested Future Investigations

Problems encountered in calibrating the electricity networks prior to the seams experiments suggest a number of fruitful topics for further investigation. Every combination of line capacity, demand and production cost assignments that led to perverse flows across the uncongested tie-line under a pre-arranged bilateral contract resulted from that bilateral arrangement being uneconomic if within-control area congestion charges were properly assessed. Thus is one value of establishing an arbitrage market across tie-lines to repair uneconomic long-term bilateral contracts? Also, in attempting to identify "proxy" busses for pricing injections and imports for two control areas when connected with two parallel AC lines, we were not able to identify any single bus that properly reflected the changes in prices that were experienced over the range of line flows that occurred in these experiments.

Continued strategic behavior in the seams experiments and the persistence of perverse flows suggest that longer trials, although difficult and expensive to conduct, may be warranted to understand both the learning process and the likely behavior of suppliers. The impact of independent arbitragers who don't own generation might also be explored.

Insights about the effects of forward markets might also benefit from longer trials and more participants in each market. The failure of arbitragers to make any money was troubling, as was the chronic over-investment by physical suppliers (presumably as a strategic device to drive competitors out). In the long run, the important question remaining is if forward markets are put in place and held before investment must be committed, will anyone participate voluntarily?

Table of Contents

List of Figures

List of Tables

Section 1. Introduction

Reducing barriers to trade across two neighboring electricity control areas, the "seams problem", and designing markets to facilitate efficient investment in electric supply infrastructure may seem like two very different problems, but the underlying economic concepts to be applied have much in common. Both issues require the proper "spacing" of markets, the one over geographic space and the other over time. And in both cases, the physical characteristics of the system and its components and the availability of adequate information should influence the structure of efficient markets. As an example, separate distinct markets usually emerge over space only when scale economies concentrate the production and/or consumption of a good or service at finite locations and when the cost of transporting the product from provider to user is significant. For many goods and services, it is the huge leap in the cost of transport across a significant physical barrier like a mountain, river or ocean that delineates a market's geographic size. In other instances, political borders and the rules or charges for crossing them define the markets' boundaries. In the case of electricity, all of these factors pertain, with an additional public concern about maintaining the reliability of service. When electricity is supplied over a large, complex network, there may be a size at which existing administrative skills and computational technology limit the ability to manage and operate these complex dynamic systems reliably (See Schuler, "Float Together; Sink Together - - " [1]). Because most electricity exchanges are "smart" markets (e.g. cleared through a cost-minimizing algorithm that is subject to all of the physical, reliability-based constraints of the power network), separate markets over space make sense when there are physical limits to the span of feasible monitoring, coordination, optimization and control. If so, operational security limitations may place additional bounds on the desired size of individual power markets and/or the conditions for exchange across those boundaries.

The dominant characteristic that governs electricity market design over time is the fact that it cannot be stored economically in appreciable quantities, mandating that adequate generation and transmission facilities are in place to meet the demand in real time or else selective shortages, or worse, unintended interruptions of service will occur. Again, a physical factor, the need for adequate facilities and the minimum gestation period for completing new investments of at least three to five years establishes the need for market mechanisms well before the time when the electricity is actually consumed (the spot market). In this case the physical limitation that helps to define the desired market structure is how rapidly facilities can be developed and made available to furnish energy. Those minimum physical lead times suggest the desired inter-temporal market structure. As an example, in existing electricity markets, shorter-term physical limits that restricts the speed of starting up and taking down some generating units (their maximum ramp rates) is the reason for having both day-ahead and real-time markets. It also forms the rationale for distinguishing between different categories of operating reserves (regulation, spinning and non-spinning reserves) based upon the different ramp rates of various types of generation. Each provides different value to the system operator in responding to unanticipated events.

The question addressed in this report is: could longer-term forward markets facilitate decisions to commit to the construction of new facilities, and if so, would those markets be efficient? While a long-term bilateral contract between one buyer and one seller may be adequate to arrange the construction of a new generator, the cost of finding the right coincidence of interests (each person wanting to buy and supply the precise identical amounts) is extremely expensive. By comparison, a market mechanism that accumulates the expressed interests of many different buyers and sellers in smaller increments can reduce the costs of exchanging information and provide assessments about the future at much lower costs. Equally important, if those estimates about the future are expressed in the forms of bids and offers into a properly-designed market, they are highly likely to be truthful, since each individual participant's potential gain is enhanced by acquiring the best possible information in developing their position. And since market outcomes are publicly known, whereas the terms of bi-lateral contracts are usually confidential and known only to the contracting parties, all market participants, plus the market overseers and regulators, might have access to better information about estimated future needs and costs through a forward market structure (Note: if no one chooses to participate in a well-structured forward market, that too is valuable information.).

1.1 Existing Market Structures for Electricity: How We Got Here

Wholesale electricity markets in the United States are a relatively recent phenomenon having emerged after over a century's operation by the industry in a very different, highly regulated environment as coordinated, frequently vertically-integrated, utilities in most locations. It is not surprising, then, that the initial market systems reflected many of those previous institutional and operational boundaries, since reliability-wise, the regulated systems had worked well. Building markets upon existing institutional structures and geographical boundaries also eased the transition to market-based operations. As an example, the first areas to establish ISO/RTOs were those that previously coordinated the bulk power operations of several utilities through integrated power pools (e.g., PJM, NYISO, ISO-New England and California), and the subsequent markets were defined within the overall spatial boundaries of those previous control areas. Furthermore, many of those existing power pools had dispatched generation according to a cost-minimizing algorithm that mirrored what a market might have done. What was different was that voluntary offers and bids into a market were substituted in the optimization for regulatory-approved, cost-based numbers furnished by regulated utilities. In most cases, pre-existing power pools defined the early spatial structure of electricity markets; although in many instances those broad markets were subdivided into smaller spatial zones to reflect further barriers created by congestion on transmission lines (a physical barrier).

With respect to structure over time, most of the initial markets also adapted the existing decision-making framework and procedures that had been widely-used under the regulated-utilities' coordinated power pool regimes. A primary emphasis was to demonstrate that the reliability of the system could be maintained under a market regime. Day-ahead markets were considered essential at the outset to be sure that sufficient time was available to anticipate the next day's needs, to be sure the necessary generation was committed and to be certain there was sufficient time to “clear” the market software. Markets for “ancillary” services that are essential tools used by system operators to maintain reliability, like operating reserves and regulation, were also rapidly introduced. Furthermore, since the physical system always clears in real time, real time markets and hour ahead balancing mechanisms (between day-ahead and real time) were also quickly adopted in many jurisdictions to coincide with these physical and decision-making constraints. In those jurisdictions that had mandated the separation of generation from owners of transmission and distribution, a mechanism was needed to ensure the availability of adequate generation capacity for the load serving entities that retained the obligation to provide reliable service. Short term capacity markets (e.g. ICAP, LICAP) were initiated in many jurisdictions, anywhere from one to six months ahead of the real time when electricity flowed, as one mechanism to make those prior commitments, but none of these initial forward or options markets were conducted more than six months prior to real time. Since the elapsed time required to plan for and construct new generating facilities is at least three to four years (most demand reduction investments require less lead time), a commitment to invest has to be formulated on the basis of estimates of prices further into the future than six months. Those long range forecasts can be made by the people committing to the construction and/or their financing, but it is also helpful if forward markets are in place to reflect the expectations of other parties about future prices. Relying upon a long term contract to initiate construction merely shares the burden of who’s taking the risk of those forecasts of future prices with a buyer; it does not introduce a larger number of players who may have better or different information into the decision-making process, as might occur under longer-term forward markets.

By comparison, before the introduction of markets under regulated regimes the adequacy of installed capacity was ensured as a result of periodic utility planning, and/or followed by state-level integrated resource planning that was enforced by command and control mechanisms. Under these government-backed systems or vertically-integrated regulated regimes, it was clear at whom the public could point their fingers (the utilities, their regulators and/or government agency heads) if adequate new facilities were not constructed in time to meet demand. In fact available installed capacity margins usually exceeded requirements (at an added cost to customers), and after 1970 the construction costs of many new utility-built generation facilities began to escalate well above historic rates [2]. Those sharp increases in per-unit capital costs in many locations, coupled with substantial excess capacity, was one motivation for substituting market-based-incentives for traditional utility price-regulated or government contracting methods for developing new generation facilities. The difference is the sequencing between electricity price-determination and cost causation; in the prior regulated regime, new facilities were completed and then their costs were rolled into the recomputed price. Under markets, the price comes first, as set in the market-place, and it’s up to developers to determine whether and how to build facilities at lower costs so they can make a profit. But under a market-based system, the coordination problem still exists of fitting individual investment pieces into a complex electricity supply network. Some form of integrated planning by an oversight regulatory or governmental authority is still required under a market-based regime as discussed by industry executives in a PSERC executive forum, Project M-16[3]. The questions addressed in this report assume the requisite overview system planning is in place that defines the potential needs and/or benefits and costs of improvements and that the projections from this analysis are widely publicized. Given that planning, what type of market-based instruments would help to provide adequate incentives for the development of needed generation under a market regime, what should the structure and timing of those markets be and what obligations and penalties for non-performance should be associated with those market-based commitments?

1.2 Motives for Evolving Electricity Market Structures

What was not available from the previous regulated-utility and/or government agency supply environment in which investment decisions were made administratively and/or politically was a decision-making apparatus that could be easily converted into a market design to facilitate the same purpose. With most other market-supplied goods and services, particularly those whose production is capital-intensive, it is the shortage of supply and the consequent rise in prices that motivates the investment in new production capacity. Were it possible politically to operate spot electricity markets in the U.S. without price caps, as in Australia, it is possible that the periodic enormous price-spikes that would occur in periods of inadequate capacity might provide a sufficient incentive for investors to risk their capital on constructing new facilities. But in the U.S. where every jurisdiction has placed caps on real-time and day-ahead spot market energy prices, additional financial incentives appear to be necessary to spur investment in a smooth, orderly fashion that is sufficient to provide comfort to system operators and public officials. That is why most ISO/RTOs have currently adopted or are exploring and developing market-like mechanisms to facilitate the long run investment in generation and/or transmission capacity. Both ISO-NE and PJM have instituted three to four year forward "procurement" markets that require their load-serving entities to purchase sufficient additional generation capacity and/or demand reduction procedures in order to meet the projected future demand plus the mandated reserve margin specified by the Electric Reliability Organization (ERO, a regulatory authority authorized by FERC to impose penalties for non-compliance). And as Mount, Schuler and Schulze [4] have emphasized, the installed capacity that is required in addition to what's needed to meet the forecast peak energy demand (to ensure a high probability of continuous supply when generators are being maintained and/or unforeseen but statistically predictable contingencies occur) is a public good because all customers connected to the electricity grid share the added reliability provided by that spare capacity. Because the benefits of that installed capacity margin are received by all buyers, although different customers may value that reliability differently, its demand must be established and assured by a regulatory agency. A market cannot provide an accurate signal of the value of reliability for these customers because of the tendency of all users to "free-ride" on each other. (If when my neighbor buys reliability, I get the benefit too, I have little incentive to reveal my own preference; otherwise my neighbor might buy less and rely on my purchases.). Therefore, the cost-benefit analysis to determine the proper installed capacity reserve margin must be performed by a central authority acting in the public interest. But once that "public" determination about required reserve margins is made, and the parties responsible for securing those margins are assigned, the question addressed here is: can and what type of markets might be used to acquire the needed installed capacity plus the required margin?

New York and New England have operated short term (six months forward or less) installed capacity markets for some time (the NYISO for ten years, ISO-NE over a more recent span) where the load-serving-entities (LSEs) are required to purchase sufficient capacity to meet their peak demands plus the required margin, and to have that capacity committed one month ahead, or else face penalties in the form of likely high-priced acquisitions made on their behalf by the ISO/RTOs. But if there is not sufficient generation capacity in the ground anywhere in the region, no one will be able to buy adequate capacity even at an infinite price. In contrast, if there is a glut of capacity the price will fall to zero. And since investments in new generation capacity is lumpy, there are bound to be periods of excess supply, followed by deficiencies, unless the target is set higher than the actual desired capacity margin. That is why in an attempt to smooth out these prices over time, some jurisdictions have applied an administratively determined demand curve in their installed capacity markets. In addition, in order to provide a fair incentive for investment in demand-side-management initiatives that might reduce the requirements for installed capacity in the future, these demand-reduction programs have also been allowed to participate in installed capacity markets.

What's different about both PJM's and ISO-NE's recently implemented forward "procurement" markets that allow both new installed capacity and demand-management programs to participate is that they are conducted three or more years before the capacity is forecast to be needed. However, these structures are not full forward markets where both physical and financial buyers and sellers may participate. The primary difference between these forward procurement markets (in PJM and ISO-NE) and installed capacity markets (like the ICAP markets conducted by the NYISO) is that the mandatory date for having contracted for projected future needed capacity without incurring a penalty is moved forward from one-month to three years. This much longer lead time provides greater comfort to planners and those responsible for system reliability that adequate capacity will be available when it's needed because new facilities and systems can be put in place within the three year time frame. But these mandatory procurement markets accommodate only contracts for physical delivery and not for financial bets, thereby restricting the variety of information that might inform the market's outcome.

The other difference under a voluntary forward capacity market where the LSEs are still obligated to have arranged for the target installed capacity margins in real time or else face significant penalties, is that suppliers and the buying load-serving entities (LSEs) bear the risk of errors in their respective demand forecasts and therefore have a strong incentive to develop the best possible forecasting tools. Under the mandatory forward procurement markets implemented by PJM and ISO-NE, the ISO/RTO develops the forecast and so the customers pay directly for any errors in those forecasts, either in terms of the cost of carrying excess capacity or the cost of less reliable service. In that sense, mandatory forward procurement markets provide little difference from capacity arranged under regulated rate-of-return rate-making where the customers also pay directly for errors in forecasts by the regulatory authority. The advantage of forward procurement markets over traditional regulation is that they offer suppliers the incentive to keep their construction costs down since the price comes first (before construction is complete).

Another substantial motivation for implementing market-based exchanges for electricity, initially, was to allow buyers and sellers to benefit from wholesale price differences that exist across neighboring power pools. Following the inauguration of markets, substantial steps have been taken to reduce these price differentials across control areas through the elimination of multiple, separate transmission charges (rate "pancaking") and by more frequent re-computation of transfer limits on connecting tie-lines to reflect real-time conditions. Nevertheless, substantial pressure remains to expand the geographic scope of electricity markets beyond existing control areas (ISO/RTOs) because in many instances large price differences remain at the borders. In some cases the direction of power flow is perverse in the sense that it moves from high- to low-priced areas. Thus FERC, and several ISO/RTOs continue to place the elimination of these "seams" high on their list of needed market improvements. In addition, there is a continued concern about the physical and price vulnerability of electricity supplies in particular areas that rely too heavily on a single fuel source. These areas are exposed adversely to both fuel supply disruptions and international political and terrorist activities, and it is thought that a broader geographic market might increase the diversity of suppliers. Tapping renewable-based electricity supplies that are economical in many regions also means reaching out to neighboring markets, or beyond, and all of these factors combine to provide an impetus to develop broader, regional electricity exchanges on an ongoing basis. But given the physics of feasible electricity supply over a network, when, where and how are these aspirations economical and/or realistic?

Section 2. Economic Principles for Guiding the Structure of Markets

Nearly all markets fall short of the economist's textbook efficiency ideal of equating marginal benefits to marginal costs for all incremental transactions, particularly where the cost of arranging each different transaction is appreciable. In the case where the costs required to serve different buyers vary, rather than assigning a different price to every buyer, those costs are normally "averaged" over some market segment in which a uniform market price is assessed, primarily to reduce administrative costs and to avoid confusing customers with complicated price schedules. This averaging can be across a variety of dimensions: space, where typically transportation costs are averaged ("postage stamp" and "license plate" pricing are examples); product quality, where for most mass-produced products every customer does not receive a unique customized version; and time, although if the product is storable and the decision is left to each buyer to provide their own inventories, those timing costs may be individualized. But with the current trend toward just-in-time manufacture to reduce the cost of inventories, if the product's availability on-demand is essential for some customers, then the existence of forward markets to increase the likelihood that delivery will take place when the product or service is needed may be important. Under these circumstances, the frequency and duration of these forward markets may also have substantial efficiency implications.

It is particularly important to think through these pricing principles when a new type of commodity or service is marketed for the first time. By comparison, markets for long-established goods and services that have evolved over many years through many "experiments of the whole" may embody many of these efficiency-enhancing considerations about their market structure. Clear examples of the difficulties that can arise when structural details of markets for similar (but slightly different) commodities are simply transferred to a new product are provided by problems that arose in the early years of operating electricity markets in the U.S. Although the level of reliability (quality) of electric service is (and should be) set by a regulatory body for a service provided over a network, that reliability (and other shared quality characteristics like voltage and frequency) is a public good because everyone in a neighborhood receives the identical reliability regardless of their individual preferences, even though costs do vary appreciably by location (See Mount, Schuler and Schulze [4] for a discussion of this issue). And since electricity is not storable in large quantities, buyers want some assurance that they will receive their electricity when they want (need) it and as much of it as they want.

The general principle that must be considered when devising the optimal segmentation, or "grain", of any new market is to balance four different costs: 1) the inefficiencies of not precisely matching marginal benefit with marginal cost for incremental transactions as the market becomes larger, 2) the greater transactions costs incurred by having a large number of separate market segments (both price-posting and marketing costs for suppliers and decision-making costs for buyers), 3) arbitrage costs across the borders of market segments where substantial price-differences may exist (generally these differences become larger as the market segments become larger), and 4) the effect of the size of market segments on the ability of buyers and/or sellers to behave strategically and exercise market power. Generally the smaller the market segments are, the better the results under the first and third criteria, but this comes at the expense of greater administrative and decision-making costs, plus fewer buyers and sellers in each segment may lead to a greater exercise of market-power (and therefore an offsetting decline in allocative efficiency). Thus trade-offs must be made across these ideal components of market grain in the practical design of nearly every market. Oren [5] discusses some of these tradeoffs in the context of proposing forward markets for the electric industry. These considerations have become particularly applicable to large industrial economies that are characterized by scale economies in production which means that producers are concentrated at finite locations for discrete product groups. By comparison, in a locally "self-sufficient" society, every buyer provides everything they require so there is no need for markets, spatial or otherwise; although if those individuals couldn't produce all of the goods that they required instantaneously, on-demand, and the commodities were not storable, or if each individual produced a slightly different variety of each product, markets might still spring up. But when scale economies in production and transportation costs are added to buyers and producers who are spread across the landscape, the question of market-grain again becomes relevant. Neo-classical economic principles presume all human activity transpires on the head of a pin; market realities are otherwise.

In fact, many market boundaries are defined by physical constraints that limit transactions. In the spatial context, geographic barriers like mountains, deserts, rivers and oceans (as well as socio-political gaps) define many market boundaries, but many technological advances over the years have allowed us to span many of these barriers. Indeed, in many cases it was the price/quality differences across those boundaries that provided the incentive to breach them. As an example, spatial markets for electricity are often defined by a series of junctions where congested transmission lines limit flow across these boundaries. These physical barriers have been used to define pricing zones, and it is thought that the differences in locational marginal prices (LMP) across these zones provides an incentive for the efficient location of new production, as well as a signal for the needed construction of additional transmission facilities. So too where there are physical constraints on the ability of suppliers to meet demand and/or on the ability of buyers to adjust their usage, a proper spacing of forward markets over time may enhance efficiency. In the case of electricity supply, there are a range of decisions from a short run choice of committing a unit for the next hour (or day) that results in start-up costs, to decisions on scheduling prolonged maintenance, on through the long-term choice of building additional capacity. Since each of these decisions has some minimum physically-limiting lead time before electricity is actually generated, as do decisions by customers on economical ways of reducing their consumption in particular intervals, the existence of a forward market within a similar time-frame to those physical limitations on changes in supply and demand should provide additional information to assist in the decision, as well as offering opportunities to hedge risks and/or engage in strategic posturing. But the emphasis in structuring market segments is to identify these physical limits on transactions, on the ability to make decisions and on feasible rates of change in these physical limits as the starting points for establishing market boundaries.

2.1. Markets over Space

When customers arrange for a product's transportation, as an example by driving to the shopping mall to make purchases, the price is customarily quoted at the supply location and all buyers pay the same price at the same time (Mill Pricing (MP)). In this case, each buyer who utilizes the product somewhere other than at the store incurs the transportation cost and effectively considers and pays a different delivered price at the point of utilization. However, when it is most effective to have the supplier also deliver the product to each buyer, then because the supplier can specify the delivered price to each customer, spatial discriminatory patterns of pricing (SDP) may emerge, or uniform delivered pricing (UDP) may be employed as a simplified variant. In both of these cases differences in delivered prices do not reflect the differences in transportation costs that are incurred to reach each customer (See the paper by Holahan and Schuler [6] for a survey of spatial pricing options and the consequences of competition). Technological factors frequently determine who arranges transportation, as in the case of electricity supply where there is only one effective way to haul the product and there are substantial scale economies to be achieved by concentrating that transportation in one provider. SDP and UDP are particularly likely to be applied where production and transportation are vertically integrated within the same entity (See Schuler and Holahan [7] for a discussion of this tendency), as they have been under most regulated- or government-run electricity supply systems. Under both SDP and UDP pricing structures, few customers pay the actual marginal cost of manufacture plus delivery. As an example, “postage stamp” pricing used by the U.S. postal service is an example of UDP where cost differences are not reflected in the prices paid by customers. In many cases where UDP is employed by competitive firms, the cost differences between transactions to different customers is small when compared to the administrative costs of assessing individualized prices, as an example, by origin-destination, weight and volume for every letter and parcel sent in the postal system. Particularly where there is a substantial regulatory involvement in the provision of service, UDP is frequently used because it seems on the surface to be equitable, again as in the case of postal service.

Because there are substantial physical barriers to transporting electricity where inadequate transmission line capacity exists, it is reasonable to tolerate spatial price differences that in some instances are large. Typically in electricity markets, different prices are allowed to emerge in different locations that are effectively separated by line congestion (e.g. location-based marginal pricing (LMP)), but within an un-congested region all buyers may face the identical wholesale price even though the line-losses may differ slightly depending upon location (so UDP within an un-congested zone). Furthermore, where different regions have different operating entities (Independent System Operators (ISO) or Regional Transmission Organizations (RTO)) that are responsible for maintaining service reliability within each of their separate regions, their operators dispatch power to minimize overall region-wide cost, subject to reliability constraints, within their own jurisdictions. The physical necessity of coordinating operations within a set geographic area in order to maintain reliability creates boundaries that might inhibit buyers and sellers who attempt to transact electricity across the borders of those jurisdictions. In fact buyers and sellers across RTOs/ISOs confront many of the same problems encountered in international trade, but with the compounding problem of just-in-time delivery. Until the entire nation (or continent) can be served reliably with confidence by a single power supply entity (a daunting technological challenge at present) it makes sense to separate markets in accordance with this overriding physical operational constraint. Nevertheless, mechanisms also need to be established to facilitate exchanges across these boundaries (efficient arbitrage) and reduce what are called “seams” issues.

As an illustration of how competition works when transportation costs are appreciable, consider two un-congested power systems located on a line (e.g., in a valley) as shown in Figure 1 where there are transmission losses that result in different short-run transportation costs for serving different customers who are distributed throughout each system. Note, that allocating the capital costs of those lines, economically, presents an even more difficult issue that can effect short-run efficiency and therefore will be discussed subsequently. But for this short-run analysis without capital cost recovery, suppose that each ISO contains a single generator, but each has a different marginal cost (as represented by the height of the MC curve at each generator in Fig. 1) and that all customers are distributed uniformly across space with identical demand curves. Under UDP without trade across the border between the control areas, suppose generator #1 charges P1A in its region and generator #2, with the higher marginal cost, charges the slightly higher UDP of P2B in its own, ISO-B, jurisdiction. The marginal cost of delivered power at each location is represented by the rising line (because of losses) as customers are located further away from each generator. Note, under UDP, few customers are paying the actual delivered cost of supplying power. Even if the price were pushed down by a regulatory authority in this example so that generators were earning only a normal level of total profit, under a UDP structure some customers would pay less, while other customers paid more than their full marginal cost of production and delivery. Furthermore, the generator would have a disincentive to serve customers beyond the point where the price received was less than the marginal cost of supply.

Now, open the borders in this example to bi-lateral transactions between the control areas. In this case generator #1 might offer a lower price across the border, P1B, than she charges within her own ISO-A, in order to try to serve all of B's customers up to point R1 where her price P1B just covers her marginal cost of production plus transport. As a result, generator #2 may be induced to lower his UDP below P2B in order to forestall further incursions into his market by generator #1, but supplier #2 is unable to compete against #1 between the border and R1 because of his higher combined marginal production and delivery costs in that territory. Furthermore if ISO-B sets the pricing rules in Region B and requires each generating company to offer the same uniform price to all of the customers it wants .to serve in Region B, if there is no congestion, then generator #2 may be reluctant to compete with generator #1, even between R1 and R2 were generator #1 to try to extend its range further into ISO-B. And, how would generator A try to compete within the market segment between R1 and R2? **It would compete by raising its price slightly above P1B.** (Note this example of the reversal of traditional micro-economic concepts when transportation costs are important; with UDP, generator #1 **competes by raising**, not lowering its price!). In this case generator #2 might be reluctant to meet the competition under UDP since it would have to lower its price to all customers in ISO-B's region, including those to the right of R2. Thus generator #2 might find it more profitable to lose its customers between R1 and R2 than to compete and give up the profitability from its remaining customers. Also note, **this efficiency-enhancing benefit of competition across the border by generator #1 has led to power flowing from a higher price in Region A (P1A) to a lower price in Region B (P1B)** - - a so-called perverse flow! (See the paper by Schuler and Hobbs [8] that analyzes many of the seemingly anomalous results to be expected with spatial competition under UDP.) To be sure with arbitrage and reselling allowed back across the border between Regions A and B, some of the customers in B between the border and R1 might resell power they bought at P1B at a price below P1A to customers in region A. But if the resellers also were required to pay for their transportation costs, the extent of their reselling would be limited and the price would still decline across the border from A to B in the direction of electricity flow. So with UDP, arbitrage would reduce the perverse price difference across the border, but it might not eliminate it completely. (See Hadsell [9] for an analysis of the effectiveness of arbitrage in the NYISO's markets.)

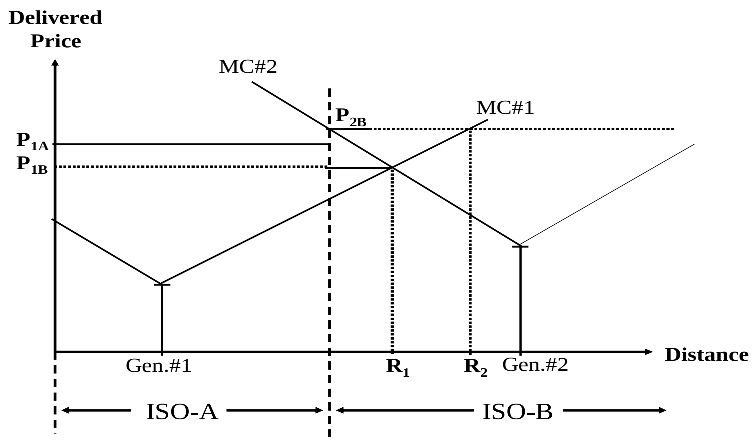


Figure 1. Spatial Competition at a Border under Uniform Delivered Pricing

If a key attribute that determines the borders for segmented markets is the existence of some physical barrier that might be reduced by an additional expenditure, then a congested transportation network can always be improved by building more facilities (e.g. roads, terminals, or transmission lines, etc.), and those price differences across the boundary provide a clear signal of the benefits to be derived by reducing that congestion. But, the market boundary needs to be specified at the point of congestion. Similarly, if for managerial and reliable operation purposes, the entire U.S. is not operated as a single power grid, then it may make sense to have separate markets in each operating jurisdiction as in the illustration of Figure 1, and to the extent that price differences exist across those "seams", they may be warranted, in part, by added internal congestion costs that might result from attempts to arrange large transfers across those boundaries. And, to the extent that those price differences exist, they provide powerful incentives to make the arrangements and investments to reduce those physical and operational procedure barriers.

But the third factor to consider in designing a market structure is to minimize the cost of arbitrage across borders, so the possible gains from further trade need to be explored in the example in Figure 1. In this case only a lack of information or impenetrable market rules would inhibit customers in Region B who are located in-between the ISOs' boundary and R1 from reselling their purchases back to customers in ISO-A at a price lower than P1A, thereby achieving potential profits up to the difference between P1A and P1B. The assignment of additional transportation costs for back-hauling from ISO-B to A would limit the extent of resale, but it is that arbitrage capability, when it exists, that normally causes discriminatory price structures to disappear. In fact, since electricity flows according to the laws of physics, not of commerce, the actual commodity would never leave ISO-A were customers in B to buy at price P1B and then resell to customers in A at some price below P1A. The arrangement would be purely financial, and the physical flow would stay in ISO-A. That's a primary reason why electricity traders have been so keen to eliminate rate "pancaking" (charging for transmission access by each jurisdiction across which a contract is negotiated, including adding charges for a reverse flow contract when in fact no reverse physical flow occurs). However in practice, even with the elimination of rate "pancaking" and in the presence of arbitrage markets at the borders, this type of perverse flow continues to be observed in many instances, and the subsequent experimental analysis is designed to explore this phenomenon, and the ability to reduce it through financial arbitrage arrangements. With a finite number of generators in these repeated markets for electricity, oligopolistic behavior is likely to persist, and the analytic tools do not exist to predict the outcome with certainty; that is why experimental methods are used to gain insights about likely outcomes.

A final conceptual point about the effect of the structure of transmission charges that are designed to recover full long-term economic costs (line losses, maintenance expenses and capital costs), particularly when those power transfers cross a number of jurisdictions: it may be most economical to require every jurisdiction to impose a uniform per-unit-distance transport charge ($/MWh/mile). Recent regulatory initiatives to encourage transfers across control areas by reducing rate "pancaking" (the assessment of separate transmission fees by each jurisdiction crossed) may have reduced the administrative cost of contracting for transfers, thereby enhancing short-run arbitrage to reduce perverse flows, but it also may result in an uneconomic long-run spatial pricing structure. As illustrated in Figure 2, if the long-term transport costs (losses, maintenance and capital costs) increase in direct proportion to the distance transported for a constant quantity of power transferred, then eliminating "pancaking" by assessing only a single "access" fee to be paid to the jurisdiction in which the buyer of the transfer is located can be uneconomic. In this illustration, suppose that the transmission access fee is in the form of a uniform per MWh payment for all transfers terminating in a particular control area. Then those uniform within-jurisdiction fees might vary greatly between jurisdictions depending upon the load density and miles of transmission line in a particular region. In the example of Fig. 2, the uniform transportation charges to be assessed in each of three ISOs for purchases of power by that particular ISO's customers, but that is generated in ISO-A, are shown as the different horizontal lines within each region. These charges are shown to vary both because of different line costs in each jurisdiction, but also because of different spatial densities of demand. In this illustration, the per MWh transmission charge is about right, on average, if purchased in ISO-A, too high if purchased in ISO-B, and too low if purchased in ISO-C. Thus this form of eliminating rate "pancaking" leads to highly inefficient pricing in two of the three jurisdictions, encouraging too little power to be purchased from A in ISO-B and too large a purchase in ISO-C. Furthermore, through this simplified method of charging for transmission, customers in C are imposing costs on whomever had to finance and pay for the construction of transmission lines in the sparsely-settled, in-between jurisdiction B. The off-setting benefit of assessing the transmission charge only in the jurisdiction where the ultimate buyer is located, as an example, is that in the short-run a buyer in C might be able to resell in B (arbitrage) if they aren't assessed the transmission charge twice.

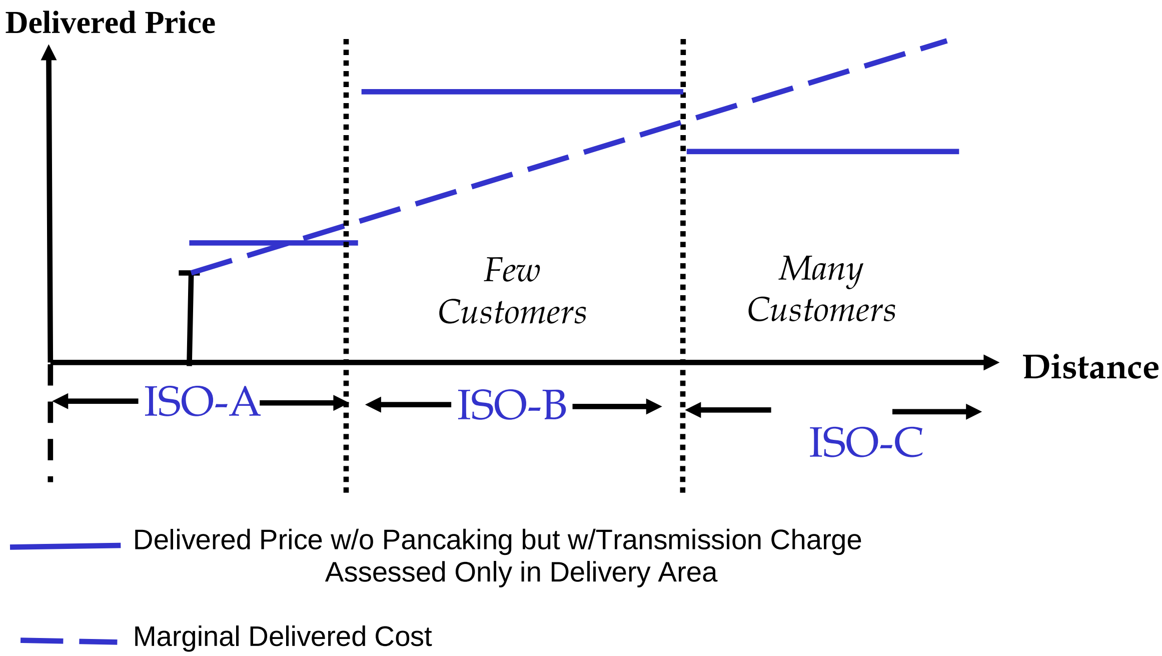


Figure 2. Comparison of Delivered Cost of Electricity Including Recovery of Transmission Capital Costs

2.2. Markets over Time

For reasons similar to those that explain why some markets accumulate many buyers and sellers over space into a single exchange with a single clearing price, in cases where physical impediments restrict commodity flows over time, buyers and sellers may find that a sequence of distinct markets may facilitate transactions and improve efficiency. A salient example is capital intensive industries that have a long gestation period between the time when a capacity addition is begun and it is available to produce. In this case the added information provided by a forward market that takes place close to the decision time when a commitment must be made to invest in physical resources may prove helpful in promoting economic efficiency.

The existence of forward markets and their positive effect upon investment can be easily explained by market participants' reluctance to take risks. However, Allaz and Vila [10] also suggest strategic reasons for the existence of forward markets and argue that firms with market power engage in forward contracts to enhance their market share in spot markets. Allaz and Vila conclude that more frequent forward markets make firms worse off and drive the spot prices down. Models that adopt the Allaz and Vila framework suggest that forward markets decrease spot prices and enhance efficiency as well (See Green [11], Ferreira [12], Lien [13], Le Coq and Orzen [14] and Newbery [15]). A crucial assumption in their analyses is that firms are underutilizing their capacity levels in the absence of forward markets or that firms can adjust their production levels without incurring additional cost. A model developed by Adilov [16] includes a firm's decision to invest in capacity levels in the analysis so that the effects of the timing of forward markets on competition and efficiency can be examined.

For this conceptual analysis, three types of players are considered in the marketplace: firms, an intermediary who can arbitrage between forward and spot markets and buyers (see Figure 3). Firms produce and sell the product in forward and spot markets. There are a small number of firms, and thus the firms have some market power. In the electricity markets, firms are represented by generators that produce and sell electricity. Buyers buy the product in spot markets for consumption purposes. It is assumed that buyers are many and/or regulated, always bidding their marginal valuation, and that there is some randomness to the aggregated buyers' valuation (demand). In the electricity markets, buyers are represented by load serving entities (LSE) buying for large numbers of smaller customers and large, primarily industrial, electricity consumers. The intermediary buys forward contracts from firms in forward markets and resells the product in the spot market. It is assumed that the intermediary earns zero profits due to free entry and exit. In a regulated electricity industry, an Independent System Operator that buys forward contracts and effectively sells the electricity at a spot market price in the spot market can represent the intermediary. Two classes of forward markets are considered, shorter-term and longer-term, based upon their occurrence with respect to the lead time required to complete investment. Shorter-term forward markets take place after investment decisions, and longer-term forward markets take place before investment decisions that commit capital (but before the spot market).

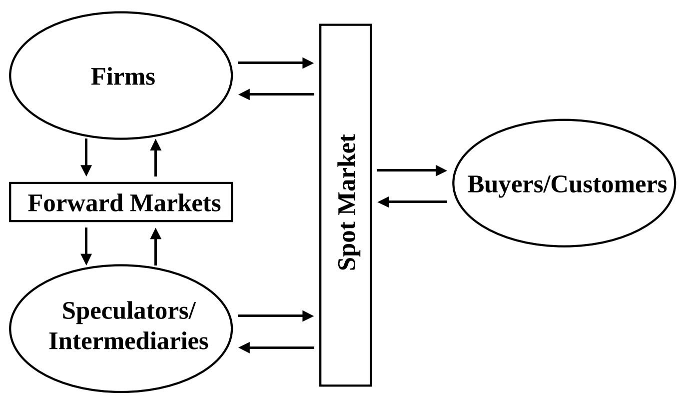


Figure 3. Market Structure and Market Participants

The game consists of three repetitive stages. Graphical representation of these stages is given in Figure 4. The object of the analysis is to determine whether the introduction of a forward market into a sequence of investment and subsequent spot market decisions affects the level of investment and spot market prices. In the first stage, the firms and the intermediary simultaneously present their longer-term forward market supply functions and longer-term forward market demand schedules, respectively. The forward market price and quantities are determined. In the second stage, after observing the forward

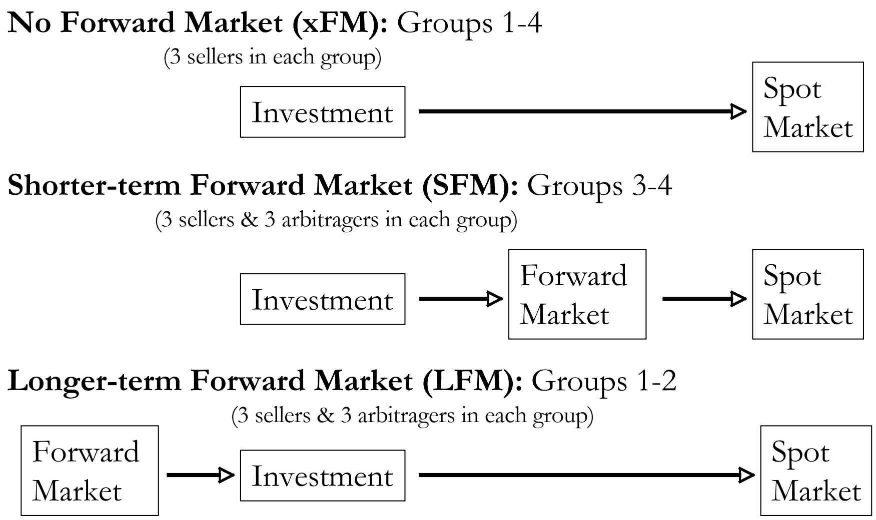


Figure 4. The Timing of Events

market price and quantities, the firms simultaneously choose their new capacity levels. In the third stage, the producers and the intermediary simultaneously choose shorter-term forward market supply functions and shorter-term forward market bid schedules. Then shorter-term forward market price and quantities are determined. In the fourth and final stage, the firms and the intermediary simultaneously choose spot market supply functions. The spot market price and firms' profits are realized. Note that the firms compete in spot and forward markets by choosing price-quantity schedules, i.e., supply functions. An equilibrium price in the forward (spot) market is determined by the intersection of forward (spot) market supply and demand. Firms' maximum quantity sales in the spot market are subject to capacity constraints that are chosen simultaneously by the firms prior to the spot market.

The theoretical analysis considers the consequences of three alternative market structures: 1) no forward markets (stages two and four, only), 2) shorter term forward markets added, only, (stages two, three and four), and long term forward markets added, only, (stages two, three and four). Technical derivations of the following summary of results are presented in Adilov [16].

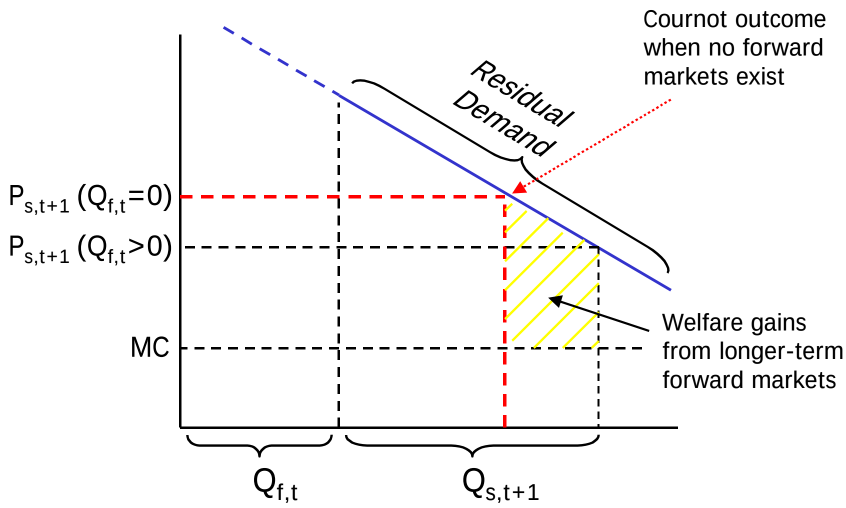


Figure 5. Residual Demand Competition

2.2.1 Implications of Longer-Term Forward Markets

Longer-term forward markets decrease spot market prices and enhance efficiency. This result is consistent with the existing literature because capacity levels are flexible in the long-run. The intuition here is similar to that underlying the two-period durable goods monopolist's problem and the Stackelberg leader game. In the durable goods monopolist's problem, higher product sales in the first period reduce the price in the second period. In the model presented here, after longer-term forward market commitments are signed, firms compete for residual demand in the spot market (see Figure 5 where the outcome under a single-shot market is also shown as the Cournot solution). Since forward market prices are fixed after that market clears, firms behave aggressively and are more inclined to cut the price in the spot market. Firms cannot keep spot prices high by restraining themselves from participating in forward markets although firms are jointly better off by not participating in longer-term forward markets. Similar to the Stackelberg leader logic, each firm is trying to increase its market share by increasing its forward market commitment levels; each tries to gobble up market share before its competitor can. Thus, higher longer-term forward market commitments reduce spot market prices by encouraging more aggressive spot market behavior, which, in turn, encourages larger capacity investments. In effect, a prisoners' dilemma-like situation arises in the forward market as each firm tries to co-opt the competition by securing a larger number of forward sales. And each firm has minimal risk in doing so when the forward market occurs before the lead time in which they can support their contractual obligations by building more generation capacity.

2.2.2 Implications of Shorter-Term Forward Markets

Similar to longer-term forward markets, shorter-term forward markets push spot markets prices down, however in this case, firms cannot respond to this anticipated price pressure by altering their capacity since investment commitments take place prior to the forward market. The overall effects of shorter-term forward markets on prices and efficiency depend on the degree of demand uncertainty. When the demand uncertainty is small or absent, the price-reducing effects of shorter-term forward markets on the spot market disappear because capacity investment serves as a prior commitment device. The firms commit to capacity levels that fully eliminate the firms' possible undercutting behavior in the spot market. The intuition behind this result is similar to that of Kreps and Scheinkman [17] in that if firms simultaneously choose quantity production levels before engaging in Bertrand competition, then the Cournot outcome prevails. In this model, firms choose capacity levels before engaging in shorter-term forward markets. Introducing shorter-term forward markets puts downward pressure on spot market prices subject to the capacity constraints. This implies that the firms' total capacity levels determine the spot market price. Therefore, the unique outcome for optimal capacity choices in the absence of demand uncertainty is the Cournot outcome.

Under uncertainty, the investment in capacity choice becomes an imperfect commitment device because the firms might choose to underutilize their capacity levels during the periods of low demand. Similar to the certainty case, shorter-term forward markets induce more aggressive behavior in the spot market, forcing the firms to decrease spot market prices. However, the possible decrease in spot market price has a lower bound that is determined by the firms' overall capacity levels. Thus, from the firms' perspective, the introduction of shorter-term forward markets imply that the firms utilize their capacity levels more often at lower spot market prices. To counteract this spot market price-reducing effects of shorter-term forward markets during excess capacity periods, the firms decrease capacity investments more in the presence of shorter-term forward markets. In sum, shorter-term forward markets under uncertainty increase capacity utilization, but decrease capacity investment. The overall effect of the two factors – higher capacity utilization and lower capacity investment – on social welfare depends on the shape of demand and the firms' marginal cost curves. With linear demand and constant marginal costs, the presence of shorter-term forward markets results in a Pareto inferior outcome reducing both consumer and producer surplus. The intuition why shorter-term forward markets might decrease social welfare can be explained by observing the spot market prices. Lower capacity levels and high spot market price volatility in the presence of shorter-term forward markets contribute to lower expected social welfare because social welfare is concave with respect to spot prices.

2.2.3 Forward versus Futures Contracts

The implications of the model are the same whether one considers forward contracts for a physical delivery of the commodity at a specified time in the future or futures contracts that are solely financial transactions with no physical commitments. The intuition behind this is the following. Consider a firm that holds one unit of a futures contract to buy, i.e., “short” futures contract. If the spot price is above the futures price, then the firm suffers a financial loss equal to the price difference from holding this futures contract. When the amount of financial loss is subtracted from the revenue received from the physical delivery of one unit of commodity in the spot market, the net revenue equals the futures price. On the other hand, if the spot price is below the futures price, then the firm has a financial gain equal to the price difference from holding one unit of a futures contract. When this financial gain is added to the revenue received from physical sales of one unit of commodity, the net revenue for that unit equals the futures price. Thus, from the firm’s perspective, holding one unit of a futures contract to buy is just like selling one unit of a forward contract. Similarly, holding one unit of a futures contact to sell, i.e., “long” futures contract, is just like buying one unit of a commodity in the forward market.

2.2.4 Policy Implications

The existing literature on strategic use of forward markets suggests that forward markets either enhance efficiency or make producers better off if they collude. While these welfare-enhancing effects of longer-term forward markets are well known, the effects of shorter-term forward markets in relation to firms’ investment decisions have not been analyzed in depth previously. The findings here imply that under some circumstances, a regulator can make both consumers and firms better off by eliminating shorter-term forward markets. In many existing electricity markets in the United States, forward capacity markets take place one month to three years prior to the spot market, whereas investment commitments are made at least three years in advance. Therefore, it seems crucial to develop longer-term forward markets in the electricity industry to maintain adequate investment levels and to sustain low spot market prices. One of the difficulties a regulator faces when introducing longer-term forward markets is the inability of some market participants to commit to specific long-term physical consumption levels. Then, a regulator might develop financial futures markets, since the analysis indicates that financial futures markets have the same effects on prices and social welfare as forward markets do.

It is realistic to assume that firms choose supply schedules in forward and spot markets, yet the findings hold for both supply function and Cournot quantity competition. This implies that the Cournot framework is a good approximation for studying analytical implications of forward markets. The multiplicity of equilibria under the supply function competition, however, yields a rich variety of outcomes. Also note that these analytical results hold both for risk-neutral and risk-averse market participants. One key assumption in this body of theoretical literature is that because of easy entry and exit for parties who arbitrage, those markets are competitive. This important assumption will be tested in the subsequent experimental analysis.

Section 3. Overview of Existing Electricity Market Structures in the U.S.

Since electricity markets were originated separately in different regions of the U.S., in most cases coincident with existing control areas, it is not surprising that details of those market structures also differ from place to place. The summary of those structures in Tables 1 and 2 show that while the broad dimensions of those designs are similar (e.g. all clear through prices that can vary by location depending upon transmission line congestion), the details frequently differ (e.g. the " grain" of spatial differentiation varies from one jurisdiction to another). Some markets clear at different prices by node; whereas others aggregate those nodal prices to zones, and others specify trading hubs. These markets are typically comprised of day-ahead and real-time energy markets, a market for energy reserves (reliability), transmission congestion hedging instruments, and generation capacity. Table 1 provides a summary of the specific types of markets operated by the seven ISO/RTOs in the United States while Table 2 provides details on the specific characteristics of these markets. The geographic spans of these entities is displayed In Fig. 6. Power procured in regions In the U.S. not covered by the seven ISO/RTOs occurs via bilateral contracts.

Table 1: Wholesale Electricity Markets (2006)

	Real-time Market	Day-ahead Market	Virtual Bidding	Ancillary Services Market	Financial Transmission Rights	Capacity Markets
ISO-NE	Y	Y	Y	Y	Y	Y
NYISO	Y	Y	Y	Y	Y	Y
PJM	Y	Y	Y	Y	Y	Y
MISO	Y	Y	Y	P	Y	
SPP	Y					
ERCOT	Y	P		Y	Y	
CAISO	Y	P	P	Y	Y	P

Key: Y = Market; P = Projected For the Future.

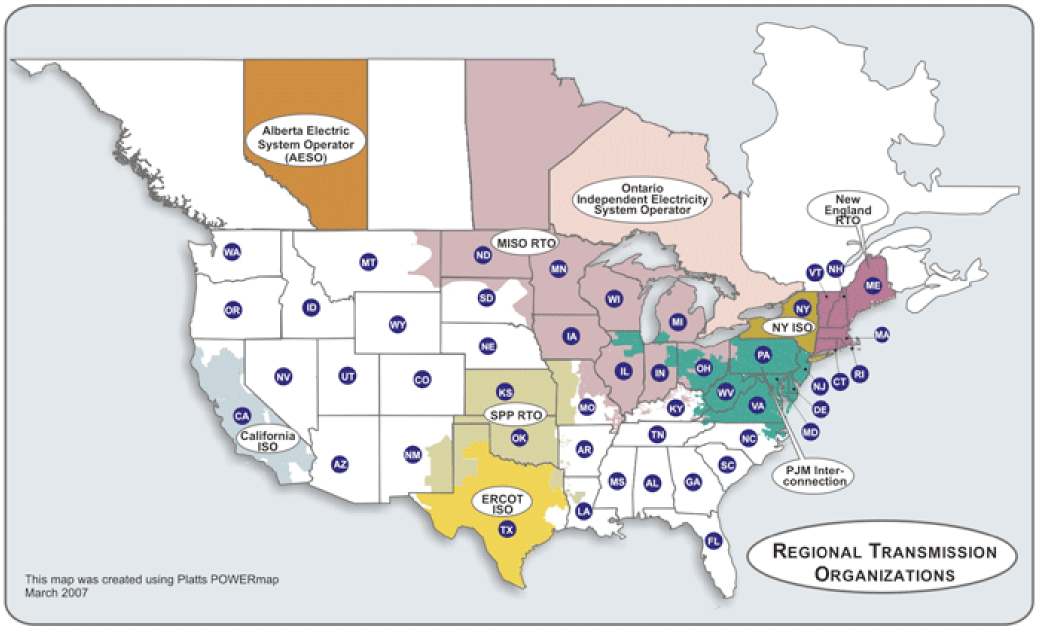
Source: FERC, *2006 State of the Market Report* [18].

Currently all seven ISOs operate real-time power markets, while only four (ISO-NE, NYISO, PJM, and MISO) have day-ahead markets in place. As part of its Market Redesign and Technology Upgrade proposal, the CAISO plans to introduce a day-ahead energy market with location based marginal prices and co-optimization of energy and ancillary services. By 2009, ERCOT has stated that it intends to implement a day-ahead market with nodal pricing and raise its energy and ancillary services offer limit to $3,000/MWh. Note that these different market grains over time can create seams over space, since a generator preferring to commit its units day-ahead may be inhibited from trading in a neighboring control area that only conducts real-time auctions. One spur to FERC's attempt to introduce a standard market design (SMD) throughout the United States was the desire to facilitate exchange across jurisdictions by having uniform market segments in all regions, but this effort failed, in part due to states' rights concerns and debates over which of the existing market structures is "best" and therefore ought to form the basis for the standard.

Table 2: ISO/RTO Market Characteristics in 2006

	ISO-NE	NYISO	PJM	MISO	SPP	ERCOT	CAISO
Bilateral Transactions	Y	Y	Y	Y	Y	Y	Y
Active Online Physical Trading	Y		Y	Y		Y	Y
Active Online Financial Trading	Y	Y	Y	Y			Y
Real-time Energy Markets	Y	Y	Y	Y	Y	Y	Y
Locational Energy Price	Y	Y	Y	Y	Y	Y	Y
Hourly Energy Price	Y	Y	Y	Y	Y	Y	Y
Congestion Price	Y	Y	Y	Y	Y	Y	Y
Losses Price	Y	Y	P	Y	Y	Y	Y
Day-ahead Energy Market	Y	Y	Y	Y		P	P
Locational Energy Price	Y	Y	Y	Y		P	P
Hourly Energy Price	Y	Y	Y	Y		P	P
Congestion Price	Y	Y	Y	Y		P	Y
Losses Price	Y	Y	P	Y			P
Ancillary Services Market	Y	Y	Y	P	C	Y	Y
Regulation Service Market	Y	Y	C	C	C	Y	Y
Operating Reserves Market	Y	Y	Y	C	C	Y	Y
Reactive Power Market	O	O	C	C	C	C	C
Black Start Market		C	C	C	C	O	C
Financial Transmission Rights	Y	Y	Y	Y		Y	Y
Capacity Market	Y	Y	Y			O	C
Regional Transmission Scheduling	Y	Y	Y	Y	Y	Y	Y
Regional Economic Dispatch	Y	Y	Y	Y	Y	Y	Y
Regional Transmission Planning	Y	Y	Y	Y	Y	Y	Y
Regional Interconnection Process	Y	Y	Y	Y	Y	Y	Y
Independent Market Monitor	Y	Y	Y	Y	Y	Y	Y
Mitigation	Y	Y	Y	Y	Y		Y

Key: Y = Market; P = Projected For the Future; C = Cost-Based; O = Other.
Source: FERC, *2006 State of the Market Report* [18].



Source: http://www.ferc.gov/market-oversight/st-mkt-ovr/som-rpt-2006.pdf [19]

Figure 6. Map of ISO/RTOs in North America

Where there are both day ahead and real time markets, virtual bidding allows market participants to place buy and sell orders that are settled financially instead of through physical delivery. Through this financial arbitrage opportunity, it is believed that day-ahead and real-time prices will converge and overall price volatility will be reduced. Currently all of the four ISOs that operate both real-time and day-ahead energy markets allow virtual bidding.

Ancillary service markets enable the procurement of services such as voltage support, black start capability, and spinning and non-spinning reserves needed to ensure the secure and reliable operation of the electricity system. These services are generally differentiated based upon response time with faster response times representing higher quality support. The CAISO, ISO-NE, NYISO, PJM and ERCOT have developed separate ancillary service products that are procured by the ISOs through an auction on behalf of market participants to meet NERC reliability standards.

Financial Transmission Rights (FTRs) are a mechanism by which market participants can hedge against transmission congestion costs. They typically represent a financial instrument that entitles the holder to receive compensation for Transmission Congestion Charges that arise when the transmission grid is congested. In PJM, the NYISO, ISO-NE and ERCOT the FTRs represent purely financial contracts while the CAISO has introduced contracts that have both a physical and a financial element.

Three of the seven ISOs have taken steps to encourage generation investment by implementing capacity markets. ISO-NE has shifted from its current installed capacity market to an annual auction to procure generating resources to meet needs three years in the future as part of its Forward Capacity Market. PJM has enacted a market based on its

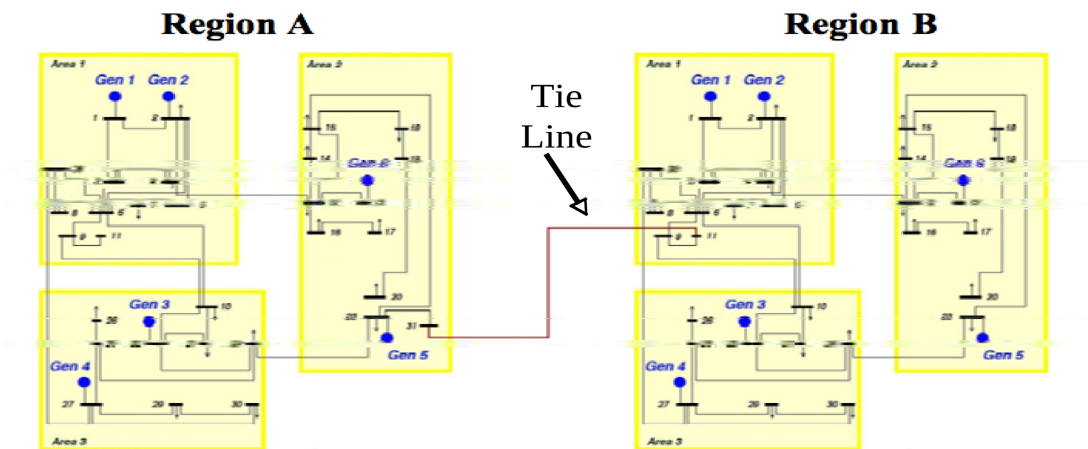
Reliability Pricing Model that includes a sloped demand curve, a forward commitment requirement for capacity, and differentiates between generators based upon their location. A key difference between a forward market and procurement auctions such as those developed in ISO-New England and PJM Is that a forward market can be either physical or financial; whereas a procurement auction is for the specific action of developing new physical resources, only. Also, these procurement auctions require mandatory participation by the buying LSEs. Both of these procurement markets have not satisfied initial expectations, in part because they have cleared at quite low prices that reflect the offers of demand-response investments, and so they have yet to demonstrate that they will encourage investment in new generation capacity when needed. The NYISO's installed capacity market (ICAP, held monthly six to one months prior to the spot energy market) has been in place the longest (ten years), and thus far needed new capacity in the region has been developed. But the NYISO's ICAP market too is a forward market with mandatory purchases by LSEs one-month prior to the spot energy market, and so conceptually it may face the shortcomings of a short-run forward market that were outlined in the theoretical analysis above.

Section 4. Experimental Tests of Efficiency Effects of Spatial Arbitrage on a Simulated Electricity Market ("Seams" Reduction)

The validity of the implications of the previous theoretical discussions of improved designs for electricity markets are tested through laboratory and classroom experiments. These experiments are conducted on simplified representations of electricity markets and their underlying physical delivery systems in order to provide insights into the validity of human behavioral assumptions that are implicit in any theoretical analysis. The alternative of conducting experiments on the actual full-scale electricity system are simply too risky (and therefore costly) to undertake without the benefits of the insights from these "bench-scale" tests.

4.1 Simulated Physical System

Consider two simulated IEEE thirty-bus electricity networks, each with six generator locations, that are connected by a single tie-line between two additional busses, numbers 31 in Region A and number 11 in Region B, as illustrated in Figure 7. Customers may be sited at any of the thirty locations in each simulated network. This combined network is calibrated (details of the calibration are provided in Appendix A) for this exercise with demand and individual generator cost characteristics so that in isolation, the demand and the generation costs are higher in Region B, as compared to Region A. In particular generators 1 and 2 in Region A are assigned the lowest production costs, and generators 5 and 6 in the high demand Region B have the highest costs. Thus, with the connection of the tie-line between Regions A and B, we would expect the predominant flow of electricity to be from A to B. However, by placing the generators with the largest cost differentials at the extreme opposite ends of each region, and assigning parameters to transmission line so that flows within each region may become congested as more energy is attempted to be pushed across regions, then that within-region congestion may place severe limits on the actual physical transfer of power between regions. This type of complication is not unusual on power networks where flows may be constrained because of thermal or voltage limits on lines within each system, even though the tie-line connecting the two systems may not be congested.



Note: (1) Demand Region A < Demand Region B

(2) Avg. Production Cost Region A < Avg. Production Cost Region B
(Gens. 1&2 Lowest Cost) (Gens. 5&6 Highest Cost)

Figure 7. System Model, Two ISO/RTO Electricity Networks

4.1.1 Generator Dispatch and Market Clearing

Using this simulated electrical network, the normal operation of the system and its economic consequences can be tested under various circumstances. Normally, a system operator in each control area (ISO) arranges the offer curves (price and quantity) of each generator in ascending order in each market period and accepts sufficient offers to ensure that demand is met at the lowest combined cost to the system, subject to all line and generator capacity constraints. Here the PowerWeb non-linear A.C. optimal power flow (OPF) algorithm is used. Unless supply equals demand in real time, the entire electric system will collapse since large-scale storage is economically infeasible. Furthermore, since these markets are repeated many times per year, it has been shown in other experimental work that for repeated, multi-unit auctions, the best market settlement method is to pay all generators required to meet demand the same uniform price that is equal to the last accepted offer (LAO) leads to the most efficient overall market outcome (see Bernard, et. al.[20] for an experimental illustration). As an example, the populist proposal of not paying any supplier more than the price they offer into the market simply causes all suppliers to estimate the market clearing price and to raise their offers to that uniformly high level; whereas with a uniform price being paid to all suppliers based upon the last accepted offer, no infra-marginal supplier has the incentive to raise their offer prices above their costs. The exception may be the generator who thinks they might be the last one selected and therefore capable of setting the market price with her offer, but in that case, she risks everything (the chance of selling nothing) as a result of her high offer.

In order to be sure the market clears, and that there is a feasible solution to the least-cost optimized power flow (OPF) solution, subject to all non-linear voltage and thermal constraints on lines and the maximum capabilities of each generator, it is assumed that additional generation capacity is always available from outside of either region, but at a very high price that is approximately double the highest cost of generation by any seller within either region.

4.1.2 Scenarios Examined

The experimental structure is designed to test the effectiveness of the tie-line and various market structures to facilitate exchange across it in improving overall economic efficiency. Three benchmarks for optimality are examined: the average prices paid by customers, consumers' surplus, and producers' profits. In the first scenario, each region's system operator minimizes its own total cost of supplying its own customers' demand. These demand schedules vary between normal and high demand periods and are subject to further random fluctuations. Generation costs are derived from the offers of each supplier in the region where they are located, and the cost minimization is subject to satisfying all reliability criteria. In the second scenario the two regions are connected with a tie line, and the welfare-maximizing solution for both regions, combined, is computed, subject to all of the constraints listed above, as if the combined regions could be operated as an integrated control area. The third set of scenarios recognize the political and physical operational realities that usually make this type of global economic dispatch impossible, in part due to limitations in existing computational technologies to permit a timely global solution. In this scenario, both a predetermined bilateral contract for firm transfers, and a market-based set of arbitrage rules are developed and tested against the theoretical, globally optimal benchmarks. In these last two cases, each ISO performs its own locally-optimized dispatch, subject to the exchanges accepted across the tie line through bilateral contracts and/or the arbitrage market. In one case, the arbitrage market is overlaid upon predetermined bilateral commitments to explore whether bi-lateral inefficiencies can be reduced by arbitrage.

4.1.3 Structure of the Arbitrage Market

Each arbitrager chooses the quantity she wishes to transfer from one region to the next, and the direction. If scheduled, she receives or pays that quantity times the actual realized price spread. Since there is no guarantee that the total quantity of energy that arbitragers wish to transfer is within the tie-line capacity, the arbitragers also specify the maximum per MW charge they are willing to pay to the ISO/RTOs in order to be accepted in the case where the tie-line flow is constrained. In the case of a congested tie-line, these bids to use the line are arranged from highest to lowest in descending order and the associated quantities to be transferred are added up, algebraically, and accepted up until the estimated congestion limit of the line is reached. In this way, all accepted transfers must pay the same uniform price equal to the lowest bid that is accepted for the quantity that reaches the line's transfer limit (thermal or voltage), and bids made to transfer in opposite directions cancel each other out, allowing more bids to be accepted.

This methodology is identical to the economist's theoretical ideal for pricing services provided from a fixed capital investment (the classic bridge-pricing problem). If the facility is not congested, then going forward the short run marginal cost of permitting another customer to use the facility is zero, since no other customer is inconvenienced. When, however, the usage is sufficient to cause congestion, each additional user inflicts an externality on every other user, and so all customers should be assessed a congestion charge in order to sort out usage priorities efficiently. Furthermore, as those congestion charges rise, they provide an excellent guide to when the construction of additional capacity is warranted. But collecting usage fees merely to cover the capital costs of a project, while essential if the project is arranged and managed by the private sector, is inefficient as a general practice in a social welfare-maximizing sense.

In the cross-border markets analyzed here, if a prearranged bilateral contract is in place, its magnitude is added to the transactions scheduled through the arbitrage process. If the capacity constraint is reached, the bilateral contract is always accepted before the arbitragers' bids. Bilateral contract holders are therefore given priority, as has been common practice in this industry, and they are not assessed congestion charges on the tie-line.

4.1.4 Calibration of Cost, Demand and Line Constraint Parameters.

Values were assigned to each of the system parameters in order to induce electricity flows from Region A to Region B in a majority of periods. Furthermore, the tie-line capacity was set at 50 Mw and the bilateral contract was set for 25 Mw of firm capacity. When combined with other line parameter settings within each region, in periods with the bilateral contract for transfers in force, the system was calibrated so that perverse flows of electricity from a higher-priced export bus in one region to a lower-priced input bus in the adjacent region would likely be observed under normal demand conditions. Thus the systems were calibrated to be able to test whether or not the addition of an arbitrage market on top of bilateral contracts would eliminate the perverse flows and improve overall economic efficiency. Calibration details are provided in Appendix A.

The potential benefits to be gained by overlaying an arbitrage market were identified in the numerical simulation process that was used to calibrate the system with one tie-line (Appendix A1). In the simulations for the case where only bilateral contracts govern transfers across the tie-lines, a contract between low-cost generators 1 and 2 in Region A to sell to customers located in the portion of Region B where high cost generators 5 and 6 are located normally increases overall economic efficiency. However, if the within region transmission lines become congested when moving power to and from the borders, perverse pricing patterns between nodes 31 and 11 can occur (see Fig. 7). If, however, the proper within ISO-congestion charges are assigned to these bilateral transfers, those contracts always proved to be uneconomical in our numerical simulations during conditions when perverse flows developed across the border. Thus an important consequence of adding an arbitrage market may be to permit the adjustment of transfer quantities in periods when those bilateral contracts turn out to be uneconomic, and the prior knowledge that spot market arbitrage is available may enhance the willingness of parties to invest in firm contracts knowing that they may be able to receive some relief through the arbitrage market should those contractual flows prove to be uneconomic subsequently.

An attempt was made to calibrate this simulation model with two connecting tie-lines, modeling AC operation of those lines and allowing for parallel and loop flows. Under independent operation within each region, however, that requires the identification of a "proxy" bus to be used by each region to identify the representative effect combined of an increased injection into or outflow from each region. In Appendix A2, the inability to identify busses that would serve as good proxies over the range of flows that might be experienced, even with marginal cost offers and an efficient combined dispatch, was not possible. The identity of the best representative proxy bus changed as the system demand, and therefore desired transfers, changed. Thus, the experimental results presented here represent arbitrage markets only over a single AC tie-line; although they should apply to multiple lines where the others are DC or are controlled by phase angle regulators, provided there is a separate market for each transfer path.

4.2 Description of Experiments to Estimate the Effects of Spatial Arbitrage

The six cases (treatments) that were examined and compared with the results from a theoretically optimal combined dispatch for both ISOs are as follows:

3 Sellers in Each Region (Groups 1 and 2):

- *Treatment 1:* No Arbitrage/No Power Flow on Tie-line
- *Treatment 2:* Arbitrage is Allowed/No Pre-Scheduled Transfer
- *Treatment 3:* No Arbitrage/25 MW Pre-Scheduled Transfer from A to B
- *Treatment 4:* Arbitrage is Allowed/25 MW Pre-Scheduled Transfer from A to B

6 Sellers in Each Region (Group 3 and 4):

- *Treatment 5:* No Arbitrage/No Power Flow on Tie-line
- *Treatment 6:* Arbitrage is Allowed/No Pre-Scheduled Transfer

Each treatment was conducted for 16 periods during which participants who were graduate students who were enrolled in a power systems seminar, and therefore were familiar with the operation and markets for electricity, were paid money proportional to their earnings from selling power in the experiments, including in some treatments, arbitraging across the border. Participants made offers to sell power through the PowerWeb 30 bus simulated electric power grid shown in Figure 7. Their earnings from generation were based upon the market clearing price in each period times the quantity of their generation accepted by the ISO because it was offered at a price equal to or less than the clearing price needed to meet demand, minus the pre-assigned production cost of each MWh sold and minus a standby cost that was incurred for every MWh offered into the market (regardless of its being accepted). This standby cost reflects the expenses necessary to keep the associated generating unit ready to run, were its offer accepted. In treatments where arbitraging was possible, participants also entered their bids to transfer power between regions in an interface designed in Excel that works interactively with Power Web via a market administrator. Each treatment was tested twice with different groups of students. Groups 1, 2 and 3 were comprised of students from October 2005 sessions while group 4's exercises were conducted in May, 2006. Since fewer participants than required to staff all positions volunteered in May, two positions were played by computer-simulated agents that submitted marginal cost offers for generation and did not participate in arbitrage activities. All arbitrage positions were filled by humans.

No regulatory restrictions were placed on the behavior of the participants as suppliers, other than they were not allowed to communicate with one another. However, there were no restrictions on the prices they offered, and they were free to offer different blocks of power at different prices or withhold completely some or all of their supplies from the market.

4.2.1 Assumed ownership roles and welfare accounting

Each of the market participant groups and their welfare accounting is described according to the following notation:

- p_t^A and p_t^B are the Locational Marginal Prices (LMP) at the tie-line ends in period *t* in region *A* and *B* respectively;
- $T_t \in [-K, K]$ is the net tie-line flow from region A to B. *K* is the tie-line capacity;
- λ_t is the net revenue generated from the tie-line capacity congestion charge in period *t*;
- $p_{i,t}^A$ and $p_{i,t}^B$ are the LMPs for generator *i* in region A and B respectively in period *t*;
- $q_{i,t}^A$ and $q_{i,t}^B$ are the dispatched quantities for generator *i* in region A and B respectively in period *t*;
- $c_{i,t}^A$ and $c_{i,t}^B$ are generator *i's* total operating costing in period *t* in region A and B respectively;
- τ_t^A and τ_t^B are the consumer prices of electricity at time *t* in regions A and B respectively;
- d_t^A and d_t^B are the consumer demand (MWs) at time *t* in regions A and B respectively;
- I_t^A and I_t^B are the imports (in MWs) in period t in regions A and B respectively that are brought in from outside regions A and B, if needed, to meet generation requirements;

- $\overline{p}$ is the cost/MW of imports.

When losses are small, $d_t^A \approx \sum_i q_{i,t}^A - T_t$ and $d_t^B \approx \sum_i q_{i,t}^B + T_t$. On average, total losses average around 2.5 percent of system load.

4.2.2 System operators (ISO/RTOs)

ISO/RTO A and B are responsible for determining the least cost dispatch of generators in their own service territory, taking demand and the tie-line flow as given. Each ISO has the potential to collect revenue from its consumers and from flows of power exiting their service territory via the tie-line. The revenue for tie-line flows are presumed to be collected from the Arbitrage Market Administrator. Offsetting these revenues are costs associated with payments to generators and payments to the Arbitrage Market Administrator for flows into the region over the tie-line.

Each generator is assumed to receive his LMP on each MW that he is called upon to generate by the ISO. The ISO is required to pay the LMP at the tie-line for any tie-line flows leading into their service territory and will be compensated at the LMP rate for any exiting flows. Consumers are assumed to pay a flat cost/MWh to the ISO regardless of location within their control area. The consumer price in each region in each period is calculated as:

$$\tau_t^A = \frac{\sum_i p_{i,t}^A q_{i,t}^A - p_t^A T_t + \overline{p} I_t^A}{d_t^A},$$

$$\tau_t^B = \frac{\sum_i p_{i,t}^B q_{i,t}^B + p_t^B T_t + \overline{p} I_t^B}{d_t^B}.$$

Under this consumer price setting scheme, each ISO runs a balanced budget in every period.

4.2.3 Arbitrage market administrator (AMA)

The AMA is responsible for clearing the tie-line arbitrage market, facilitating the transfer of power between regions, and imposing a tie-line capacity congestion charge (credit) when the requested net flow exceeds the tie-line capacity. In order to do this, it collects offers from arbitragers and in treatments 3 and 4 also takes as given the 25 MW pre-scheduled power transfer over the tie-line. In the process of clearing the market, the AMA is responsible for purchasing and selling power in the respective ISOs depending on the direction of the tie-line flow. The AMA runs a balanced budget in periods where the tie-line flow is less than the tie-line capacity and runs a positive budget of λ_t through the collection of the market-clearing bid per MW from all users (except the holders of the rights for a 25MW pre-arranged bilateral transfer) for the right to use the tie-line in periods where it is loaded to capacity.

Table 3. Average Congestion Value λ_t by Treatment

	Average Capacity Limit Congestion Charge Revenue
Treatment 1	NA
Treatment 2	\$60.92
Treatment 3	NA
Treatment 4	\$6.73
Treatment 5	NA
Treatment 6	\$20.35
Social Optimum	\$0.00

Table 3 reports the average value of λ_t for each treatment. In general, the AMA runs a small positive budget in the treatments which allow arbitrage. These revenues are derived from the bids those parties interested in arbitraging price spreads across the tie-lines pay in order to acquire capacity on the line when it is congested. These revenues could presumably be put towards maintaining the tie-line or to offset other efficiency-enhancing expenditure associated with dispatch or running the markets.

4.2.4 Consumers

In every treatment, demand is stochastic but unresponsive to real time power costs (no demand elasticity is considered in this short-run analysis). Consumer welfare in each region will therefore vary with the aggregate consumer cost of power in regions A and B, calculated as $d_t^A \tau_t^A$ and $d_t^B \tau_t^B$, respectively. Therefore, consumers in each region will be made better or worse off depending on the cost of power. As a result, consumer prices, τ_t^A and τ_t^B, are the measures of consumer welfare for each treatment.

Table 4. Average Consumer Price (τ_t^A and τ_t^B) by Treatment

	Average Consumer Price per MW		
	Region A	Region B	Combined
Treatment 1	\$74.72	\$89.04	\$82.37
Treatment 2	\$56.03	\$81.46	\$69.61
Treatment 3	\$85.61	\$99.60	\$93.06
Treatment 4	\$78.54	\$100.26	\$90.14
Treatment 5	\$46.56	\$71.88	\$60.09
Treatment 6	\$42.32	\$67.11	\$55.56
Social Optimum	\$39.89	\$52.89	\$46.87

Table 4 shows the average consumer price/MW after pooling the results across groups. It suggests that the introduction of the arbitrage market generally reduces consumer prices in both regions.

4.2.5 Generators

Aggregate generator profits in each of the respective regions are calculated as:

$$\pi_t^A = \sum_i \left[p_{i,t}^A q_{i,t}^A - c_{i,t}^A \right],$$
$$\pi_t^B = \sum_i \left[p_{i,t}^B q_{i,t}^B - c_{i,t}^B \right].$$

Table 5 reports the average total generator profits per period in each region separately, and combined, for each treatment. In most cases those profits are lower with arbitrage across the tie-line because the market is more competitive. In treatments 1 through 4 the presence of the tie-line increases the number of potential competitors in each region from three to six, and in treatments 5 versus 6 the number of potential competitors increases from six to twelve, large enough to suggest that the markets should become competitive in this case.

Table 5. Average Generator Profits (π_t^A and π_t^B) Per Period by Region and Treatment

	Average Aggregate Generator Profits per Period		
	Region A	Region B	Combined
Treatment 1	$7,043	$8,192	$15,236
Treatment 2	$4,677	$6,308	$10,985
Treatment 3	$10,331	$8,967	$19,298
Treatment 4	$8,723	$9,436	$18,159
Treatment 5	$1,767	$4,491	$6,258
Treatment 6	$2,108	$3,713	$5,821
Social Optimum	$1,018	$1,027	$2,045

4.2.6 Arbitragers

Arbitragers make money depending on the direction of their accepted transfer, the tie-line border prices, and any capacity congestion charge (or credit) levied. Aggregate arbitrager earnings can therefore be calculated as

$$\psi_t = (p_t^B - p_t^A)T_t - \lambda_t$$

in periods where no bilateral is in place and

$$\psi_t = (p_t^B - p_t^A)(T_t - 25) - \lambda_t$$

in treatments when the 25 MW bilateral contract is in effect.

Table 6 shows the average total arbitrager earnings per period by treatment. On average, arbitragers lost money in every treatment. The losses from arbitrage activities were greatest in treatments 5 and 6 when the individual ISOs were most competitive. Since individual generators were also the only participants in the arbitrage market, it is important to examine their combined operating (generating) profits with these arbitrage losses, as shown in the second column of Table 6. In all cases, the addition of the tie-line including arbitrage opportunities across the two regions leads to lower combined producer profits; although they remain above the socially optimal level.

Table 6. Average Total Arbitrage Earnings (ψ_t) per Period by Treatment plus the Sum of Average Generator (π_t^A and π_t^B) and Arbitrage Profits per Period

	Average Arbitrager Earnings per Period	Sum of Average Gen. + Arbitrage Profits
Treatment 1	NA	$15,236
Treatment 2	-$476	$10,509
Treatment 3	NA	$19,208
Treatment 4	-$419	$17,740
Treatment 5	NA	$6,258
Treatment 6	-$1,246	$4,575
Social Optimum	$0	$2,045

4.2.7 Owner of pre-scheduled bilateral transfers

To keep the analysis simple, the pre-scheduled bilateral transfer is viewed as a financial contract for a 25 MW flow from A to B over the tie-line. The owner of the pre-scheduled bilateral, like the arbitragers, makes or loses money based on the border price differences but is assumed able to avoid paying tie-line capacity congestion charge. As such, the bilateral contract owner will experience profits in treatments 3 and 4 equal to:

$$\eta_t = (p_t^B - p_t^A)25$$

Table 7 reports the average earnings that accrued to the owner of the bilateral contract in each period. Here arbitrage across the tie-line enhances the profitability of a fixed bi-lateral contract.

Table 7. Average Bilateral Earnings (η_t) per Period

	Average Bilateral Earnings per Period
Treatment 1	NA
Treatment 2	NA
Treatment 3	$39
Treatment 4	$196
Treatment 5	NA
Treatment 6	NA
Social Optimum	$0

4.2.8 Total welfare

A measure of total welfare can be calculated by adding together the participant group welfare measures described above. Formally, we add together:

1. ISO A and B Welfare: 0 (due to balanced budget consumer pricing by this not-for-profit entity).
2. Arbitrage Market Administrator: λ_t (only when arbitraging is allowed).
3. Consumer Surplus: $-d_t^A \tau_t^A - d_t^B \tau_t^B$.
4. Generator Surplus: $\pi_t^A + \pi_t^B$.
5. Arbitrager Surplus: ψ_t (only when arbitraging is allowed).
6. Bilateral Contract Holder: η_t (only when bilateral is in place).

The addition of the welfare measures shows that aggregate social welfare is maximized when the aggregate cost of production in both regions is minimized, since with a vertical demand curve, the change in consumers' surplus resulting from a price change is measured by the change in their total expenditures. So a reduction in the price of electricity merely shifts revenue (surplus) from sellers to buyers, and total combined surplus remains the same. The only gain in total welfare that can arise in these cases, then, is from reductions in cost of supply:

$$\sum_i c_{i,t}^A + \sum_i c_{i,t}^B + \bar{p}I_t^A + \bar{p}I_t^B$$

Table 8 provides the average cost of power production per period for each treatment.

Table 8. Average Cost of Power Production per Period ($\sum_i c_{i,t}^A + \sum_i c_{i,t}^B + \bar{p}I_t^A + \bar{p}I_t^B$)

	Average Cost of Power Production per Period		
	Region A	Region B	Combined
Treatment 1	\$7,298	\$11,536	\$18,834
Treatment 2	\$8,290	\$11,329	\$19,619
Treatment 3	\$8,700	\$10,439	\$19,138
Treatment 4	\$8,676	\$10,712	\$19,388
Treatment 5	\$7,192	\$11,341	\$18,533
Treatment 6	\$7,515	\$11,684	\$19,199
Social Optimum	\$7,408	\$9,762	\$17,170

Table 8 suggests that the even numbered treatments that involve the arbitrage market (2, 4 and 6) always perform worse in terms of aggregate social welfare (aggregate average generation costs are the highest), even though buyers may be better off. A partial explanation for this outcome might be that arbitrage activity across the tie-line introduces more uncertainty for generators about whether or not any particular offered block of power might be accepted in the market. Because they face a standby cost in each period for each block of power they offer into the market, regardless of whether that offer is accepted, the total MWs offered into each region by its own generators may become more sporadic with the introduction of arbitrage. If insufficient supply is offered into the market in each region to meet demand and/or is unable to reach customers because of tie-line-congestion, each region's demand is met in these exercises and the market is cleared by importing generation at a pre-specified high price from outside of either region A or B. These imports are assumed to cost $110/MW which is about twice as much as the incremental cost of the most expensive internal generator. Thus greater speculation and withholding by generators located within either region can lead to higher average production costs as a result of heavier reliance on imports to clear the markets.

Table 9 shows the average MWhs imported each period for each of the treatments. In all pair-wise comparisons, the imports from outside of regions A and B are greater when the tie-line is connecting them and the arbitrage market is in operation (Treatments 2, 4 and 6). But this increase in imports is the smallest between treatment 3 and 4 where the tie-line is connected and transfers a 25MW pre-arranged bilateral both before and after the arbitrage market is introduced.

Table 9. Average Imports by Treatments

	Average MW Imported Per Period		
	Region A	Region B	Combined
Treatment 1	0.25	1.38	1.62
Treatment 2	11.64	2.43	14.07
Treatment 3	1.78	0.00	1.79
Treatment 4	0.93	1.75	2.68
Treatment 5	0.36	0.24	0.60
Treatment 6	8.09	0.42	8.51
Social Optimum	-	-	-

4.3 Statistical Tests of Effects of Market Design and Structure on Welfare and Efficiency

In order to provide tests of the validity of inferences made above about differences in average values between treatments, regressions were run to measure the statistical significance of these effects of different market treatments on consumer, producer, and total welfare. The regression specification includes group specific demand effects (e.g. different effects for high versus low demand periods and for each different experimental group of participants to account for unusual within-group dynamics) as well as the treatment effects that are the focus of these experiments. A chi-squared test was performed of the null hypothesis that treatment effects are equal for each regression. Thus a rejection of the hypothesis implies that the outcomes between treatments are significantly different. In particular, pair-wise comparisons are examined for treatments that are identical except for the presence of a connecting tie-line between regions with market arbitrage across it. The regressions include 16 observations for each treatment that was experienced by each of the four groups of participants in these exercises, plus, as the base case, 16 hypothetical observations were computed for each group representing outcomes that would have been obtained in the economist's perfectly competitive world with efficient transfers between the regions (equivalent to the social optimum). In total there are 256 observation used for each regression.

In addition to these welfare measures, the analysis also investigates the effect of different market structures on price differences at either end of the tie-line. Under an efficiently operating arbitrage market, these border prices would be equal. Also, with these same regressions, other inferences can be made about the relative efficiency of a variety of forms of increased competition (e.g. adding a greater number of suppliers within each region, providing access to a larger number of suppliers via a tie-line, etc.).

The detailed results of these statistical tests are reported in the tables in Appendix B, and a summary of statistically supported inferences are provided below.

4.3.1 Consumer welfare: consumer cost per MWh (Table 4)

Regressions were run on the average consumer price of power in each region for each period, as well as the quantity-weighted average consumer cost of power across both regions. The regressions control for the demand state and individual group effects. A summary of the results that are tabulated in Appendix B-1 is:

- Comparisons with Social Optimum: In all market experimental treatments, consumers were worse off (paid higher prices) than they would have been at the theoretical social optimum. The exception was in region A under treatment 6, the most competitive case (six suppliers in each region plus a connecting tie-line with arbitrage between), where it could not be concluded statistically that the customers' prices were different from the social optimum. This finding suggests a background inference for all of these spatial experiments in which the number of suppliers in each of the two regions ranged from three to six (a maximum of twelve available suppliers in some treatments): attempts to and the ability of suppliers to exercise market power was substantial, and/or with only sixteen market repetitions for each treatment, substantial learning may have been ongoing and reflected in the participants' behavior.
- Impact of Introducing the Arbitrage Market: A comparison of treatments 1 with 2, 3 with 4, and 5 with 6 provides mixed results.
    - When comparing treatments 1 and 2 (three suppliers in each region), consumers were better off in both regions after introducing the arbitrage

market. This finding is statistically significant at the 99 percent level in region A and for the combined markets, and at the 96 percent significance level in region B.

  - When comparing treatments 3 with 4 (three suppliers in each region, plus a pre-arranged bilateral transfer), we find that consumers in region A are made better off at the 98 percent level by allowing arbitrage over the tie-line. The treatment effect for 3 and 4 are not statistically different from one another in region B or both regions combined, however.
  - A comparison of treatments 5 and 6 (six suppliers in each region) suggests that consumers in both regions are made better off by the arbitrage market. This result is statistically significant (at the 97 percent level) only when examining the combined consumer surplus for both regions, however.

- Impact of Introducing Bilateral Contracts, but without Allowing for Arbitraging: Comparing treatments 1 and 3 suggests that this particular 25MW firm bilateral contract between low-cost control area A and high-cost control area B tends to make consumers worse off in all regions with a high level of statistical significance (99 percent in both regions individually and combined).
- Impact of Introducing both an Arbitrage Market and the Bilateral Contract: Comparing treatments 1 and 4 suggests that consumers in both region B and for the two regions combined still are not as well off with the exchanges across regions as they would be without these transfers (significant at the 99 percent level). But by introducing the arbitrage market on top of the firm bilateral transfer, consumers in region A are made somewhat better off in the sense that it cannot be concluded that they are any worse off (cannot reject the null hypothesis that prices are the same) than they would have experienced without either the bilateral contract or the arbitrage across the border. In sum, the arbitrage market tends to offset some of the adverse effects on consumers through the price they pay that arise from the imposition of a firm bilateral contract.
- Impact of Introducing More Competition in Both Regions: Comparing treatments 1 with 5 and 2 with 6, consumers are made better off in both regions, both with and without an arbitrage market between the two regions, when the number of generators is doubled from 3 to 6 in each region. This result holds with high (99 percent) statistical significance in all cases.

4.3.2 Producer welfare (profits) (Table 5)

The regression analysis of producer welfare, summarized in Appendix B-2, measures how each of the treatments affects the combined profits of generators after controlling for systematic group-specific effects and differences in forecast demand. The key findings about producer surplus are:

- Comparison with Social Optimum: Producers earned larger combined profits than they would have in the socially optimal condition under every treatment, and this difference was statistically significant at the 99 percent level. The exception was in region A under treatment 5 (six low-cost producers in a low-demand region without a tie-line or arbitrage between regions). These results reinforce the inference made under the consumer welfare analysis that producers attempted/were able to exercise some market power in all treatments.

- Impact of Introducing the Arbitrage Market: The same pair-wise comparisons were made between treatments 1 and 2, 3 with 4, and 5 with 6, as was done for the consumer surplus, and they yield varied results depending upon other market structure factors.
    - When comparing treatments 1 and 2 (three producers in each region), producers are always made worse off when the arbitrage market is introduced (more competition means lower profits), and this loss of profitability is significant at the 99 percent level for the combined regions and in region A and at the 98 percent level in region B.
    - When comparing treatments 3 with 4 (three producers in each region plus a firm bilateral transfer over the tie-line), a statistically significant decline in profits is seen in region A as a result of adding arbitrage (99 percent level) but not in region B or overall in the two regions combined. Because profits actually increase for generators in region B as a result of adding the arbitrage market (less power is actually transferred between regions), although not by a statistically significant amount, the overall competitive effects of adding the arbitrage market are muted in this case.
    - A comparison of treatments 5 and 6 (six producers in each region) yields no statistically significant differences in profits as a result of introducing arbitrage. (How can average prices fall significantly between these two treatments, but profit declines not be significant? The answer lies in part by using "average" price as a measure, since Table 5 shows profits in region A actually rise, while falling in region B, with the implementation of the tie-line and arbitrage, while average prices fell in both regions. The other factor is the level of arbitrage profits/losses which is discussed separately).
- Impact of Introducing a Bilateral Contract without Allowing for Arbitrage: Comparing treatments 1 and 3, the introduction of bilateral contracts increase seller profits in Region A and in both regions combined (at the 99 percent level) but not in region B alone. This is consistent with economic principles that the bilateral contract transferring generation from region A to B should increase the market power in the lower-cost, excess-supply region A, and reduce market power in region B, where available capacity is tight, by introducing more, lower-priced supplies.
- Impact of Introducing an Arbitrage Market together with the Bilateral Contract: While in comparing treatments 3 and 4 above, it was shown that the introduction of the arbitrage market on top of the fixed bilateral contract reduced producers' profits in region A (it increased slightly, but insignificantly, in region B), this effect is not strong enough to overcome completely the profit enhancement derived from adding the fixed bilateral contract. This is shown by comparing

treatments 4 with 1 where profits increase by a statistically significant amount both in region A and in the two regions combined (99 percent levels), but not in region B alone, as a result of adding the tie-line with the fixed bilateral contract plus an arbitrage market. This comparison with the no tie-line case shows a similar overall upward effect on profits as does merely adding the fixed bi-lateral contract without an arbitrage market.

- Impact of Introducing More Competition in Both Regions: Comparing treatments 1 with 5 and 2 with 6 shows that producers in aggregate are made worse off by a statistically significant amount of lost profits (at the 98-99 percent levels), in every case, when the number of competitors is increased from three to six in each region. This result holds regardless of the presence of arbitrage on the tie-line.

4.3.2 Total supplier (combined generation and arbitrage) profits (Table 6)

Since all generators in these experiments were allowed to participate in the arbitrage market across the two regions, and there were no additional demand-side or pure financial arbitragers, it is important to analyze the total profits earned by these suppliers both from their physical generation and their arbitrage activities, as shown in Table 6. Of interest is the fact that the suppliers always took an aggregate loss on their arbitrage activities, and depending upon the identity of the counter-party, that loss could have flowed back to consumers as a benefit (e.g. if the ISO/RTO were on the other side of the contract, or were the owners of tie-line with regulated rates-of-return the counter-party), or the loss might have been absorbed as a profit by other production sectors if they held the rights to the arbitrage payments. What is shown in the additional regressions summarized in Table B-3 in Appendix B that combine generator and arbitrage earnings for both regions as the dependent variable is that these losses increased slightly over time. After controlling both for demand and group effects as before, and in testing for the effects of various market treatments explicitly, a time-sequence variable was added to also detect any systematic learning by the participants in subsequent market periods. In this case, since the suppliers were losing money on their arbitrage efforts, one might expect that those losses would decrease over time as the participants learned how to avoid them. In fact, the losses grew by a small amount over time, although not with a statistically significant coefficient. Therefore, one might surmise that the suppliers willingly took losses in the arbitrage market in order to support their generator profits. That is also why it is instructive to analyze the treatment effects on combined profits as presented in Table B-3.

Taking arbitrage losses into account, the combined profits for suppliers in both regions do drop by a statistically significant amount (94 percent level) when comparing treatment 5 and 6 (introduction of the arbitrage market). Here a more subtle explanation emerges of how customers' average prices could have fallen by a statistically significant amount, as shown in Table 4, while generators' profits did not, as described above. The answer may lie in the statistically significant drop of combined generator and arbitrage profits: in order to try to hold some prices (generator profits) up in Area 1 of Region B, the generators may have taken arbitrage losses in trying to keep the tie-line between regions congested. If this type of speculative behavior was the reason for these arbitrage losses, it also may explain some of the results in the total welfare analysis below.

One final observation on learning effects: it is possible that some learning may have occurred across treatments since the various treatments were always conducted in the same sequence, but there is no clear way of testing for this effect without having conducted many more trials over all possible sequences of treatments. Without those additional trials it is impossible to distinguish cross-treatment experiential learning from the treatment-specific effects, but the fact that there is no discernable trend in combined profits across the sequence of treatments suggests that this effect may have been small. Furthermore, events do unfold sequentially in the real world, and the evolution of power systems has been from separate to interconnected operating systems and from fewer to more suppliers. This is the overall sequence of treatments tested in these experiments.

4.3.3 Aggregate social welfare (Table 8)

Aggregate social welfare is maximized in these experiments when the total cost of power production used to supply both regions combined is minimized because changes in profits merely reflect identical offsetting transfers in consumers' surplus with a fixed (inelastic) quantity demanded in each period (different, but vertical demand curves for each demand period). The regressions for the aggregate cost of production in both regions combined, controlling for demand and treatment effects are summarized in Appendix B. Table B-4, and suggest:

- Comparison with Social Optimum: The cost of power production is statistically greater in every treatment, compared with the social optimum, at the 99 percent level. Although the suppliers are able to exercise some degree of market power in all treatments, they do not produce efficiently (indeed, that inefficient behavior may contribute to their ability to exercise market power by relying on external suppliers to set the market-clearing price).
- Impact of Introducing the Arbitrage Market: In every case when the arbitrage market is added (moving from treatments 1 to 2, 3 to 4, or 5 to 6), the combined cost of power production goes up, resulting in a decline in aggregate social welfare. This result is statistically significant at the 99 percent level in both of the cases where there is no pre-arranged bilateral contract (comparing treatments 1 to 2 and 5 to 6). Only where the arbitrage market is added on top of a bilateral contract (comparing treatment 3 to 4) is there no statistically significant loss in welfare (but there is also no gain) resulting from the introduction of arbitrage. This counterintuitive impact of the arbitrage market on total production costs may be related to the more sporadic output by individual suppliers in each region under those treatments. Since sellers face standby costs only in periods when they offer a block of capacity into the market, they may have an incentive to make fewer offers in periods when the effective number of competitors may increase (and their chances of being accepted decreases) as is the case with the addition of the arbitrage markets. As a result, the overall cost of production by all sellers combined is likely to be greater when there is an arbitrage market between

regions. (Note that a periodic withholding of some capacity is likely to explain why combined production costs are always significantly greater in these exercises than the socially optimal level where all capacity blocks are assumed to be available at marginal cost.)

- Impact of Allowing Bilateral Contracts but without Arbitrage: The treatment effects comparing 1 and 3, although showing higher costs (lower welfare), are not large enough to be statistically distinguishable from one another. This result suggests that the 25 MW bilateral transfer does not have a significant impact on overall welfare in these experiments.

- Impact of Introducing Both Bilateral Contracts and the Arbitrage Market: Comparing treatments 1 and 4, bilateral contracts in combination with the arbitrage market leads to a statistically significant decline in total surplus (increase in total production cost). This finding holds at the 99 percent level.

- Impact of Introducing More Competitors in Both Regions: Comparing treatments 1 and 5 and 2 and 6 shows that the combined cost of production does fall slightly as more competitors are added in each region, but by a statistically insignificant result. Once again, this result suggests that increasing competition only shifts surplus from producers to consumers without affecting the combined total surplus; although it doesn't seem to have the negative impact on total production costs that arises when competition is increased by allowing arbitrage across the tie-line with its resulting variability in offers.

4.3.4 Further inferences about market structure, competitiveness and efficiency

With only three suppliers in each region and no tie-line, as is the case under treatment 1, the exercise of substantial market power is to be anticipated and was demonstrated in these experiments. But these regressions that can identify statistically significant differences between treatments can be used to infer which methods for introducing more competitors into the market might be most effective at reducing market power and increasing efficiency. By adding a tie-line and arbitrage market between regions A and B, treatment 2 provides a measure of the effectiveness of this method of increasing the number of potential suppliers in each region from three to six. By comparison, treatment 5 simply doubles the number of suppliers within each region (by breaking up each existing supplier into two separate entities). Comparing treatments 2 and 5 for both regions combined in Tables B-1 through B-4 shows that by all measures, the average outcomes between these two treatments are significantly different at the 99 percent level.

What are those significant outcomes? As shown in Tables 4, 5, 6 and 8, compared to building a tie-line and allowing arbitrage, if it were possible to merely double the number of suppliers at each existing generator location in each region (without any increase in total supply capacity), average prices would fall from \$69.61 to \$60.09/MWh, average generator profits would fall from \$10,985 to \$6,258 per period, combined arbitrage and generation profits would decline from \$10,509 to \$6,258 and total average per period production costs would fall from \$19,619 to \$18,533 (overall welfare (total surplus) would rise). Thus where electricity is supplied over a network (in this case a specific 30 bus network shown in Figure 7), the benefits of competition are introduced more effectively by doubling the number of suppliers at each location (here by breaking existing suppliers in half) than by extending a tie-line between two of the neighboring regions and introducing an arbitrage market between them (another way of doubling the number of potential competitors).

However, as described previously, in further comparisons between treatments 5 and 6, even greater gains in competitiveness can be obtained in these exercises by adding the tie-line to treatment 5 (six generators in each region). In that case average prices fall further from \$60.09 to \$55.56 (99 percent confidence level), average generator profits fall from \$6,258 to \$5,821 per period (not statistically significant), combined generator and their arbitraging profits fall from \$6,258 to \$4,575 per period (significant at the 94 percent level), and only total welfare achieves no further significant gains or losses (total average production costs per period actually increase slightly, but not significantly).

4.4 Border Price Differences (Table 10)

In a fully efficient dispatch of generators across both regions, the prices at each end of the tie-line connecting regions would be approximately equal in demand periods when the line is not congested and losses on transfers across the tie-line are minimal. Regressions were run on the absolute value of the price differences at the border where the prices are determined by each ISO independently via their OPF. A comparison of treatment 1 and 2 in Table 10 suggests that with substantial market power in each region the arbitrage market works as expected; namely, it reduces price differences at the border. However, comparisons of treatments 3 with 4 and 5 with 6 suggests that the arbitrage market did not perform as expected in these cases. At the end of treatment 4 it was clear from the individuals who were conducting the experiments and clearing the arbitrage market that at least one participant was acting strategically to manipulate the tie-line flow, presumably to improve his earnings from generation. And with greater competition in each region (comparing treatments 5 and 6 vs. 1 and 2) Table 10 emphasizes that on-average, the arbitrage market across the tie-line does little to reduce price differences (in fact they increase), even though overall customer prices fall together with combined generation and arbitrage profits. The effect of the tie-line is to increase the competitiveness in both regions, even though it falls far short of achieving the economist’s theoretical ideal of equal prices across the arbitraged line.

Table 10. Regression of Border Price Differences

	Coef.	Std. Err.
Dependent Variable: Absolute Value of Border Price Difference		
Treatment Effects		
Treatment 1	34.43	2.49
Treatment 2	23.40	2.49
Treatment 3	15.80	2.49
Treatment 4	20.94	2.49
Treatment 5	13.01	2.49
Treatment 6	30.87	2.49
R-squared		0.695
Obs		256

In order to gain some understanding of the dynamics of the participant behavior during these experiments, an illustration in Figure 8 compares price differences across the border, with and without arbitrage, for one of the experimental groups where there were six generators in each ISO (treatments 5 and 6). Two observations emerge: 1) in many instances, the price gaps are wider with arbitrage across the tie-line, and 2) tie-line flows persisted in a perverse direction (from higher to lower priced bus, as shown by the red arrows); although there was some modest improvement over time. The length of the arrows represents the magnitude of the price-spread under treatment 6 with arbitrage, and in many cases that spread is greater than under treatment 5 without arbitrage, as represented by the spread between the colored dashes. Are these continued price gaps representative of slow learning, or do they represent persistent attempts to exercise market power? Our speculation is that both types of behavior are represented here. Moving from left to right in Figure 9, many of the price spreads begin to fall, and in fact several flows begin to reverse and follow the anticipated direction (from low to high priced area as represented by the short blue arrows). But following this more predictable and calm behavior in the normal demand periods 11 through 13, the erratic perverse behavior once again emerges in the high demand periods 15 and 16, perhaps because of speculative behavior by suppliers seeking higher profits in the periods with greatest demand.

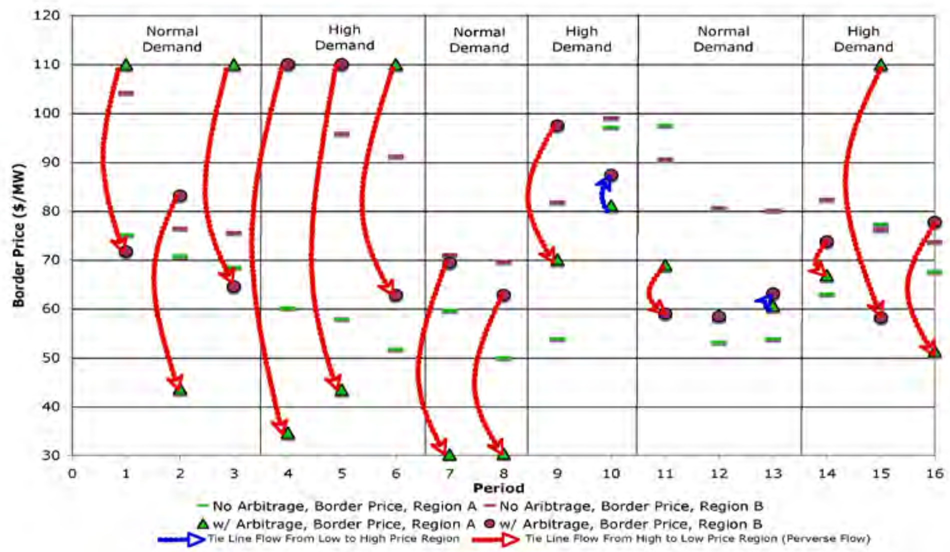


Figure 8. Border Prices and Tie Line Flows (Six Sellers in Each ISO, with and without Arbitrage)

4.5 Overall Conclusions: Spatial Arbitrage

The statistical analysis of the seams experiment suggests the following results:

- Competitive Effects: The introduction of an arbitrage market between control areas has the potential to reduce the system-wide average prices customers pay and the combined generator and arbitrage profits earned by suppliers. This effect was found to be statistically significant in most cases where generators had initial market power (treatments 1 – 4). The implication is that the introduction of real-time markets for transfers between control areas can improve market conditions by increasing the number of effective competitors. This was shown to benefit consumers by reducing the retail price of electricity in the experiments.

- Efficiency: While the incorporation of an arbitrage market for flows between ISOs was found to increase competition in many cases, it also had a tendency to increase the average production cost/MWh experienced by generators. This is at least partially due to increased uncertainty faced by generators for production needs within their service territory brought about by speculative behavior through participation in the arbitrage market. As a result, overall economic efficiency tends to be lower with the arbitrage market in place in these experiments, despite its ability to reduce market power.

- Perverse Flows: Although the economists' theoretical ideal goal of eliminating perverse flows across the tie-line, let alone totally eliminating price differences, was not realized in these exercises, as was illustrated in the conceptual discussion of spatial competition in Figure 1, that does not mean that substantial improvements in competitiveness, as reflected by lower prices and reduced producers' profits, were not achieved by introducing arbitrage. These results show that although that arbitrage was not perfect, it did move overall market outcomes in each region in the desired direction with more competitive outcomes.

- Reliability: At times, market clearing in the arbitrage market generated large power swings on the tie-line between control areas. In many cases, these swings cannot be attributed to forecast events such as changing demand within each ISO. This suggests that reliability may be compromised by enabling market participants to speculate on flows into or out of their own service territory.

- Distributional Implications: When trading is possible between control areas, these exercises show that in general, consumers gain and generators lose because of increased competition. But in some cases consumers or generators in one region may end up benefiting at the expense of parties in the neighboring region. As an example in the case where there is a prearranged bilateral contract for transfers across the tie-line (treatment 3), the introduction of arbitrage on that line leads to fall in customer prices in the lower cost region A and an increase in customer prices in the higher-cost region B, because some of the uneconomic transfers under the bilateral contract were negated by the arbitrage (predictably generator profits moved in the reverse direction increasing in region B and declining in A). Thus as illustrated in these exercises, a well-designed market that generates system wide operational and economic efficiencies may not be sufficient to ensure political feasibility.

- Additional Policy Inference: Although not the intent of these experimental designs, a statistically significant consequence for the future structural emphasis of the industry became evident. When competition takes place over space and is restricted to flow through a network (here the 30 bus PowerWeb construct), the introduction of competitive pressure is more effective as a result of splitting up existing suppliers into a larger number of competitors at their existing locations (adding additional suppliers that adds to total available supply capacity would probably have an even more pronounced effect), than does arranging potential access to the same added number suppliers by connecting a tie-line with arbitrage across two such networks. Having said that, connecting the tie line always led to even further competitive improvements.

Section 5. Experimental Tests of Forward Markets and Capacity Investment

This set of experiments was conducted to explore the effects that the existence and timing of forward markets might have on the investment in new facilities – and ultimately on spot market prices. In addition, a crucial assumption made in all theoretical analyses of the efficiency of forward markets was tested experimentally: that arbitrage markets over time are perfectly competitive and therefore that forward and spot prices are always equal. The sellers and the arbitragers in these experiments were allowed to trade freely in the forward and spot markets without price mitigation or regulatory intervention. Furthermore, arbitragers were allowed to withhold some of their forward purchases from the spot market tested here by hypothetically selling the residual of their forward purchases to some external market at a modest, pre-specified price. In this way, the arbitragers were free to try to raise spot market prices in the market tested by withholding supplies, just as could physical generators do with their capacity that had not been previously committed through the forward markets.

5.1 Experimental Structure

The three treatments illustrated in Figure 4 of the theoretical analysis about the effect of forward markets on investment and spot market prices are tested and compared through exercises using humans who are compensated in proportion to their earnings in the simulated markets. Experimental treatment A represents the benchmark, without forward markets, where the sellers adjust their capacity investment levels and then participate in the spot market. This sequence of events is repeated in each cycle. Since forward markets are not included in this treatment, there are no individuals representing the role of arbitragers. In treatment B a shorter-term forward market is introduced that takes place after the capacity investments are made but before the spot market occurs. In this treatment the intermediaries purchase the electricity in the forward market and resell it in the spot market (or to an external ISO at a pre-set low price). In treatment C there are no shorter-term forward markets but there is a longer-term forward market that takes place before capacity investments are made. This sequence of treatments used in the exercises follows a historic development of forward markets in the electricity industry.

We conducted these experiments over three days with four groups of subjects who were students in a professional engineering management program at Cornell. Each group consisted of three physical suppliers/sellers in treatment A, and of three sellers and three arbitragers in treatments B and C. The sellers in groups 1-2 played the role of arbitraging intermediaries in treatment B, and the sellers in groups 3-4 played the role of intermediaries in treatment C.

5.1.1 Sequence of decisions and markets:

Day 1: Treatment A (Groups 1-4). Capacity investments and spot markets.

Day 2: Treatment B (Groups 3-4). Capacity investments, shorter-term forward markets, and spot markets.

Day 3: Treatment C (Groups 1-2). Longer-term forward markets, capacity investments, and spot markets.

5.1.2 Calibration and participants

Capacity denotes the maximum amount of electricity that the seller can produce each period. Each seller began with 32 units of capacity that corresponds to the average socially optimal capacity level for this exercise design. The sellers could increase their plant capacity during the capacity sub-period by purchasing extra capacity, but they could decrease their capacity holdings only by allowing them to depreciate. Each unit of additional capacity cost $30, and each seller's total capacity depreciated at a rate of 20 percent per period. Production costs were a constant $5 per unit for all producers in the market. Production and capacity costs were calibrated to match the relative sizes of these costs in real electricity markets. The International Energy Agency estimates that total fuel costs represent from 10% to 80% of total plant costs over the lifetime of a power plant. Experimental per unit production costs represented from 32% to 46% of total costs [21].

The demand schedule used was stochastic with modest price responsiveness. The market participants were provided the forecasted demand schedule before the physical sellers made their decisions on capacity investments or any forward markets were conducted. While the forecast demand schedule remained constant throughout, the actual demand schedule that was realized in the spot market varied randomly by up to 10 percent of the forecast demand, drawn from a uniform distribution. Figure 12 illustrates the expected, maximum and minimum demand schedules that were applied in the spot market. In most real electricity markets demand is increasing steadily over time and firms invest in additional capacity in order to meet the growing demand; in these exercises with a constant but random demand the participants have an incentive to invest in capacity due to the steady depreciation of existing generation.

The experiments were conducted with advanced undergraduate and masters of engineering management students. Each treatment consisted of 5 learning periods and 15-27 trial periods. The students were not told in advance in which period each experimental treatment would end (nor was there a predictable time at which each experiment was ended) in order to reduce end game effects. Students received extra grade credit proportional to their experimental earnings in addition to a $10 show up fee per experiment.

5.2 Summary of Experimental Results

5.2.1 Prices/Rational expectations hypothesis

The theoretical analyses of the effects of forward markets are based on a rational expectations assumption (implying no residual arbitrage opportunities so that effectively, these arbitrage markets are perfectly competitive and their prices are the same). Thus, we analyze whether the forward market prices predict spot market prices rationally and whether three intermediaries can consistently collect a forward market premium. The time trends of the forward and spot prices are shown in Figures 10 and 11. A visual inspection of the prices does not suggest any systematic differences between forward and spot prices in either treatment. However price volatility seems to be lower in the presence of longer-term forward markets (treatment C). This observation is consistent with the theoretical predictions that a shorter-term forward market exhibits larger variations due to increased speculation in the markets. On the other hand, longer-term forward markets might be expected to help coordinate investment decisions and result in "self-fulfilling" price expectations.

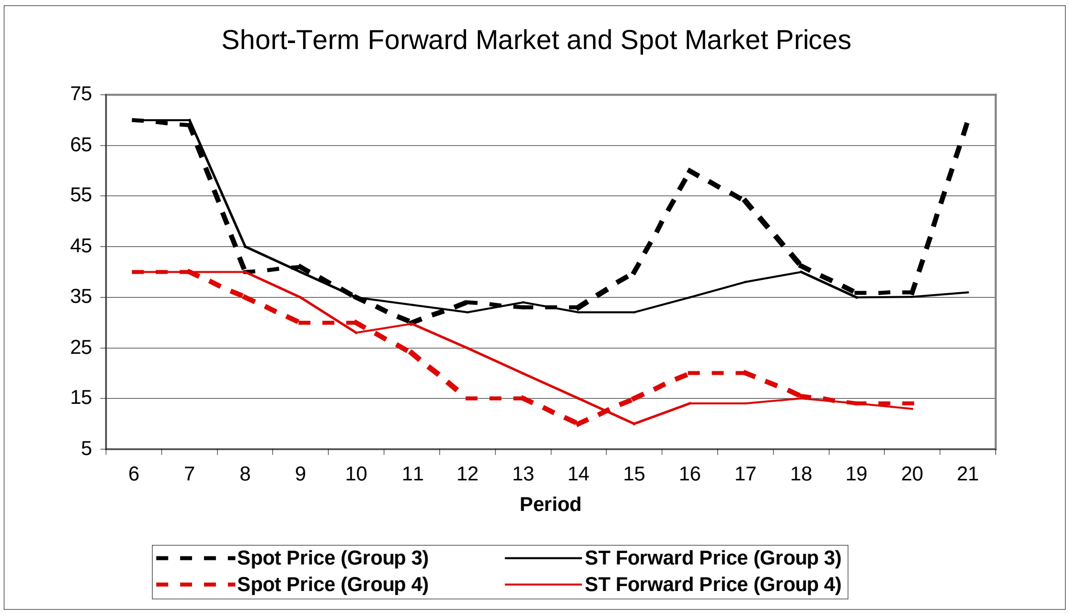


Figure 10. Rational Expectations, Forward Price = Spot Price?

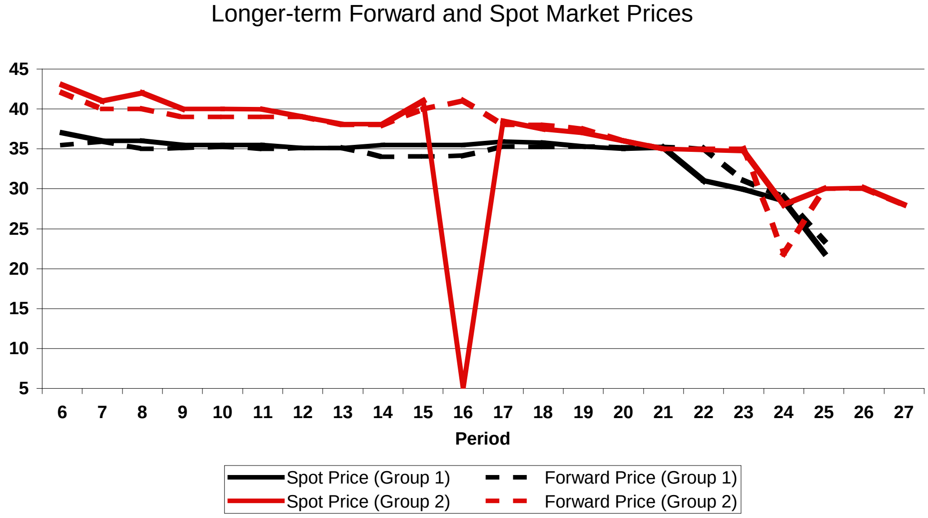


Figure 11. Rational Expectations, Forward Price = Spot Price?

The rational expectations hypothesis about the behavior in forward markets is tested by considering a model of log-lagged differences in prices, controlling for different treatment effects. The price differences are considered in order to diminish the effects of statistical interpretation due to time series correlation effects, and the logs are used to study relative (forward market rates) instead of absolute price deviations. The following model specification was used in the econometric regression:

$$(\ln S_{t+1} - \ln S_t) = \alpha_i + \beta_i*(\ln F_{t+1} - \ln S_t) + \varepsilon_{it}$$

where S_t denotes the spot price at time t, F_t denotes the forward price at time t, and i denotes the particular experimental treatment. The regression results are summarized in Table 11. The intercept and slope coefficients seem reasonable and support the unbiased forward rate hypothesis (UFRH) except for the slope coefficient under shorter-term forward markets. However, this particular slope coefficient has a high standard error, so we cannot reject UFRH either in treatment B or in treatment C. Under the UFRH, the constant term should be zero and the slope coefficient should be one. Specifically, we test and cannot reject the joint hypothesis that $\alpha_i = 0$ and $\beta_i = 1$ for each i. Testing hypotheses on these parameter values separately yields similar conclusions.

Although we did not detect any statistically significant forward rates, some arbitragers withheld capacity in the spot market (behavior inconsistent with a perfectly competitive market), but they were not able to improve their profitability or significantly affect market outcomes. Thus, the arbitragers could not exercise market power even though we had only three arbitragers in the market. But it is possible that we also could not reject the hypothesis that forward markets are unbiased (particularly in the case of shorter-term forward markets) simply due to the limited number of observations that may have contributed to the large standard error of the estimated slope coefficient in this case.

Table 11. Regressions Comparing Forward and Spot Prices Using Log Price Differences

```
                                                  Number of obs =      69
                                                  F(4, 65)      =   11.86
                                                  Prob > F      =  0.0000
                                                  R-squared     =  0.4220
                                                  Adj R-squared =  0.3864
                                                  Root MSE      =  .30793

------------------------------------------------------
      d_lnSt |      Coef.   Std. Err.      t    P>|t|
-------------+----------------------------------------
     alpha_B |   .0020455   .0645435     0.03   0.975
     alpha_C |  -.0431121   .0487739    -0.88   0.380
      beta_B |   .4736861   .3707524     1.28   0.206
      beta_C |   .9883338   .1470411     6.72   0.000
------------------------------------------------------

(1)  alpha_B = 0
(2)  beta_B  = 1
     F(2, 65) = 1.31
     Prob > F = 0.2766

(1)  alpha_C = 0
(2)  beta_C  = 1
     F(2, 65) = 0.40
     Prob > F = 0.6724
```

5.2.2 Investment and capacity levels

Under all treatments, capacity levels were volatile and exhibited cycles that are frequently observed in capital-intensive industries (see Figure 12). The theory predicts that investment levels should be the highest under longer-term forward markets and the lowest under shorter-term forward markets because the sellers would use capacity levels as commitments to try to raise spot market prices. Consistent with the theory, investment levels are the highest with longer-term forward contracts (see Table 12). Capacity levels under treatment B with short term forward markets are not much different from the capacity investment levels under treatment A. In general, however, the sellers had trouble using capacity levels as commitment devices and over-invested in capacity. Over-investment in capacity occurred mostly in groups 2 and 4. This systematic behavior of over-investment in capacity might be improved in future exercises with increased training of experimental subjects.

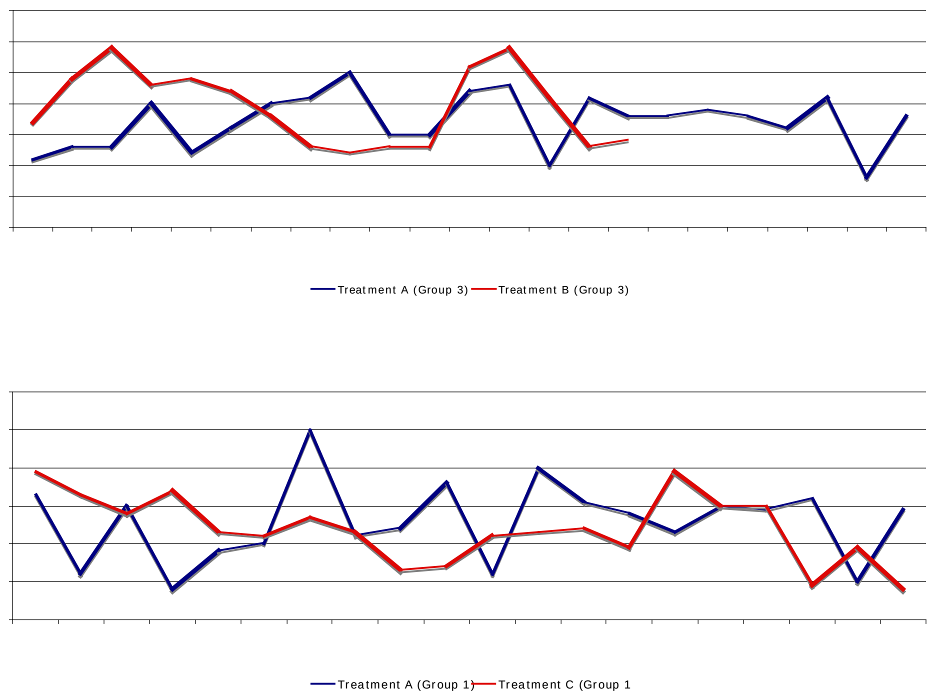


Figure 12. Investment Cycles

Table 12. Comparisons of Average Capacity Investment per Period

	Group 1	Group 2	Group 3	Group 4	Combined
Treatment A	25.850	27.714	17.188	33.222	25.968
Treatment B	---	---	19.438	32.400	25.919
Treatment C	23.950	51.318	---	---	37.634

5.2.3 Speculative behavior in the spot market

Possible speculative behavior by sellers and arbitragers in the spot market is analyzed by constructing an aggregate supply function for each treatment. The aggregate supply function is derived by associating the total quantity offered into the spot market by all participants with each respective offer price, and then fitting those data statistically. The following linear relationship was estimated where i denotes the treatment subscript:

$$Quantity\_Offer_t = a_i + b_i*Price\_Offer_t + e_{it}$$

Higher positive values of b_i would indicate that the spot market is more competitive (a flatter offer function), and b_i values close to zero would indicate that the spot market is the least competitive. In this case the market participants' offers could be characterized as representing the sharply upward-sloping portion of a “hockey-stick”. The estimated aggregate supply curves are shown for each treatment in Figure 13. This estimated supply curve is the most competitive (the flattest) in the presence of longer-term forward markets. The aggregate supply curve is the least competitive (the steepest) in the presence of shorter-term forward markets. We test if these slope coefficients are statistically significantly different between treatments, as summarized in Table 13. The slope in treatment B is significantly different than from treatments A and C, but we cannot reject the hypothesis that slopes A and C are the same. Also note that all coefficients are statistically different from zero except the slope coefficient under treatment B. These graphs and statistical tests are consistent with the theoretical predictions that shorter-term forward markets increase spot market speculation. By comparison, longer-term forward markets increase competition in the spot market.

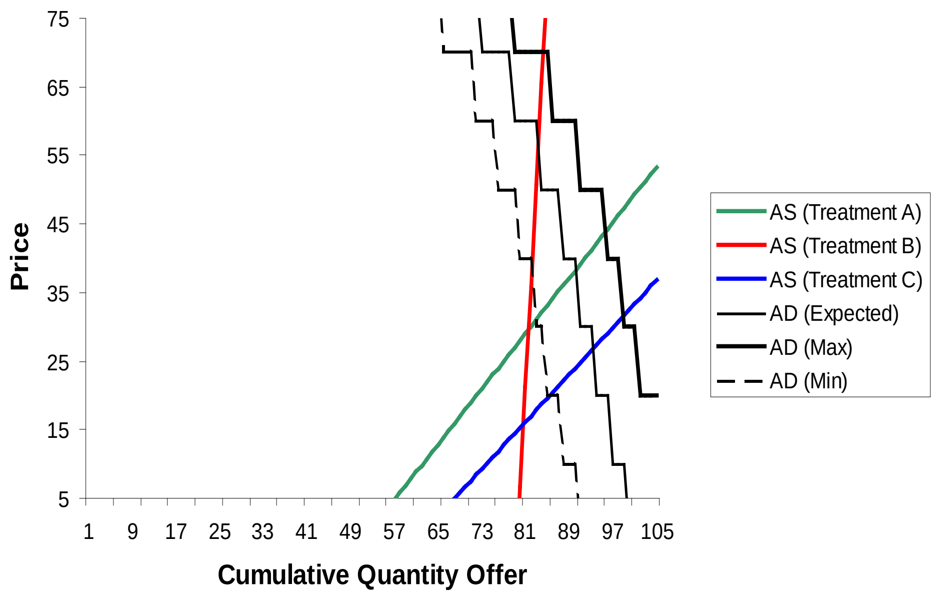


Figure 13. Expected Aggregate Supply and Demand

Table 13. Aggregate Supply Curve Regressions

Number of obs	= 935
F(6, 929)	= 418.17
Prob > F	= 0.0000
R-squared	= 0.7298
Adj R-squared	= 0.7280
Root MSE	= 56.098

Quantity_Offer	Coef.	Std. Err.	t	P>|t|
b_A	.9889442	.1287428	7.68	0.000
b_B	.0678314	.0425029	1.60	0.111
b_C	1.16312	.1990056	5.84	0.000
a_A	52.28698	4.574799	11.43	0.000
a_B	79.5554	4.266299	18.65	0.000
a_C	62.22907	8.550414	7.28	0.000

```
b_A = b_B
F(1,929) = 46.16
Prob > F = 0.0000

b_B = b_C
F(1,929) = 28.97
Prob > F = 0.0000

b_A = b_C
F(1,929) = 0.54
Prob > F = 0.4626
```

5.2.4 Social surplus

Social surplus (the sum of consumers and suppliers surplus) levels vary widely by groups (see Table 14). This volatility could be the result of the sellers' inability to manage their capacity levels properly. The theoretical analysis implies that longer-term forward markets improve social welfare whereas the implications of shorter-term forward markets are ambiguous. Social surplus in groups with reasonable capacity investment levels (groups 1 and 3) are consistent with the theoretical predictions: social surplus under treatment A is lower than the social surplus under treatment C, i.e., longer-term forward markets increase social surplus, and social surplus under treatment A is higher than the social surplus under treatment B (i.e., shorter-term forward markets decrease social welfare), but a large number of additional experimental trials would be required in order to make statistically significant inferences about overall welfare levels on these group-specific observations.

Table 14. Social Surplus as Percent of Socially Optimal Level

	Group 1	Group 2	Group 3	Group 4	Combined
Treatment A	0.948	0.948	0.964	0.922	0.945
Treatment B	---	---	0.951	0.936	0.943
Treatment C	0.960	0.823	---	---	0.891

5.3 Overall Conclusions about Forward Markets and Investment

Both theoretical and experimental analyses suggest that it is important to time forward markets so they can aid with the investment decision. In these examples, physical suppliers use the forward markets to guide their investments, and it is shown how important it is to have those markets placed prior to the investment decision if additional capacity is desired. Forward markets that occur after the investment commitment must be made, like with most installed capacity markets for electricity generating capacity, have little effect on the level of the investment. Furthermore, spot market prices are generally lower with properly-placed forward markets because of the increased capacity that is available. Since these experimental results were obtained with only three physical suppliers and three intermediaries acting in an arbitrage capacity, it is interesting how competitive the spot market became with the well-placed forward market and how the intermediaries played the arbitrage role in a competitive manner.

<u>Effect of Forward Market and Its Placement on Investment</u>: The experiments are designed to test the theoretical results that suggest that forward markets conducted prior to the lead time needed to begin construction on new facilities will result in greater investment and more competitive behavior in the spot markets, but if those forward markets are only conducted after investment decisions have been made, like existing installed capacity (ICAP) markets in the electric industry, they may have little positive effect. The experiments confirm most of these theoretical findings. The investment levels were higher in the presence of forward markets with a longer lead time, but forward markets with shorter lead times did not significantly affect the investment levels. In addition, the sellers increased their profits under longer-term (lead time) forward contracts because longer-term forward contracts reduce the uncertainty associated with investment decisions. Shorter-term forward markets, however, decreased seller profits. This is consistent with intuition that shorter-term forward markets do not reduce investment uncertainty but might introduce strategic response from the sellers. Under all treatments, the investment levels exhibited investment cycles that are frequently observed in capital-intensive industries.

<u>Effect of Forward Market and Its Placement on Spot Markets:</u> Forward markets with shorter lead times increased spot market speculation yielding steeper supply curves in the spot market. On the other hand, longer-term forward markets increased spot market competition by shifting the supply curve outward. This result is consistent with the theoretical predictions that longer-term forward markets reduce market power. Spot market prices are less volatile in the presence of longer-term forward markets.

Efficiency of Forward Markets: The theoretical results regarding the effects of forward markets are based on rational expectations hypothesis that the forward and spot prices are equal. The tests using the experimental data could not reject this unbiased forward rate hypothesis. Some arbitragers withheld some of their capacity in the spot markets but were not able to improve their profitability or significantly affect market outcomes. Thus, the arbitragers could not exercise market power even though we had only three arbitragers in the market. Nevertheless, substantial additional testing is warranted on the structure and effects of forward markets for the wholesale exchange of electricity, since in these exercises few arbitragers ever made positive profits. In longer trials, therefore, some intermediaries would be expected to discontinue their arbitrage activities, and that further potential impact upon market efficiency needs to be tested.

Section 6. Concluding Observations

Throughout this analysis and the subsequent experimental simulations, a small number of competitive suppliers and/or arbitragers were utilized specifically in the anticipation that some attempts to exercise market power might be expected and observed. In this way the effects of increased market connections both over space and time might be able to highlight the competitive effects of the proposed market improvements. Furthermore, no regulatory restrictions were imposed on participant behavior in the experimental analysis, other than they were not allowed to communicate with each other. But, participants were allowed to make as large price offers as they wanted to and/or to withhold capacity blocks from the market. In the case of the forward markets, we explicitly allowed for non-competitive behavior by the intermediaries by permitting them to sell all or some of their forward purchases into an external market in order to test the perfect arbitrage assumption that is prevalent in the financial literature. These calibrations were successful since some speculative behavior was exhibited throughout the various experimental trials.

What is evident from this analysis and experimental results is that establishing formal markets across the seams between separate electricity controls areas (ISOs/RTOs) and putting forward markets in place for capacity that are held prior to the latest advance date necessary to begin construction on new facilities (if demand is to be met in real time) should both have substantial beneficial effects on the overall competitiveness of electricity markets. Furthermore the results suggest that these markets should be open to both financial arbitragers as well as physical suppliers, not just to physical suppliers/ demand reduction as in the cases of ISO-NE's and PJM's forward procurement markets, and some details of preferred market structures are provided in the body of the report.

As with most advances for the market design of this unique commodity (electricity) what appear to be perverse outcomes according to the norms of traditional economic theory may in fact prove to be substantial improvements. Thus both a conceptual analysis and experimental support are provided of the likely persistence of periodic "perverse" flows from high- to low-priced regions following the introduction of competition and arbitrage across the boundaries of two separate control areas, but the competitiveness of those markets, as reflected by the average prices paid by buyers, also improved in many instances despite those perverse flows. To be sure, further efficiencies might be achieved in theory by completely eliminating those perverse flows through even greater arbitrage, but the issue not addressed here is how much additional transmission capacity and operational expense might be required, and is that investment cost-effective?

An unintended but statistically valid observation that emerged from the experimental analyses on seams reduction was the ability to compare two different methods for enhancing the competitiveness of electricity markets. The first method doubles the number of suppliers by breaking up the existing generators within each control area into twice the number at the same locations (with total capacity remaining the same). The second method doubles the number of potential suppliers by connecting a tie-line between two ISOs, each with the same number of generators and capacity, and includes an arbitrage market available only to the physical generators. To be fair, the tie-line approach involved some impediments to its competitiveness like potential within-ISO line congestion and different cost structures of generation across the ISOs. Nevertheless, without examining the cost-effectiveness of the investments (political or physical) needed to implement either policy, the larger number of smaller suppliers in each ISO was more effective than connecting the tie-line in enhancing competitiveness (reducing average prices to buyers and per seller profits). Having said that, after doubling the number of physical suppliers in each ISO, connecting a tie-line with a full arbitrage market added further to the competitiveness of the markets, but not by as great an increment.

Numerous detailed inferences about market structure and the potential benefits of enhancing arbitrage across both spatial and inter-temporal boundaries are laid out in Section 4.5 for markets across control areas (seams), including the demonstration of an efficient arbitrage market, and in Section 5.3 for forward markets. An underlying principle for efficient market structure that is common to both cases is that it is beneficial to have those markets span all key decision-making by operators and investors. Thus, although electricity only flows in real time and so it is essential to have a well-designed spot market in operation, since the planning for the construction of new large generation can take at least three to four years, it is helpful to have a forward market in place at least that far in advance (note: since demand response investments can be arranged in a shorter time-frame, they require only a one to two year forward market). Furthermore, these markets should be open to financial arbitrage, both to convey the information and insights held by a larger, more diverse set of entities who are willing to put their money where their mouths are, and also to provide additional liquidity to physical suppliers who make their arrangements through bi-lateral contracts with physical buyers, but who may want to hedge their bets and/or be able to alter some of their commitments through subsequent market activity.

This last insight is illuminated by both the system calibration effort for the spatial market across two control areas, and the subsequent experimental results when there was a pre-arranged bilateral contract in place across the tie-line. Through the calibration process, we were unable to establish a bilateral exchange between low- cost generators on the far side of region A (see Figure 7) to buyers at busses with much higher prices on the far side of region B that generated perverse flows across the uncongested tie-line between these two regions that would have been economic had the participants in the bilateral had to pay all of the within-region congestion charges that contract created. What the subsequent experimental trials demonstrated is that adding an arbitrage market on that tie-line undid some of those uneconomic bilateral transfers; prices fell in region A by a statistically significant amount; although they increased slightly in the high-cost region B (but not statistically significantly). Two counteracting policy inferences can be drawn from these results. First, one advantage of instituting a spatial arbitrage market across a seam is that it may encourage the signing of longer term bilateral contracts (that could encourage investment in expanded transmission capacity) if the parties believe they can undo partially through arbitrage any subsequent congestion payments that had not been anticipated at the contract's signing. The opposite implication is since in this case, without paying for congestion, buyers in region A lost and buyers in region B gained through the bilateral contract; implementing the arbitrage market partially reversed some of those welfare exchanges. Thus, depending upon the starting point, implementing a pro-competitive arbitrage market can have substantial distributional consequences and therefore face severe political opposition.

Finally, although the introduction of arbitrage markets over both space and through forward markets was nearly always shown to improve the competitiveness of spot markets (reduce prices and profits) in the experimental treatments, still those participating in the arbitrage always lost money on average. This result was partially understandable in the case of the arbitrage across the tie-line between regions A and B, since the generators were the arbitragers and they may have been pursuing combined profits. But in the forward markets, the arbitragers were independent agents who could have cut their losses simply by withdrawing from the market. Nevertheless the arbitrage between the forward and spot markets was shown to be efficient, statistically, in the sense that those prices could not be shown to be different; whereas, significant price spreads always remained across the tie-line, and in many instances those spatial price differences were perverse, flowing from high to low priced busses. Throughout these experiments that demonstrated how arbitrage markets might enhance the competitiveness of electricity supply, those sellers were shown to continue to engage in speculative behavior throughout.

These results suggest that although wholesale electricity markets can be made more competitive through the introduction of arbitrage markets over space and time, they may never reach the economist's ideal of perfect competition so long as transportation costs matter and/or substantial lead times are required between the initiation of investment and the completion of that new supply. But what is also suggested is that the addition of more, smaller suppliers near the buyers can have a significant pro-competitive effect.

Appendix A.1 Calibration of the Spatial Experiments (one tie-line)

Demand and generator costs in each region were calibrated in such a way that the optimal transfer between region A and B varied considerably depending on whether it was a normal or high demand period. For the experiment, the average demand varied between 180 and 200 MW in region A and from 200 to 240 MW in region B depending on whether the period was a normal or high demand period respectively. Actual demand varied by an additional 10 percent around the average forecast demand for each type of demand condition.

Table A1 reports the generator cost assignments used in treatments 1 through 4. In these treatments, each participant was assigned two generators in either region A or region B. In treatments 5 and 6, the capacity of each generator was divided into two 30 MW blocks with the same variable and standby costs. Participants in treatments 5 and 6 were assigned two 30 MW blocks in either region A or region B, but the blocks were not necessarily located at the same node location. As a result, we would expect more intense competition in treatments 5 and 6.

Table A1: Generator Costs Structure Used in Treatments 1 through 4

	Generator	Variable	Standby	Fixed Cost/Period
Region A				
Generator 1	60	$15.00	$5.00	$45.00
Generator 2	60	$25.00	$5.00	$45.00
Generator 3	60	$35.00	$5.00	$45.00
Generator 4	60	$35.00	$5.00	$45.00
Generator 5	60	$45.00	$5.00	$45.00
Generator 6	60	$55.00	$5.00	$45.00
Region B				
Generator 1	60	$35.00	$5.00	$120.00
Generator 2	60	$35.00	$5.00	$120.00
Generator 3	60	$35.00	$5.00	$120.00
Generator 4	60	$55.00	$5.00	$120.00
Generator 5	60	$55.00	$5.00	$120.00
Generator 6	60	$55.00	$5.00	$120.00

Table A2 shows the resulting power transfers and border prices obtained under marginal cost offers, and under alternative dispatch and transfer assumptions. The results obtained from the combined OPF represent a global social optimum where total generator costs in both regions are minimized. The separate OPF outcomes, with and without a transfer, are representative of what would occur in a highly competitive environment when each region is independently optimized on generation cost. Notice that the generation capacities and costs and the within-region line constraints were set in an attempt to generate perverse flows across the tie-line (flow from higher to a lower-priced bus) under normal demand conditions in treatments where the 25 MW bilateral transfer is in effect.

In fact the system calibration used for all experimental treatments was able to create this perverse flow, as summarized in Table A2.

Table A2: Tie-Line Flows and Border Prices Under Alternative Dispatch Conditions

	Tie Line Flow From Region A to B (MW)	Border Price	
		Region A	Region B
Average Normal Demand Period			
Combined OPF	4.14	$50.00	$50.00
Separate OPFs w. 25 MW Transfer	25.00	$50.00	$41.07
Separate OPF w. no Transfer	0.00	$50.00	$52.97
Average High Demand Period			
Combined OPF	28.58	$52.02	$52.03
Separate OPFs w. 25 MW Transfer	25.00	$51.63	$54.57
Separate OPF w. no Transfer	0.00	$50.00	$58.09

Note: Assumes all generators make marginal cost offers.

An interesting numerical insight that derived from this calibration process, however, was that in all dispatches where the bilateral transfer was mandated and perverse flows across the tie-line were observed, had the contracting parties on the bilateral contract also been required to pay the within-region congestion charges that resulted from that dispatch, the bilateral transfer would have proved to have been uneconomic.

Appendix A.2 Multiple Tie-lines

Since in most cases, multiple AC tie-lines connect neighboring control areas, attempts to increase the flow of power across the boundaries by increasing the generation in one ISO and lowering generation in the other needs to be simulated to ensure that operating constraints (thermal, voltage or dynamic) on any individual tie-line are not exceeded. Unlike the simulated connection through a single tie-line between bus 31 in region A and bus 11 in region B that was used for the seams experiments described in this report (see Figure 7) where the injection and rejection busses are clear in each region, with multiple tie lines, depending on the spatial distribution of demand and generator dispatch in each of the ISOs, attempts to transfer power across the border may be divided in different ways between the two lines under different patterns of supply and demand.

Were these DC tie-lines, or if they had phase angle regulators associated with each AC tie-line, then the flow on each line could be controlled at the order of the system operators and a separate arbitrage market created for each line. In this case, the outcomes might be expected to be similar to those from these exercises with a single tie-line. With multiple, parallel AC lines however, uncertainty over which line the power might flow requires that the arbitrage market for transfers consider the net flow over the two (or more) parallel. But if a single price spread is to be allocated to the transfer, which prices should it be? One could try some average price (weighted sum across all busses) in each ISO, or as in actual practice in most cases, the system operators in each region identifies a bus in the neighboring control area whose deviation in power flow represents the magnitude of the change in power flow across the interface. Thus, each control area represents its neighbor as a single “proxy” bus for purposes of modeling the transfers across the boundary. The transfer across the interface (seam), then, is priced at the price difference between the two ISOs’ proxy busses.

We attempted to identify a proxy bus for each region using our simplified thirty-bus, six generator representations for each region, as shown in Figure A1, with a second AC tie-line connecting bus 32 in region A (near the terminal of the first tie-line at bus 32) and bus 31 in Region B (again, near the terminal bus 11 of the first tie-line). Because of the simplified network, we searched for generator busses whose deviation in power flow (increments and decrements) corresponded to increases and decreases in net transfers between the two regions, using the actual variations in demand patterns that were imposed on the experimental analyses with a single tie-line, but assuming that all generator offers were efficient at their marginal cost of production. Furthermore, to ensure efficient transfers across the interface, both regions were dispatched simultaneously with a full AC least-cost, optimal power flow (OPF) algorithm. But, even under these simplified, idealized dispatch conditions, no single bus in either control area emerged as the incremental bus for the full range of transfers over these two tie-lines with a combined capacity of 100 MW.

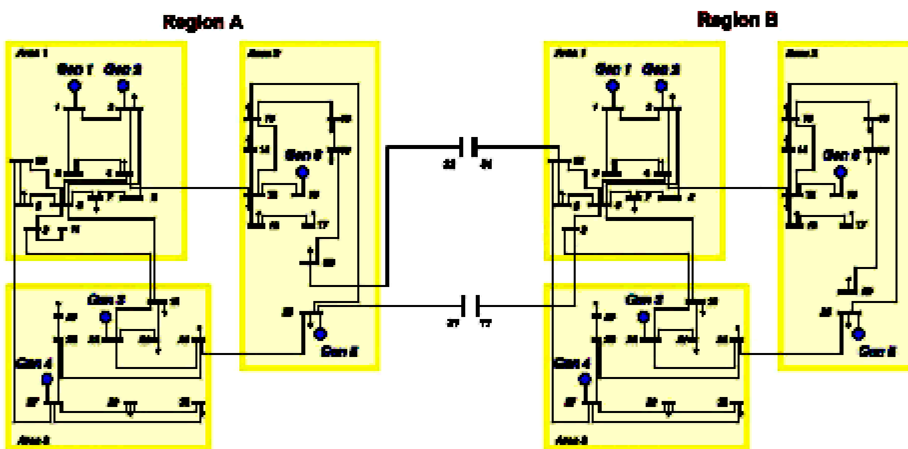


Figure A1. System Model with Two AC Tie-lines connecting the Two Regions

As shown in Figure A2, the marginal units, and therefore the busses experiencing increases or decreases in injections, changes when transfers across the tie-line exceed a 20MW increment or a 30MW decrement in flow in Region A. Since the combined tie-lines have a 100MW capacity, the alterations in dispatch must be caused in part by within region transmission congestion; although in the case of seller 4, its generation capacity limit of 60MW also alters the dispatch mix at transfers out of region A in excess of 20MW. Nevertheless, this simple exercise illustrates the difficulty in identifying a proxy bus that is sufficiently robust to represent a realistic range of transfers. And since suitable busses could not be identified for these two thirty bus networks, no experiments were conducted with simulated multiple AC tie-lines.

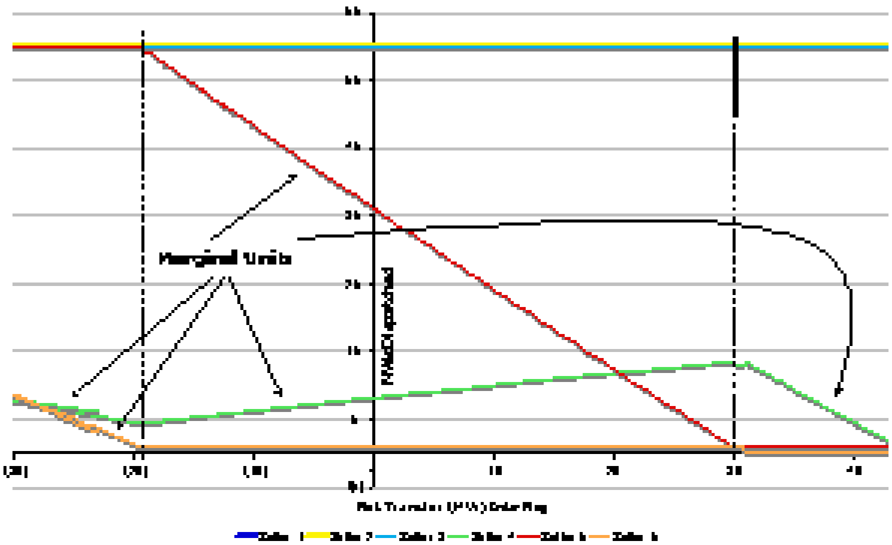


Figure A2. Identifying Proxy Busses : Changes in Generator Dispatch In Region A with Transfer Level across the Tie-line

Appendix B Statistical Analyses of Spatial Arbitrage Experimental Results

Table B1: Consumer Surplus Regressions (Dependent Variable: Price per MW)

	Both Regions		Region A		Region B	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Dependent Variable: ISO Cost/MW per Period						
Demand Effects by Group						
Normal (Group 1: Treatments 1 to 4)	49.70	1.86	37.97	2.62	60.34	3.14
High (Group 1: Treatments 1 to 4)	49.70	1.86	39.79	2.62	57.82	3.14
Normal (Group 2: Treatments 1 to 4)	42.66	1.86	38.84	2.62	46.17	3.14
High (Group 2: Treatments 1 to 4)	45.12	1.86	42.91	2.62	46.76	3.14
Normal (Group 3: Treatments 5 and 6)	49.79	2.08	43.27	2.93	55.82	3.51
High (Group 3: Treatments 5 and 6)	55.73	2.08	43.78	2.93	65.63	3.51
Normal (Group 4: Treatments 5 and 6)	39.46	2.08	35.34	2.93	43.31	3.51
High (Group 4: Treatments 5 and 6)	42.20	2.08	37.11	2.93	46.33	3.51
Treatment Effects						
Treatment 1	35.51	2.08	34.73	2.93	36.32	3.51
Treatment 2	22.74	2.08	16.01	2.93	28.83	3.51
Treatment 3	46.17	2.08	45.71	2.93	46.74	3.51
Treatment 4	43.42	2.08	38.58	2.93	47.86	3.51
Treatment 5	13.13	2.08	6.62	2.93	18.83	3.51
Treatment 6	8.50	2.08	2.36	2.93	13.86	3.51
R-squared	0.99		0.97		0.97	
Obs	256		256		256	

P-Value of Chi-Squared Test that Treatment Effects Are Equal (Both Regions)

	Soc. Opt.	Treat 1	Treat 2	Treat 3	Treat 4	Treat 5	Treat 6
Soc. Opt.	NA						
Treat 1	0.000	NA					
Treat 2	0.000	0.000	NA				
Treat 3	0.000	0.000	0.000	NA			
Treat 4	0.000	0.000	0.000	0.189	NA		
Treat 5	0.000	0.000	0.001	0.000	0.000	NA	
Treat 6	0.000	0.000	0.000	0.000	0.000	0.027	NA

P-Value of Chi-Squared Test that Treatment Effects Are Equal (Region A)

	Soc. Opt.	Treat 1	Treat 2	Treat 3	Treat 4	Treat 5	Treat 6
Soc. Opt.	NA						
Treat 1	0.000	NA					
Treat 2	0.000	0.000	NA				
Treat 3	0.000	0.000	0.000	NA			
Treat 4	0.000	0.190	0.000	0.016	NA		
Treat 5	0.025	0.000	0.024	0.000	0.000	NA	
Treat 6	0.421	0.000	0.001	0.000	0.000	0.146	NA

P-Value of Chi-Squared Test that Treatment Effects Are Equal (Region B)

	Soc. Opt.	Treat 1	Treat 2	Treat 3	Treat 4	Treat 5	Treat 6
Soc. Opt.	NA						
Treat 1	0.000	NA					
Treat 2	0.000	0.034	NA				
Treat 3	0.000	0.003	0.000	NA			
Treat 4	0.000	0.001	0.000	0.749	NA		
Treat 5	0.000	0.001	0.045	0.000	0.000	NA	
Treat 6	0.000	0.000	0.003	0.000	0.000	0.158	NA

Table B2: Producer Surplus Regressions (Dependent Variable: Generator Profit per Period)

	Both Regions		Region A		Region B	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Dependent Variable: Generator Profit per Period						
Demand Effects by Group						
Normal (Group 1: Treatments 1 to 4)	2,399	709	449	505	1,951	688
High (Group 1: Treatments 1 to 4)	4,680	709	1,628	505	3,052	688
Normal (Group 2: Treatments 1 to 4)	(894)	709	176	505	(1,070)	688
High (Group 2: Treatments 1 to 4)	1,996	709	1,821	505	174	688
Normal (Group 3: Treatments 5 and 6)	2,616	793	1,448	564	1,168	769
High (Group 3: Treatments 5 and 6)	6,133	793	2,031	564	4,102	769
Normal (Group 4: Treatments 5 and 6)	(950)	793	(13)	564	(937)	769
High (Group 4: Treatments 5 and 6)	382	793	607	564	(226)	769
Treatment Effects						
Treatment 1	13,191	793	6,025	564	7,166	769
Treatment 2	8,939	793	3,658	564	5,281	769
Treatment 3	17,253	793	9,313	564	7,940	769
Treatment 4	16,114	793	7,704	564	8,409	769
Treatment 5	4,213	793	749	564	3,464	769
Treatment 6	3,776	793	1,090	564	2,686	769
R-squared	0.94		0.87		0.82	
Obs	256		256		256	

P-Value of Chi-Squared Test that Treatment Effects Are Equal (Both Regions)

	Soc. Opt.	Treat 1	Treat 2	Treat 3	Treat 4	Treat 5	Treat 6
Soc. Opt.	NA						
Treat 1	0.000	NA					
Treat 2	0.000	0.000	NA				
Treat 3	0.000	0.000	0.000	NA			
Treat 4	0.000	0.000	0.000	0.152	NA		
Treat 5	0.000	0.000	0.000	0.000	0.000	NA	
Treat 6	0.000	0.000	0.000	0.000	0.000	0.582	NA

P-Value of Chi-Squared Test that Treatment Effects Are Equal (Region A)

	Soc. Opt.	Treat 1	Treat 2	Treat 3	Treat 4	Treat 5	Treat 6
Soc. Opt.	NA						
Treat 1	0.000	NA					
Treat 2	0.000	0.000	NA				
Treat 3	0.000	0.000	0.000	NA			
Treat 4	0.000	0.003	0.000	0.005	NA		
Treat 5	0.186	0.000	0.000	0.000	0.000	NA	
Treat 6	0.054	0.000	0.002	0.000	0.000	0.546	NA

P-Value of Chi-Squared Test that Treatment Effects Are Equal (Region B)

	Soc. Opt.	Treat 1	Treat 2	Treat 3	Treat 4	Treat 5	Treat 6
Soc. Opt.	NA						
Treat 1	0.000	NA					
Treat 2	0.000	0.015	NA				
Treat 3	0.000	0.315	0.001	NA			
Treat 4	0.000	0.107	0.000	0.542	NA		
Treat 5	0.000	0.001	0.096	0.000	0.000	NA	
Treat 6	0.001	0.000	0.018	0.000	0.000	0.313	NA

Table B3: Producer + Arbitrager Surplus Regressions (Dependent Variable: Profits from Power Sales and Arbitraging per Period) and Analysis of Learning Effects

	Both Regions		Both Regions	
	Coef.	Std. Err.	Coef.	Std. Err.
Dependent Variable: Generator Profit + Arbitrage Profit per Period				
Demand Effects by Group				
Normal (Group 1: Treatments 1 to 4)	2,742	781	2,702	785
High (Group 1: Treatments 1 to 4)	4,718	781	4,757	785
Normal (Group 2: Treatments 1 to 4)	(901)	781	(941)	785
High (Group 2: Treatments 1 to 4)	1,623	781	1,662	785
Normal (Group 3: Treatments 5 and 6)	2,703	874	2,670	876
High (Group 3: Treatments 5 and 6)	5,989	874	6,022	876
Normal (Group 4: Treatments 5 and 6)	(912)	874	(945)	876
High (Group 4: Treatments 5 and 6)	402	874	434	876
Treatment Effects				
Treatment 1	13,191	874	13,493	999
Treatment 2	9,043	874	9,346	999
Treatment 3	17,253	874	17,555	999
Treatment 4	16,091	874	16,393	999
Treatment 5	4,213	874	4,516	999
Treatment 6	2,530	874	2,833	999
Period x (1- Social Opt. Dummy)	NA	NA	(36)	57
R-squared		0.93		0.93
Obs		256		256

P-Value of Chi-Squared Test that Treatment Effects Are Equal (Both Regions, Excluding Time Trend)

	Soc. Opt.	Treat 1	Treat 2	Treat 3	Treat 4	Treat 5	Treat 6
Soc. Opt.	NA						
Treat 1	0.000	NA					
Treat 2	0.000	0.000	NA				
Treat 3	0.000	0.000	0.000	NA			
Treat 4	0.000	0.001	0.000	0.185	NA		
Treat 5	0.000	0.000	0.000	0.000	0.000	NA	
Treat 6	0.000	0.000	0.000	0.000	0.000	0.055	NA

Table B-4: Total Surplus Regressions (Dependent Variable: Cost of Power Production per Period)

	Both Regions	
	Coef.	Std. Err.
Dependent Variable: Cost of Power Production per Period		
Demand Effects by Group		
Normal (Group 1: Treatments 1 to 4)	15,028	185
High (Group 1: Treatments 1 to 4)	18,384	185
Normal (Group 2: Treatments 1 to 4)	15,967	185
High (Group 2: Treatments 1 to 4)	19,301	185
Normal (Group 3: Treatments 5 and 6)	15,698	206
High (Group 3: Treatments 5 and 6)	19,371	206
Normal (Group 4: Treatments 5 and 6)	15,321	206
High (Group 4: Treatments 5 and 6)	18,289	206
Treatment Effects		
Treatment 1	1,664	206
Treatment 2	2,449	206
Treatment 3	1,968	206
Treatment 4	2,218	206
Treatment 5	1,363	206
Treatment 6	2,029	206
R-squared		0.998
Obs		256

P-Value of Chi-Squared Test that Treatment Effects Are Equal (Both Regions)

	Soc. Opt.	Treat 1	Treat 2	Treat 3	Treat 4	Treat 5	Treat 6
Soc. Opt.	NA						
Treat 1	0.000	NA					
Treat 2	0.000	0.000	NA				
Treat 3	0.000	0.142	0.021	NA			
Treat 4	0.000	0.008	0.264	0.228	NA		
Treat 5	0.000	0.303	0.000	0.039	0.004	NA	
Treat 6	0.000	0.212	0.152	0.835	0.519	0.001	NA

Project Publications

Adilov, N., and Schuler, R., 2005, “Optimal Market Grain Over Space and Time”, *Rutgers Advanced Workshop in Regulation and Competition: 24th Annual Eastern Conference*, Skytop, PA.

Adilov, N. and Schuler, R., 2006, “Electricity Markets: How Many, Where and When?”, *Symposium on Electric Power Systems Reliability, Control and Markets*, Hawaii International Conference on Systems Science 39, Kauai, HI.

Schuler, R., 2006, “Optimal Market Grain Over Space and Time”, *North American Regional Science Association Meetings*, Toronto, Canada.

Adilov, N., Light, T., Mitarotonda, D., Mount, T., Schuler, R., Schulze, W., & Zimmerman, R., 2008, “Spatial Arbitrage to Reduce "Seams" Across Electricity System Control Areas: An Experimental Analysis”, *Proceedings of the Hawaii International Conference on systems Science 41,* Waikaloa, HI.

References

[1] Schuler, R., 2006, “Float Together/ Sink Together? (The Effect of Connectivity on the Power System)”, *The Economic Impact of Terrorist Attacks,* edited by Peter Gordon, Harry Richardson & James Moore II, Edward Elgar: Cheltenham, UK, 91-118.

[2] Schuler, R., 1986, "The Institutional and Regulatory Structure for Providing Electric Service: A Conceptual Basis for Change", Chapt. 25 in, Saltzman, S., and Schuler, R., eds., The Future of Electrical Energy: A Regional Perspective of an Industry in Transition, New York: Praeger.

[3] Schuler, R., 2009, “Summary of the Executive Forum on Planning, Markets and Investment in the Electric Supply Industry”, PSERC Report 09-01.

[4] Mount, T., Schuler, R. and Schulze, W., 2003, “Markets for Reliability and Financial Options in Electricity: Theory to Support the Practice”, *Proceedings of Hawaii International Conference on Systems Science 36*, Waikaloa, Hawaii.

[5] Oren, S., 2001, “Design of Ancillary Service Markets,” *Proceedings of Hawaii International Conference on System Sciences 34*, Maui, HI.

[6] Holahan, W. and Schuler, R., 1988, "Imperfect Competition in a Spatial Economy: Pricing Policies and Economic Welfare", CORE Discussion Paper 8821, Universite Catholique de Louvain, Louvain-la-Neuve, Belgium.

[7] Schuler, R. and Holahan, W., 1978, “Competition vs. Vertical Integration of Transportation and Production in a Spatial Economy” *Papers of the Regional Science Association, 41*, 209-25.

[8] Schuler, R., and Hobbs, B., 1982, “Spatial Price Duopoly under Uniform Delivered Pricing”, *The Journal of Industrial Economics, Symposium on Spatial Competition and the Theory of Differentiated Markets, 31*, 1/2, September/December, 175-187.

[9] Hadsell, L., 2007, "The Impact of Virtual Bidding on Price Volatility in New York's Wholesale Electricity Market." *Economics Letters 95*, 1, April, 66-72.

[10] Allaz, B. and Vila, J.-L., 1993, “Cournot Competition, Futures Markets and Efficiency,” *Journal of Economic Theory 59*, 1-16.

[11] Green, R., 1999, “The Electricity Contract Market in England and Wales,” *Journal of Industrial Economics 47*, 107-124.

[12] Ferreira, J., 2001, “The Role of Observability in Futures Markets,” Departamento de Economia, *Universidad Carlos III de Madrid* working paper.

[13] Lien, S., 2000, “Forward Contracts and the Curse of Market Power,” Department of Economics, *University of Maryland* working paper.

[14] Le Coq, C. and Orzen, H., 2002, “Do Forward Markets Enhance Competition? Experimental Evidence,” Department of Economics, *University of Stockholm*, Series in Economics and Finance working paper 506.

[15] Newbery, M., 1995, “Competition, Contracts and Entry in the Electricity Spot Market,” *RAND Journal of Economics 29*, 729-749.

[16] Adilov, N., 2005, *Essays on Market Design and Strategic Interaction*, Ph.D. Dissertation, Cornell University, Ithaca, NY.

[17] Kreps, D. and Scheinkman, J., 1983, “Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes,” *Bell Journal of Economics 14*, 326-337.

[18] FERC, 2006 State of the Markets Report, available at:
http://www.ferc.gov/market-oversight/st-mkt-ovr/som-rpt-2006.pdf

[19] FERC, ISO/RTO Handbook, August 14, 2006 available at:
http://www.ferc.gov/industries/electric/indus-act/rto/handbook.asp

[20] Bernard, J., Schulze, W., Mount, T., Zimmerman, R., Thomas, R., and Schuler, R., 1998, “Alternative Auction Institutions for Purchasing Electric Power: An Experimental Examination”, *Proceedings of Bulk Power Systems Dynamics and Control IV – Restructuring*, Santorini, Greece.

[21] International Energy Agency website available at:
http://www.iea.org/